HEALTH INFORMATICS MEETS eHEALTH

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 223

Recently published in this series

ISSN 0926-9630 (print)
ISSN 1879-8365 (online)

Health Informatics Meets eHealth

Predictive Modeling in Healthcare – From Prediction to Prevention

Proceedings of the 10th eHealth2016 Conference

Edited by

Günter Schreier

AIT Austrian Institute of Technology GmbH, Graz, Austria

Elske Ammenwerth

UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria

Alexander Hörbst

UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria

and

Dieter Hayn

AIT Austrian Institute of Technology GmbH, Graz, Austria

IOS Press

Amsterdam • Berlin • Washington, DC

ISBN 978-1-61499-644-6 (print)
ISBN 978-1-61499-645-3 (online)
Library of Congress Control Number: 2016938025

Publisher
IOS Press BV
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

Distributor in the USA and Canada
IOS Press, Inc.
4502 Rachael Manor Drive
Fairfax, VA 22032
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

Preface

Predictive modeling in health care – from prediction to prevention

For a number of decades, progress in medicine has relied on classical research pathways, which at some point involve the conducting of clinical trials. Ideally, a number of randomized clinical trials (RCTs) can be conducted and referred to in order to establish a high level of evidence for a given therapeutic intervention, be it a drug to be prescribed to the patient, a medical device to be used, or a particular health care process to be set-up.

RCTs are doubtless a very sound methodology for establishing both efficacy and safety, as well as cause-effect relationship, in an unambiguous way, but this process also has its limitations. Firstly, RCTs are expensive and lengthy procedures. At present, it seems that the number of potential innovative treatment options is developing much faster than can be assessed by a classical RCT approach in a reasonable amount of time. This is particularly the case in areas such as rare diseases. Secondly, RCTs usually focus on a well characterized group of patients by means of proper inclusion and exclusion criteria. This can, however, significantly reduce the generalizability of the results to other groups such as the elderly, children, or patients with several co-morbidities. Thirdly, and most importantly, the trend towards precise and personalized medicine aims at the optimal approach for an individual patient, which presents a particular challenge for the classical approach to establishing evidence.

This is where predictive modeling comes in to play. Today, partly due to the increasingly digital nature of healthcare, vast amounts of data are available which can be used as a basis for profiling patient characteristics, to find out correlations between outcomes and potential predictors, and to predict future developments based on a plethora of existing cases.

In order to approach the concept of a truly personalized and preventive care plan tailored to the particular and complex needs of the individual patient, we need to include various sources of information, and predict the future course based on patient-specific models, as illustrated in Fig. 1. Ideally, such a model would not just tell us

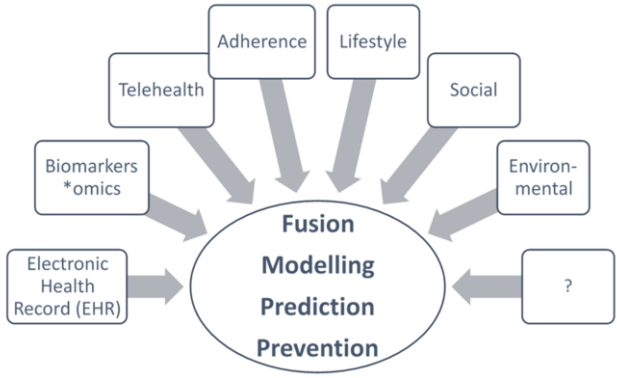


Figure 1. Concept for prevention unwanted events, based on diverse sources of information.

what could potentially happen to the health status of the individual, but also which options could best influence and modulate the person's situation so as to prevent unwanted events.

The question mark on the right hand side of Fig. 1 indicates that this is not an exhaustive list. There are many more potential sources of valuable information about an individual patient and his/her context which – taken into account – can help to tailor a healthcare strategy to the patient's particular needs.

Of course, correlations cannot be used to directly establish cause-effect relations, but what finally counts is whether a data driven approach works better than, or at least as well as, the state-of-the-art approach. And machine learning approaches usually improve when the amount of available data increases. This reflects the way in which humans become more experienced the more cases they have seen and managed. When it comes to "Big Data" environments however, the human mind faces limits in scalability which predictive modeling tools can potentially help to overcome.

Today, most predictive modeling approaches target non-medical goals in the fields of logistics and resource allocation, not least because predictive modeling concepts with a direct impact on patient treatment face huge regulatory hurdles. This seems to be a major factor in explaining why, so far, most of the data-driven predictive approaches directly related to patient care remain at a research level and in a retrospective setting. Only a few have made it into daily routine – but hasn't this initially been the case with almost all healthcare innovations? Nevertheless, we thought that now was the right time to make this topic "Predictive modeling in health care – from prediction to prevention" our theme for the 2016 issue of the annual scientific conference on "Health Informatics meets eHealth" in Vienna. It is time to give these new possibilities some additional visibility in the framework of the realm of eHealth as a whole.

Günter Schreier
Elske Ammenwerth
Alexander Hörbst
Dieter Hayn

March 19, 2016
Graz, Hall in Tirol, and Vienna

Scientific Program Committee

Univ.-Prof. Elske Ammenwerth, UMIT – University for Health Sciences, Medical Informatics and Technology
Univ.-Doz. Dr. Günter Schreier, MSc, AIT Austrian Institute of Technology GmbH

Univ.-Prof. Dr. Klaus-Peter Adlassnig, Medizinische Universität Wien
Prof. Dr. Josef Altmann, FH Oberösterreich, Wels
Univ.-Prof. Dr. Andrea Berghold, Medizinische Universität Graz
Prof. Dr. Britta Böckmann, Fachhochschule Dortmund
Prof. Dr. Oliver Bott, FH Hannover
Univ.-Prof. Dr. Ruth Breu, Universität Innsbruck
Prof. Dr. Harald Burgsteiner, FH Joanneum, Graz
Univ.-Prof. DDr. Wolfgang Dorda, Medizinische Universität Wien
Prof. Dr. Stephan Dreiseitl, FH Hagenberg
Univ.-Prof. Dr. Georg Duftschmid, Medizinische Universität Wien
Univ.-Prof. Dr. Martin Dugas, Universität Münster
Univ.-Prof. Dr. Walter Gall, Medizinische Universität Wien
Dr. Holger Gothe, UMIT, Hall in Tirol
Prof. Dr. Martin Haag, Hochschule Heilbronn
Prof. Dr. Peter Haas, FH Dortmund
Prof. Dr. Anke Häber, FH Zwickau
Dr. Dieter Hayn, AIT, Graz
Dr. Mira Hercigonja-Szekeres, University of Applied Sciences, Zagreb, Croatia
Univ.-Prof. Dr. Alexander Hörbst, UMIT, Hall in Tirol
Prof. Dr. Ursula Hübner, Hochschule Osnabrück
Prof. Dr. Guido Kempter, FH Vorarlberg
Prof. Dr. Peter Klutke, Hochschule Kempten
Prof. Dr. Petra Knaup, Universität Heidelberg
Prof. Dr. Werner Kurschl, FH Hagenberg
Univ.-Prof. Dr. Richard Lenz, Universität Erlangen-Nürnberg
Prof. Dr. Nicos Maglaveras, Aristotle University, Thessaloniki, Greece
Prof. Dr. Michael Marschollek, Medizinische Hochschule Hannover
Prof. Dr. Christian Menard, FH Kärnten
Prof. Dr. George Mihalas, University of Medicine and Pharmacy, Timisoara, Romania
Ass.-Prof. Dr. Ivana Ognjanovic, Donja Gorica University, Podgorica, Montenegro
Univ.-Prof. Dr. Barbara Paech, Universität Heidelberg
Univ.-Prof. Dr. Karl-Peter Pfeiffer, FH Joanneum, Graz
Univ.-Prof. Dr. Ulrich Prokosch, Universität Erlangen
Dr. Stefan Sauermann, FH Technikum Wien
Prof. Dr. Paul Schmücker, Hochschule Mannheim
Dr. Michael Shifrin, N.N. Burdenko Neurosurgical Institute, Moscow, Russia
Prof. Dr. Martin Stämmler, FH Stralsund
Prof. Dr. Christian Stingl, FH Kärnten
Prof. Dr. Lăcrămioara Stoicu-Tivadar, Polytechnic University of Timisoara, Romania
Prof. Dr. Jana Zvarova, Charles University, 1st Faculty of Medicine, Prague, Czech Republic

Reviewers

We would like to thank all reviewers for their significant contribution to the proceedings of the eHealth2016:

Adlassnig, Klaus-Peter
Aerts, Jozef
Altmann, Udo
Altmann, Josef
Ammenwerth, Elske
Beck, Peter
Böckmann, Britta
Bott, Oliver
Burgsteiner, Harald
Dreiseitl, Stephan
Duftschmidt, Georg
Dugas, Martin
Fortwengel, Gerhard
Gall, Walter
Ganslandt, Thomas
Garde, Sebastian
Göbel, Georg
Gothe, Holger
Haag, Martin
Haas, Peter
Häber, Anke
Hackl, Werner
Hayn, Dieter
Helm, Emmanuel
Heß, Michael
Holm, Jürgen
Hörbst, Alexander
Hübner, Ursula
Kastner, Peter
Kempter, Guido
Klutke, Peter
Kollmann, Alexander
Krefting, Dagmar
Kriegel, Johannes
Kurschl, Werner
Marschollek, Michael
Mense, Alexander
Miksch, Silvia
Mischak, Robert
Müller-Mielitz, Stefan
Neuhaus, Philipp
Neururer, Sabine
Oberbichler, Stefan
Paech, Barbara
Panek, Paul
Rinner, Christof
Sabutsch Stefan
Sauermann, Stefan
Schlieter, Hannes
Schreier, Günter
Sedlmayr, Martin
Skonetzki, Stefan
Spreckelsen, Cord
Stämmler, Martin
Stingl, Christian
Umbach, Nadine
Werner, Franz
Zagler, Wolfgang

Contents

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-1

Cardiovascular Dysautonomias Diagnosis Using Crisp and Fuzzy Decision Tree: A Comparative Study

Ilham KADI[a,1] and Ali IDRI[a,1]
[a]*Software Project Management Research Team, ENSIAS, Mohammed V University in Rabat, Morocco*

Abstract. Decision trees (DTs) are one of the most popular techniques for learning classification systems, especially when it comes to learning from discrete examples. In real world, many data occurred in a fuzzy form. Hence a DT must be able to deal with such fuzzy data. In fact, integrating fuzzy logic when dealing with imprecise and uncertain data allows reducing uncertainty and providing the ability to model fine knowledge details. In this paper, a fuzzy decision tree (FDT) algorithm was applied on a dataset extracted from the ANS (Autonomic Nervous System) unit of the Moroccan university hospital Avicenne. This unit is specialized on performing several dynamic tests to diagnose patients with autonomic disorder and suggest them the appropriate treatment. A set of fuzzy classifiers were generated using FID 3.4. The error rates of the generated FDTs were calculated to measure their performances. Moreover, a comparison between the error rates obtained using crisp and FDTs was carried out and has proved that the results of FDTs were better than those obtained using crisp DTs.

Keywords. Cardiovascular dysautonomias, autonomic nervous system, fuzzy logic, fuzzy decision tree, C4.5 algorithm.

1. Introduction

Autonomic nervous system is the part of the nervous system that controls involuntary actions, such as the beating of the heart and the widening or narrowing of the blood vessels. It also controls body temperature, digestion, metabolism, the production of body fluids, and other processes [13]. However, the ANS is frequently subject to malfunctions that are called dysautonomias. This disorder can cause serious problems, including: Blood pressure problems, Heart problems, Trouble with breathing and swallowing and others. Doctors can check for signs of dysautonomias during the physical examination. They measure blood pressure (BP) and heart rate (HR) while a person is lying down or sitting and after the person stands.

Data Mining (DM) is a set of techniques designed to explore large datasets in order to discover hidden and previously unknown patterns, relationships and knowledge. DM can therefore be considered as the kernel of a knowledge discovery process from data [7]. Classification is one of the main tasks of DM. In fact, classification techniques are able of predicting categorical class labels and classify data based on a training set [4]. DTs are considered as one of the popular classification techniques. The goal of DTs is creating a model that predicts the value of a target variable by learning simple decision

[1] Corresponding Authors: Software Project Management Research Team, ENSIAS, Mohammed V University in Rabat, Morocco, Email: ilham.kadi@um5s.net.ma, ali.idri@um5.ac.com.

rules inferred from data. There are many algorithms to construct DTs such as ID3 [16], C4.5 [17], and CART [3]. A FDT is a tree structure where every edge is annotated with a condition, and every leaf is annotated with a fuzzy set. It is a generalization of a crisp DT to handle attributes either with numerical or linguistic values.

In our previous work, we performed a case study in the ANS unit of the hospital Avicenne. In this unit, a set of ANS tests is performed to diagnose patients with cardiovascular dysautonomias and provide them the appropriate treatment. These tests include: deep breathing test, hand grip test, mental stress test and orthostatic test. For each test, the specialists analyze deeply the obtained results of HR and BP and provide an interpretation for each result. These interpretations are called preliminary conclusions. Thereby, a decision support system (DSS) was developed to automate the generation of the preliminary conclusion and make it easier for specialists [8, 11]. The developed model was constructed using C4.5 algorithm. However, since our case study is directly related to human lives, we were motivated to extend our research and apply FDTs in order to understand and increase the accuracy rates of decision support systems. Indeed, unlike crisp DTs, FDTs deal with numerical values by transforming them into linguistic ones. According to Zadeh, the use of linguistic values instead of (or in addition to) numbers serves many purposes [21]: (1) They are easier to understand than numerical values; (2) They make allowance for imprecision; (3) They generalize numbers (only when precise information is available) and (4) They accept the finite ability of the human mind to resolve detail and store precise information. For that, we apply in this study FDT techniques to the same dataset used in our previous work. Thereby, a comparison between the results obtained using crisp DTs and FDTs is provided and discussed.

This paper is organized as follow: Section 2 presents an overview of the existing studies in literature applying crisp DTs and FDTs in cardiology. Section 3 provides details about the techniques and the experiment design used in this study. Section 4 presents and discusses the results obtained. Finally, conclusion and future work are presented in Section 5.

2. Related work

DTs are known as one of the most popular classification techniques in medical DM [4]. As a result, data miners have used DM algorithms in different disciplines of medicine including cardiology. To the best of our knowledge, there is no work carried out in applying DTs in an ANS domain. Thus, we have performed a case study in this context using C4.5 DT algorithm [8, 11]. Therefore, we developed a classification model using a dataset collected from the ANS unit of the hospital Avicenne. The classifier obtained presented a high level of accuracy up to 98.5%. On the other hand, since ANS is related to cardiovascular system, a summary of some studies conducted in cardiology using DTs and FDTs is presented in this section. Pecchia et al. proposed a platform to enhance effectiveness and efficiency of home monitoring using CART classifier for early detection of any worsening in patient's condition [15]. The developed system achieved accuracy of 96.39% in detecting heart failure. Karaolis et al. developed a DT based on C4.5 for the assessment of Coronary Heart Disease (CHD) related risk factors targeting in the reduction of CHD events. The DT was applied on a dataset collected from a hospital including 528 cases and has reached an accuracy rate of 75% [12].

On the other hand, FDTs were also getting more interest by researchers in this field. Bohacik and Kambhampati presented a method based on a FDT for prediction of the

death of a patient with heart failures. The results showed that this method is a useful technique by reaching a sensitivity of 67.3% and a specificity of 62.6% [2]. Behadada and Chikh implemented a classifier based on a FDT to extract decision rules and classify some cardiac abnormalities. The best classifier in the experiments achieved an accuracy of 71%, and a sensitivity of 100% [1]. Overall, the results obtained by studies applying DTs and FDTs algorithms in cardiology were satisfactory and encouraging.

3. Materials and Methods

In this section, brief presentations of FDTs, experiment design and the medical dataset used in this study are introduced.

Fuzzy decision tree: FDTs combine the DT paradigm with the fuzzy sets theory. They are a generalization of crisp DTs to handle imprecise and uncertain attributes with numeric-linguistic values. FDTs differ from crisp DTs by [9]: 1) They use splitting criteria based on fuzzy restrictions; 2) Their inference procedures are different; and 3) The fuzzy sets representing the data have to be defined. A FDT induction has two major components: a procedure for FDT building and an inference procedure for decision making [14]. In fact, the incorporation of fuzzy logic in building DTs requires a fuzzy partitioning for fuzzy subsets of each input variable. In the FDTs, each node is associated with a variable, and each branch is associated with a fuzzy subset of this variable. Therefore, every path leading to a leaf of the tree will match with a fuzzy rule.

Medical dataset description: The dataset used in this study is the same one adopted in our previous research to provide a comparison between the results obtained by FDTs and crisp DTs. Thereby, a total of 178 records were collected from the ANS unit. This dataset includes records of patients suffering from cardiovascular dysautonomias who went to ANS unit in the period between January 2013 and May 2014. 11 attributes were identified to be required for the generation of the preliminary conclusion and were selected to perform our experiment, including: age, VR_DB, VR_HG, PSR α, CSR α, CSR β, VR_Ort, HRmin, HRmax, BPmin, BPmax. Table 1 provides a brief description and details about input attributes for each ANS test. As can be noticed, the age attribute is considered as the main factor that influences the results interpretation obtained in each ANS test.

Experiment design*:* In the ANS unit, a patient's diagnosis is based on several dynamic tests. In each test, specific metrics are calculated by means of measuring continuously HR and BP values and a set of mathematical equations. According to the tests results, a set of preliminary conclusions is deduced for each ANS test. These conclusions are analyzed deeply by the specialists to provide a global synthesis and diagnosis of the patient's state. However, the measurement and analysis of the test results are done manually by the specialists which makes it more difficult for the specialists. In this study, the same input data identified in our previous paper were used to generate the FDTs [8, 11]. Thus, for each ANS test, one or two or three FDTs were generated to automate the obtaining of preliminary conclusions. For each FDT, three classes were identified based on the expert's guidelines namely: low, normal and high. As a result, eight FDTs were generated and tested.

Table 1. Description and details about input attributes for each ANS tests

ANS tests	Measured values	Description	Input attributes
Deep Breathing	Vagal response (VR_DB)	Vagal response measured using HR values in Deep Breathing test (DB)	Age VR
Hand Grip	Vagal response (VR_HG)	Vagal response measured using HR values in Hand Grip test (HG)	Age VR
	PSR α	Peripheral sympathetic response α measured using BP values in HG test	Age PSR
Mental stress	CSR α	Central sympathetic response α measured using BP values in Mental Stress test (MS)	Age CSR α
	CSR β	Central sympathetic response β measured using BP values in MS test	Age CSR β
Orthostatic test	Vagal response (VR_Ort)	Vagal response measured using HR values in Orthostatic test (Ort)	Age VR
	SP_HR	heart rate measured in Orthostatic test using supine position (SP)	Age HRmin HRmax
	SP_BP	blood pressure measured in Orthostatic test using supine position	Age BPmin BPmax

In order to build FDTs, we adopt the experiment design presented in Figure 1. This process includes several major phases namely:

- Data pre-processing: It is a critical step which deals with the preparation and transformation of the initial data [6]. Since the dataset adopted in this study is the same as that one used in our previous work, the pre-processing phase was already done and presented in [8, 11]. In fact, a data-cleaning process was performed to fill the missing values that did not exceed 4% of the whole set. Besides, real numbers of the dataset were all transformed to integer values to simplify the generation of crisp DTs.
- Fuzzification process: It aims at transforming numeric values to linguistic ones [20]. For this, membership functions must be built to define the degree of membership of a numeric value to fuzzy sets of linguistic variables. In fact, for each linguistic value, a membership function is defined [20]. The attributes are then described by fuzzy sets and the tree arcs are associated with these fuzzy sets. Moreover, the choice of shapes of membership functions may be subjective. Trapezoidal and triangular shapes for membership functions are generally used.

FID fuzzy decision tree: It is a classification system, implementing an efficient recursive partitioning technique of DTs, while combining fuzzy representation and approximate reasoning for dealing with noise and language uncertainty [10]. FID has three major components related to partitioning continuous attributes, building an

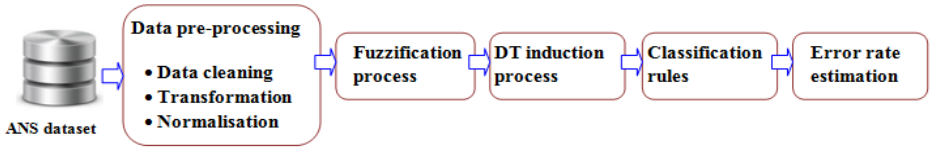


Figure 1. Experiment design process

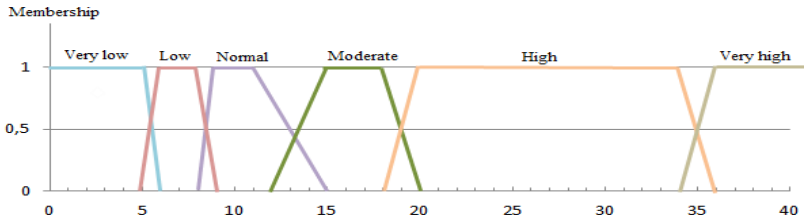


Figure 2. Fuzzy sets defined for VR_HG attribute

explicit tree, and inducing knowledge from the tree [5]. Furthermore, FID includes a pruning algorithm to avoid over-fitting, which balances a tradeoff between the size of the tree and its predictive accuracy.

- Classification rules: With a fuzzy context, an example may belong to several sub-nodes with different degrees of membership [9]. FDTs can be interpreted by a set of fuzzy rules. Each path from the root to the leaf can be converted to a rule where the condition part represents the attributes and the conclusion part represents the classes. The weights of the conclusion of each rule is calculated by means of the membership degrees using the AND operator.

4. Results and Discussion

In order to generate the FDTs, we used the FID 3.4 version which is one of the FID programs [19]. It uses three input files namely: attribute, data, and parameter files. The attribute file contains information about the attributes, definitions of partitioning sets, the corresponding names and the definition of the decision classes. A data file includes information about training examples: the values of each attribute (numeric, linguistic, or missing), the decision value (numeric or linguistic), and weight value (weight of the event). A parameter file contains default values for all of the configurable options. Thereby, these files were generated and were changed in each trial. In fact, in order to measure the efficiency of the generated FDTs, the data set was randomly partitioned into training and independent test sets by means of a 10-fold cross-validation process. Thus, the data set was split into ten equal sized blocks with similar class distributions [18]. Moreover, the error rate metric was used to evaluate the performance of the generated classifiers [4]. Figure 2 presents the membership functions associated to the VR_HG attribute. As can be seen, six fuzzy sets were defined for this attribute. These fuzzy sets were identified based on the empirical knowledge of ANS experts since the data used in this study are very critical and each system error may risk a human life. For this reason, we preferred in this work to rely on the specialists expertise when identifying the fuzzy sets rather than using fuzzy algorithms such as fuzzy C-means.

Table 2 presents the results obtained when measuring the mean error rate and standard deviation for each generated crisp DT and FDT using the training and test sets and 10-fold cross-validation method. The error rate was computed by FID3.4 depending on the degrees of membership. In fact, an example is considered to be incorrectly classified when the “Delta:Best” feature (which contains the decision class having the highest membership) is different from “Class” feature (which contains the actual classwhere the error rates obtained by FDTs were better than those obtained by crisp DTs. Thereby, we notice a reduction in error rates obtained using FDT techniques in

Table 2. Comparison of mean error rate obtained by crisp DTs and FDTs on training and test sets

ANS tests	Phase	Learning phase				Test phase			
		Mean error rate		Std deviation		Mean error rate		Std deviation	
		Crisp DT	FDT	Crisp DT	FDT	Crisp DT	FDT	Crisp DT	FDT
Deep Breathing	Vagal response	3.05%	**2.69%**	1.97	0.89	2.43%	**1.75%**	2.33	1.62
Hand Grip	Vagal response	3.72%	**2.99%**	1.53	1.13	2.98%	**2.07%**	1.18	1.05
	PSR α	1.11%	**1.02%**	0.89	1.22	3.21%	**2.72%**	2.19	1.79
Mental stress	CSR α	1.02%	1.13%	0.76	0.91	0.93%	1.31%	0.88	0.84
	CSR β	0.34%	**0.11%**	1.88	0.83	1.99%	**1.08%**	1.21	1.58
Orthostatic	Vagal response	0.27%	0.41%	0.91	0.97	1.07%	1.91%	0.91	0.99
	SP_FC	1.68%	**1.19%**	1.21	1.42	7.81%	**5.02%**	0.94	1.03
	SP_TA	2.14%	**1.83%**	2.36	1.65	2.31%	**1.94%**	0.80	0.78

most ANS tests (6 of 8 tests) which contributes to the improvement of the generated classifiers performance. Overall, the error rates obtained using FDTs techniques in learning phase were very low reaching a maximum value of 2.99%. Thus, the accuracy rates recorded in all ANS tests exceeded 97.01% which is very promising and encouraging. As for the test phase, the maximum average value of error rates recorded was 5.02%. Thus, the minimum average value obtained for accuracy rates is 94.98% which is a very acceptable accuracy. Moreover, the comparison between the results obtained in the test phase showed an improvement in the error rates of FDTs. Six of eight generated FDTs have reached error rates lower than those of crisp DTs. On the other hand, we can notice that the standard deviation values are close to 0 which indicates that the error rates of the different trials tend to be very close to the average error rate.

Using the 10 fold cross-validation method, 10 trials were carried out for each ANS test. Thus, since eight FDTs were generated in this study, a total of 80 trials were performed. According to the results obtained for both techniques FDTs and crisp DTs, we noticed that using the training sets, FDTs recorded low error rates comparing to crisp DT in 92% of cases (74 of 80 trials). Regarding the test sets, in 89% of cases (71 of 80 trials), the error rates obtained by FDTs were lower than those obtained by crisp DTs. These results may be explained by the fact that FDT techniques are based on modeling uncertainty around the split values of the features which results in soft splits instead of hard splits of crisp DTs techniques; which contributes to the generation of a set of rules with simple conditional parts and relative high accuracy rates.

Moreover, the classification rules generated by FDTs facilitate the results interpretation by the experts. In fact, as aforementioned, the numerical data of the dataset were transformed into linguistic values. By this way, the extracted rules included linguistic forms instead of numerical ones which made them more readable and interpretable. As an example, the rule 1 and rule 2 present two classification rules extracted respectively from the generated FDT of Figure 3 and the crisp DTs regarding Hand grip test and expressing the same decision. Thus, we notice that rule 1 is easier to interpret than rule 2. Besides, based on our observations in the ANS unit, we noticed that the specialists prefer using linguistic values when interpreting the obtained results which can be suitable to the results provided by the generated FDTs. Therefore, the use of FDT techniques was beneficial and more appropriate in this case (*Rule 1:* **IF** Age=”Young” **AND** VR_HG=”High” **THEN** Class=”High”; *Rule 2:* **IF** 25<Age <= 35 **AND** VR_HG=25 **THEN** Class=”High”)

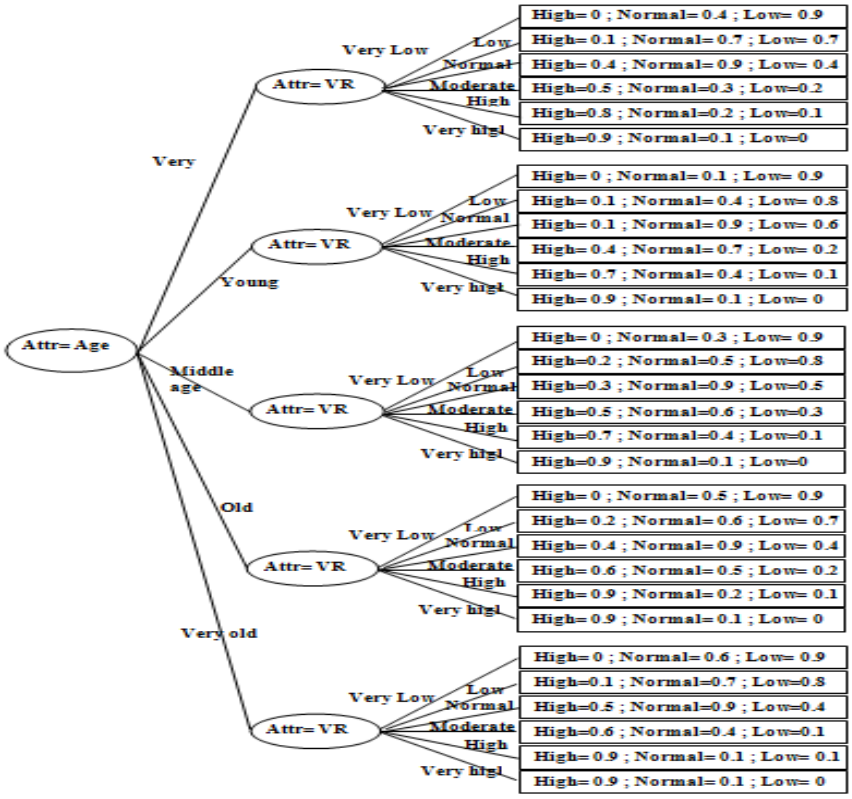


Figure 3. Example of a generated FDT regarding Hand grip test

Overall, this study has shown that FDT classifiers were more efficient and have achieved low error rates. These results may be explained by the fact that combining comprehensibility of DTs with the expressive power of fuzzy sets allowed to handle uncertainties and thus increase the performance of the generated classifiers. However, the complexity of FDTs was slightly higher than the complexity of crisp DTs. In fact, the number of nodes and leafs was higher in FDTs than crisp DTs while the depths of FDTs and DTs were slightly the same. For example, the generated FDT regarding HG test includes 6 nodes and 30 leafs (Figure 3) while the generated crisp DT for the same test contains only 4 nodes and 9 leafs.

5. Conclusion and future work

In this paper, FDT techniques were applied on a dataset extracted from the ANS unit of university hospital Avicenne. This study aimed at providing a making decision system to help doctors in the analysis procedure of the ANS's test results and improving the accuracy rates. This work used the results obtained in a previous case study so as to provide a comparison between the results obtained using crisp DTs and FDTs. Thereby, we used FID 3.4 algorithm to generate FDTs. The error rates recorded in each trial were very low which contributes to the increase of accuracy rates. These results were compared with the error rates obtained using the crisp C4.5 algorithm. Thus, the

generated FDTs were proved to be more accurate than those obtained using crisp DTs. However, as a limitation of this research, we may mention the small size of the dataset used which requires performing more validation tests over large data sets. For future work, we intend to conduct a research on evaluating the quality of DM based DSS in medical fields, especially in cardiology, since the data processed by these systems are very important and critical so as not to threaten the patient lives.

Acknowledgements

This research is part of the project "mPHR in Morocco" financed by the Ministry of High education and Scientific research in Morocco 2014-2016.

References

[1] Behadada O, Chikh MA, An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree. Artificial Intelligence Research 2 (2013), 45-58.
[2] Bohacik J, Kambhampati C, Classification in a heart failure dataset with a fuzzy decision tree. Advanced Research in Scientific Areas 1 (2012), 1981-1985.
[3] Breiman L, Friedman JH, Olshen RA, Stone CJ, Classification and regression trees. in: Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1984. pp. 246-280.
[4] Esfandiari N, Babavalian MR, Moghadam AE, Tabar V, Knowledge discovery in medicine: Current issue and future trend. Expert Systems with Applications 41 (2014), 4434-4463.
[5] Fajfer M, Janikow CZ, Bottom-up partitioning in fuzzy decision trees. Proceedings of the 19th International Conference of the North American Fuzzy Information Society (2000), 326-330.
[6] Familia A, Shenb WM, Weberc R, Simoudis E, Data preprocessing and intelligent data analysis. Intelligent Data Analysis, 1 (1997), 3-23.
[7] Fayyad U, Piatetsky-Shapiro G, Smyth P, From data mining to knowledge discovery in databases. in: AI Magazine, 1996. pp. 37‑54.
[8] Idri A, Kadi I, Benjelloun H, Heart disease diagnosis using C4.5 algorithms: a case study. International Conference on Health Informatics (2015) 397-404.
[9] Janikow CZ. Fuzzy Decision Trees: Issues and Methods. IEEE Transaction on Systems, Man, and Cybernetics -Part B, 28 (1998), 1-14.
[10] Janikow CZ, Kawa K, Fuzzy decision tree FID. Annual Meeting of the North American Fuzzy Information Processing Society (2005), 379-384.
[11] Kadi I, Idri A, A decision tree-based approach for cardiovascular dysautonomias diagnosis: a case study. IEEE Symposium on Computational Intelligence (2015).
[12] Karaolis MA, Moutiris JA, Hadjipanayi D, Pattichis CS, Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees. IEEE Transactions on Information Technology in Biomedicine 14 (2010), 559-566.
[13] Kreibig SD, Autonomic nervous system activity in emotion: A review. Biological Psychology 84 (2010), 394-421.
[14] Lee KM, Lee KM, Lee JH, Hyung LK, A Fuzzy Decision Tree Induction Method for Fuzzy Data. IEEE International Fuzzy Systems Conference Proceedings 1 (1999), 16-21.
[15] Pecchia L, Melillo P, Bracale M, Remote Health Monitoring of heart failure with data mining via CART method on HRV features. IEEE Transaction on Biomedical Engineering 58 (2011) 800-804.
[16] Quinlan JR, Induction of Decision Trees. Machine Learning 1 (1986), 81-106.
[17] Quinlan JR, C4.5 Programs for Machine Learning. in CA: Morgan Kaufmann Series in Machine Learning, 1993. pp. 1-302.
[18] Salzberg SL, On comparing classifiers: pitfalls to avoid and a recommended approach. Data Mining and Knowledge Discovery 1 (1997), 317–327.
[19] www.cs.umsl.edu/~janikow/fid/ (accessed January 2015)
[20] Zadeh LA, Fuzzy sets. Information and Control 8 (1965), 338-353.
[21] Zadeh LA, From computing with numbers to computing with words—From manipulation of measurements to manipulation of perceptions. IEEE Transa

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-9

Data Driven Methods for Predicting Blood Transfusion Needs in Elective Surgery

Dieter HAYN [a,1], Karl KREINER [a], Peter KASTNER [a], Nada BREZNIK [a], Axel HOFMANN [b], Hans GOMBOTZ [c] and Günter SCHREIER [a]
[a] AIT Austrian Institute of Technology GmbH, Austria
[b] Department of Anesthesiology, University Hospital Zurich, Switzerland
[c] *Department of Anesthesiology and Intensive Care, General Hospital Linz, Austria*

Abstract. Research in blood transfusions mainly focuses on Donor Blood Management, including donation, screening, storage and transport. However, the last years saw an increasing interest in recipient related optimizations, i.e. Patient Blood Management (PBM). Although PBM already aims at reducing transfusion rates by pre- and intra-surgical optimization, there is still a high potential of improvement on an individual level. The present paper investigates the feasibility of predicting blood transfusions needs based on datasets from various treatment phases, using data which have been collected in two previous studies. Results indicate that prediction of blood transfusions can be further improved by predictive modelling including individual pre-surgical parameters. This also allows to identify the main predictors influencing transfusion practice. If confirmed in a prospective dataset, these or similar predictive methods could be a valuable tool to support PBM with the ultimate goal to reduce costs and improve patient outcomes.

Keywords. Predictive Modelling, Ensemble Model, Bagged Trees

1. Introduction

Blood transfusion (BT) is a lifesaving procedure with various indications in surgery, intensive care, cardiac disease and many other fields. However, BT is characterised by a high inter-hospital variability of transfusion rates (TR, number of transfused patients per patient population) and a high variability of the amount of product used per transfused patient. Additionally, BT causes not only substantial direct and indirect costs but it can also lead to several adverse events (infections, immunologic reactions, etc.) [1]. Therefore, an increasing number of clinical trials and observational studies try to re-evaluate current BT practices.

While donor blood management focuses on processes related to donation, such as screening, storage and transport of blood products, Patient Blood Management (PBM) concerns procedures applied *to the potential recipient* with the main focus on reducing the amount of blood products which need to be transfused. PBM builds on three pillars:

- Detection and treatment of anaemia *before* elective surgeries
- Minimization of blood loss *during* surgery
- *Exploitation of individual anaemia tolerance* and rational use of blood products according to relevant guidelines [1,2]

[1] Corresponding Author: Dieter Hayn, AIT Austrian Institute of Technology, Reininghausstr. 13, 8020 Graz, Austria. E-Mail: dieter.hayn@ait.ac.at

In Austria, benchmark studies for blood use in elective surgery were commissioned by the Austrian Federal Ministry of Health and were conducted from 2004 to 2005 [2] and from 2009 to 2010 [3]. The aims were to measure the key variables of transfusion practice in elective surgery to chart the current situation, to identify predictors of transfusion, and to use the data for developing strategies to optimise transfusion practices across Austrian hospitals. After completion, the investigators of these studies provided a final report to the contracting authority and individual benchmark reports to the participating hospitals. This report is available online from the public information portal of the Austrian Federal Ministry of Health [4], including detailed and overall findings, as well as all relevant data. In terms of predictive aspects, the data were analysed using logistic regression on all collected variables so as to identify the main BT drivers, which were found to be lost red blood cell volume, relative preoperative haemoglobin, lowest relative postoperative haemoglobin, and sex in all types of surgeries analysed. In some kinds of surgery, also age, regional anaesthesia, American Society of Anaesthesiology (ASA) score, body mass index and platelet aggregation inhibitors were identified as independent predictors of BT [3].

According to current PBM guidelines, prior to elective surgeries with a high risk for significant blood loss, physicians are required to order a reasonable number of units of red blood cells (RBC). Based on the Mercuraili algorithm [5], this number is calculated from a) the individual patient's current erythrocyte volume ($EV_{preoperative}$ [L]), b) the individual lower limit of EV that the patient is expected to tolerate ($EV_{min\ acceptable}$ [L]) and c) the expected loss of erythrocyte volume (LEV) during the surgery ($LEV_{anticipated}$ [L]).

$$TEV_{required} = LEV_{anticipated} - (EV_{preop} - EV_{min\ acceptable}) \quad (1)$$

EV_{preop} is estimated from the individual patient's current blood volume (BV [L]), the preoperative venous haematocrit (Hct [1]) and the empirically determined correction factor 0.91 [1] (not used in original formula but in later studies) according to Equation 2

$$EV_{preop} = BV * Hct * 0.91 \quad (2)$$

where BV [L] is calculated from the patient's body weight (BW [kg]) and body height (BH [m]), using gender dependent empirical factors as shown in Equation 3 and Equation 4.

$$BV_{female} = 0.3561 * BH^3 + 0.03308 * BW + 0.1833 \quad (3)$$

$$BV_{male} = 0.3669 * BH^3 + 0.03219 * BW + 0.6041 \quad (4)$$

$EV_{min\ acceptable}$ is also estimated on an individual level. Based on clinical assessment, the physician evaluates, which EV / which Hct the patient is expected to tolerate without significant symptoms. However, the third input parameter for calculating the required EV according to Equation 1, i.e. $LEV_{anticipated}$, is only estimated from data of previous similar surgeries within the respective hospital, not taking into account any individual parameters of the patient [1].

Various attempts have been made to provide clinicians with estimates and predictions of BT needs [6]. However, when it comes to elective surgeries, less efforts have been made, although information on BT needs for specific surgeries could be of

high value. Recently, following the success in different industries, the application of data driven business intelligence and decision support has gained momentum in healthcare settings in general and in the BT topic as well [7].

The present paper re-evaluated the data from the Austrian Benchmarking Studies by going beyond previous analysis attempting to predict BT related outcomes, with the objective to evaluate statistical models to predict LEV as well as TEV based on different feature sets formed by individual pre-, intra and post-surgical parameters.

2. Methods

2.1. The data set

All analyses were based on the data collected in the course of the two Austrian Benchmarking studies, which were comprised of a total of 6,530 case records from 16 centres (408 ± 222 records per centre, min 164, max 907), obtained for elective surgeries of one of the following procedures: total hip replacement, total knee replacement and coronary artery bypass grafting. PBM guidelines require that for estimation of loss and transfusion of blood, historical data of the respective centre are considered. Since no historical data were available within the dataset, these historical data were estimated from all data available within the dataset (features *Historic mean LEV* and *TEV per centre and type of surgery*). Table 1 shows all features which were included in six feature sets:

- Historical data as recorded for a specific centre and type of surgery (Hist. LEV and Hist. TEV)
- Data available prior to the surgery excluding the amount of blood units ordered (Pre, Pre + ordered)
- Data available prior to the surgery including the amount of blood units ordered
- Data recorded during the surgery (Intra)
- Data recorded post-surgery (Post)

2.2. AIT Predictive Modelling Pipeline

The AIT PM pipeline is based on MATLAB R2015b (The Mathworks, Inc, Natik, NE) and consists of the following main modules:

Feature Set Compiler – an extract, transform, load (ETL) module for importing data from a variety of sources (databases, EXCEL or CSV files, output of preprocessing components e.g. for Biosignal analysis, …) governed by a Source Data Definition File and converting the data into a MATLAB datasets object for memory efficient computing based on a feature set definition.

Model Generator – utilizing the MATLAB Statistics and Machine Learning Toolbox and a modelling definition file, a variety of different models can be generated from the Feature sets, including General Linear Models, Bagged Trees, etc. Observations in the feature set can be arbitrarily divided into subsets for training, testing and validation with corresponding predictions being computed automatically.

Model Evaluator – allows visualising and evaluating model based predictions using methods like Receiver Operating Characteristics (ROC) and a variety of standard key performance indicators.

Table 1. The Feature Set used for the different types of predictions. LEV...Lost Erythrocyte Volume, TEV...Transfused Erythrocyte Volume, Pre, Intra, Post ... data available pre, intra and postoperatively. Features pre surgery are analysed including and excluding the number of blood units ordered. (Hb = Haemoglobin concentration, Hct = Haematocrit)

Description	Hist. LEV	Hist. TEV	Pre	Pre + ordered	Intra	Post
center number	1	1	1	1	1	1
type of surgery	1	1	1	1	1	1
Historic mean LEV per center and type of surgery	1		1	1	1	1
Historic mean TEV per center and type of surgery		1	1	1	1	1
gender			1	1	1	1
aggregation inhibitors			1	1	1	1
type of aggregation inhibitors			1	1	1	1
surgical technique			1	1	1	1
preoperative anemia			1	1	1	1
preoperative Hb > normal value			1	1	1	1
preoperative Hb category			1	1	1	1
preoperative Hb			1	1	1	1
Hbpre as percentage of WHO-anemia limit			1	1	1	1
preoperative Hct			1	1	1	1
preoperative circulating ery-volume			1	1	1	1
age			1	1	1	1
body mass index			1	1	1	1
body weight			1	1	1	1
body surface area			1	1	1	1
blood volume			1	1	1	1
total number of PRBC ordered				1	1	1
tranexamic acid					1	1
ASA-Score					1	1
duration of surgery					1	1
Euroscore					1	1
number of bypasses					1	1
extracorporeal circulation (ECC) used					1	1
duration of ECC					1	1
cell saver volume					1	1
cell saver ery-volume					1	1
cell saver used					1	1
unwashed shed blood					1	1
type of anaesthesia					1	1
regional anaesthesia only					1	1
Hb at the end of surgery					1	1
type&screen					1	1
Hb on postoperative day 3						1
Hb on postoperative day 5						1
min of all documented perioperative Hb-values						1
Hb3 as percentage of WHO anemia limit						1
Hb5 as percentage of WHO anemia limit						1
Hbmin as percentage of WHO anemia limit						1
Hct on postoperative day 3						1
Hct on postoperative day 5						1

This pipeline has a number of additional features useful to process large scale and heterogeneous healthcare data, from clinical codes to bio signals. It has previously been utilised for predictive analytics on different healthcare data sets, e.g. to predict the number of future days in hospital based on health insurance claims [8], to evaluate the utility of groups of features in given models by applying statistical tests on a set of related models build from observational subspaces (leave 10% out) [9] and to predict future events using time series approaches [10].

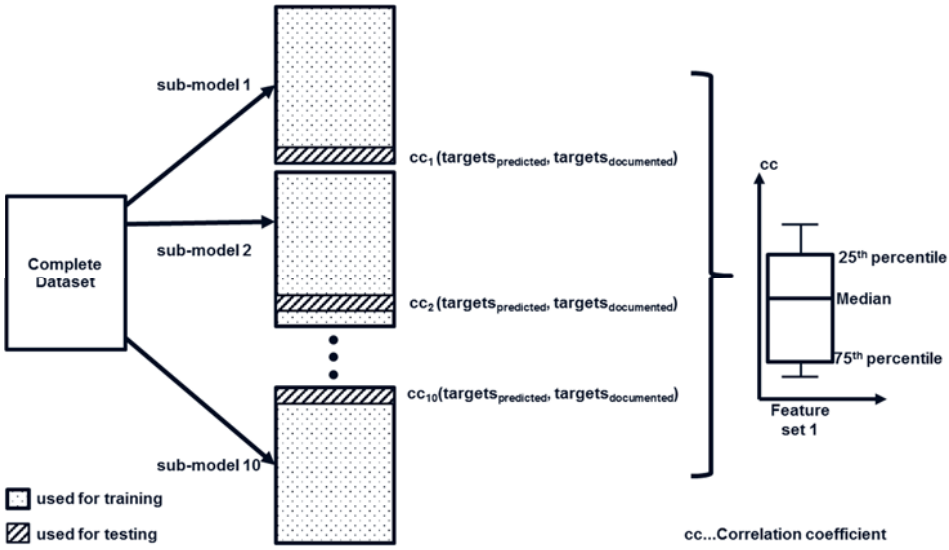


Figure 1. Leave 10% out approach used for training, prediction and statistical evaluation of each model

2.3. Training, Testing and Evaluation

A leave 10 % out approach was used for training and testing our models. This resulted in 10 different sub-models based on a training set of 90 % of the whole data set, which was applied to the remaining 10 % as a test set. Each model was trained with a random forest approach [11] using MATLABs *TreeBagger* functionality with default settings except for *OOBPred = on, NPrint = 1, MinLeaf = 10, Method = regression, Surrogate = off,* and *OOBVarImp = onPrediction.* The modelling result of each sub-model was compared to the actual target parameters and Pearson's correlation coefficient was calculated for each sub-model as a measure of the linear correlation between observed and predicted values. Statistical parameters of the correlation coefficients of the ten sub-models were visualized using boxplots. This approach is illustrated in Figure 1. Results of each feature set were compared to one another and to a baseline / reference model, which comprised of only one single feature representing a single, random number.

3. Results

Models for two prediction targets, i.e. LEV and TEV, with all feature set classes were trained and tested, applying a leave 10 % out approach. On a personal computer with an Intel® Core™ i7-3770 CPU with 3.50 GHz, 8 GB RAM and Windows 7, the overall modelling process, including training, prediction and evaluation for all feature sets and all target parameters took 5 min 11 s. Application of a trained model on an individual patient's dataset took 23 ms for the largest feature set (post-surgery).

Boxplots of the correlation coefficients of the 10 sub-models were drawn for each feature set to compare the respective model performance. The non-parametric Kruskal-Walis test was used to assess the statistical significance differences in performance of each features class, taking the 10 model runs in each group of features as an independent sample (see Figure 2 and Figure 3).

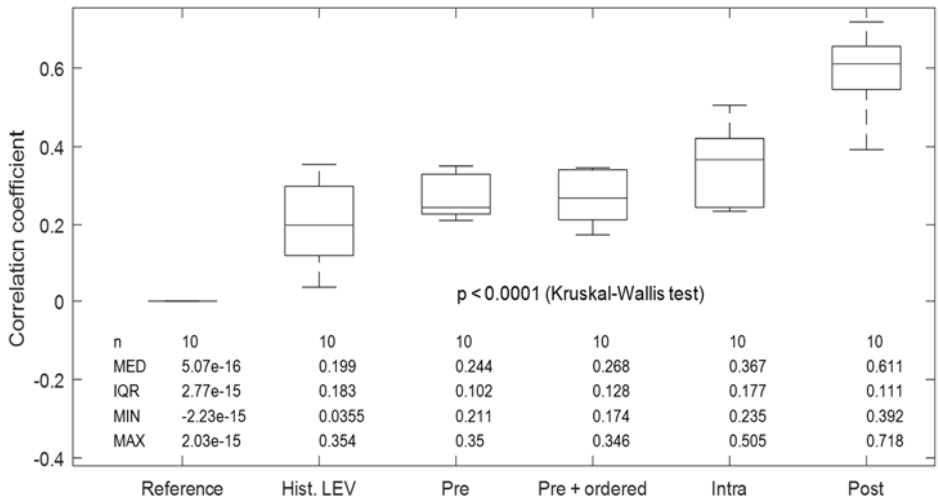


Figure 2. Boxplots of correlation coefficients of predicted loss of erythrocyte volume during surgery (LEV) as compared to actual loss, depending on the class of features used for prediction. Features classes as described for Table 1. Each box represents 10 correlation coefficients as received for different subsets when applying a leave 10 percent out approach for training and testing the respective model.

The correlation coefficient when considering the reference model based on the random feature only was zero for both LEV and TEV. As expected, predictions of LEV as well as TEV increased in precision when historical, pre-, intra- and post-surgical features were included in the model, respectively, resulting in statistically significant differences between the groups. Both parameters showed improved prediction performance even when pre-surgical, individual features were considered as compared to historical data from the respective centre only.

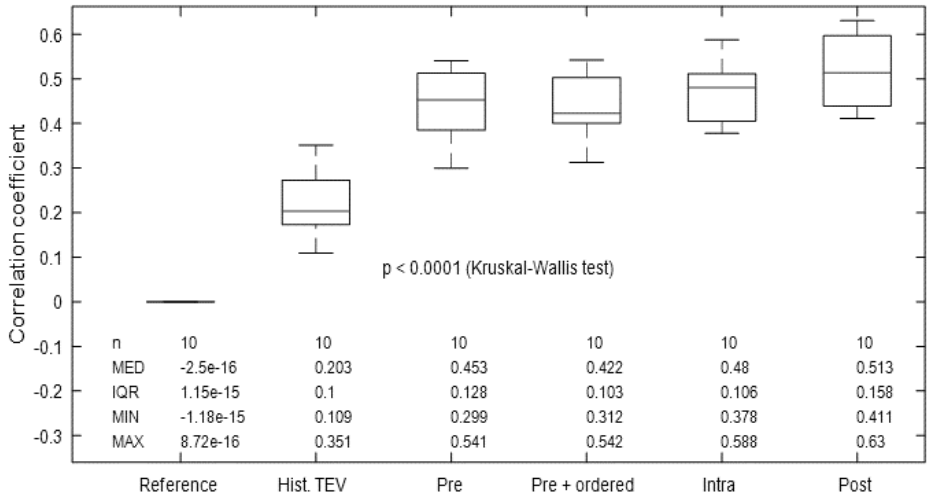


Figure 3. Boxplots of correlation coefficients of predicted transfused erythrocyte volume (TEV) as compared to actual volume, depending on the class of features used for prediction. Features classes as described for Table 1. Each box represents 10 correlation coefficients as received for different subsets when applying a leave 10 percent out approach for training and testing the respective model.

For TEV, the most distinct improvement was found when individual, pre-surgical features were added, while for LEV post-surgical features were most pivotal. Neither for LEV nor for TEV the number of units ordered significantly improved the model performance.

4. Discussion

While individual approaches are already suggested for TEV prediction in current guidelines, LEV is currently predicted without considering individual data. Our results indicate that not only TEV but also LEV can be predicted more precisely when considering individual pre-surgical parameters than when estimating these parameters from historical data of a certain centre only.

As expected, prediction of LEV was significantly more accurate if intra- and post-surgical features were considered. TEV prediction accuracy, on the other hand, already showed good correlation factors when considering pre-surgical data only, but it did not improve that much when adding intra- and post-surgical features. These results can be used to analyse the importance of single features or the performance of certain centres in a retrospective setting. However, in a real world scenario, these data will not be available for predication prior to surgeries.

Murphree et al. applied a large number of different model approaches to a related topic, i.e. the prediction of complications after BT [7]. His results indicate that most models give good results if applied alone and that combining those models with a "majority vote" strategy did not yield a significant improvement. These results have not yet been verified with our dataset.

During the two Austrian benchmarking studies [2, 3] it could be shown that different centres show significantly different blood transfusion patterns and that for some centres significant improvements could be achieved from the first to the second study by applying PBM principles. The present paper, however, did not consider, whether a centre had already applied a PBM program or not. This might have a severe influence on the resulting model which has not yet been investigated. Prediction of LEV are rather independent from PBM programs and therefore we expect that similar results would be achieved if our model would be applied to other centres. TEV prediction and number of ordered blood units, on the other hand, are highly influencing one another, they are dependent on the application of PBM processes, and on centre specific transfusion triggers. These influences could not yet be analysed sufficiently, and, therefore, further analyses are required in the future.

With the end of the EU project *EU-PBM Patient Blood Management* [12] getting closer, there will be a chance to validate these models with an independent, prospectively collected dataset. Also, this data is expected to allow to look at the impact of factors related to the three columns of the PBM strategy. Future work will also focus on aspects of providing prediction results to the physicians in a way which is easy to access and comprehend anywhere in their institution and anytime when decisions need to be made.

5. Conclusion

The results obtained in the present work indicate that predicting BT needs in elective orthopaedic and cardiac surgery is feasible based on a set of parameters which can

expected to be available in most centres. Individual parameters and more features, in particular from time points closer to the respective treatment phase, lead to better results. If these results can be confirmed with independent datasets, an additional tool to support PBM would be available.

References

[1] Patient Blood Management - Individuelles Behandlungskonzept zur Reduktion und Vermeidung von Anämie und Blutverlust sowie zum rationalen Einsatz von Blutprodukten: Georg Thieme Verlag, 2013.
[2] H. Gombotz, P. H. Rehak, A. Shander, and A. Hofmann, "Blood use in elective surgery: the Austrian benchmark study," Transfusion, vol. 47, pp. 1468-80, Aug 2007.
[3] H. Gombotz, P. H. Rehak, A. Shander, and A. Hofmann, "The second Austrian benchmark study for blood use in elective surgery: results and practice change," Transfusion, vol. 54, pp. 2646-57, Oct 2014.
[4] H. Gombotz, P. H. Rehak, and A. Hofmann, "Fortsetzung der Studie „Maßnahmen zur Optimierung des Verbrauchs von Blutkomponenten bei ausgewählten operativen Eingriffen in österreichischen Krankenanstalten“ 2008 -2010," 2012.
[5] Mercuriali F, Inghilleri G. Proposal of an algorithm to help the choice of the best transfusion strategy. Curr Med Res Opin 1996;13(8):465-78
[6] A. Kadar, O. Chechik, E. Steinberg, E. Reider, and A. Sternheim, "Predicting the need for blood transfusion in patients with hip fractures," International Orthopaedics, vol. 37, pp. 693-700, Apr 2013.
[7] D. Murphree, C. Ngufor, S. Upadhyaya, N. Madde, L. Clifford, D. J. Kor, et al., "Ensemble Learning Approaches to Predicting Complications of Blood Transfusion," presented at the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Scociety, Milan, 2015.
[8] Y. Xie, G. Schreier, D. C. Chang, S. Neubauer, Y. Liu, S. J. Redmond, et al., "Predicting Days in Hospital Using Health Insurance Claims," IEEE J Biomed Health Inform, vol. 19, pp. 1224-33, Jul 2015.
[9] Y. Xie, S. Neubauer, G. Schreier, S. Redmond, and N. Lovell, "Impact of Hierarchies of Clinical Codes on Predicting Future Days in Hospital," in 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, 2015, pp. 6852-6855.
[10] Y. Xie, G. Schreier, M. Hoy, Y. Liu, S. Neubauer, D. C. Chang, et al., "Analyzing health insurance claims on different timescales to predict days in hospital," J Biomed Inform, Jan 2016.
[11] L. Breiman and A. Cutler. (2014). Random forests. Available: http://www.stat.berkeley.edu/breiman/RandomForests/cc home.htm
[12] AIT Austrian Institute of Technology GmbH. (n.d., Feb 8). EU-PBM Patient Blood Management. http://www.europe-pbm.eu/ Accessed 2016.02.09

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-17

Clinical Decision Support for the Classification of Diabetic Retinopathy: A Comparison of Manual and Automated Results

Christoph MITSCH[a,1], Karsten FEHRE[b], Sonja PRAGER[a], Christoph SCHOLDA[a], Katharina KRIECHBAUM[a], Thomas WRBA[c] and Ursula SCHMIDT-ERFURTH[a]
[a]*Department of Ophthalmology, Medical University of Vienna*
[b]*Medexter Healthcare GmbH, Vienna*
[c]*Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna*

Abstract. The management of diabetic retinopathy, a frequent ophthalmological manifestation of diabetes mellitus, consists of regular examinations and a standardized, manual classification of disease severity, which is used to recommend re-examination intervals. To evaluate the feasibility and safety of implementing automated, guideline-based diabetic retinopathy (DR) grading into clinical routine by applying established clinical decision support (CDS) technology. We compared manual with automated classification that was generated using medical documentation and an Arden server with a specific medical logic module. Of 7169 included eyes, 47% (n=3373) showed inter-method classification agreement, specifically 29.4% in mild DR, 38.3% in moderate DR, 27.6% in severe DR, and 65.7% in proliferative DR. We demonstrate that the implementation of a CDS system for automated disease severity classification in diabetic retinopathy is feasible but also that, due to the highly individual nature of medical documentation, certain important criteria for the used electronic health record system need to be met in order to achieve reliable results.

Keywords. Diabetes complications, retina, decision support systems, clinical.

1. Introduction

Diabetic retinopathy (DR) is the leading cause of visual impairment in working-age adults worldwide [1, 2]. The disease is characterized by capillary non-perfusion and ischemia within the retina, which ultimately leads to macular edema and retinal neovascularization with the potential to severely damage visual function [3]. Although the development of sight-threatening complications of diabetes mellitus (DM) can be delayed by appropriate treatment of systemic diseases such as DM itself, high blood pressure and lipid metabolism abnormalities [4], an estimated 40% (8% for vision-threatening DR) of people with type 2 diabetes and 86% (42%) with type 1 diabetes in the USA have diabetic retinopathy [5, 6].

With the proportion of people who have a predominantly sedentary lifestyle and are overweight growing worldwide, the prevalence of type 2 DM and thus of DR is

[1] Corresponding Author: Dr. C. Mitsch, Department of Ophthalmology, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna. E-Mail: christoph.mitsch@meduniwien.ac.at

continuously increasing, burdening healthcare systems all over the world. Patients with DM and DR need to be examined regularly by various specialists: At least one general practitioner and one ophthalmologist are involved in a patient's routine check-ups. In the case of ophthalmological examinations, indirect ophthalmoscopy (examination of the patient's retinae through dilated pupils) is the only way to determine the extent of retinopathy. This examination is time-costly and requires a specialist; it is not possible to delegate it to non-medical personnel.

According to the American Academy of Ophthalmology guidelines for the management of DR [7], follow-up intervals should depend on the current stage of disease progression. Patients with normal retinae or mild DR should be re-examined annually, while those with moderate DR should be re-examined within 6 to 12 months as disease progression is common. Patients with severe DR should be re-examined within 2-4 months and patients with proliferative DR should receive treatment with appropriate follow-up intervals.

Medical informatics in general and clinical decision support systems in particular aim to assist healthcare delivery and are especially pertinent when a high patient load meets with well-defined clinical processes. With the growing incidence of DM, it seems logical to make use of computer technologies. Many efforts have been reported: Automated, fundus-photograph-based disease detection and classification (also termed automated retinal image analysis, ARIA; or computer-aided detection; CAD) is a promising technique to relieve healthcare systems' burdens of regular screening of diabetic patients. Many software products offering this technology are commercially available and could potentially improve the manner in which diabetes eye care is delivered [8]. Automated detection of more advanced diabetic lesions in the patient's eye has become conceivable with the advent of new, non-invasive optical coherence angiography. Integration of the resulting data with information about the patients' systemic disease such as diabetes duration, hemoglobin A1c, and body mass index, and demographics such as age and sex allows predictive modeling to be applied and thus prognosis of the individual disease course [9].

The Arden syntax is a standard for representing clinical and scientific knowledge in an executable format. The Arden server stores and processes details from medical guidelines in the form of so-called medical logic modules (MLMs), and provides an interface for administrative and maintenance tasks [10]. The server is used in a wide spectrum of clinical situations such as adverse drug event monitoring [11] or medical guideline implementation [12]. Although this technology is not new, there are few clinical decision support applications in the field of ophthalmology and most of them are related to automated image analysis or guides and material for preparing ophthalmology students [13].

The implementation of technologies that potentially increase efficiency in the management of chronic diseases is highly attractive, especially when it enables maintaining consistent quality of medical care.

We are not aware of any published efforts to use clinical decision support (CDS) technology in the classification of DR. Thus, the aim of this study was to compare the results of manual classification of DR with those of a CDS-based automated classification and assess the applicability and safety of the CDS-based automated with the manual method for disease classification before deploying this technology for its use in daily clinical routine at our department.

2. Methods

2.1. Classification of DR

To date, the International Clinical Disease Severity Scale for DR[14] has become the most commonly used severity classification of DR. The variables used are related to the integrity of and pathological alterations to retinal vessels which specialists observe on the retinae of the patients via ophthalmoscopy. Their presence and extension are noted and manually processed as described in Table 1. This manual approach currently represents the gold standard of DR classification.

Table 1 - The International Clinical Disease Severity Scale for diabetic retinopathy (DR). IRMA = intraretinal microvascular abnormalities

Disease Severity Level	Findings Observable upon Dilated Ophthalmoscopy
No retinopathy	No abnormalities
Mild DR	Microaneurysms only
Moderate DR	More than just microaneurysms but less than severe DR
Severe DR	Any of the following: • More than 20 intraretinal hemorrhages in each of four quadrants • Venous beading in two or more quadrants • Prominent IRMA in one or more quadrants .. and no signs of proliferative DR
Proliferative DR	One or both of the following: • Neovascularization • Vitreous/pre-retinal hemorrhage

2.2. Ophthalmological findings

Microaneurysms (MA) are small alterations to retinal capillaries which, clinically, mark the very beginning of retinal disease in patients with DM. The rupture of MA leads to *dot-blot hemorrhages*. Areas of venular dilatation and bulging are called *venous beading* and *Intra-retinal microvascular abnormalities (IRMA)* are newly grown vessels within the layers of the retina. Extra-retinal neovascularization can occur at the optic disc (*neovascularization on the optic disc*, NVD), elsewhere on the retina (*neovascularization elsewhere*, NVE) or in the anterior chamber of the eye (*neovascularization on the iris*, NVI). Rupture of newly formed vessels can result in *pre-retinal or vitreous hemorrhages*.

2.3. Documentation of relevant findings

Clinical examinations and observations at the outpatient clinic for Diabetic Retinopathy and Traumatology of the Department of Ophthalmology and Optometry of the Medical University of Vienna are performed and documented either by resident physicians working under consulting physicians' supervision or by experienced consulting physicians themselves. Information is stored in a customized electronic health record (EHR) system. This system consists of a series of forms on the Research Documentation and Analysis platform (RDA) of the Medical University of Vienna's center for medical

Table 2 - List of the variables relevant for the classification of diabetic retinopathy. P = present, P1 = present in one quadrant, P2+ = present in more than one quadrant, A = active, F = fibrotic, FT = fibrotic with traction, NP = not present, NA = not assessed, IMP = impossible to assess

Observation	Values
Microaneurysms	P, NP, NA, IMP
Rubeosis	P, NP, NA, IMP
Venous tortuosity	P1, P2+, NP, NA, IMP
Neovascularization at the optic disc	P, A, F, FT, NP, NA, IMP
Neovascularization elsewhere	P, A, F, FT, NP, NA, IMP
Intraretinal microvascular abnormalities	P, NP, NA, IMP
Retinal hemorrhages	Present less than 20x in all quadrants, present at least 20x per quadrant, P, NP, NA, IMP
Vitreous or pre-retinal hemorrhage	P, NP, NA, IMP

statistics, informatics and intelligent systems (CEMSIIS). Besides general functional and morphologic data, findings relevant for diabetic retinopathy are stored in a highly structured manner, separately for each eye. See Table 2 for a list of the variables queried.

2.4. Manual documentation of diabetic retinopathy severity

Retina specialists manually documented a retinopathy classification separately for each eye of patients in case the severity of the present diabetic retinopathy was relevant. It is important to note that our EHR does not require a disease classification, thus, it is not possible to differentiate between no present retinopathy and retinopathy not assessed. Two further possible values of the manual classification were “not graded/no retinopathy”, “retinopathy; severity not classified”. These values were not compared with automatic results. The descriptors for the different stages of diabetic retinopathy were “mild”, “moderate”, “severe”, and “proliferative” RP, the last stage further described as “active proliferative” or “inactive proliferative”, or “fibrotic proliferative RP”.

Table 3 - The algorithmic part of the Arden medical logic module used for automated diabetic retinopathy severity classification. For demonstration purposes, the header and the code for the left eye have been removed and the values translated into the English language. The suffix “_od” indicates that the variable represents a finding on the right eye (o.d. = *oculus dexter*)

Diabetic Retinopathy Medical Logic Module (right eye)

```
IF ((vitreous_hem_od = "true") OR (nvd_od = "true") OR (nvd_od = "active") OR
(nvd_od="fibrotic") OR (nvd_od = "fibrotic with traction") OR (nve_od = "true")
OR (nve_od = "active") OR (nve_od="fibrotic") OR (nve_od = "fibrotic with
traction") OR (rub_od = "true"))
THEN
        result.ScoreOD := 4; # proliferative DR
ELSEIF ((venous_tortuosity_od = "Present in more than one quadrant") or
(hemorrhage_od = "true, present at least 20x per quadrant") or (irma_od =
"true")) THEN
        result.ScoreOD := 3; # severe DR
ELSEIF (hemorrhage_od = "true, present less than 20x in all quadrants") THEN
        result.ScoreOD := 2; # moderate DR
ELSEIF (ma_od = "true") THEN
        result.ScoreOD := 1; # mild DR
ELSEIF (ma_od = "false") THEN
        result.ScoreOD := 0; # no DR
ELSE
        result.ScoreOD := -1; # not enough data for automated grading
ENDIF;
```

2.5. Automated classification of diabetic retinopathy severity

See Table 3 for the medical logic module (MLM) programmed for the automated classification. The possible results of the algorithm were "not enough data for automated grading", "no retinopathy", "mild DR", "moderate DR", "severe DR" and "proliferative DR".

2.6. Inclusion and exclusion criteria

Visit documentations which included patho-morphological findings and a manual DR grading were included. We excluded entries which had either no grading or no findings documented.

3. Results

3.1. Overview

We analyzed the documentations of 5727 visits of 1303 out-patients at the department of diabetic retinopathy and traumatology between 2012 and 2015, relating to both eyes, resulting in a dataset of 11454 eye examinations. Explicit manual classification (a diabetic retinopathy level of at least "mild") had been performed for 7530 eyes. Based on the documented morphology findings, automated classification was possible in 10293 eyes and not possible in 1161 eyes. No manual classification had been performed for 3924 eyes. This left us with an includable subset of 7169 eyes, where manual and automated classification were available and possible.

3.2. Included visits: Manual classification performed, automated classification possible

Based on the manual classification, 14.2% (n=1020) showed mild DR, 23.2% (n=1666) moderate DR, 18.6% (n=1335) severe DR, and 43.9% (n=3148) proliferative DR. The results of the automated classification were 17.8% (n=1277) with no DR, 15.6% (n=1117) with mild DR, 20.2% (n=1450) with moderate DR, 10.6% (n=762) with severe DR and 35.8% (n=2563) with proliferative DR. See the left graph on Figure 1 for a graphical representation of those distributions. The overall inter-method agreement was 47%; 29.4% in mild DR, 38.3% in moderate DR, 27.6% in severe DR, and 65.7% in proliferative DR. See the right graph on Figure 1 for a graphical representation of the agreement levels. The numbers and percentages (relative to the included cohort) of the eyes with equal and differing classifications are given in Table 4. The greatest fraction with disagreement was one group of eyes which had been manually graded as showing signs of "proliferative DR", but calculated as "no DR" disease by the automatic algorithm, constituting 6.4% (n=459) of all eyes.

The most frequent relative comparison result was by far "agreement" (difference = 0 stages, 47%, n=3373). 18.3% (n=1080) of all eyes compared were classified with a more severe DR automatically than manually, and 37.9% (n=2716) with a less severe DR. Classified automatically, this manual "less severely classified" group consisted of 5.2% (n=376) with moderate DR, 2.9% (n=208) with severe DR and 6.9% (n=496) with proliferative DR. Refer to the middle graph on Figure 1 for a graphical representation of

Table 4 – Proportions (relative to the whole cohort) of the groups with agreement (light grey cell background) and disagreement (white cell background) of all manual and all automated classification results. Rows: Manual classification M1 (mild DR) to M4 (proliferative DR). Columns: Automated classification A0 (no DR) to A4 (proliferative DR).

	A0	A1	A2	A3	A4
M1	3.8% (n=275)	4.2% (n=300)	5.2% (n=376)	0.4% (n=32)	0.5% (n=37)
M2	4.3% (n=306)	5.9% (n=421)	8.9% (n=638)	2.5% (n=176)	1.7% (n=125)
M3	3.3% (n=237)	2.4% (n=172)	3.1% (n=224)	5.1% (n=368)	4.7% (n=334)
M4	6.4% (n=459)	3.1% (n=224)	3% (n=212)	2.6% (n=186)	28.8% (n=2067)

the frequencies of the relative classification differences. An analysis of the categorization differences between the individual consulting physicians revealed the same range (-4 to 3 degrees) and a mean grading difference of -0.62±1.42 degrees for all graders.

3.3. Eyes excluded due to missing manual grading or unfeasible automated grading

No manual classification information was available for 34.3% (n=3924) of all screened eyes. This group comprised all eyes which had not presented with DR or no classification had been documented. For 27.3% (n=3124) of the eyes, no manual grading had been performed, but automated classification was possible. No findings relevant to DR had been documented, thus no DR was automatically classified for 18.5% (n=2123), mild DR for 2.1% (n=235), moderate DR for 2.6% (n=300), severe DR for 1.1% (n=127) and proliferative DR for 3% (n=339) of all eyes. Automated classification was not possible in 10.1% (n=1161) of all eyes due to missing findings documentation. In 3.2% (n=361) of all eyes automated classification was not possible, but manual classification had been performed. In this group, expressed as proportions of all screened eyes, 0.6% (n=66) had mild RP, 0.7% (n=77) moderate RP, 0.9% (n=102) severe RP and 1% (n=116) proliferative disease. No manual or automated classification was available for 7% (n=800) of all screened eyes.

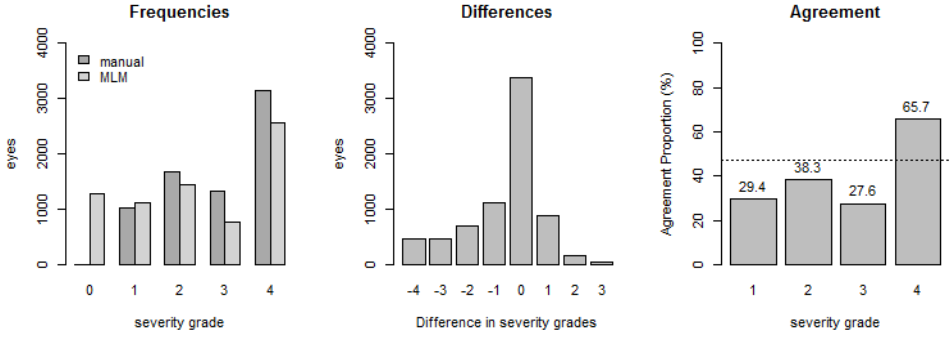


Figure 1 – Results of the manual and automated classifications (left graph), a plot of the differences and their frequencies (middle graph) and a bar-plot of the agreements in the different severity classes. Left: 0 = no DR, 1 = mild DR, 2 = moderate DR, 3 = severe DR, 4 = proliferative DR. Middle: x-Axis = classification difference between automated and manual classification. Positive values suggest a worse automated disease severity, negative values a lesser automated severity. Right: 1 = mild DR, 2 = moderate DR, 3 = severe DR, 4 = proliferative DR. The dotted horizontal line represents the overall agreement.

4. Discussion

We compared the manual and CDS-based automated DR classification of 7169 eyes examined at our diabetic retinopathy and traumatology outpatient clinic. The overall agreement between the methods was 47%, which, essentially, is not very satisfactory, given that reevaluation interval recommendations are potentially determined based on their disease classification. We used data generated by ophthalmologists in our diabetic retinopathy outpatient clinic, which, being a tertiary care center, mainly focuses on the management of more advanced disease. This may reflect in the inter-method agreement, which is the highest (65.7%) amongst proliferative cases. Retreatment intervals at this stage are much shorter and set individually, based on planned and performed interventions. Our custom EHR does not validate or check the fields required for disease staging for completeness. Morphological findings, disease classification and reexamination interval represent three independent, but, from a physician's perspective, rather redundant data elements. They are inherently linked by guideline recommendations. The high patient load in this kind of outpatient clinics does not always allow to document redundantly, so the most important piece of information remains whether an intervention is necessary or not, and, if no intervention is indicated, when an ophthalmological reexamination needs to take place. In our custom EHR, this is written as a free-text message to collaborating field ophthalmologists. Thus, the structured analysis of this information was not easily possible.

The inter-method agreement presented in this experiment has wide potential of improvement. In order to achieve sufficient reliability of automated disease classification, the input data and therefore our EHR needs to meet specific criteria. The lessons we learned are the following: First, and most importantly, redundancy needs to be eliminated. This is a general principle, not only for medical settings, and has a very specific meaning in the EHR context: If CDS is to be implemented in clinical routine and to be working with data documented during patient visits, the forms need to be reduced to only offer the one most precise specification of the relevant parameters (those in Table 2). If the classification is straight-forward and free of exceptions, which is the case in this example, the input of manual results should be similarly possible, but linked to an informative validation of its consistency with the input parameters (the documented morphological findings). The re-examination interval recommendation should be offered as a suggestion which can be altered later, as there may be other factors taken into account for this decision. If manually entered data is used for the automated calculation of patient management recommendations, it needs to be automatically validated for consistency and completeness before CDS gets to work, which, in our case, would mean checking all relevant entered ophthalmological findings data.

The main challenge seems to be preserving EHR usage flexibility for the physicians. In contrast to medical guidelines, medical documentation is a highly subjective and personal task. CDS should be implemented in a facultative manner, so that physicians can decide whether they want to use it (in which case they would have to follow certain documentation rules, see above), or if they just want to document the most relevant findings, determine the disease classification and reevaluation intervals by themselves, which is what we see in our present data.

We believe that, if those rules are met, CDS-based disease classification is certainly possible for clinical routine, where it can contribute to uniform and objective assessment and help to save valuable physicians' time, which can then be spent with the patient rather than with a computer.

To conclude, we demonstrate that the implementation of a CDS system for automated disease severity classification in an outpatient clinic treating patients with diabetic retinopathy is technically and medically feasible but that specific adaptation to the EHR structure is required in order to yield reliable results.

References

[1] J. E. Shaw, R. A. Sicree, and P. Z. Zimmet, "Global estimates of the prevalence of diabetes for 2010 and 2030," *Diabetes Res. Clin. Pract.*, vol. 87, no. 1, pp. 4–14, 2010.

[2] R. Klein, K. E. Lee, M. D. Knudtson, R. E. Gangnon, and B. E. K. Klein, "Changes in visual impairment prevalence by period of diagnosis of diabetes: the Wisconsin Epidemiologic Study of Diabetic Retinopathy.," *Ophthalmology*, vol. 116, no. 10, pp. 1937–42, 2009.

[3] D. A. Antonetti, R. Klein, and T. W. Gardner, "Diabetic retinopathy.," *N. Engl. J. Med.*, vol. 366, no. 13, pp. 1227–39, Mar. 2012.

[4] D. Wright and P. M. Dodson, "Medical management of diabetic retinopathy: fenofibrate and ACCORD Eye studies.," *Eye (Lond).*, vol. 25, no. 7, pp. 843–849, 2011.

[5] J. H. Kempen, B. J. O'Colmain, M. C. Leske, S. M. Haffner, R. Klein, S. E. Moss, H. R. Taylor, and R. F. Hamman, "The prevalence of diabetic retinopathy among adults in the United States," *Arch.Ophthalmol.*, vol. 122, no. 4, pp. 552–563, 2004.

[6] M. S. Roy, R. Klein, B. J. O'Colmain, B. E. K. Klein, S. E. Moss, and J. H. Kempen, "The prevalence of diabetic retinopathy among adult type 1 diabetic persons in the United States.," *Arch. Ophthalmol. (Chicago, Ill. 1960)*, vol. 122, no. 4, pp. 546–51, Apr. 2004.

[7] American Academy of Ophthalmology Retina/Vitreous Panel, "Preferred Practice Pattern (R) Guidelines. Diabetic Retinopathy. San Francisco, CA: American Academy of Ophthalmology; 2014. Available at: www.aao.org/ppp."

[8] D. A. Sim, P. A. Keane, A. Tufail, C. A. Egan, L. P. Aiello, and P. S. Silva, "Automated retinal image analysis for diabetic retinopathy in telemedicine.," *Curr. Diab. Rep.*, vol. 15, no. 3, p. 14, 2015.

[9] T. Aspelund, O. Thornórisdóttir, E. Olafsdottir, a Gudmundsdottir, a B. Einarsdóttir, J. Mehlsen, S. Einarsson, O. Pálsson, G. Einarsson, T. Bek, and E. Stefánsson, "Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy.," *Diabetologia*, vol. 54, no. 10, pp. 2525–32, 2011.

[10] M. Samwald, K. Fehre, J. de Bruin, and K. P. Adlassnig, "The Arden Syntax standard for clinical decision support: Experiences and directions," *J. Biomed. Inform.*, vol. 45, no. 4, pp. 711–718, 2012.

[11] K. Fehre, M. Plössnig, J. Schuler, C. Hofer-Dückelmann, A. Rappelsberger, and K. P. Adlassnig, "Detecting, Monitoring, and Reporting Possible Adverse Drug Events Using an Arden-Syntax-based Rule Engine," *Stud. Health Technol. Inform.*, vol. 216, no. 4, p. 950, 2015.

[12] A. Seitinger, K. Fehre, K.-P. Adlassnig, A. Rappelsberger, E. Wurm, E. Aberer, and M. Binder, "An Arden-Syntax-based clinical decision support framework for medical guidelines--Lyme borreliosis as an example.," *Stud. Health Technol. Inform.*, vol. 198, pp. 125–32, 2014.

[13] I. de la Torre-Diez, B. Martinez-Perez, M. Lopez-Coronado, J. R. Diaz, and M. M. Lopez, "Decision Support Systems and Applications in Ophthalmology: Literature and Commercial Review Focused on Mobile Apps," *J. Med. Syst.*, vol. 39, no. 1, 2014.

[14] C. P. Wilkinson, F. L. Ferris, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, J. T. Verdaguer, and F. Lum, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-25

Analyzing Readmissions Patterns: Assessment of the LACE Tool Impact

Christo EL MORR[a,1] , Liane GINSBURG[a], Victor-Seungree NAM[a], Susan WOOLLARD[b], and Bojay HANSEN[b]
[a] *York University, Toronto, Canada*
[b] *North York General Hospital, Toronto, Canada*

Abstract. This paper will discuss the assessment of the use of the LACE tool at North York General Hospital (NYGH). The LACE tool estimates the readmission risk of patients. This paper describes the tool and a modified LACE score implementation and use at NYGH. We also describe our statistical analysis for the LACE effectiveness in order to inform future decisions in resource allocations. We will look at suggestions for adjustments in the way the LACE tool is used as well as implications for service delivery and patients' quality of life. Our study shows that the modified LACE is a predictive tool for readmission risk in day-to-day hospital activity, but that implementation of LACE alone cannot reduce readmission rates unless coupled with efforts of those in charge of providing community-based care.

Keywords. LACE, Readmission, Health Management, Quality of Care, Quality of Service, Clinical IT, Discharge Summary, Discharge.

1. Introduction

In Canada, one in 12 patients is readmitted within 30 days of discharge. In Ontario, 9% of acute care patients returned to the emergency room and one sixth of them returned more than once within seven days of initial discharge [1]. Inpatient readmissions account for more than one in 10 dollars spent on inpatient care in Canada (excluding physician fees for services). Costs are greatest for **medical patients** who account for 64.9% of unplanned readmissions followed by **surgical** patients at 23.9% [2].

Hospital 30-days readmissions are largely unplanned and preventable. The rates of readmission are highest for clients with congestive heart failure, myocardial infarction, and pneumonia - respectively [3]. Vascular surgeries are also associated with high rates of readmission within 30 days. Research suggests that the reasons behind readmission within 30 days of discharge have to do with both the **patient characteristics** and the characteristics of the **procedure** (e.g. a 75-year-old client with diabetes was more likely to be readmitted to the hospital following an invasive vascular surgery compared to younger patients with no chronic disease [4]).

Between 2010 and 2013, the Medicine program at North York General Hospital (NYGH), Toronto, Canada, has seen an increasing trend in its 30-day readmission rate. During that period, NYGH was in excess of the corporate target for readmissions (set at 7.3%) [5]. In an effort to reduce readmissions, NYGH undertook an initiative in June 2013 to implement a risk assessment tool called LACE. LACE is "an index to predict early death or unplanned readmission after discharge from hospital to the community"

[1] Corresponding Author: Christo El Morr, York University, 4700 Keele St., HNE #415, Toronto, Canada M3J 1P3, E-Mail: elmorr@yorku.ca

[8] that is calculated based on: **L**ength of stay ("L"), **A**cuity of the admission ("A"), patient **C**omorbidity ("C"), and **E**mergency department number of visits ("E") that was developed by van Walraven et al. [6-9].

This project analyzed readmission data from NYGH to gain insight into LACE and inform future resource allocation decisions. The research project also has the potential to impact patients' quality of life since use of the LACE tool is designed for early identification of patients who are high risk for readmission and thus to start the discharge planning with the inter-professional team, in an attempt to reduce readmission rate.

NYGH is intending to dispatch new resources (e.g. teaching packs) to this project, and has already invested initiatives in order to follow up patients having a LACE score greater or equal to 10. Nevertheless, for a wise use of current and future resources, it was critical to analyze the re-admission patterns at NYGH and investigate if LACE is working as predicted or if it needs adjustment to fit NYGH patient population.

2. Methods

LACE implementation at NYGH. Before starting any data analysis, we had to understand how NYGH implemented the LACE tool in practice. In 2010, when Walraven and his colleagues developed the 'LACE' tool[8], they defined L as the current length of stay in the hospital (i.e. LOS for the index admission). Largely for practical reasons, particularly the need to use the LACE score to plan in advance of discharge, NYGH has defined L as patients' length of stay in his/her previous acute care visit within the last 30 days. The Acuity of the Admission weight indicates if the current admission is acute or not and NYGH calculated this in the same manner as Walraven et al. Comorbidity of the patient is measured by using the Charlson comorbidity index score in the original LACE work, though NYGH modified the scale used by the original authors of LACE by giving a weight of 6 instead of 5 for metastatic cancer. Using Walraven's approach, 'Emergency department use' is measured by looking at patients' total number of visits to the emergency department in the six months immediately prior to the index admission. The 'L', 'A', and 'E' and C components of LACE are calculated manually by the nurse on the floor during the index admission and are entered in the LACE software. Overall, a patient with a LACE score <10 is considered to at low risk of readmission while LACE >=10 suggests a high risk of readmission. The following figure summarizes LACE scoring methodology as has been used by NYGH staff.

Procedure for Calculating LACE. Lace was implemented at NYGH between June and October 2013 on a number of medicine units in the hospital. For each admitted patient a nurse uses a software to enter the four components of the LACE score manually, the software then calculates a LACE score for the patient. In addition to obtaining one year of LACE data (June 2013 – June 2014), we accessed data on readmission rates for each LACE unit dating back one year prior to LACE implementation at the hospital thereby allowing us to look at readmission rates in the one-year period leading up and one year following LACE implementation.

Analysis. Data were received in Excel™; then it was cleaned, imported and analyzed into SPSS™. Ethics approval was obtained from the Ethical Review Board at NYGH. In addition, each researcher completed the "Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans Course on Research Ethics" certificate (TCPS2: core).

Attribute	Value	Points	Score
Length of Stay (Prior Admit)	Less 1 day	0	
	1 day	1	
	2 days	2	
	3 days	3	
	4-6 days	4	
	7-13 days	5	
	14 or more days	6	
Acute Admission	Inpatient	3	
	Observation	0	
Comorbidity (Cumulative to a max of 6 pts)	No prior history	0	
	DM no complications, Cerebrovascular disease, Hx of MI, PVD, PUD	1	
	Mild liver disease, DM with end organ damage, CHF, COPD, Cancer, Leukemia, Lymphoma, any tumor, moderate to severe renal disease	2	
	Dementia or connective tissue disease	3	
	Moderate or severe liver disease or HIV infection	4	
	Metastatic cancer	6	
Emergency room visits during previous 6 months	0 visits	0	
	1 visits	1	
	2 visits	2	
	3 visits	3	
	4 or more visits	4	
	Take the sum of the points and enter the total ➡ ***If LACE score is 11 or greater, CM to send tool to agency/facility patient is referred to on discharge**		

Figure 1: LACE score as has been implemented by NYGH

3. Results

We have used descriptive statistics to compute the readmission rates for the low risk (LACE <10) and high risk (LACE >=10) groups and found them to be 9.7 % and 18.7%, respectively, in the one-year period following LACE implementation.

3.1. LACE predictive ability in the hospital setting

In order to conclude the predictive power of the modified LACE tool, we have conducted a logistic regression analysis that allows us to uncover and compare the odds-ratio of LACE scores greater than 10 and LACE scores lower than 10 in relation to readmission, and consequently to compare their corresponding predictive ability. The logistic regression revealed that the patients in the high risk group (LACE score ≥ 10) are **2.05** times more likely to be readmitted than those in the low risk group (LACE score < 10).

3.2. Readmission reduction

We were interested in looking into any significant difference in readmission rates for the months before LACE compared to those for the months after LACE index has been introduced. The readmission rate distribution was skewed and consequently we have used the non-parametric Mann-Whitney U test to compare readmissions rates before and after LACE implementation.

The Mann-Whitney statistical analysis showed no significant difference between the period before LACE and after LACE; consequently, LACE per say had no effect on readmission rates.

3.3. LACE threshold for risky patients

Managers at NYGH have noticed that some patients with low LACE score are being readmitted and hypothesized that a reduction in the LACE threshold to 8 would have a better discriminatory powers than 10 and allows us to capture more patients with high risk of readmission. We modified the LACE threshold to 8, in order to test whether a lower LACE score would have more predictive power. Regression results showed that for a threshold of 8 (instead of 10) LACE would have a less predictive power as the regression coefficient decreased (2.01 compared to 2.05 for threshold 10). Consequently, the LACE score threshold should be kept at 10.

3.4. Calculating LACE in the ward: data entry

Since we conducted a retrospective analysis, we were able to compute the exact 'L' and 'E' components of LACE automatically using SPSS. We compared our computed, and hence accurate, 'L' and 'E' to the manual data entered during the patients' stay in the hospital. We have conducted a Weighted Kappa Analysis to compare the agreement between the scores entered in LACE and our scores.

The data entry error rates of L and E were 33% and 49% respectively. Moreover, the level of agreement between the L and E values entered by NYGH staff compared to the correct L and E values that we have been able to compute were significantly different (Kappa values <0.7).

The data entry errors in L and E resulted in missing risky readmissions and spending time on non-risky ones. Between September 2013 and August 2014, 11% of the cases considered by the NYGH team as risky should have been considered low risk. This resulted in unnecessary resource utilization. On the other hand, between September 2013 and August 2014, 23% of patients were considered low risk while they in the high-risk range. This resulted in missing high-risk patients.

Moreover, we have conducted a logistic regression analysis that showed that our accurate LACE scores give higher odds ratio than those entered manually into NYGH system, which meant that if 'L' and 'E' were accurately entered, LACE would make a better patients' readmission prediction.

4. Conclusions

The main question was to investigate is the LACE tool is a good predictor of readmission in the real world, our data analysis shows that effectively the LACE tool is a good predictor for readmission.

The second question we had is to see if the introduction of LACE at NYGH had any influence on readmission rates; the data analysis showed that calculating LACE is not sufficient to reduce readmissions. Instead, more collaborative, cross-sectorial efforts that include those in charge of providing community-based care are needed to try to address the problem of readmissions.

As for the third question regarding the effect of any change in the LACE threshold for high risk patients (e.g. reducing the score), the data analysis showed that the threshold 10 is more appropriate than 8 and should be kept in use.

The fourth question we addressed was the accuracy of the data entry, our data analysis showed significant data entry errors which effect was to miss high risky patients and to use unnecessary resources for low risky patients. Consequently, modified approaches that reduce reliance on manual capture of LACE elements are needed. This will yield better data quality, better risk assessment and reduces data collection burden for front-line staff.

We are currently in the process of analyzing the data using Geographic Information Systems methodologies. The GIS analysis may help illuminate socio-economic and/or socio-cultural factors that may influence readmissions. We already know that geography has an impact on patient's health [10].

Finally, in our study we could not account for (remove) patients who die within 30 days of discharge, which must have introduced some bias in the data analysis.

Acknowledgment

We would like to acknowledge the "NYGH Exploration Fund" for funding our project.

References

[1] Canadian Institute for Health Information, "Data Quality Documentation for External Users: Discharge Abstract Database, 2010–2011," CIHI, Ottawa2011.

[2] Canadian Institute for Health Information, "All-Cause Readmission to Acute Care and Return to the Emergency Department," CIHI, Ottawa2012.

[3] E. H. Bradley, L. Curry, L. I. Horwitz, H. Sipsma, J. W. Thompson, M. Elma, *et al.*, "Contemporary Evidence About Hospital Strategies for Reducing 30-Day Readmissions: A National Study," *Journal of the American College of Cardiology,* vol. 60, pp. 607-614, 8/14/ 2012.

[4] B. S. Brooke, R. R. De Martino, M. Girotti, J. B. Dimick, and P. P. Goodney, "Developing strategies for predicting and preventing readmissions in vascular surgery," *J Vasc Surg,* vol. 56, pp. 556-62, Aug 2012.

[5] B. Hansen, M. Muia, E. Villar-Guerrero, and S. Woollard, "A Reduction in Readmission Rates through the Implementation of an Automated LACE Risk Assessment Tool at North York General – A Pilot Project (4N Neuro/Stroke)," North York General Hospital, Toronto2013.

[6] A. G. Au, F. A. McAlister, J. A. Bakal, J. Ezekowitz, P. Kaul, and C. van Walraven, "Predicting the risk of unplanned readmission or death within 30 days of discharge after a heart failure hospitalization," *Am Heart J,* vol. 164, pp. 365-72, Sep 2012.

[7] A. Gruneir, I. A. Dhalla, C. van Walraven, H. D. Fischer, X. Camacho, P. A. Rochon, *et al.*, "Unplanned readmissions after hospital discharge among patients identified as being at high risk for readmission using a validated predictive algorithm," *Open Med,* vol. 5, pp. e104-11, 2011.

[8] C. van Walraven, I. A. Dhalla, C. Bell, E. Etchells, I. G. Stiell, K. Zarnke, *et al.*, "Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community," *Cmaj,* vol. 182, pp. 551-7, Apr 6 2010.

[9] C. van Walraven, J. Wong, and A. J. Forster, "LACE+ index: extension of a validated index to predict early death or urgent readmission after hospital discharge using administrative data," *Open Med,* vol. 6, pp. e80-90, 2012.

[10] Health Quality Ontario Ontario, "Quality in Primary Care Setting a Foundation for Monitoring and Reporting in Ontario," 2015.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-31

Piloting the European Unified Patient Identity Management (EUPID) Concept to Facilitate Secondary Use of Neuroblastoma Data from Clinical Trials and Biobanking

Hubert EBNER[a,1], Dieter HAYN[a] , Markus FALGENHAUER[a], Michael NITZLNADER[a], Gudrun SCHLEIERMACHER[b], Riccardo HAUPT[c] , Giovanni ERMINIO[c], Raffaella DEFFERRARI[c], Katia MAZZOCCO[c], Jan KOHLER[d] , Gian Paolo TONINI[e], Ruth LADENSTEIN[f] and Guenter SCHREIER[a]
[a]*AIT Austrian Institute of Technology. Digital Safety & Security Dept., Graz, Austria*
[b]*Institut Curie, Paris, France*
[c]*Istituto G. Gaslini, Genova, Italy*
[d]*Southampton General Hospital, Southampton, UK*
[e]*University of Padua, Padova, Italy*
[f]*Children's Cancer Research Institute, Vienna, Austria*

Abstract. Data from two contexts, i.e. the European Unresectable Neuroblastoma (EUNB) clinical trial and results from comparative genomic hybridisation (CGH) analyses from corresponding tumour samples shall be provided to existing repositories for secondary use. Utilizing the European Unified Patient IDentity Management (EUPID) as developed in the course of the ENCCA project, the following processes were applied to the data: standardization (providing interoperability), pseudonymization (generating distinct but linkable pseudonyms for both contexts), and linking both data sources. The applied procedures resulted in a joined dataset that did not contain any identifiers that would allow to backtrack the records to either data sources. This provided a high degree of privacy to the involved patients as required by data protection regulations, without preventing proper analysis.

Keywords. paediatric oncology, clinical trial, biobanking, secondary use, pseudonymization,

1. Introduction

Although neuroblastoma is one of the most common cancers in infancy [1], it is still a rare disease. Therefore, data collection during clinical trials needs to be done over large networks containing many institutions. One example of such a data collecting network is the International Society of Paediatric Oncology European Neuroblastoma Research Network (SIOPEN-R-NET) [2]. Due to regulatory aspects, personalised data, which has previously collected in a clinical trial, may not be used in a different context without asking the patient for consent again. On the other hand, such secondary use of clinical data has the potential to provide significant new findings from already existing data,

[1] Corresponding Author: Hubert Ebner, AIT Austrian Institute of Technology, Reininghausstr. 13, 8020 Graz, Austria. E-Mail: hubert.ebner.fl@ait.ac.at

which could potentially improve future treatment of children suffering from cancer. To overcome this regulatory barrier, trial data may be anonymized or pseudonymized within an appropriate governance framework [3]. However, anonymization inhibits data linking completely, and linking of pseudonymized data requires novel concepts in order to simultaneously provide high levels of usability and data protection.

In the course of the project *European Network for Cancer Research in Children and Adolescents* (ENNCA) [4] the *Advanced Biomedical Collaboration Domain for ENCCA* (ABCD-4-E) [5] has been established to share data between different sources in a distributed computing system. The ABCD-4-E features the European Unified Patient IDentity Management (EUPID) [6] concept that allows registration and pseudonymization of patients and linkage of the different datasets without the need of using directly identifying data element.

We describe the implementation, set-up and application of a data processing pipeline to facilitate linking of existing data from the European Unresectable Neuroblastoma (EUNB) trial [7] and results from array comparative genomic hybridisation (CGH) analyses from corresponding tumour samples [8] in a secondary use scenario.

This paper deals with the overall concept and the implementation of the pipeline needed for the main tasks, i.e. the registration of the included patients, and the detection of patients that exist in both studies and the merging of their datasets.

2. Methods

Figure 1 shows an overview of the whole process, comprising of patient registration and data uploading using the ABCD-4-E network for different contexts. The EUNB trial represented the context *clinical trial* (CT) while the results from the CGH array analyses represented the context *biobank* (BB). Each context consisted of patients and their raw data with context specific pseudonyms (PSN_{INRG} and PSN_{GEO}) and the coordinating researchers (**Figure 1a**).

The genomic profiles from the CGH analysis will eventually be deposited to the NCBI Gene Expression Omnibus [9, 10] and accessible through GEO series accession number GSE26494 [11].

The clinical data will eventually be stored in the database of the International Neuroblastoma Risk Group (INRG) [12, 13].

2.1. Pseudonymization of patients

In both contexts the patients were identified using their identification properties (IDAT) (**Figure 1a**). Using an additional feature of EUPID, the IDAT provided for generating the PSNs consisted of the patients SIOPEN-R-NET codes (SRN, instead of first and last name) and their dates of birth for this feasibility study). Pseudonymization was done using the EUPID service [6] (**Figure 1b**). As result, two additional pseudonyms PSN_{BB} and PSN_{CT} were assigned to each patient.

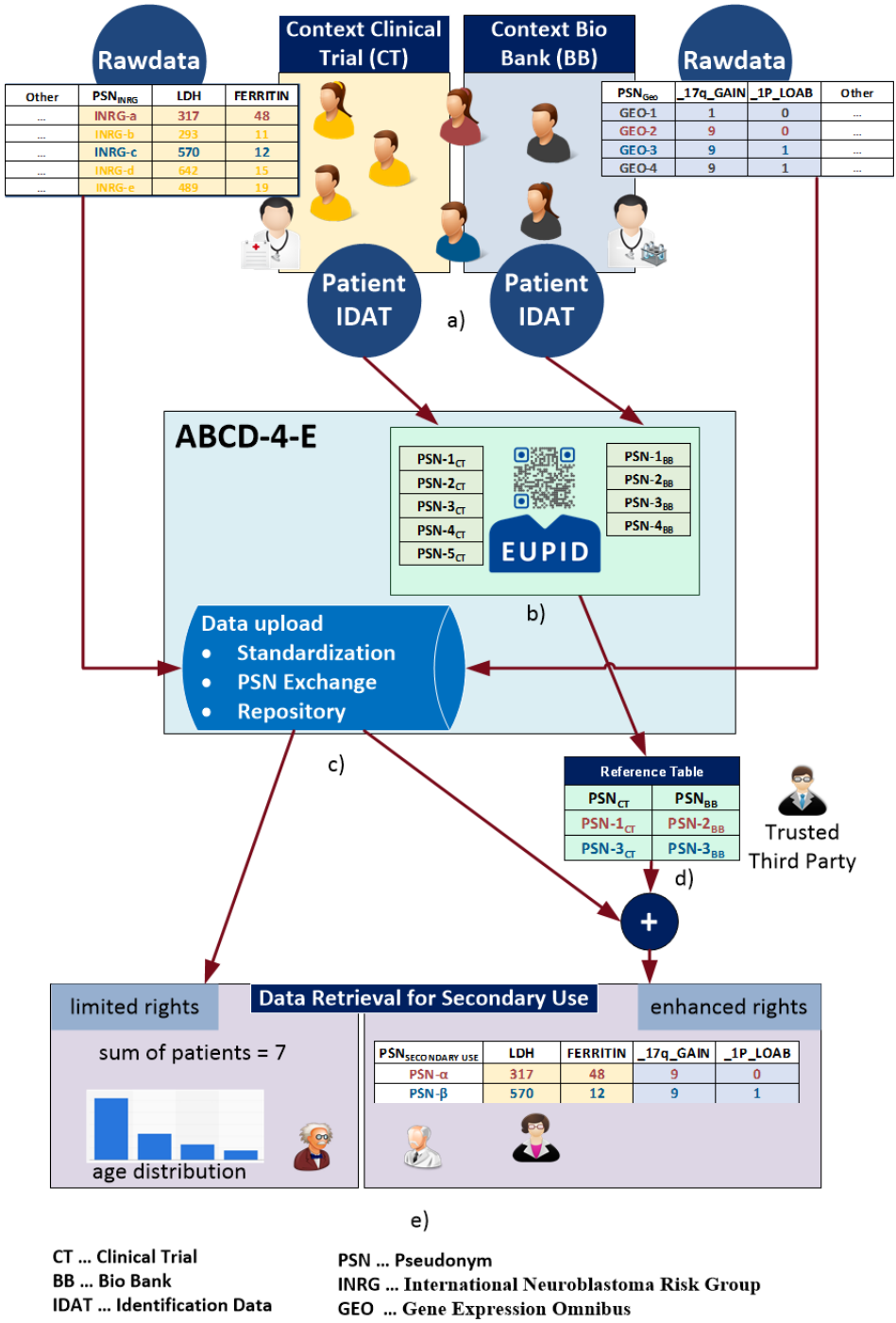


Figure 1. Workflow for patient registration and data upload of clinical trial and biobank data for combined secondary use

2.2. Data upload

After patient registration users with the role of the coordinators could upload the collected data to the ABCD-4-E system (**Figure 1c**). Therefore, three major services have been implemented:

- Standardization of the raw data
- Replacing the old pseudonyms by the EUPID generated pseudonyms
- Creation and upload of the core datasets

2.3. Standardization of Data

The raw data that were available as Excel/CSV files did not follow any interoperability standards. To provide the necessary interoperability between data sources from different domains the data needed to be standardized. Therefore, the raw data were converted into the standardized Operational Data Model (ODM) data format [14].

2.4. PSN exchange

The ODM-standardized raw data still contained the old pseudonyms used in their originally domains (PSN_{INRG} and PSN_{GEO}). These pseudonyms were exchanged with the newly generated EUPID pseudonyms (PSN_{BB} and PSN_{CT}).

2.5. Core dataset creation and upload

To facilitate semantic standardization the data needed to be transformed into core datasets, using an XSLT based transformation pipeline. The core datasets defined which information from the data was extracted from the raw data [15]. After successful transformation into core datasets the data were stored to the ABCD-4-E repository.

2.6. Data retrieval for secondary use

Depending on the permissions of the invoking researchers, queries could be generated and executed leading to correspondingly adjusted results. **Figure 1e, left,** shows results which could be computed from strongly aggregated data as provided to users with limited rights (e.g. number of patients and age distribution) while **Figure 1e, right,** shows results which could be obtained from more detailed data as provided to users with more extensive rights, allowing to merge patients' datasets from different data sources by means of an additional reference table (**Figure 1d**).

3. Results

The ABCD-4-E network was extended with web-based apps to provide the needed functions for patient registration and pseudonymization, data standardization and data upload to the network. For data retrieval access to aggregated data was provided via a querying app or – in cooperation with the trusted third party – to a full-featured merged data set, including details from both contexts. To facilitate this, the following components/extensions had to be implemented to ABCD-4-E.

3.1. ABCD-4-E pipeline for data production

Figure 2 shows the ABCD-4-E app pipeline for data production where coordinators from the two context where able to register their patients and upload data.

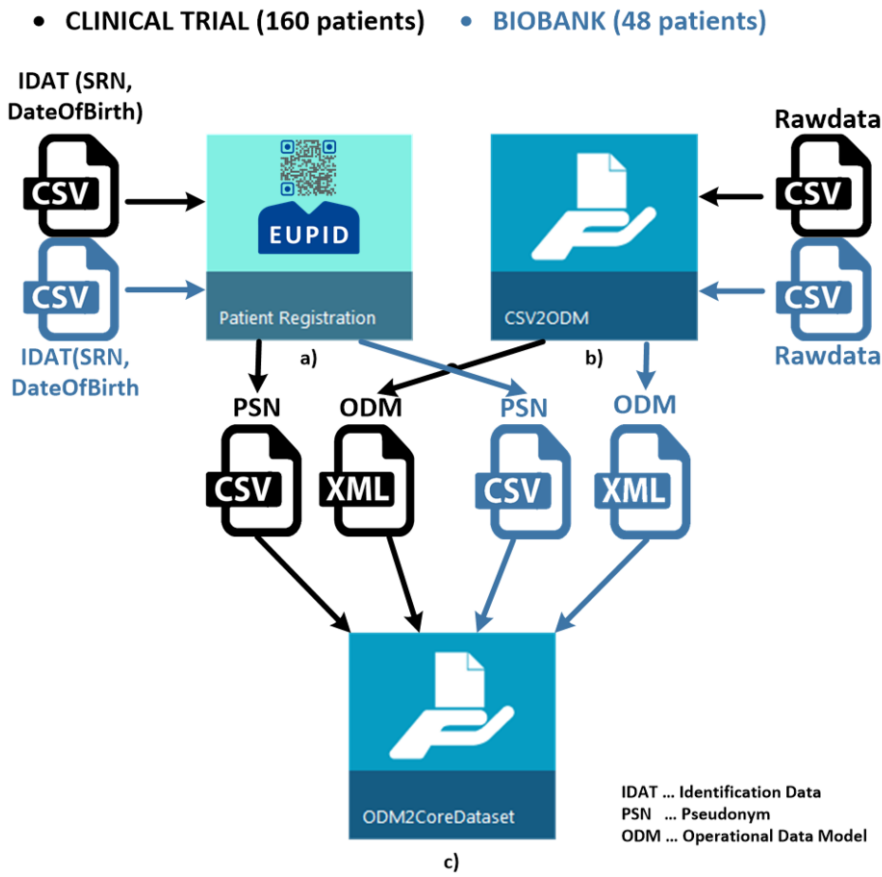


Figure 2. ABCD-4-E pipeline for patient registration and data upload using the apps in the colored boxes

3.2. Patient Registration

Every participating patient was registered in the ABCD-4-E system using the EUPID *Patient Registration* app (**Figure 2a**).

Typically, coordinators register patients using their IDAT (first name, second name, date of birth). In this special case the patients from both contexts had already been registered in SIOPEN-R-NET. Therefore, instead of using their IDAT, the patients were registered based on their SIOPEN-R-NET IDs (SRN) and their date of birth.

Registration could either be done manually or by using a newly developed feature of the EUPID, the so-called *Bulk Registration*. Therefore, a CSV list of patients' IDAT could be uploaded and a list of newly generated EUPID pseudonyms was automatically created.

A demo version of the EUPID *Patient Registration* app can be accessed at [16].

160 patients were registered by the coordinators of the clinical trial and 48 patients from the bio bank.

3.3. Data upload

To enable data sharing over the ABCD-4-E network, two apps (*CSV2ODM* and *ODM2CoreDataset)* have been developed.

The app *CSV2ODM* was developed to fulfil the requirement of data interoperability. It converted raw data from CSV format to ODM format (**Figure 2b**).

The ODM standardized raw data still contained the old pseudonyms used in their originally trials. The *ODM2CoreDataset* app [15] used the generated list of EUPID generated pseudonyms and the ODM standardized raw data from the *CSV2ODM* app and automatically replaced the old pseudonyms by the EUPID generated pseudonyms (**Figure 2c**).

The second task of the *ODM2CoreDataSet* app was to create a core dataset as described in chapter 2.2.3 and to upload that dataset to the ABCD-4-E repository.

3.4. ABCD-4-E pipeline for data retrieval

The pipeline for data consumption was used to access uploaded data from the ABCD-4-E repository. Depending on the users' rights, apps for queries and reference tables have been prepared for this purpose.

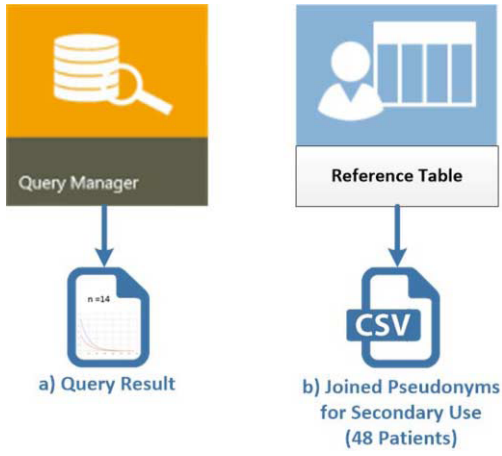


Figure 3. ABCD-4-E pipeline for data consumption

3.5. Query Manager

The current development of the *Query Manager* allowed to answer predefined queries shown in **Figure 3a**.

3.6. Reference Table

After registration of patients the EUPID system automatically linked the EUPID pseudonyms of identical patients that occurred in different contexts. This was done within the system and not transparent to the coordinators.

The *Reference Table* was an app that allowed a trusted third party to access the internal linkage of the patients (**Figure 3b**).

Within the ABCD-4-E network the EUPID was able to link each of the 48 patients from the biobank to their corresponding pseudonym from the clinical trial. The final result was a joined dataset containing linked pseudonyms for 48 patients. In the pilot test, AIT took over the role of the trusted third party to access the *Reference Table*.

4. Discussion

Secondary use of personalized health data needs patient consent, which is often difficult to obtain after some time has passed. Without patient consent, datasets need to be anonymized. Anonymization, however, prevents linkage of records, which is essential for longitudinal or cross domain studies. The EUPID concept has been designed to overcome these barriers and can leverage its full potential once registration of paediatric patients from all research activities using EUPID has started on a routine basis.

The usage of the *Reference Table* will enable not only linkage of patients' pseudonyms as shown in the results. Its purpose is to link complete datasets across domains, too. This joined datasets are delivered with newly internally generated EUPID pseudonyms (**Figure 1e,** $PSN_{SECONDARY\_USE}$). No other identifiers are used. So backtracking of the patients using the joined dataset and comparing their pseudonyms with e.g. pseudonyms of one of the source datasets is impossible.

However, to avoid patient re-identification by comparing source datasets not only via the original pseudonyms but also via the data from the joined dataset, additional precaution measures have to be implemented. First of all the joined dataset could be delivered in a new randomized order avoiding simple comparison of subsequent rows. Secondly, coarsening of typical individual information (e.g. size and postal code) could be needed especially in trials with low patient numbers [17]. Thirdly, concepts using advanced features like k-anonymity [18] and l-diversity [19] in an appropriate way could be included.

After adding these precaution measures the EUPID system would strongly meet the concepts for data protection defined by the *Technologie- und Methodenplattform für die vernetzte medizinische Forschung* (TMF) [20].

If re-identification should become necessary, the trusted third party is needed. This safeguards the privacy to the involved patients as required by data protection regulations.

Acknowledgement

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7 2007-2013) under the project ENCCA grant agreement no. 261743.

References

[1] P. P. T. E. Board., "Neuroblastoma Treatment (PDQ®): Health Professional Version.," in PDQ Cancer Information Summaries, ed: National Cancer Institute, 2016.

[2] (2016-01-19). SIPEN-R-NET. Available: http://www.siopen.org

[3] BMA. (2014, 2016-01-19). requests for disclosure of data for secondary purposes. Available: http://www.bma.org.uk/-/media/files/pdfs/practical%20advice%20at%20work/ethics/releasingdataforsecondaryuses.pdf

[4] (2016-01-19). ENCCA - European Network for Cancer Research in Children and Adolescents. Available: www.encca.eu

[5] M. Nitzlnader, M. Falgenhauer, C. Gossy, and G. Schreier, "Architecture for an advanced biomedical collaboration domain for the European paediatric cancer research," Stud Health Technol Inform, pp. 167-74, 2015.

[6] M. Nitzlnader and G. Schreier, "Patient identity management for secondary use of biomedical research data in a distributed computing environment," Stud Health Technol Inform, pp. 211-8, 2014.

[7] J. A. Kohler, H. Rubie, V. Castel, K. Beiske, K. Holmes, C. Gambini, et al., "Treatment of children over the age of one year with unresectable localised neuroblastoma without MYCN amplification: Results of the SIOPEN study," European Journal of Cancer, vol. 49, pp. 3671-3679, Nov 2013.

[8] G. Schleiermacher, J. Michon, A. Ribeiro, G. Pierron, V. Mosseri, H. Rubie, et al., "Segmental chromosomal alterations lead to a higher risk of relapse in infants with MYCN-non-amplified localised unresectable/disseminated neuroblastoma (a SIOPEN collaborative study)," British Journal of Cancer, vol. 105, pp. 1940-1948, Dec 6 2011.

[9] NCBI. (2015, 12.01.2016). Gene Expression Omnibus. Available: http://www.ncbi.nlm.nih.gov/geo/

[10] R. Edgar, M. Domrachev, and A. E. Lash, "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository," Nucleic Acids Research, vol. 30, pp. 207-210, Jan 1 2002.

[11] (2016-01-19). array-CGH profiling of human neuroblastoma samples obtained from infants included in the INES99.1, INES99.2 and INES99.3 trials. Available: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26494

[12] (2016-01-19). The International Neuroblastoma Risk Group (INRG). Available: http://inrgdb.org

[13] S. L. Cohn, A. D. J. Pearson, W. B. London, T. Monclair, P. F. Ambros, G. M. Brodeur, et al., "The International Neuroblastoma Risk Group (INRG) Classification System: An INRG Task Force Report," Journal of Clinical Oncology, vol. 27, pp. 289-297, Jan 10 2009.

[14] C. D. I. S. Consortium. (2016-01-19). Operational Data Model. Available: http://www.cdisc.org/odm

[15] N. Hochedlinger, D. Nitzlnader, M. Falgenhauer, S. Welte, D. Hayn, L. Koumakis, et al., "Standardized data sharing in a paediatric oncology research network--a proof-of-concept study," Stud Health Technol Inform, pp. 27-34, 2015.

[16] A.-A. I. o. T. GmbH. (2015-01-19). European Patient Identity Management EUPID. Available: https://eupid.eu

[17] G. Danezis, J. Domingo-Ferrer, M. Hansen, J.-H. Hoepman, D. Le Métayer, R. Tirtea, et al., "Privacy and Data Protection by Design – from policy to engineering," ed: European Union Agency for Network and Information Security (ENISA), 2014

[18] L. Sweeney, "K-anonymity: A model for protecting privacy," International Journal on Uncertainty, fuzziness and knowledge-based systems, vol. 10, pp. 557 - 570, 2002.

[19] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," Proceedings of the 22nd International Conference on, pp. 24 - 36, 2006.

[20] K. Pommerening and J. Drepperm. (2011, 2016-03-09). TMF, Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V. - Datenschutz in der medizinischen Forschung. Available: https://www.tmf-ev.de/DesktopModules/Bring2mind/DMX/Download.aspx?Method=attachment&Command=Core_Download&EntryId=13308&PortalId=0.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-39

Interoperability Architecture for a Paediatric Oncology European Reference Network

Michael NITZLNADER [a, 1], Adela CANETE NIETO [b], Antonio Juan RIBELLES [b], Barbara BRUNMAIR [c], Ruth LADENSTEIN [c] and Günter SCHREIER [a]
[a]AIT Austrian Institute of Technology GmbH, Austria
[b]Hospital Universitari i Politècnic La Fe, Valencia, Spain
[c]CCRI, Children's Cancer Research Institute, Vienna, Austria

Abstract. With the Directive 2011/24/EU on patients' rights in cross-border healthcare and the related delegated decisions, the European Commission defined a legal framework on how healthcare shall be organised by European Union (EU) member states (MS) where patients can move beyond the borders of their home country. Among other aspects, Article 12 of the directive is concerned with supporting MS with the development of so called European Reference Networks (ERN), dedicated to the treatment of "patients with a medical condition requiring a particular concentration of expertise in medical domains where expertise is rare". In the "European Expert Paediatric Oncology Reference Network for Diagnostics and Treatment" (ExPO-r-Net) project, the establishment of such an ERN in the domain of Paediatric Oncology is currently piloted. The present paper describes the high level use cases, the main requirements and a corresponding interoperability architecture capable to serve as the necessary IT platform to facilitate cross-border health data exchange.

Keywords. Integrating the Healthcare Enterprise (IHE), Secondary Use, Standardisation, Cloud Computing, Rare Diseases, Virtual Tumour Boards, Pseudonymisation

1. Introduction

Rare Diseases (RD) are illnesses which affect less than 5 in 10.000 persons. Due to the low number of cases, pooling resources and expertise from different centres is of particular concern for RD.

With the Directive 2011/24/EU on patients' rights in cross-border healthcare [1] and the related delegated decisions [2], the European Commission (EC) defined a legal framework on how healthcare shall be organised by Member States (MS) of the European Union (EU) for patients moving beyond the borders of their home country. Among other aspects, Article 12 of the above directive is concerned with the support of MS in developing so called European Reference Networks (ERN). ERNs are dedicated to the treatment of "patients with a medical condition requiring a particular concentration of expertise in medical domains where expertise is rare" [1]. In the "European Expert Paediatric Oncology Reference Network for Diagnostics and Treatment" (ExPO-r-Net) project, the establishment of such an ERN in the domain of Paediatric Oncology (PO) is

[1] Corresponding Author: Michael NITZLNADER, AIT Austrian Institute of Technology, Eduard-Wallnöfer-Zentrum 1, 6060 Hall in Tirol, Austria, E-Mail: michael.nitzlnader@ait.ac.at

currently piloted. The present paper describes the high level use cases, core requirements, the technological state of the art and a corresponding interoperability architecture. This architecture foresees capabilities and services needed by ERNs in general and by the ExPO-r-Net in particular, so as to accomplish their mission, including cross-border health data exchange.

2. Methods

The conception of the interoperability architecture for the ERN started with a detailed analysis of the objectives of the EC regulatory framework followed by a comprehensive literature search on the status quo of relevant IT solutions. The literature search was done utilising the Web of Science[2] and Science Direct[3] databases using combinations of the following keywords: European Reference Networks, Interoperability, Integrating the Healthcare Enterprise, IHE, epSOS, open NCP, Virtual Tumour Board, IT Infrastructure. Next, non-functional as well as functional requirements were derived from unstructured interviews of different experts representing the ExPO-r-Net community. Unstructured interviews and therefore open and no predefined questions were used because of the differences between the members of the ExPO-r-Net community. Subsequently, more detailed capabilities were defined and the extracted requirements adjusted accordingly to the use cases of the ExPO-r-Net community. The whole process of requirements analysis was performed accordingly to well-known standards in requirements engineering [3]. The evolving solution concept was presented to the ExPO-r-Net executive committee several times and refined accordingly based in the received feedback. This process lead to the following results:

1. technological state of the art
2. high level (most relevant) use cases
3. core requirements and finally the
4. solution architecture

which was verified concerning feasibility and the degree to which the requirements were met. Key results from these steps are presented in the following.

3. Results

3.1. State of the Art in Cross Border Health Data Exchange

A literature search for papers related to the topic “European Reference Networks” did yield only a few relevant hits, none of which dealt with IT infrastructure aspects in particular (beyond mentioning that such a dedicated IT infrastructure is needed). One paper proposed a minimum dataset definition for rare diseases, which may serve as a starting point in the definition of the data elements in the registries, supposed to be associated with each ERN [4].

As a consequence, we extended our search to publications that were concerned with cross-border healthcare data exchange in general, omitting the context of ERNs. This led

[2] http://apps.webofknowledge.com
[3] http://www.sciencedirect.com

to the following categories of solution approaches beyond those already discovered in [5]:

- ABCD-4-E - Secondary use of Healthcare data for research - IHE-based concepts that deal with a Secondary Use Approach as designed and piloted in ENCCA [5]
- Cross-border exchange of medical images for continuity of care in proton beam therapy [6]. In this paper, the cross border requirement was managed by IHE's Cross Community Access (XCA) profile.

3.2. High level Use Cases of the ExPO-r-Net

The requirements of the EC on ERNs and in particular the needs of the ExPO-r-Net community lead to the following high level use cases:

3.2.1. Survivorship Passport

The Survivorship Passport (SP) is a document intended to provide survivors in Paediatric Oncology with all essential information needed for optimal long-term care. For each child and adolescent cured of cancer an online tool shall provide instant access to his/her medical history. A self-generating document shall contain all details of the survivor's disease and treatment. Within the ExPO-r-Net project the SP shall be made available accordingly to Article 12 of the Directive 2011/24/EU on patients' rights in cross-border healthcare.

3.2.2. Virtual Tumour Board

Within the ExPO-r-Net project a Virtual Tumour Board (vTB) in the Paediatric Oncology domain shall be developed based on interoperability concepts, in particular IHE. The VTB shall support sharing of radiology images, radiotherapy treatment plans and support the discussion of cases within online clinical conferences. The purpose of a vTB is to provide expert advice independent of geographical location.

3.2.3. Data Provision to Clinical Research

Provision of data collected for healthcare purposes to clinical research is one of the objectives for ERNs. Whereas professionals always need to be identified directly with personal identifiers, i.e. by their names, for patients, two different types of identification methods need to be provided:

- Healthcare: Personal identifiers (names, social insurance numbers, patient codes, …),
- Clinical research: Pseudonyms, generated and managed by a central pseudonymisation service.

3.2.4. Multimodal communication

Communication is key to facilitate the aim of ERN, i.e. optimised patient treatment, potentially in a multicentre setting. Therefore, a variety of different communication capabilities need to be provided and tailored to the task and group of participants.

- Telephone (for general ad hoc communication)
- Email and instant messaging (for administration collaboration within teams)

- Discussion forums in dedicated groups (for store-and-forward case evaluation)
- Video Conferencing (for vTB clinical conferences)
- Social networking (for teaching and organising events)
- Health data transfer (for preparing vTB and in case of referrals)

Some of the channels need to be sufficiently secured to allow for the communication of sensitive data, e.g. health data.

3.3. Requirements and how they will be met

The non-functional requirements are similar to those already discovered in [5]. **Table 1** gives an overview on the main functional requirements considering ExPO-r-Net, respectively, and how they can be met by the proposed concept.

3.4. ERN solution architecture

The ERN solution architecture was designed to meet the given objectives and requirements, as well as to link healthcare and research.

Figure 1 schematically illustrates the main components. The national contact points (NCPeH) of each country where a participant of the ExPO-r-Net community resides are connected via epSOS and allow the exchange of patient summaries. Documents like the

Table 1. The functional requirements and their fulfilment

		Requirement Description	**Met by …**
Functional requirements in the Healthcare as well as in the Clinical Research Domain	F1	ERN development appropriate to the specifications of the EC	... a development accordingly to Article 12 of the Directive 2011/24/EU on patients' rights in cross-border healthcare
	F2	Utilisation of standards and terminologies for a structured and consistent exchange of data	… the application of IHE and agreed dataset types.
	F3	Cross-patient as well as cross-domain request for documents concerning the ExPO-R-Net high level use cases (3.2)	… the application of epSOS respectively IHE including appropriate adaptations.
	F4	Additional privacy protection in the Clinical Research Domain	… the application of a context-specific pseudonymisation provided by the EUPID-Service [7], separation of metadata and patient data as well as restrictions on the aggregation and output of data.
	F5	Definition and control of transparent policies for secondary use of health data	… a governance policy that restricts registration of documents to previously registered patients. Data manipulation is generally audited.
	F6	Distributed solution approach	… the ERN solution architecture (see 3.4).
	F7	Linking clinical research systems as well as healthcare systems using standardised interfaces	… the application of IHE, epSOS and concepts of ABCD-4-E.
	F8	Respecting data protection regulations	… the security concepts of epSOS and ABCD-4-E [7, 8].

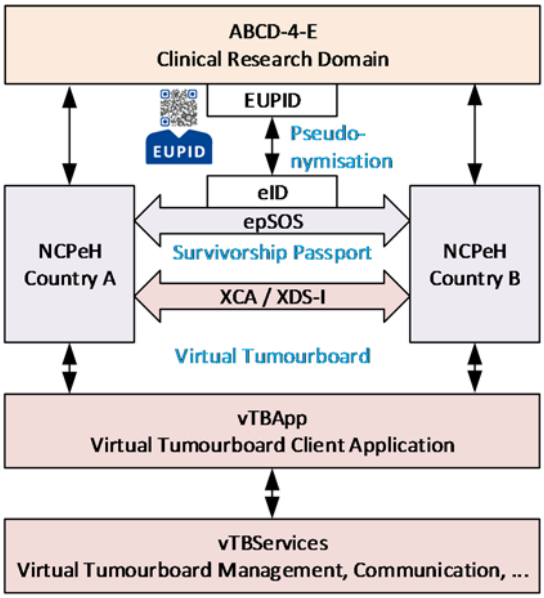


Figure 1. ERN solution architecture.

Survivorship Passport (SP) are stored within one of the national electronic health record systems and transferred to another country on demand. On its way to the recipient, the document is translated by the epSOS terminology services. As the vTB also handles DICOM images, Radiotherapy plans etc. a cross-border communication cannot be facilitated based on epSOS alone. Therefore, additional XCA / XDS-I connectors are used to provide vTB materials also for vTB participants on the remote site. The vTB itself is facilitated by a web-based application (vTBApp) and associated services for the management, the communication etc. of the vTB (vTBServices). The vTBApp implements the workflow defined by the IHE XTB-WD profile [9]. The workflow document and the referenced documents are stored in one of the repositories connected to the NCPeH (Figure 2).

To provide the data of the ERN ExPO-r-Net also for clinical research a platform for secondary use of healthcare data in clinical research developed in the ENCCA project [10] [5] is connected to the NCPeH of each country so a simple transfer of documents to ABCD-4-E is possible. In the course of transferring data to ABCD-4-E the patient identifier (eID) is exchanged with an EUPID-based pseudonym [7] that is associated to eID. The generated pseudonym is used for data registration. Before a registration can be done, a data clearance step is required. All information contained in the data giving a hint on the patient's identity will be removed prior to the registration and patient consent will be required.

4. Discussion

Some elements of the presented IT infrastructure and tools have already been conceived in the context of the **E**uropean **N**etwork for **C**ancer Research in **C**hildren and **A**dolescents (ENCCA) project [10]. Based on the tight link between research and

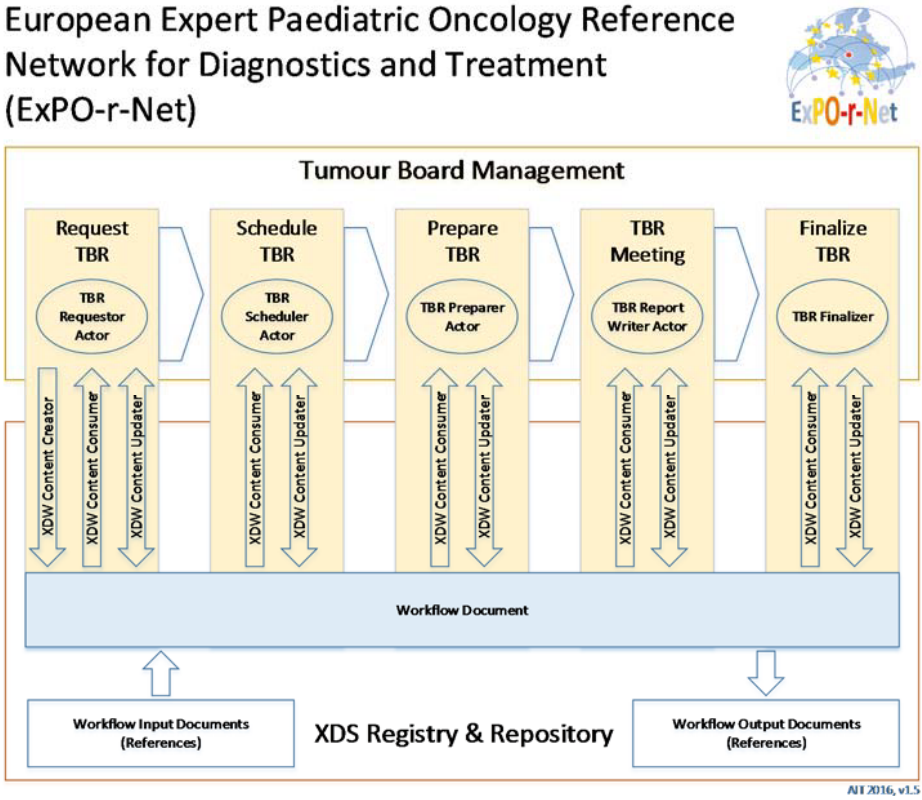


Figure 2. The vTBApp workflow and associated actors and transactions

healthcare in the PO community, some of the concepts and developments of this so called Advanced Biomedical Collaboration Domain for Europe (ABCD-4-E, [5]) were evaluated for their relevance and usefulness for the PO-ERN as well. In this respect, we did not have to start from scratch, but could draw on previously gained expertise related to PO network interoperability concepts (IHE) and tools like the EUPID pseudonymisation service [7] and the Survivorship Passport [11] as well as on a previously developed solution for cross border ion beam radiotherapy [6].

The present approach utilises well established, industry supported and standards-based technologies for the main building blocks of the IT infrastructure.

A still ongoing debate refers to the necessity to link the different ERNs - which are supposed to be established in the next couple of years – with each other. There is the general notion that some resources and – in particular – IT tools could be useful for some if not all ERNs, probably with some customisation. In addition, the need for cooperation between ERNs may arise from the expectation that a significant number of patients may come into contact with more than one ERN. This may be the case, for example, for the PO ERN and a potential ERN for rare cancers in adults. Furthermore, physicians and researchers in larger centres of expertise may also be associated with more than one ERN in their professional activities. Finally, all ERNs will share some regulatory and governance aspects on the national as well as the European level. All those aspects require to foresee ways to link ERNs and, as a consequence, the corresponding IT

infrastructures. Here, an IHE based interoperability solution can undoubtedly provide a versatile solution by means of the Cross Community Access concept [12].

Acknowledgement

The research leading to these results has received funding from the European Union's Health Programme (2008-2013), grant agreement Nr. 2013 12 07.

References

[1] Directive 2011/24/EU of the European Parliament and the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare, (2011).
[2] Commission Delegated Decision of 10 March 2014 setting out criteria and conditions that European Reference Networks and healthcare providers wishing to join a European Reference Network must fulfil (2014).
[3] K. Pohl, C. Rupp. Requirements engineering fundamentals : a study guide for the certified professional for requirements engineering exam, foundation level, IREB compliant. Second edition. ed. xvii, 163 pages p.
[4] R. Choquet, M. Maaroufi, A. de Carrara, C. Messiaen, et al. A methodology for a minimum data set for rare diseases to support national centers of excellence for healthcare and research. J Am Med Inform Assn. 2015 Jan;22(1):76-85. DOI: 10.1136/amiajnl-2014-002794.
[5] M. Nitzlnader, M. Falgenhauer, C. Gossy, G. Schreier. Architecture for an advanced biomedical collaboration domain for the European paediatric cancer research community (ABCD-4-E). Studies in health technology and informatics. 2015;212:167-74.
[6] M. Drobics, G.u. Rauchegger, R. Mayer, G.u. Schreier, editors. Basiskonzept eines transnationalen Bilddatenmanagements fur die Ionentherapie. Tagungsband der eHealth2011 26-27Mai 2011; Wien; 2011. publications\2011\2011-05-26-27_eHealth2011\RegIonCo_final\drobics2011a.pdf: OCG.
[7] M. Nitzlnader, G. Schreier. Patient identity management for secondary use of biomedical research data in a distributed computing environment. Studies in health technology and informatics. 2014;198:211-8.
[8] M. Nitzlnader. Konzept eines IHE-basierten Medizinischen Forschungsnetzwerkes in der Cloud [Master Thesis]: UMIT - The Health & Life Sciences University; 2014.
[9] I. International. IHE Patient Care Coordination (PCC), Cross-Enterprise Tumor Board Workflow Definition (XTB-WD): IHE International, http://www.ihe.net/uploadedFiles/Documents/PCC/IHE_PCC_Suppl_XTB-WD.pdf, last access: 2016-02-10.
[10] ENCCA - European Network for Cancer Research in Children and Adolescents, http://www.encca.eu, last access: 2013-11-05.
[11] R. Haupt, C. Dellacasa, S. Karner, L. Kremer, et al. The Survivorship Passport for Long-Term Care of Childhood Cancer Survivors. An Initiative of the European Network for Research on Cancer in Children and Adolescents (Encca). Pediatr Blood Cancer. 2014 Dec;61:S306-S.
[12] I. International. IHE IT Infrastructure (ITI) - Technical Framework. Volume 2b (ITI TF-2b) - Transactions Part B: IHE International; 2013. p. 296.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-46

An IT-Supported Evaluation Tool for Biobanks Based on International Guidelines to Improve the Biosample Quality

Sabrina B. NEURURER[a,1], Philipp HOFER[a] and Georg GÖBEL[a]
Department of Medical Statistics, Informatics and Health Economics, Medical Universit of Innsbruck, Innsbruck, Austria

Abstract. Background: The quality of samples stored within a biobank relies on the specimen collection, the transportation, the pre-analytical processing and the long-term storage. Standard Operating Procedures (SOPs) are essential tools to guarantee the quality of samples. Objectives: The aim of this paper is to present an IT-supported tool (Pre-An Evaluation Tool) that allows assessing the compliance of current pre-analytical procedures (defined in SOPs) of a biobank with international guidelines. The Pre-An Evaluation Tool was implemented based on CEN technical specifications for pre-analytical procedures using REDCap. Results: The data collection instrument of the Pre-An Evaluation tool consists of more than 250 items related to the CEN technical specifications. In order to create a dynamic questionnaire, items following a branching logic were implemented. Conclusion: The Pre-An Evaluation tool is a user-friendly tool that facilitates the assessment of the coverage of the CEN technical specifications by specific SOPs. This tool can help to identify gaps within SOPs and therefore contribute to the overall quality of biological samples stored within a biobank.

Keywords. Biobank, sample handling, evaluation

1. Introduction

Biobanks store biological samples together with related clinical data, informed consent declarations and information regarding the pre-analytical processes and storage conditions. Biological samples are used for research purposes that aim to integrate biological findings, genomic data, molecular technologies and phenotype data in order to improve the knowledge of human diseases and to develop new diagnostic and therapeutic approaches [1]. High quality samples are an essential quality indicator for (bio-) medical research outcomes [2]. The quality of samples relies on the specimen collection, the pre-analytical processing as well as the long-term storage. Bad sample quality can result in inaccurate data and, as a consequence, in compromised research outcomes [3]. For instance, transport delays or deviations can lead to low quality DNA or RNA and therefore to low quality data. Also, degradation of enzymes and nucleic acids might be influenced by the duration of tissue fixation [4]. Standard Operating Procedures (SOPs) that provide a strong standardization of sample handling help to

[1] Corresponding Author: Sabrina B. Neururer, Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Schöpfstraße 41/1, 6020 Innsbruck, E-Mail: sabrina.neururer@i-med.ac.at.

assure quality in a laboratory environment. However, SOPs are not only defined to describe sample handling procedures but also the conception of experiments and analysis [5]. Depending on the size, the background and the certification of a laboratory, SOPs are often documented at different levels of detail. The implementation of harmonized standard operating procedures for biobanking is a key objective of the Austrian national node (BBMRI.at) of the pan-European Biobanking and BioMolecular Resource Research Infrastructure (BBMRI) [6]. In order to trigger such a harmonization process, the current state had to be assessed. Therefore, it was necessary to evaluate the compliance of the current pre-analytical procedures in Austrian biobanks with internationals standards and guidelines, such as the WHO/IARC, OECD, or the CEN guidelines.

The aim of this paper is to present an IT-supported tool (Pre-An Evaluation Tool) that allows assessing the compliance of current pre-analytical procedures of a biobank with international guidelines. This tool should enable to gather information about to which extent the currently installed SOPs cover the requirements defined within international guidelines in a quick and useful manner using a questionnaire. The tool should provide a framework to assess information at different levels of granularity in order to subsequently identify commonalities and differences between SOPs and to put a step towards harmonization and standardization of SOPs.

The remainder of this paper is organized as follows: Section 2 includes a description of the material and methods involved in the IT tool development. In Section 3, the results are described in detail. Section 4 discusses the obtained results and provides an outlook and a conclusion.

2. Material and Methods

This section describes the material and methods used to implement the proposed Pre-An Evaluation Tool.

2.1. International Standards and Guidelines

Currently, several different international standards and guidelines exist that provide recommendations regarding the collection, reception, processing, storage and retrieval of high-quality samples in biobanks. Examples for such standards and guidelines are the WHO/IARC guidelines, the OECD guidelines or the CEN technical specifications. After having analyzed all of these standards, it was decided to use the CEN technical specifications as the basis for the implementation of our Pre-An Evaluation Tool. These provide concrete guidelines for handling, documenting and processing samples of high quality for the following sample types: (1) venous whole blood, (2) serum, (3) plasma, (4) urine, (5) snap-frozen tissue, (6) formalin-fixed and paraffin-embedded (FFPE) tissue, and (7) PAXgene-fixed and paraffin-embedded tissue. Specifically, the following guidelines listed in Table 1 were used to assess the compliance of current pre-analytics based on important recommendations to enhance the quality of fluid and tissue samples.

The guidelines were analyzed in detail in order to identify a common structure that could be used for comparison with the current pre-analytical procedures established within a biobank and the implementation of the Pre-An Evaluation Tool.

Table 1. List of the CEN Technical Specifications used for the implementation of the Pre-An Evaluation Tool

CEN Technical Specifications
CEN/TC 140 Molecular in-vitro diagnostic examinations - Specifications for pre-examination process for blood - genomic DNA (Version 2013/10)
CEN/TS 16835-1 Molecular in-vitro diagnostic examinations - Specifications for pre-examination process for venous whole blood – Part 1: Isolated cellular RNA (Version 2015/07)
CEN/TC 140 Molecular in-vitro diagnostic examinations - Specifications for pre-examination process for metabolomics in urine, venous blood serum and plasma (Version 2015/01)
CEN/TC 140 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Isolated genomic DNA (Version 2014/12)
CEN/TC 140 Molecular in-vitro diagnostic examinations - Specifications for pre-examination process for metabolomics in urine, venous blood serum and plasma (Version 2015/01)
CEN/TC 140 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Isolated circulating cell free DNA (Version 2014/12)
NVN-CEN/TS 16826-1 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for snap frozen tissue – Part 1: Isolated RNA (Version 2015/09)
CEN/TS 16827-3 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for FFPE tissue – Part 2: Isolated DNA (Version 2015/08)
NVN-CEN/TS 16826-2 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for snap frozen tissue – Part 2: Isolated proteins (Version 2015/09)
CEN/TS 16827-1 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for FFPE tissue – Part 1: Isolated RNA (Version 2015/08)
NVN-CEN/TS 16827-2 Molecular in-vitro diagnostic examinations - Specifications for pre-examination processes for FFPE tissue – Part 2: Isolated proteins (Version 2015/09)

2.2. Research Electronic Data Capture (REDCap)

The basic structure of the Pre-An Evaluation tool was developed by extracting the most important steps and requirements for pre-analytical processes in biobanking from the CEN technical specifications (e.g. documentation of information about the sample donor ID, documentation of protocol deviations). These requirements were first listed in a Microsoft Excel matrix in order to identify a common structure for the several different material sub-types from the various CEN technical specifications. In a workshop, this matrix was discussed by stakeholders from different disciplines (e.g. quality managers, computer scientists, medical experts). The stakeholders approved of the matrix. Thereafter a suitable framework for the implementation needed to be identified. Such a framework had to fulfill the following requirements, which were pre-defined by the work-package leader and his team and committed by the management committee of the project (consisting of different stakeholders such as medical experts, computer scientists, quality managers):

- web-based
- secure
- dynamic/dependent fields
- easy to implement
- user-friendly
- scalable
- flexible

In order to facilitate the evaluation and the quick analysis of the results avoiding media disruption, the evaluation tool needed to be IT-based. As we wanted to use a tool that needs to be installed and implemented in one place while granting access for all national partners of the BBMRI.at project, we decided to use a web-based tool. Whenever a web-based tool is selected, a strong focus has to be put on security aspects. To develop and implement a user-friendly evaluation tool, which is an import prerequisite for IT-based tools, the usage of dependent fields (fields that depend on the input given by another (previous) field) was strongly recommended. The evaluation tool should also be flexible and scalable in order to allow and motivate re-usage by non-IT specialists for similar purposes.

The Research Electronic Data Capture (REDCap) framework was used for the implementation of our "Pre-An Evaluation Tool". REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing (1) an intuitive interface for validated data entry, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for importing data from external sources [7]. REDCap was developed by Paul Harris and his colleagues at the Vanderbilt University [7]. Today the REDCap consortium consists of more than 1,500 active institutional partners in more than 90 countries. The Department of Medical Statistics, Informatics and Health Economics of the Medical University of Innsbruck is one of the members of this REDCap consortium and runs this system in order to facilitate several different studies, where data is collected and managed (e.g. using electronic case report forms). REDCap is a PHP-based framework using a MySQL database. It offers basic functionalities that are important for the Pre-An Evaluation Tool in order to assess the compliance of current pre-analytical procedures with international guidelines. It is easy to implement and allows an individual design of data collection instruments by a point-and-click approach. Apart from common option for data collection fields, such as text boxes, drop-down fields, radio buttons, sliders, it also offers calculated fields and facilitated a so-called branching logic, which enables the implementation of dependent fields and therefore supports the dynamic compilation of the final questionnaire. For example, if a participant is male (sex), there is no need to answer questions concerning the participant's pregnancy. These questions are not displayed when using a branching logic that activates a field (e.g. concerning pregnancy) only if the variable sex is female ([sex]=female).

3. Results

The Pre-An Evaluation tool based on the aforementioned CEN technical specifications was implemented using REDCap and consisted of one data collection instrument. The Pre-An Evaluation Tool was logically subdivided into three parts: (1) affiliation and material type information, (2) steps performed outside the laboratory (e.g. documentation of information about the sample donor, transport requirements), and (3) steps performed inside the laboratory (e.g. storage requirements, selection of storage containers). The first part collects basic information about the partner completing the questionnaire in order to facilitate a proper data analysis for each partner. Additionally, the partners have to select the specific material types and sub-types for which the questionnaire was completed (see Fig. 1). The second and third part of the Pre-An Evaluation tool contained information according to the CEN technical which also offer a subdivision according to processes

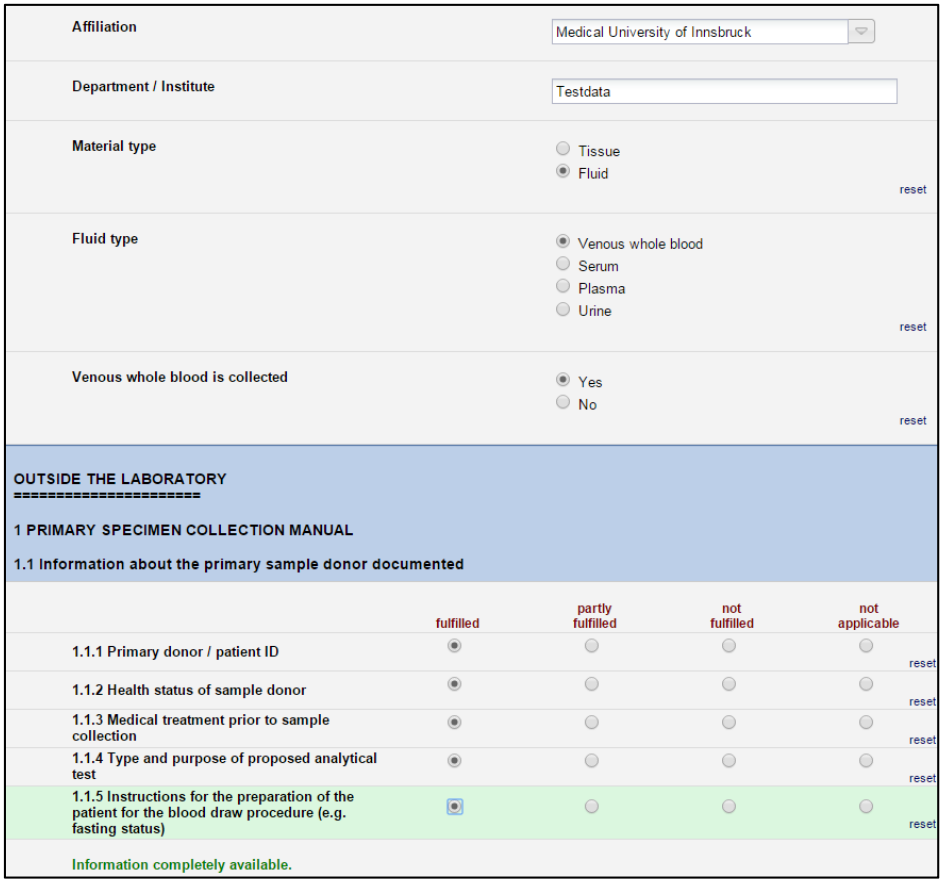


Figure 1 Excerpt from the implemented Pre-An Evaluation Tool using REDCap

outside and inside the laboratory. All material types (tissue, fluid) and sub-types covered by the CEN technical specifications (see 2.1) were taken into account for the Pre-An Evaluation tool. This selection of the type and the material sub-type influenced the content of the subsequent question items as the tool provides flexible content depending on the material sub-types selected. This allowed us to bring only such questions into the focus that are relevant for the specific material sub-type. We implemented the data collection instrument of the Pre-An Evaluation Tool by providing four pre-formulated answers (single choice): (1) fulfilled, (2) partly fulfilled, (3) not fulfilled, and (4) not applicable. These answers described the extent to which a specific recommendation of the CEN technical specification is fulfilled by a biobank (see Fig. 1).

For example, the “Testdata” Department of the Medical University of Innsbruck completed the questionnaire stating that the Primary donor / patient ID of a specimen is documented. As all required information regarding the “Information about the primary sample donor” is documented, we wanted to provide the user, who is completing the questionnaire, with immediate feedback to which extent the requirement of the CEN guidelines are covered by his/her SOPs. Therefore, we implemented a mechanism for several batteries of questions that provides such feedback based on their input (fulfilled,

Information completely available.
([q1] <> "") and ([q2] <> "") and ([q3] <> "") and ([q4] <> "") and ([q1] = 1) and ([q2] = 1) and ([q3] = 1) and ([q4] = 1)

Information partly available.
([q1] <> "") and ([q2] <> "") and ([q3] <> "") and ([q4] <> "") and !((([q1] = 1) and ([q2] = 1) and ([q3] = 1) and ([q4] = 1)) or (([q1] = 3) and ([q2] = 3) and ([q3] = 3) and ([q4] = 3)) or (([q1] = 4) and ([q2] = 4) and ([q3] = 4) and ([q4] = 4)))

Information not available.
([q1] <> "") and ([q2] <> "") and ([q3] <> "") and ([q4] <> "") and ([q1] = 3) and ([q2] = 3) and ([q3] = 3) and ([q4] = 3)

Information not applicable.
([q1] <> "") and ([q2] <> "") and ([q3] <> "") and ([q4] <> "") and (([q1] = 4) and ([q2] = 4) and ([q3] = 4) and ([q4] = 4))

Figure 2 Branching logic for a battery of four questions (1…fulfilled, 2…partly fulfilled, 3…not fulfilled, 4…not applicable)

partly fulfilled, not fulfilled or not applicable). We used separate fields that indicate, whether the information is (1) completely available, (2) partly available, (3) not available or (4) not applicable. For the implementation of this mechanism, the branching logic was used. An example for the branching logic related to the information summary for an array of four questions is given in Fig. 2.

Altogether the questionnaire of the Pre-An Evaluation Tool consisted of more than 300 items. The number of items for each material sub-types related to part 2 and part 3 of the questionnaire (outside and inside the laboratory) are listed in Table 2. Within this table, the items related to the affiliation and the material types as well as the special items for the immediate user feedback (information completely available/partly available/not available/not applicable) were not displayed.

Validation of the Pre-An Evaluation Tool: In order to guarantee the validity, usefulness and good usability of the Pre-An Evaluation Tool, it was iteratively tested and improved. First, it was tested by the implementers. Then, a pilot version was sent to the major BBMRI.at partners (Medical University of Innsbruck, Medical University of

Table 2. Number of items of the Pre-An Evaluation Tool for each material sub-type

Material sub-type	Number of items	
	Part 2: Outside the laboratory	Part 3: Inside the laboratory
Venous whole blood	18	8
Serum	17	15
Plasma	20	18
Urine	15	13
Snap-frozen tissue	14	27
FFPE tissue	15	33
PFPE tissue	15	30

Vienna, Biobank Graz, University of Veterinary Medicine Vienna, Paracelsus Medical University). They were asked to provide feedback on the content as well as on the usability of the tool. No training or expertise is required to use the tool. The users were provided with short instructions on the tool and were able to use it without any problems. The pilot version of the Pre-An Evaluation Tool was thereafter revised according to feedback of the participants. The reported issues (e.g. splitting singe question items into two items) lead to minor revisions of the tool. Then, a second pilot evaluation round was triggered including the same partners which lead again to a revision of the tool (mainly typo fixing). Finally, the approved version of the Pre-An Evaluation Tool was provided and will be used for future evaluation. The tool also offers a reporting mechanism, which facilitates the export of the data collection to various different statistical packages (e.g. SPSS, STATA, R, SAS). This allows performing an assessment and comparison of the different participating institutions after finalizing the evaluation phase.

4. Discussion

This paper describes the implementation of the Pre-An Evaluation Tool that facilitates the assessment of the compliance of current pre-analytical procedures within a biobank with international guidelines in biobanking. The basic structure of the Pre-An Evaluation Tool is pre-defined by the CEN technical specifications. We decided to use the (upcoming) CEN guidelines as they are concrete guidelines for handling, documenting and processing samples, and since they are going to become international ISO standards within the next few years. They provide important recommendations to ensure high quality for biological samples which is a major prerequisite for high-quality research outcomes.

The Pre-An Evaluation Tool itself was implemented using REDCap. We decided to use this framework as it fulfills all of our pre-defined requirements. It is an internationally used framework for capturing and managing data. REDCap allowed us to implement the Pre-An Evaluation Tool in a highly dynamic and customized manner, which offers the future survey participants maximum comfort, when completing the survey.

A major point of discussion was not related to the implementation of the Pre-An Evaluation Tool itself, as it is flexible and can be adjusted and adapted easily but to the pre-defined answers of the questionnaire. One could state that - according to a specific SOP - for example the sample donor's name is either documented (fulfilled), not documented (not fulfilled) or, for certain reasons, not applicable (e.g. for animal biobanks) but it certainly cannot be partly fulfilled. Even though we were aware of this point, we decided to provide such an answer option for the following reason: Within the BBMRI.at project, there exist several different types of biobanks. There are several centralized biobanks having a common quality management system and SOPs, but there are also many decentralized sample collections established for specific research study purposes, which don't have shared SOPs or quality management systems. Therefore, it is not always possible to answer "yes" or "no" for a partner, sometimes it is "partly fulfilled" (some collections' SOPs provide the required information, others don't).

Altogether we conclude that the Pre-An Evaluation tool is a user-friendly tool that facilitates the assessment of the coverage of the CEN technical specifications by specific SOPs. This tool can help to identify gaps within SOPs and therefore contribute to the overall quality of biological samples stored within a biobank.

Currently, we are conducting an assessment of the compliance of current pre-analytical processes with international guidelines for Austrian biobanks participating in the BBMRI.at project using the proposed Pre-An Evaluation Tool. The presented tool is highly flexible and can be adapted in order to meet other requirements related to the quality assurance of pre-analytical procedures in biobanking.

The usage of the presented Pre-An Evaluation Tool as a basis for a self-assessment tool for quality management in biobanking is discussed on an international level within the BBMRI-ERIC[2] infrastructure.

Acknowledgment

This work was supported by the Austrian Biobanking and BioMolecular resources Research Infrastructure (BBMRI.at) funded by the Austrian Federal Ministry of Science, Research and Economy (BMWFW GZ 10.470/0016-II/3/2013).

References

[1] H. U. Prokosch, a Beck, T. Ganslandt, M. Hummel, M. Kiehntopf, U. Sax, F. Uckert, and S. Semler, "IT Infrastructure Components for Biobanking.," Appl. Clin. Inform., vol. 1, no. 4, pp. 419–29, 2010.

[2] D. Simeon-Dubach and A. Perren, "Better provenance for biobank samples," Nature, vol. 475, no. 7357, pp. 454–455, 2011.

[3] D. Simeon-Dubach, A. D. Burt, and P. a. Hall, "Quality really matters: The need to improve specimen quality in biomedical research," J. Pathol., vol. 228, pp. 431–433, 2012.

[4] P. J. T. a Groenen, W. a M. Blokx, C. Diepenbroek, L. Burgers, F. Visinoni, P. Wesseling, and J. H. J. M. van Krieken, "Preparing pathology for personalized medicine: Possibilities for improvement of the pre-analytical phase," Histopathology, vol. 59, no. 1, pp. 1–7, 2011.

[5] G. Marko-Varga, Á. Végvári, C. Welinder, H. Lindberg, M. Rezeli, G. Edula, K. J. Svensson, M. Belting, T. Laurell, and T. E. Fehniger, "Standardization and utilization of biobank resources in clinical protein science with examples of emerging applications," J. Proteome Res., vol. 11, no. 11, pp. 5124–5134, 2012.

[6] C. Viertler and K. Zatloukal, "Biobanken und Biomolekulare Ressourcen Forschungsinfrastruktur (BBMRI). Bedeutung für die pathologie," Pathologe, vol. 29, no. SUPPL. 2, pp. 210–213, 2008.

[7] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support.," J. Biomed. Inform., vol. 42, no. 2, pp. 377–81, Apr. 2009.

2 www.bbmri.eu

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-54

Towards a Holistic Biomedical Information Platform for Primary and Secondary Use Settings

Gerd SCHNEIDER[a,1], Björn SCHREIWEIS[a], Raluca PAHONTU[a], Theresia EICHNER[a] and Björn BERGH[a]
[a]*Heidelberg University Hospital, Center for Information Technology and Medical Engineering, Heidelberg, Germany*

Abstract. Background: Clinical information is often used for biomedical research. Data warehouses can help providing researchers with data and the opportunity to find eligible participants for clinical trials. Objectives: To define an information platform for healthcare and biomedical research based on requirements by clinicians and researchers. Methods: Interviews with clinicians, researchers, data privacy officers, IT and hospital administration combined with a questionnaire sent to 60 medical departments at our hospital were conducted. Results: Resulting requirements were grouped and a platform architecture was designed based on the requirements. Conclusion: Requirements lead to a single platform supporting both, patient care and biomedical research.

Keywords. Hospital Information System, Big Data, In-Memory Database, Analytics, Primary Use, Secondary Use, Information Platform Architecture.

1. Introduction

Available clinical information within the healthcare environment strongly increased recently [1, 2]. This amount of data offers enormous potential to support various decision processes along the clinical value chain. Both the primary use (patient care) and the secondary use (e.g. biomedical research) of data will profit from real-time data preparation and data analysis. Additionally, a shift from hypothesis-driven to data-driven research including literature-based evidence is emerging in the direction of practice-based evidence which has to be considered [3]. With new technological possibilities arises an increasing need for the integration of gene expression data with clinical data intended for improvements in personalized medicine [4]. Due to the current developments in personalized medicine and collaborative research the competence of professional provision and visualization of clinical data has become a critical success factor [5]. New methods and tools are required as data collections are becoming so large and complex that traditional methods and tools can no longer process them [2]. The challenge is to establish a holistic biomedical information platform based on a generic architecture by concurrently fulfilling several requirements respective to the information technology. The main difficulty of classical source systems like Hospital Information System (HIS), Laboratory Information System (LIS) and Radiology Information System

[1] Corresponding Author: Gerd Schneider, Heidelberg University Hospital, Center for Information Technology and Medical Engineering, Speyerer Str. 4, 69115 Heidelberg, Germany, E-Mail: gerd.schneider@med.uni-heidelberg.de.

(RIS) is the transactional design. These systems are neither offering an integrated view on patient data nor are they providing capabilities for (real-time) analyses and aggregation of data (e.g. sensor data from the operating room) [6].

Thus, we introduce a new platform providing a holistic view on the patient by being 'primary use enabled' and 'secondary use compliant'. This platform allows physicians to easily compare a current patient to similar cases and to draw conclusions for patient care (practice-based evidence [3]). Furthermore, the strategy of our hospital is to have a single point of truth for all kinds of requests (e.g. data for quality management, biomedical research, patient care, process management and innovations in general). Additionally, this single point of truth has to preserve patient data privacy and allow for integration with current developments in the field of patient empowerment. Usually a data warehouse is established for realizing this scope [2]. Such a clinical data warehouse integrates different data sources and enables physicians to get a holistic view of the patient or otherwise to compare the patient´s course of disease to similar courses of diseases to improve the therapy. In addition, delivering information of appropriate patient cohorts for recruitment is an important functionality.

The main goal of this work is to provide a platform architecture utilizable in primary and secondary use scenarios. First, the functionalities and components required are identified and described as part of a generic platform architecture. Afterwards, our initial steps towards implementing the architecture and the first prototype are displayed.

2. Methods

First, requirements were analyzed and the derived functionalities were categorized. Afterwards, these categories were used to define the modules of a generic architecture for a patient care and biomedical research platform.

2.1. Requirements analysis

A combination of methods was used to identify requirements for an information platform used in patient care and biomedical research.

First, a literature analysis and unstructured interviews were performed to obtain demanded functionalities from clinical and research users, IT and hospital administration, and data privacy regulations. Clinicians from the departments of cardiology, oncology, pathology, surgery, radiotherapy and radiology were interviewed. Researchers working in the same departments were interviewed as well. For data privacy requirements national, regional and special hospital laws were analyzed. The local data privacy officers' expert opinion was taken into account, too.

Additionally, a questionnaire was sent to all 60 medical departments of the hospital. The questionnaire included questions on data storage and management (Where are data currently stored? Centralized, internally decentralized, externally decentralized), data sources (What data sources are currently used? patient care, biomedical research, animal models or "omics"), data analytics (In which department are data analyzed? Which partners help analyzing data?), cooperative data usage (Are you using certain data for cooperative research?), experiences in biomedical research IT infrastructure (What are the experiences concerning a biomedical research IT infrastructure that your department could share?), requirements regarding IT, medical informatics and bioinformatics (What would your department expect from IT, medical informatics and bioinformatics?), and

future research topics relevant for a biomedical research platform (Can you list future research topics relevant for a biomedical research IT platform?).

2.2. Requirements categorization and architecture modularization

Five experts in the fields of computer science, biomedical informatics and business informatics grouped the functionalities identified by the requirements analysis into categories. Initially, each expert performed the categorization independently. Afterwards the individual results were compared and differences were discussed to find a common categorization all five experts could agree on. The resulting categories of requirements are:

- Data sources: where clinical and research relevant patient information is stored
- Data extraction: to retrieve the data from source systems
- Data storage: to store retrieved data integrated in a single system
- Data analysis: to allow for analysis on the integrated data pool
- Data Distribution: to use data for specific research questions
- Research Support Services: to be conform with legislation and regulation
- Miscellaneous

Based upon these categories and the associated requirements the generic architecture of the biomedical information platform was derived in iterative software architecture design workshops. The previous categorization fostered the modularization of the architecture as we aimed for a highly modular information platform to ensure exchangeability of individual functionality.

3. Results

3.1. Requirements

One fundamental requirement resulting from unstructured interviews is a single hospital-wide clinical data warehouse. Because of different data sources a technical infrastructure is required to integrate and harmonize data during the process of data extraction. Therefore, various methods and tools are required to extract and interpret data automatically and to improve data quality. Furthermore, it is necessary to offer possibilities of big data storage e.g. like a research cloud, and archiving options e.g. a central archive for “omics” data. Another requirement is the possibility to analyze unstructured documents like medical reports. In most cases the main information is only located in unstructured documents, which makes it necessary to extract structured data by using text-mining algorithms.

A mandatory requirement for biomedical research emerging from literature is patient data privacy in compliance to the German data privacy guidelines. As a consequence, different components for consent management, identity management and pseudonymization/anonymization are required.

The questionnaire was filled in and returned by 50 of the 60 medical departments (response rate 83.33%). Most needed by the clinicians of these different departments (n=50) were technical (80%) und methodological (66%) support for data management, central data management (76%) and methodological (74%) support for data analysis. Responses included for example a storage of large data sets, a technical infrastructure for

Table 1. Entire list of categorized requirements

Requirement	Category
Storage for large data sets	Data storage
Central data warehouse	Data storage
Central "omics" archive	Data storage
Data integration	Data sources, extraction and storage
IT infrastructure for research networks	Data distribution, research support services
Consent management	Research support services
Identity management	Research support services
Pseudonymization/Anonymization	Research support services
High performance computing	Analytics
Machine learning	Analytics
Genome sequencing and gene expression	Analytics
Computational image analytics	Analytics
Data / text mining	Data extraction, analytics
Harmonization of different data formats	Data extraction
Design of experiments	Miscellaneous

integration of data, an IT infrastructure for research networks, and access to high performance computing. Furthermore, machine-learning algorithms for improvement of research, analysis of genome sequencing and gene expression, harmonization of different data formats, support for designing experiments, and computational analysis of images were requested.

The resulting categorization of all requirements is listed in Table 1.

3.2. Generic Platform Architecture

Based upon the described requirements and requirement categories a generic and modular architecture for a holistic biomedical information platform was designed. As a central and hospital-wide infrastructure the platform should serve both clinical care and biomedical research in a data privacy compliant way. The main focus is on an integrated, digital view on patient and non-patient related data to support personalized care. A wide variety of data sources distributed across the entire hospital has to be merged into a central data pool. Based upon this data pool, data analytics and visualization for decision support with regard to clinical care can be achieved. Furthermore, this data pool can be used to provide anonymized or pseudonymized clinical data for biomedical research meeting strict legal and data privacy regulations. The generic platform architecture is shown in Figure 1. It is divided into three main subplatforms: Data and Integration Platform, Analytics Platform, and Service Platform.

The *Data and Integration Platform* encompasses the components Data Harvesting, Support Systems and Data & Processing Platform. Data Harvesting enables data extraction from and data access to various clinical/medical data sources (Data Provider). These data can be transferred or linked to the Data & Processing Platform. During transfer/linkage the data can be passed through diverse steps of refinement. As part of this process, methods of text extraction, semantic annotation, and data or information fusion can be applied. According to the type of a connected data source and the possibility and authorization to replicate the provided data three different connection modes are feasible. Data sources allowing physical data replication (Materialized Sources) can be connected either via an Extract-Transform-Load (ETL/Batch Layer) or via a real-time data stream procedure (Speed Layer). Conventional ETL sources are e.g. HIS, LIS and RIS along with other database-based and transactional systems. On the

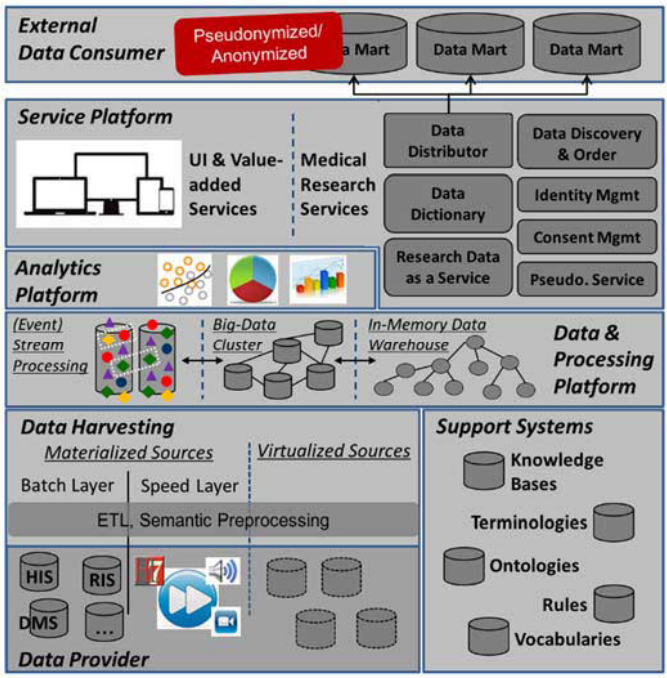


Figure 1. Generic Platform Architecture

other hand, real-time or data stream sources encompass bedside monitoring, tracking systems, audio and video systems as well as other sensor technology. In case physical data replication from the data source is not possible a database view can be established (Virtualized Sources). In doing so, data remain at the source and are accessed only on a temporary basis as needed. The process of data extraction and data access is assisted by the so-called Support Systems component. These systems include knowledge bases, terminologies/ontologies (like ICD-10-GM, OPS (German version of ICPM), SNOMED CT or MeSH) and rules/vocabularies used for refinement and harmonization of the extracted/accessed data and text mining respectively.

Finally, the Data & Processing Platform component is the core of the Data and Integration Platform. It is the central data storage and processing capability containing all extracted and refined data. The Data & Processing Platform is subdivided into an in-memory based data warehouse (In-Memory Data Warehouse), a storage and computing cluster for mass data (Big-Data Cluster), and an infrastructure for (Event) Stream Processing. Mutual access is possible. The in-memory based data warehouse contains data extracted by the conventional ETL procedure and allows for high performance data provision and data analysis of structured data. In contrast, the big data cluster stores medical images, documents, real-time data and data streams, e.g. monitoring or tracking data, and allows for mass data based analytics to discover new insights. Essential results can be accessed by the in-memory based data warehouse. Finally, the stream-processing infrastructure enables real-time responses to events based upon complex event processing and event analytics utilizing basic analytics capabilities provided by the Analytics Platform. This subdivision enables the proposed information platform to support various requirements regarding data analysis, data storage and data processing.

The *Analytics Platform* is based upon the Data & Processing Platform and enables the realization of methods and tools regarding data analytics accessing both the in-memory based data warehouse and the big data cluster. Various techniques can be applied either by using open source or commercial software including methods like exploratory data analysis, descriptive statistics, statistical inference, regression analysis, and machine learning. These analysis techniques or rather the analysis results can finally be provided to an end-user via the Service Platform.

The *Service Platform* is divided into two components both able to access the Analytics Platform as well as the Data & Processing Platform. The component User Interface (UI) & Value-added Services provides user interfaces and specific value-added services for clinical care according to the demands of clinical/medical departments. A broad variety of interfaces and services can be covered ranging from mobile applications via web-based applications to desktop applications. The second component of the Service Platform is Medical Research Services encompassing functionality to provide data - captured during clinical routine - for biomedical research in a data privacy compliant way. An essential aspect is that clinical routine data must be either anonymized or pseudonymized to be used for research purposes. Thus, in Germany the presence of an informed consent authorizing to pass on patient's clinical data to biomedical research is decisive. In case of an informed consent, clinical routine data can be passed on pseudonymized still allowing for patient re-identification. This is of crucial importance as clinical care and biomedical research are closely linked and findings and recommendations originating from research should be applied directly in clinical care. If, in contrast, no informed consent exists the transfer of clinical routine data to biomedical research is only allowed in an anonymized way. Thus, the patient's identity cannot be deduced. Within the Service Platform the following subcomponents ensure the described data privacy compliant functionality: Identity Management, Consent Management, and Pseudonymization Service. Additional subcomponents/services related to the usage of clinical routine data in biomedical research are a Data Dictionary, a Data Discovery & Order component to discover and order clinical data for research purposes, and a component to provide Research Data as a Service. The Data Dictionary provides a metadata catalog describing the data objects and their relationships stored within the Data & Processing Platform. Thus, interested researchers can get a quick overview regarding the available data structures and their possible exploitation by biomedical research projects. In case clinical routine data is appropriate and should be applied for biomedical research the component Data Discovery & Order must be used. This component allows browsing the Data & Processing Platform for relevant clinical routine data according to researcher's specific criteria. Patient identifying data has been already removed for this procedure to ensure data privacy. If clinical routine data matches the given search criteria these data can be requested by a researcher for a specific research purpose. After successful request evaluation clinical routine data can be provided for further research either via the subcomponent Research Data as a Service or as data transfer/export. Research Data as a Service provides an individualized and within the Service Platform hosted environment for further usage and analysis of the requested clinical data. The provided data are stored isolated from the remaining data pool and are either anonymized or pseudonymized. By this service a researcher or research group can avoid setting up an own, mostly expensive and time-consuming infrastructure for data processing and storage. If, for any reason, Research Data as a Service is not appropriate to provide the requested clinical data, these data can also be forwarded and exported respectively. This service is offered by the subcomponent Data Distributor. After

successful application of clinical routine data for research purposes the appropriate data flow/export can be configured by the Data Distributor. By this means, previously anonymized or pseudonymized data can either be transferred directly to an external research database or provided as an export file.

The final receivers of these data are the so-called External Data Consumers residing outside of the entire information platform encompassing all external users/consumers of clinical routine data for biomedical research.

3.3. Interoperability and standardization

Interoperability and standardization of the designed biomedical information platform is addressed by the components Data Harvesting, Support Systems and Medical Research Services. As these components act as external interfaces to data sources, terminology and knowledge sources, and research data sinks, they have to support common healthcare interoperability standards to enable reuse and smooth interconnection. Nevertheless, the components Data Harvesting and Support Systems must also facilitate proprietary connections to external systems, as many data sources do not comply to standardized interfaces and data formats. E.g. decentralized databases for detailed therapy documentation developed individually by medical departments.

The Data Harvesting component is based upon profiles of the Integrating the Healthcare Enterprise (IHE) initiative. Thus, common healthcare standards like Health Level 7 (HL7), Digital Imaging and Communications in Medicine (DICOM) and CEN ISO/IEEE 11073 Health Informatics-Medical/Health Device Communication Standards are supported. Exemplary IHE profiles used by the Data Harvesting component are the Cross Enterprise Document Sharing (XDS) profile of the IHE domain IT Infrastructure (ITI) and the Device Enterprise Communication (DEC) profile of the IHE domain Patient Care Device (PCD). The XDS profile is used to access clinical records provided as documents, and the DEC profile is used to receive data originating from patient monitoring systems (vital signs) and surgical devices (e.g. insufflation pressure).

The Support Systems component is designed to access vocabularies, terminologies and classifications, like ICD-10-GM, based upon the HL7 specification Common Terminology Services Release 2 (CTS2). Additionally, Resource Description Framework (RDF) and Web Ontology Language (OWL) are supported to access knowledge bases, ontologies and other linked data sources.

Finally, the Data Distributor component - responsible for providing biomedical research data to external consumers and located within the Medical Research Services - is designed to provide data according to the HL7 standard Clinical Document Architecture (CDA) and based upon the IHE profile XDS.

3.4. Initial platform implementation

The initial platform implementation focuses on the development of the components Data Harvesting and Data & Processing Platform and should support two preliminary scenarios. The first scenario is a conventional ETL/batch process transferring clinical data from a HIS, a cancer registry (CR) and an Enterprise Resource Planning (ERP) system to an in-memory based data warehouse. The clinical data sources are based upon various software products. The HIS is a combination of the SAP product IS-H and the Cerner product i.s.h.med (formerly Siemens). The cancer registry is based upon the AGFA Healthcare product ORBIS and as ERP system the product SAP ERP is used. As

in-memory database we chose the product SAP HANA as we have this technology on hand for research and testing purposes. Thus, SAP HANA represents the in-memory part of the proposed component Data & Processing Platform. The implementation of the ETL/batch based part of the Data Harvesting component is also based upon a product of SAP called SAP Data Services. The second scenario that should be supported by the initial implementation is a real-time/data streaming scenario handling data arising from medical devices in operating rooms (OR). These data can be patients' vital parameters or data originating from surgical equipment like light sources, operating tables or high-frequency surgical units. All these data have to be handled in real-time, thus, the speed layer part of the Data Harvesting component and the (event) stream processing part of the Data & Processing Platform have to be implemented. For realization we decided to use a combination of two software products provided by the Apache Software Foundation: Apache Camel and Apache Storm. Apache Camel is an integration framework used in our setting to establish the basic connection to various data sources (speed layer part). The received data are forwarded to an Apache Storm topology for distributed real-time computation (event stream processing part). Based upon this topology we are able to process received data in a parallel and scalable way. As a final data sink for these mass data we have chosen Apache Hadoop in combination with the Hadoop-related projects Apache HBase to store structured data and Apache Hive as data warehouse infrastructure.

4. Discussion

In this paper we present a concept and initial implementation of a holistic biomedical information platform. It is a novel approach as it provides a single platform for data storage and analytics with regard to patient care as well as for data dissemination related to biomedical research. Thus, we can avoid setting up two separate data storages with surrounding software systems and achieve synergy by sharing data and system capabilities across the two application areas care and research. Furthermore, our concept goes beyond the typical data warehouse approach as proposed by [7, 8, 9] and addresses upcoming big data and real-time data processing aspects as proposed by [10] as well. Therefore, the presented information platform is influenced by the lambda architecture as proposed by [11].

The initial platform implementation is just a first step towards our vision of a holistic biomedical information platform. As an ongoing research and development project it is still far from being holistic. Nevertheless, first tests showed a sustainable, scalable and flexible solution appropriate for the entire hospital. The test was based upon a surgical use case combining batch data from transactional systems and real-time data originating from an operating room to analyze surgical processes.

Our current research and development effort focuses on broadening the range of connected data sources (Data Provider) as well as on realizing the components Analytics Platform and Service Platform. Further data sources are LIS, RIS, pathology information system, biobank information system, Picture Archiving and Communication System (PACS), and genome databases. Especially the last three systems mentioned are challenging. Handling of biobank and genome data requires very strict compliance to data privacy and data security guidelines, whereas PACS and genome data are ambitious with regard to the required storage and computing resources.

The near future research and development plan encompasses the implementation of the component Medical Research Services - based upon the pattern proposed by [12] - as well as the integration of a text mining functionality and the capability for semantic annotation.

As personalized medicine through integration of patient care and biomedical research is of growing significance [5] the connectivity of our platform with regional and national research projects and networks has to be established as well.

References

[1] F. Kopcke, B. Trinczek, R.W. Majeed, B. Schreiweis, J. Wenk, T. Leusch, T. Ganslandt, C. Ohmann, B. Bergh, R. Rohrig, M. Dugas and H.U. Prokosch. Evaluation of data completeness in the electronic health record for the purpose of patient recruitment into clinical trials: a retrospective analysis of element presence. BMC medical informatics and decision making 13(2013): 37.

[2] N. Peek, J.H. Holmes and J. Sun. Technical Challenges for Big Data in Biomedicine and Health: Data Sources, Infrastructure, and Analytics. IMIA Yearbook 9(2014): 42-7.

[3] F. Martin-Sanchez and K. Verspoor. Big Data in Medicine Is Driving Big Changes. IMIA Yearbook 9(2014): 14-20.

[4] R. Bellazzi. Big Data and Biomedical Informatics: A Challenging Opportunity. IMIA Yearbook 9(2014): 8-13.

[5] H. Aamot, C.D. Kohl, D. Richter and P. Knaup-Gregori. Pseudonymization of patient identifiers for translational research. BMC medical informatics and decision making 13(2013): 75.

[6] H.G. Kenngott, M. Wagner, M. Gondan, F. Nickel, M. Nolden, A. Fetzer, J. Weitz, L. Fischer, S. Speidel, H.P. Meinzer, D. Bockler, M.W. Buchler and B.P. Muller-Stich. Real-time image guidance in laparoscopic liver surgery: first clinical experience with a guidance system based on intraoperative CT imaging. Surgical endoscopy 28(2014): 933-40.

[7] B.D. Athey, M. Braxenthaler, M. Haas, Y. Guo. tranSMART: An Open Source and Community-Driven Informatics and Data Sharing Platform for Clinical and Translational Research. AMIA Joint Summits on Translational Science proceedings AMIA Summit on Translational Science. 2013;2013: 6-8.

[8] G.M. Weber, S.N. Murphy, A.J. McMurry, D. Macfadden, D.J. Nigrin, S. Churchill, I.S. Kohane. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. Journal of the American Medical Informatics Association : JAMIA. 2009;16(5): 624-30.

[9] G. De Moor, M. Sundgren, D. Kalra, A. Schmidt, M. Dugas, B. Claerhout, T. Karakoyun, C. Ohmann, P.-Y. Lastic, N. Ammour, R. Kush, D. Dupont, M. Cuggia, C. Daniel, G. Thienpont, P. Coorevits. Using electronic health records for clinical research: The case of the EHR4CR project, Journal of Biomedical Informatics, Volume 53, February 2015: 162-173.

[10] B. Stein and A. Morrison. The enterprise data lake: Better integration and deeper analytics. PwC Technology Forecast: Rethinking integration, Issue 1, 2014, http://www.pwc.com/us/en/technology-forecast/2014/cloud-computing/assets/pdf/pwc-technology-forecast-data-lakes.pdf, last access: 9.3.2016.

[11] N. Marz and J. Warren. Big Data: Principles and best practices of scalable realtime data systems. Manning Publications Co., Place 2015.

[12] B. Schreiweis, G. Schneider, T. Eichner, B. Bergh and O. Heinze. Health Information Research Platform (HIReP) - An Architecture Pattern. Stud Health Technol Inform 205(2014): 773-7.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-63

Automated Transformation of openEHR Data Instances to OWL

Birger HAARBRANDT[a,1], Thomas JACK[b], Michael MARSCHOLLEK[a]
[a] *Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig and Hannover Medical School, Germany*
[b] *Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School*

Abstract. Standard-based integration and semantic enrichment of clinical data originating from electronic medical records has shown to be critical to enable secondary use. To facilitate the utilization of semantic technologies on clinical data, we introduce a methodology to enable automated transformation of openEHR-based data to Web Ontology Language (OWL) individuals. To test the correctness of the implementation, de-identified data of 229 patients of the pediatric intensive care unit of Hannover Medical School has been transformed into 2.983.436 individuals. Querying of the resulting ontology for symptoms of the systemic inflammatory response syndrome (SIRS) yielded the same result set as a SQL query on an openEHR-based clinical data repository.

Keywords. Electronic Health Records*/standards, openEHR, Biomedical Ontology

1. Introduction

Standard-based data integration is one of the premises for leveraging clinical data for secondary use. This requirement has driven the development of a set of methods that are often referred to as detailed clinical models (DCM) [1]. DCM aim for unambiguous, formalized, computable and manageable representation of domain content models. The evolving two-level modelling approach utilized by openEHR is layered as Archetype Model (AM) and Reference Model (RM) [2]. DCM can be complemented by the use of standardized medical terminologies. Comprehensive nomenclatures like SNOMED CT[2] address the need for standardized and unambiguous identifiers of clinical concepts, but also encompass ontological characteristics [3]. These characteristics allow utilization of semantic technologies to infer knowledge from medical data, merge data from diverse repositories, to calculate semantic similarities between patients and to identify cohorts for clinical research. Additionally, first studies have demonstrated the combined use of semantic web representations of archetypes and the Semantic Web Rule Language[3] (SWRL) to support clinical decision support and the calculation of quality indicators [4][5]. However, mapping of clinical data from source systems had to been done manually in these studies due to proprietary and non-standardized data structures and information models. With this work, we wanted to investigate if clinical data that is based

[1] Corresponding Author: Birger Haarbrandt, Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig and Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany, E-Mail: birger.haarbrandt@plri.de

[2] http://www.ihtsdo.org/snomed-ct

[3] https://www.w3.org/Submission/SWRL/

on the openEHR Reference Model, can be automatically transformed into instances of corresponding Web Ontology Language [4] (OWL) representations of archetypes to facilitate the secondary use of this data.

2. Methods

As a proof-of-concept study, we developed a software program to automatically transform real-world clinical data from an openEHR-based clinical data repository (CDR) into OWL individuals. The openEHR specification describes an open, interoperable electronic health record, built upon the two-level modelling approach [2]. By offering flexible standard-based data integration, the use of openEHR allows the development of systems that are to a certain degree independent from the clinical data models which are expressed as archetypes using the Archetype Definition Language (ADL). Each archetype represents a comprehensive collection of attributes, forming a maximum data set, to be able to address the requirements for medical documentation of any thinkable use-case [2].

The CDR is part of the Hannover Medical School Translational Research Framework (HaMSTR), a research data warehouse to investigate the use of DCM for secondary use. HaMSTR follows an Inmon architecture (i.e. data is normalized in an enterprise data warehouse before being delivered to customers through data marts). It contains data from two patient data management systems and the electronic medical record. All clinical data stored in the database is represented by serialized objects of the openEHR reference model. To store openEHR-based data, it uses a hybrid approach, utilizing both relations and *Extensible Markup Language* (XML) documents for persistence. HaMSTR uses openEHR compositions as data input. Compositions can be thought of as clinical documents composed of a set of archetypes to cover specific clinical use cases. They are represented in openEHR XML, which is a commonly used representation of serialized instances of the openEHR Reference Model. These XML documents are analysed and sections of particular archetypes are extracted and stored in corresponding tables having columns with xml data fields and an ID. This approach preserves the hierarchical structures of the templates while allowing a certain degree of optimization on the column level.

To build an openEHR reference model to OWL converter (RM2OWL), we utilized and enhanced the work previously done by Lezcano et al. [4], who developed the ADL2OWL translator and introduced the ArchOnt framework. The latter is the conception of a toolchain to allow the semantic enrichment of clinical data. While the ADL2OWL translator permits the automated translation of openEHR archetypes into OWL, it does not address the need for *automated* transformation of real-world clinical data into adequate OWL individuals.

The RM2OWL converter and involved parts of HaMSTR have been implemented using Java 8, SQLite and a Microsoft SQL Server 2012 database. To build the software, we utilized rapid prototyping development methodology. The libraries that were mainly used are the OWLAPI[5] and the Java Reference Implementation of openEHR[6].

[4] https://www.w3.org/TR/owl2-overview/

[5] http://owlapi.sourceforge.net/

[6] https://github.com/openEHR/java-libs

To test the correctness of the transformation, we transformed and loaded de-identified data from the CDR of HaMSTR into a triple store (*Stardog 4.0.3*[7]) and used SPARQL to query for paediatric patients matching symptoms of systemic inflammatory response syndrome (SIRS).

3. Results

Figure 1 shows the basic components of the RM2OWL converter and their role in the transformation and semantic enrichment process. Basically, the architecture is a implementation of the proposed workflow of the ArchOnt framework [4]. We slightly modified the ADL2OWL translator to introduce consistent naming convention between archetypes and their OWL representations, to check created ontologies for consistency and satisfiability (using the pellet reasoner[8]) and to make it part of a newly developed *Knowledge Store*. The *Knowledge Store* (based on a SQLite database) stores all archetypes used in the clinical data repository in both, ADL and OWL. Whenever an archetype is added to the store, a corresponding OWL representation gets automatically created. Our RM2OWL converter introduces functions to (1) query the CDR for archetype qualifiers, (2) obtain and merge OWL representations of archetypes from the *Knowledge Store*, (3) resolve terminology bindings to establish relations to terminologies/ontologies for semantic enrichment, (4) automatically transform openEHR data to OWL individuals, and (5) export the result to a triple store.

The first step of each transformation is the provision of a set of encounter numbers as input parameters. Optionally, a list of archetype qualifiers can be stated to limit the transformation to a defined subset of clinical concepts. Following, the converter generates a list of all archetypes that have been incorporated in matching templates. The list is used to retrieve respective ADL und OWL representations of archetypes from the *Knowledge Store*. Then, the OWL representations are merged into a single ontology and automatically checked for consistency (again using pellet).

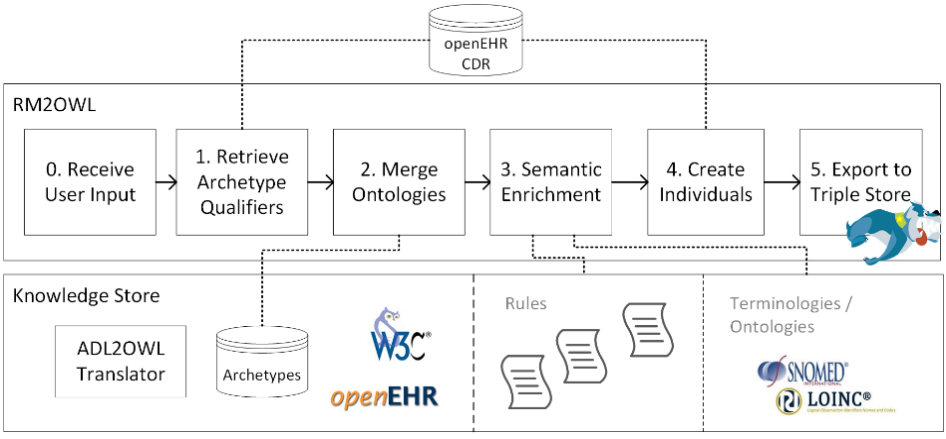


Figure 1: Overview of the transformation steps of the RM2OWL converter and the involved components.

[7] http://stardog.com/

[8] https://github.com/clarkparsia/pellet

Additionally, for *semantic enrichment* of the data, the ontology section of each archetype is inspected for available terminology bindings. If a binding exists and the terminology is available in the *Knowledge Store,* the system adds an OWL *equivalent class* axiom to the ontology and subsequently merges the archetype ontology with the target terminology.

After building the ontology, the system is set to *create individuals* from openEHR instance data. The RM2OWL converter retrieves all matching XML documents that are associated with the given encounter numbers. The transformation is done template-wise to be able to preserve hierarchical relationships that have been defined in a template. As the transformation of templates is not yet support by the ADL2OWL translator, we circumvent this shortcoming by only creating individuals of the generic *Composition* concept with a name property holding the template's name. Additionally, we added concepts to the ontology to conveniently represent subject IDs and encounters.

The transformation of each XML document is done by depth-first search using the Document Object Model representation. Generally, the transformation is rule based, i.e. each structure within the XML document triggers a function to create individuals of corresponding concepts. All information required to create an OWL individual is retrieved from the corresponding AM by using the *archetype_node_id*. Through applying the same naming convention as the ADL2OWL Translator, including a newly

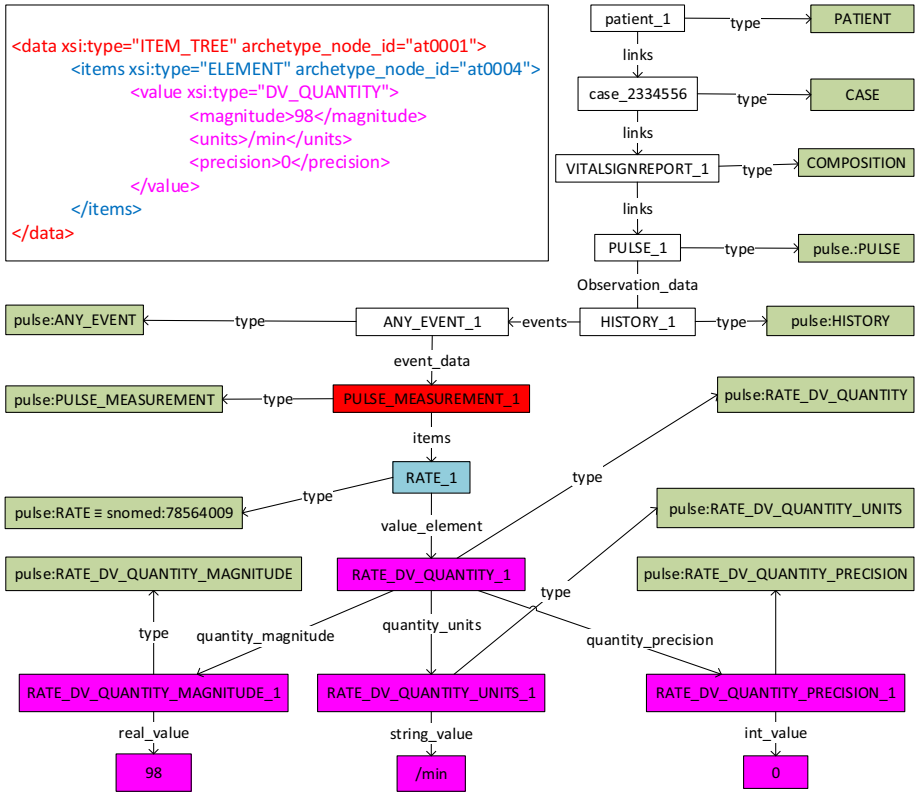


Figure 2: Example of an openEHR instance data transformation. Green indicates concepts originating from the OWL representations of the pulse archetype and concepts of the ontology. The colors used in the code listing (top left) and the elements indicate which individuals result from which element in the openEHR XML file.

added namespace based on the qualifier of the respective archetype, an individual can unambiguously be created and assigned to a concept. In order to establish object properties between individuals, a reference of each parent individual is passed to the function that is responsible to create a child. As every data point in openEHR XML is encapsulated by a defined data type of the reference model, it is possible to automatically generate literals for each value at the document's leaf nodes.

Figure 2 illustrates the interrelationship between instances of clinical data represented in openEHR XML, the OWL representation of archetypes and created individuals on the example of a pulse archetype. To simplify and make the figure more understandable to readers, some elements have been omitted. The colours indicate how a fragment looks before and after the transformation.

The example starts after the creation of individuals representing the patient, an encounter number and a stub for the template has been completed. To represent relationships between archetypes and templates, we use the *links* object property as proposed by Dentler et al. [5]. At first, the *data* node within the openEHR XML document is reached. It is checked for its type attribute, which in this case is an *ITEM_TREE*. With the help of the *archetype_node_id* ("at0001") the converter is able to perform a lookup of the element's name defined in the ontology section of the archetype. In the case of the pulse archetype, the ITEM_TREE has been constrained to represent a *PULSE_MEASUREMENT* (i.e. a tree structure holding data and state information of a single measurement of a pulse).

In the next step, an items element of type ELEMENT is found as a child of the data node. Through the *archetype_node_id* ("at0004") it can be identified as the RATE element of the archetype which contains the rate of a pulse, measured in beats per minute. As the measurement of a pulse rate results in a numeric value, it is represented by the DV_QUANTITY data type, enclosing data fields for magnitude, precision and units. The corresponding values get extracted using XPath expressions and are used to create OWL literals.

The binding between concepts in the archetype ontology and a terminology (e.g. SNOMED CT) is shown on the example of the Rate. An equivalent class axiom binding the concept of the OWL archetype and the concept having the SNOMED CT ID 78564009 has been introduced during the *semantic enrichment*. This binding allows queries and rules to be applied across merged ontologies.

Finally, the created ontology including archetypes, individuals and terminologies is loaded into a triple store (e.g. Stardog 4.0.3). By applying this approach, we have been able to successfully create OWL representations and individuals for all observational archetypes from the Knowledge Store. Table 1 provides a list of all incorporated archetypes.

Table 1: overview of archetypes used in HaMSTR and tested for transformation

Archetype	
OBSERVATION.blood_pressure.v1	OBSERVATION.respiration.v1
OBSERVATION.body_temperature.v1	OBSERVATION.body_weight.v1
OBSERVATION.pulse.v1	OBSERVATION.indirect_oximetry.v1
OBSERVATION.glasgow_coma_scale.v1	OBSERVATION.height.v1
OBSERVATION.braden_q_scale.v1	EVALUATION.problem_diagnosis.v1
OBSERVATION.lab_test.v1	EVALUATION.health_risk.v1

4. Evaluation

To test the correctness of the RM2OWL converter, we transformed and queried data of a subset of patients from the pediatric intensive care unit of Hanover Medical School. As an early step to develop a decision support system that alerts physicians about suspected presence of the systemic inflammatory response syndrome (SIRS), we conducted a retrospective cohort identification of patients matching the defined criteria for SIRS. Early detection of SIRS is of high clinical relevance, as it is closely related to sepsis (defined as SIRS with proven infection), organ dysfunction and organ failure which are among the most common reasons for morbidity and mortality in pediatric intensive care medicine [6].

In pediatric patients, age-specific vital signs and laboratory values are used for clinical diagnosis of SIRS. As SIRS criteria were not derived by population studies but defined by experts at the International Consensus Conference on Pediatric Sepsis (IPSCC), there is a lack of evidence for the exact ranges of vital signs for the different ages [6]. Two or more of the following values need to be present (of which one must be an abnormal temperature or leukocyte count): (1) fever (>38.5°C) or hypothermia (<36°C), (2) tachycardia, (3) tachypnea or (4) an abnormal blood leucocyte count. Except for the body temperature, all ranges of the variables depend on the patient's age. The definition of the IPSCC divides pediatric patients into six different age groups (<1 week, 1 week to 1 month, 1 month to 1 year, 2-5 years, 6-12 years, 13-<18 years).

We took a sample of 229 patients, which is all patients that had an encounter at the pediatric ICU in the first quarter of 2015 (01.01.2015 – 31.03.2015). All data has been de-identified; therefore, no demographic information has been available for querying. Hence, we applied the least rigid criteria for adolescent patients (13-<18 years) in order to retrieve a subset with a high recall and low precision for further processing.

```
PREFIX bt:<http://mh-hannover.de/openEHR-EHR-OBSERVATION.body_temperature.v1#>
PREFIX pulse: <http://mh-hannover.de/openEHR-EHR-OBSERVATION.pulse.v1#>
PREFIX resp: <http://mh-hannover.de /openEHR-EHR-OBSERVATION.respiration.v1#>
PREFIX oer: <http://klt.inf.um.es/~cati/ontologies/OpenEHR-SP-v2.0.owl#>
PREFIX lab: <http://mh-hannover.de /openEHR-EHR-OBSERVATION.lab_test.v1#>
PREFIX patient: <http://klt.inf.um.es/~cati/ontologies>

SELECT DISTINCT ?indiv  WHERE
  {
 …
 OPTIONAL{
 ?indiv oer:has/cati:observation_data/oer:events ?event .
 ?event a lab:ANY_EVENT .
 ?event oer:event_data/oeri:items/oer:value_element/oer:quantity_magnitude/oer:real_value ?lab .
 ?event oer:event_data/cati:items/oer:value_element/oer:value_text/oer:string_value ?testName .
 FILTER ((?testName = 'Leukozyten'^^xsd:string)
 } BIND(MIN(?lab) AS ?minLab)

FILTER ((?maxLab >= 11.0 || ?minLab <= 4.5) || (?maxBT >= 38.5 || ?minBT <= 36.0)) .
  FILTER ( ((?maxLab >= 11.0 || ?minLab <= 4.5) && (?maxBT >= 38.5 || ?minBT<= 36.0)) ||
((?maxLab >= 11.0 || ?minLab <= 4.5) && (?maxPuls >= 110)) || ((?maxLab >= 11.0 || ?minLab <=
4.5) && (?maxResp >= 20)) || ((?maxBT>= 38.5 || ?minBT <= 36.0) && (?maxPuls >= 110)) ||
((?maxBT >= 38.0 || ?minBT<= 36.0) && (?maxResp >= 20)) || ((?maxPuls >= 110) && (?maxResp
>= 20)) )
   }
 }
```

Figure 3: Excerpt of a SPARQL query to identify patients suspected of SIRS.

In sum, the converter created 2.983.436 individuals. To test the accurate transformation of the data, we performed semantically identical queries using both, SQL and SPARQL. The SQL query was conducted on the relational tables of respective data in HaMSTR. Figure 3 presents an excerpt of the SPARQL query, showing the pattern to retrieve the leucocyte count from the OWL representation of the *lab_value* archetype and the filter statement to retrieve only patients for which at least two criteria are matching. In total, the query returned the surrogate IDs of 153 patients that have met the stated inclusion criteria. We were able to retrieve the exact same result set as by using an equivalent SQL statement on the relational database of HaMSTR.

5. Discussion

Our work contributes a novel method to enable automated transformation of clinical data into OWL individuals. The converter uses a generic and fully automated transformation of openEHR instance data which exploits the strength of two-level modelling: changing requirements of an application system can be addressed by the definition or the versioning of archetypes without making alterations to the underlying reference model. Therefore, we are optimistic that the source code of the converter will not need to be modified when new clinical concepts are integrated in the CDR (or any other openEHR-based application system). By enabling such a generic transformation, the use of openEHR to represent clinical data can help to save resources for maintenance and to allow fast and correct provision of OWL representations of patient data to researches. Thereby, the RM2OWL converter might help to promote research on semantic technologies in real-world clinical environments. For example, investigation on automated semantic consistency checking and case-based semantic similarity measurements can be conducted more efficiently when customized transformation scripts from source data to OWL instances are not needed anymore.

Moreover, even in cases where these advanced research questions are not targeted, the representation of openEHR instance data in OWL can complement data retrieval based on SQL or the Archetype Query Language (AQL) by allowing semantic queries using SPARQL. For example, while AQL currently lacks operators to join data from diverse templates, this is possible with SPARQL [9].

Related work has been conducted in the SHARPn project [7] (using the Clinical Element Models (CEMs) approach instead of openEHR) and by Fernández-Breis et al. [8]. The former describes the transformation of the CEMs (which can be compared to archetypes) to OWL representations. However, there is a lack of explanation if and how instance data is actually transformed to individuals. As openEHR and CEMs developed independently but are quite similar instances for two-level modelling approaches, our methodology might in principle be applicable for CEM instances as well.

Fernández-Breis et al. have demonstrated a (semi-)automatic transformation within a rich framework for utilization of clinical data for secondary use employing a hybrid approach that clearly separates data level and clinical knowledge by linking XML and OWL representations. In contrast, our transformation of data instances, based on the archetype transformation of Lezcano et al. [4], expresses all types of clinical data as individuals of OWL representations of archetypes. While this might limit the use of certain aggregation functions (though, SPARQL 1.1 shows improvement regarding this

matter), it allows exploiting the ontological characteristics of the openEHR reference model within semantic queries.

Limitations of this work exist regarding the transformation of *action* and *instruction* archetypes. At the time of this study, the CDR of HaMSTR only contained data according to *evaluation* and *observation* archetypes. Hence, we decided to develop the first version without the above mentioned entry level classes. Moreover, the representation of templates is not yet supported. Therefore, some metadata elements as the healthcare facility or participations of clinical personnel were not available for querying.

Further development of the converter will concentrate on three aspects: First, the transformation of templates and further entry classes needs to be implemented in the ADL2OWL translator to represent the full set of metadata represented in document headers. Second, the introduction of a more granular way of pre-selecting particular data elements from templates is needed to avoid the transformation of unwanted data. For example, if only the result of a laboratory measurement (represented by a lab value archetype) is needed for research, only this element should be transformed. Third, the creation of a graphical user interface to select the above mentioned elements and to help enter the parameters would make the software more easy to use. Simultaneously, there will be investigations addressing the need to learn about the practical implications of this approach for secondary use.

References

[1] Goossen W, Goossen-Baremans A, van der Zel M. Detailed clinical models: a review. Healthc Inform Res. 2010 Dec;16(4):201-14. doi: 10.4258/hir.2010.16.4.201. Epub 2010 Dec 31.

[2] Beale T. Archetypes, constraint-based domain models for future-proof information systems. In: Eleventh OOPSLA Workshop on Behavioral Semantics: Serving the Customer (Seattle, Washington, USA, Nov 4, 2002). 2002, pp. 16-32.

[3] Rector AL, Brand S. Why do it the hard way? The case for an expressive description logic for SNOMED. J am Med Inform Assoc. 2008 Nov-Dec15(6):744-51.

[4] Lezcano, L., Sicilia, M. A., & Rodríguez-Solano, C. (2011). Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. Journal of biomedical informatics, 44(2), 343-353.

[5] Dentler K, ten Teije A, Cornet R, de Keizer N. Semantic Integration of Patient Data and Quality Indicators based on openEHR Archetypes. ProHealth 2012/KR4HC 2012, LNAI 7738:85–97, 2013.

[6] Goldstein B1, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. Pediatr Crit Care Med. 2005 Jan;6(1):2-8.

[7] Fernández-Breis JT1, Maldonado JA, Marcos M, Legaz-García Mdel C, Moner D, Torres-Sospedra J, Esteban-Gil A, Martínez-Salvador B, Robles M. Leveraging electronic healthcare record standards and semantic web technologies for the identification of patient cohorts. J Am Med Inform Assoc. 2013 Dec;20(e2):e288-96.

[8] Tao C1, Jiang G, Oniki TA, Freimuth RR, Zhu Q, Sharma D, Pathak J, Huff SM, Chute CG. A semantic-web oriented representation of the clinical element model for secondary use of electronic health records data. J Am Med Inform Assoc. 2013 May 1;20(3):554-62.

[9] openEHR Foundation. Archetype Definition Language ADL2. Available from: http://www.openehr.org/releases/1.0.2/architecture/am/adl2.pdf [last accessed January 26, 2016].

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-71

Extraction of UMLS® Concepts Using Apache cTAKES™ for German Language

Matthias BECKER[a,1] and Britta BÖCKMANN[a,b]
[a] *Department of Medical Informatics, University of Applied Sciences and Arts, Dortmund, Germany*
[b] *IMIBE, University Hospital Essen, Germany*

Abstract. Automatic information extraction of medical concepts and classification with semantic standards from medical reports is useful for standardization and for clinical research. This paper presents an approach for an UMLS concept extraction with a customized natural language processing pipeline for German clinical notes using Apache cTAKES. The objectives are, to test the natural language processing tool for German language if it is suitable to identify UMLS concepts and map these with SNOMED-CT. The German UMLS database and German OpenNLP models extended the natural language processing pipeline, so the pipeline can normalize to domain ontologies such as SNOMED-CT using the German concepts. For testing, the ShARe/CLEF eHealth 2013 training dataset translated into German was used. The implemented algorithms are tested with a set of 199 German reports, obtaining a result of average 0.36 F1 measure without German stemming, pre- and post-processing of the reports.

Keywords. Natural Language Processing, Information Extraction, Data Mining, Medical records, UMLS.

1. Introduction

Almost all healthcare institutions use Medical Information Systems. These systems tend to replace medical documentation on paper [1]. So often diagnoses, symptoms and medications are recorded and documented in a structured manner. However, a variety of clinical data is still stored in different unstructured formats like text and images. The development of new tools for intelligent IT-based text analysis can make it possible to access this knowledge and use it - for the electronic patient record itself or even for research purposes. Most of the available tools and also terminologies are in English language and cannot be used directly for German text.

There are various approaches internationally. One promising approach is machine learning natural language processing (NLP) [2] with information extracting techniques [3]. Extracting information from clinical notes has been the focus of a growing body of research these past years. There are existing solutions such as an advanced Text Mining pipeline based on Apache UIMA for German language by Averbis [4] and extensive researches on this topic by ID Berlin [5] and Semfinder [6]. This approach is consistent to the mentioned approaches with the difference of the focus on the identification of SNOMED CT codes using open source technology. The Challenges of this approach are to combine and adapt existing English standards to the German language and to evaluate the results without an existing German gold standard corpus.

[1] Corresponding Author: Matthias Becker, University of Applied Sciences and Arts, Dortmund, Emil-Figge-Str. 42 44227 Dortmund, Germany, Matthias.Becker@fh-dortmund.de

One of the fundamental tasks in clinical natural language processing research is to extract clinically relevant entities (e.g. diseases and drugs) using semantic standards such as Concept Unique Identifier (CUI) defined in the Unified Medical Language System (UMLS) [7] and Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT) [8]. SNOMED CT is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world [9], but it is not available in German language.

Therefore, the idea is to use the English SNOMED-CT by combining it with UMLS, which is available in German language. By extracting the UMLS concepts it is possible, to map this concepts to other non-German speaking classifications and ontologies [10]. To extract the concepts out of German clinical notes, a natural language processing pipeline for the German language is needed.

One of the most proven natural language processing tools is the open source natural language processing system for extraction of information from electronic medical record clinical free-text Apache cTAKES. Apache cTAKES already offers a variety of algorithms for text analysis and information extraction. This system was deployed at the Mayo Clinic and is currently an integral part of their clinical data management infrastructure and has processed over 80 million clinical notes [11]. It can normalize to domain ontologies such as SNOMED-CT using UMLS concepts. The aim of this project is to adjust an open source natural language processing pipeline to the German language to extract the concepts out of German clinical notes to map them with SNOMED-CT.

2. Methods

2.1. Architecture Overview

For the extraction of UMLS concepts from German clinical notes, a natural language processing pipeline with a mapping to the UMLS database is necessary. Once the concepts are identified, it is possible to map to domain ontologies and international terminologies. So Apache cTAKES was adapted to German clinical notes. Figure 1 shows the overview architecture of the system.

To configure the natural language processing pipeline for the objectives, various German OpenNLP models and German UMLS database were integrated into the cTAKES database. The cTAKES standard pipeline *AggregatePlaintextFastUMLSProcessor* was applied to extract the German concepts from the unstructured German clinical notes. For the mapping of the concepts to SNOMED-CT codes, the *ctakessnorx* database of cTAKES was expanded.

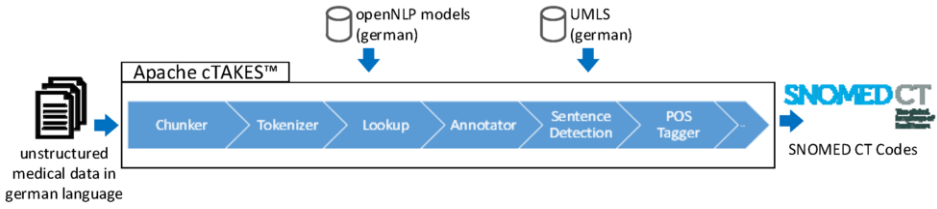


Figure 1. Architecture Overview for German language

2.2. Dataset and Evaluation

In 2013, the ShARe/CLEF eHealth Evaluation Lab (SHEL) organized three shared tasks on natural language processing and information retrieval (IR): 1) clinical disorder extraction and encoding to SNOMED-CT, 2) acronym/abbreviation identification, and 3) retrieval of web pages based on queries generated when reading the clinical reports. To test and evaluate the system, a ShARe/CLEF eHealth 2013 shared task 1 training set of 199 notes was used. Table 1 shows the amount of concepts within the training dataset which is used as gold standard.

Using these notes the pipeline was evaluated with English as well as German clinical notes to compare both results with the gold standard. The 199 English notes have been translated into German using the Google Translator initially [12]. The translated notes were then processed manually. The translated notes were analyzed by the same pipeline, but with an integrated German UMLS database and German OpenNLP models.

Table 1. Statistics of the training dataset:

Type	#Note	#CUI
ALL	199	2798
DISCHARGE	61	1969
ECHO	42	479
RADIOLOGY	42	257
ECG	54	93

The evaluation of the adapted pipeline follows the standard metrics of evaluation for the task using F1 (Equ. 3), i.e. the harmonic mean of recall (Equ. 2) and precision (Equ. 1). This is the same metric used by participants of the ShARE/CLEF 2013 Task 1.

$$P = \frac{true\ positive}{true\ positive + false\ positive} \tag{1}$$

$$R = \frac{true\ positive}{true\ positive + false\ negative} \tag{2}$$

$$F1 = \frac{2PR}{(P+R)} \tag{3}$$

Precision (P) is the number of correct positive results divided by the number of all positive results, and Recall (R) is the number of correct positive results divided by the number of positive results that should have been returned. The F1 score can be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst at 0.

2.3. Semantic Standards

The Unified Medical Language System (UMLS) integrates and distributes key terminology, classification and coding standards, and associated resources to promote creation of more effective and interoperable biomedical information systems and services, including electronic health records. For the natural language processing pipeline, the English and German concepts from the UMLS database were used. The

existing mapping between the concepts and SNOMED-CT codes of cTAKES was applied. For the German-language expansion of the system the UMLS version 2015AB was integrated into the cTAKES database.

The primary purpose of SNOMED CT is to encode the meanings that are used in health information and to support the effective clinical recording of data with the aim of improving patient care. SNOMED CT provides the core general terminology for electronic health records and contains more than 311,000 active concepts. It is comprehensive coverage includes: clinical findings, symptoms, diagnoses, procedures, body structures, organisms and other etiologies, substances, pharmaceuticals, devices and specimens. In this test series only the diseases where considered.

2.4. Natural Language Processing Pipeline

Natural language processing is an area of research and application that discovers how computers can be used to understand and manipulate natural language text or speech. Most common approaches to natural language processing are based on machine learning, a type of artificial intelligence that examines and uses patterns in data to improve a program's own understanding. For natural language processing and UMLS concept extraction on German clinical notes, different analysis steps in the form of an Analysis Engine (AE) are required. The applied Analysis Engine consists of the steps Chunker, Tokenizer, DictionaryLookupAnnotator, SentenceDetectorAnnotator, SimpleSegmentAnnotator and POSTagger. The Lookup Annotator is linked to the German UMLS database for matching concepts [13].

In this approach, the NLP-Pipeline was implemented in Apache cTAKES. It processes clinical notes and identifies types of clinical named entities – drugs, diseases/disorders, signs/symptoms, anatomical sites and procedures using OpenNLP models [11]. For the preprocessing of German notes, three TIGER data [14] trained OpenNLP models were integrated into the natural language processing pipeline [15]. The TIGER project aims to produce a large syntactically annotated corpus of German newspaper text. This includes a German Maxent Part-of-Speech tagger, Tokenizer and Sentence Detector.

3. Results

The results show the number of UMLS concepts extracted for English and German clinical notes compared to the gold standard.

Table 2 shows the performance of the pipeline with the ShARe/CLEF eHealth 2013 training dataset on English notes. The pipeline achieved an average recall of 0.69 with a moderate average precision of 0.25. The moderate precision has the consequence that the F1 value is attenuated. The best F1 could be achieved with discharge notes, which are the most complex notes with an average CUI of 32.28.

Table 2. Statistics of NLP-Pipeline with English notes:

Type	Recall	Precision	F1
DISCHARGE	0.71	0.27	0.39
ECHO	0.56	0.26	0.35
RADIOLOGY	0.69	0.23	0.35
ECG	0.81	0.25	0.38

Table 3. Statistics of NLP-Pipeline with German notes:

Type	Recall	Precision	F1
DISCHARGE	0.37	0.30	0.33
ECHO	0.47	0.51	0.49
RADIOLOGY	0.37	0.28	0.32
ECG	0.39	0.25	0.30

Table 3 shows the performance of the pipeline with the ShARe/CLEF eHealth 2013 training dataset on translated German notes. The results of the German pipeline have a much lower average recall of 0.4, but a slightly higher average precision of 0.34 compared to the pipeline with English notes. Thereby the average F1 value of 0.36 is almost equivalent to the English average F1 value of 0.37. Especially with echo notes, a much better F1 value could be achieved with German notes.

For the evaluation of the system, only concept codes were used from the dataset. diseases with no CUI Code (CUI-less) were not considered. Consequently, the experimental results are not directly comparable with those achieved by systems participating in the ShARE/CLEF Tasks.

4. Discussion

In this research a clinical disorder recognition and encoding system combining a machine learning based approach for entity recognition with UMLS concept mapping for German language was developed. The dictionary lookup approach on German clinical notes with terminological content from the UMLS for detecting disease was successful with a moderate F1 value.

The first results presented here seem promising even without German stemming, good results could be achieved and English SNOMED-CT codes could be derived. There are several reasons why the recall value of the German pipeline is significantly lower than the English pipeline. The UMLS database is not as extensive for the German Language (196842 entries) like the English (5571374 entries) UMLS database. That is the reason why the English pipeline is much better with complex notes such as discharge and receives more CUIs. Nevertheless, the German pipeline has a better average precision because it also identifies less false positive concepts. This also explains why the F1 value on German echo notes is higher than the value of the English pipeline. The Echo notes with an average CUI of 11.4 are not complex and have less different CUIs compared to the discharge notes. The German pipeline identifies less CUIs but with a higher precision. One more reason why the English pipeline achieved a much higher recall is that the applied cTAKES Analysis Engine is optimized for the English language. The AE contains stemming only for English language and the machine learning OpenNLP models are trained on English expert-annotated gold standard data.

In addition, no German stemming has been integrated into the pipeline at the time of the results publication and the German OpenNLP models have not yet been trained to medical notes. Generally, both results can be improved by pre- and post-processing of the notes. The first Analysis of the result dataset and the translated German notes show that some incorrectly or not identified concepts can be caused by translation errors. The quality of the translation has a significant impact on suitability for evaluation and has to be analyzed and optimized.

The next steps are to implement a Context Analyzer and German stemming to improve the results. Even an extension of the UMLS database and training of the models will be accomplished to see whether and how the results can be improved. Despite the fact that the models were used out-of-the-box without extra training the first results are promising. The pipeline for the German notes also lacks Negation Detection, which is also important to improve the results. Future attempts shall focus on optimizing the natural language processing pipeline for German language.

References

[1] A. Winter, R. Haux, E. Ammenwerth, B. Birgl, N. Hellrung, F. Jahn, Hospital Information Systems, in: Health Information Systems: Architectures and Strategies (Health Informatics). Springer, 2010. pp. 33-36

[2] G. Heyer, U. Quasthoff, T. Wittig, Wissensverarbeitung gestern und heute, in: Text Mining: Wissensrohstoff Text. W3L, Witten, 2012. pp. 12-18

[3] Franke J., Nakhaeizadeh G, Renz I., XML Retrieval and Information Extraction, in Text Mining – Theoretical Aspects and Applications, Physica-Verlag, Germany, 2003. pp. 29-32

[4] Averbis text analytics, https://averbis.com/branchen/pharma/, last access: 18.03.2016.

[5] ID SEMGuide, http://www.id-berlin.de/de/products/1-codierung/2-id-semguide/, last access: 18.03.2016.

[6] Semfinder DE, http://www.semfinder.com/produkte/semfinder-de/eigenschaften.html, last access: 18.03.2016.

[7] U.S. National Library of Medicine, Unified Medical Language System (UMLS), https://www.nlm.nih.gov/research/umls/, last access: 18.01.2016.

[8] Varghese J., Dugas M., Frequency analysis of medical concepts in clinical trials and their coverage in MeSH and SNOMED-CT. PubMed 54(1) (2015),83-92

[9] T. Benson, Clinical Terminology, Principles of Health Interoperability HL7 and SNOMED (Health Information Technology Standards), Springer, 2012, pp. 201-212

[10] McInnes BT, Pedersen T, Carlis J. Using UMLS Concept Unique Identifiers (CUIs) for word sense disambiguation in the biomedical domain. In: AMIA Annual Symposium Proceedings, Volume 2007. American Medical Informatics Association: (2007), 533

[11] Apache, Apache cTAKES, http://ctakes.apache.org/, last access: 18.01.2016.

[12] Pecina P., Dušek O., Goeuriot L., Hajič J., Hlaváčová J., Jones G.J., Kelly L., Leveling J., Mareček D., Novák M., Popel M., Rosa R., Tamchyna A., Urešová Z., Adaptation of machine translation for multilingual information retrieval in the medical domain, Artif Intell Med 61(3) (2014),165-85

[13] Divita G, T Zeng Q., Gundlapalli A., Duvall S., Nebeker J., Samore M., Sophia: A Expedient UMLS Concept Extraction Annotator, Journal of the American Medical Informatics Association (2014), 467–476 developed to extract UMLS concepts from clinical text. Among them, cTAKES

[14] Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. Journal of the American Medical Informatics Association (2010), 17(5):507–13

[15] Institut für Maschinelle Sprachverarbeitung, http://www.ims.uni-stuttgart.de/forschung/ressourcen/korpora/tiger.html, last access: 18.03.2016.

[16] Apache, Sourceforge OpenNLP Tools Models, http://opennlp.sourceforge.net/models-1.5/, last access: 18.01.2016.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-77

A Patient Safety Information Model for Interoperability

Jean Marie RODRIGUES[a,b,1], Neelam DHINGRA-KUMAR[c], Stefan SCHULZ[d], and Julien SOUVIGNET[a,b]
[a] *Department of Public Health and Medical Informatics, University of Saint-Etienne, France*
[b] *INSERM, U1142, LIMICS, F-75006, Paris, France; Sorbonne Universités, UPMC Université Paris 06, Paris, France ; Université Paris 13, Sorbonne Paris Cité, F-93430, Villetaneuse, France*
[c] *Patient Safety and Quality Improvement Unit, WHO, Geneva, Switzerland*
[d] *Institute for Medical Informatics, Statistics, and Documentation, Medical University of Graz, Austria*

Abstract. Current systems that target Patient Safety (PS) like mandatory reporting systems and specific vigilance reporting systems share the same information types but are not interoperable. Ten years ago, WHO embarked on an international project to standardize quality management information systems for PS. The goal is to support interoperability between different systems in a country and to expand international sharing of data on quality and safety management particularly for less developed countries. Two approaches have been used: (i) a bottom-up one starting with existing national PS reporting and international or national vigilance systems, and (ii) a top-down approach that uses the Patient Safety Categorial Structure (PS-CAST) and the Basic Formal Ontology (BFO) upper level ontology versions 1 and 2. The output is currently tested as an integrated information system for quality and PS management in four WHO member states.

Keywords. Ontology, Information Model, Semantic Alignment, Patient Safety, Interoperability

1. Introduction

Today, the management of data exchange has become one of the most crucial strategic topics within the health care domain [1, 2]. Institutions and organizations have realized the importance of data quality in medical databases for dependent applications and analytics [3] such as the assessment of Patient Safety (PS).

Risk management systems for PS, established by at national or international level, claim to integrate quality assessment phases. However, these risk assessment systems are so heterogeneous that their efficiency is mitigated by their inability to interoperate and to ensure a systematic exchange of safety-relevant information [4] across institutions and countries. Finally a lot of less developed countries have no recording system for PS risk management.

Since 2004, the World Health Organization (WHO), through its Patient Safety Programme, has campaigned for patient safety improvement by proposing advanced

[1] Corresponding Author: Jean Marie RODRIGUES, Department of Public Health and Medical Informatics, University of Saint-Etienne, France. Email: rodrigues@univ-st-etienne.fr

quality models for PS risk management [5]. In 2009, the WHO Department of Patient Safety published, in a first report, a conceptual basis for an International Classification for Patient Safety (ICPS) [6]. The report contained a conceptual framework including a list of terms and definitions of patient safety key concepts used in national PS reporting systems.

Whenever a consensus between existing national reporting systems for PS was established on definitions of PS terms, several issues were raised regarding the conceptual framework. S. Schulz and co-authors pointed that "there is no linkage at all between ICPS Key Concepts" [7], and the model was neither computable nor semantically explicit due to the lack of a formal foundation. W. Ceusters and co-authors indicated that "further effort is required [...] following [an ontological] methodology" [8].

This paper addresses these critiques and is organized as follows: 1) a recollection of the semantic alignment of ICPS Key Concepts to a Categorial Structure (CAST) semantic model named PS-CAST, adapted to national reporting for PS and to specific vigilance systems; 2) the alignment method between PS-CAST categories and Upper level Ontology BFO 2.0, with the definition of a terminology based on BFO 2.0 high levels entities for the main PS class Incident; and 3) the results in terms of the final definition of a Minimum Information Model for Patient Safety (MIMPS) common to national reporting and vigilance systems, and a terminology of types of incidents which facilitate interoperability for comparing PS risk management within and across countries and namely for less developed countries.

2. Material

In a contract with WHO PS department we aligned the ICPS key concepts to a so-called PS-CAST (Patient Safety - Categorial Structure) [9] in 2010, counting on computer science expert advices [7, 8]. PS-CAST is a semantic representation of Patient Safety terms, using the normalized Categorial Structure method [10].

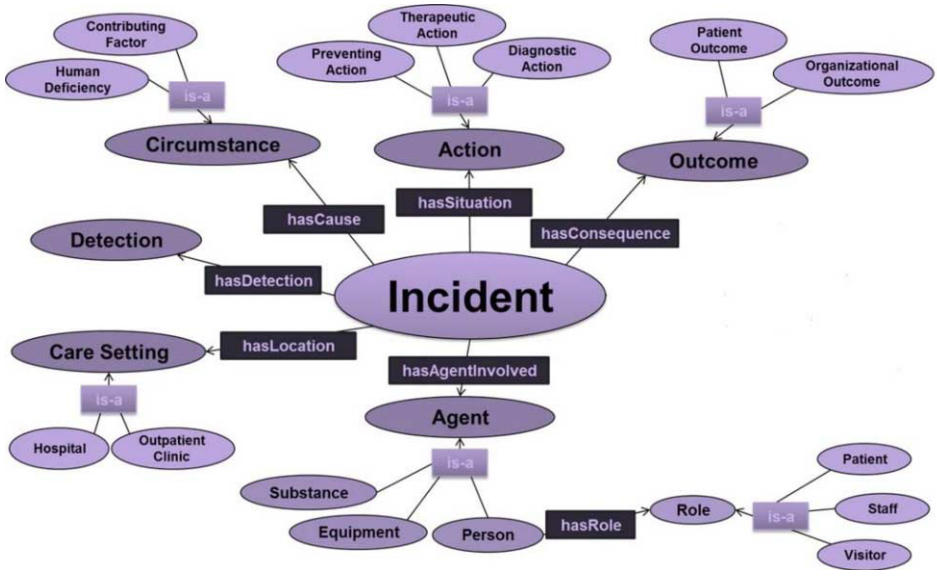


Figure 1: International Classification for Patient Safety (ICPS) key concepts in PS-CAST.

By using the semantic categories of PS-CAST we have shown common categories across national reporting systems from Australia, Japan, Belgium, Denmark, British Columbia (Canada), and the U.S. AHRQ [11]. In addition, we adapted PS-CAST by specific features as present in more than one of the national systems. Eight items present in all national reporting systems were extracted from PS-CAST as the minimum necessary named Minimum Information Model for PS (MIMPS), which was then tested on nine international or national Vigilance templates forms selected by WHO. This adapted version is depicted in Fig1.

Using the semantic categories of PS-CAST allowed us to show that these vigilance systems shared the same PS-CAST categories across them as well as with the national PS reporting systems [12]. In 2015, an international expert consultation in Warsaw approved a MIMPS consisting of ten items, as shown in the Results section.

3. Methods

In order to strengthen the semantic interoperability of PS reporting and vigilance systems we chose a two-level ontology approach as summarized in [13]. PS-CAST entities and relations in general and of the ICPS/PS-CAST class Incident in particular were aligned with Basic Formal Ontology BFO 2.0 [13] [14]. A philosophically inspired upper ontological framework, BFO provides universal categories as a basis for domain ontologies. It is based on a central dichotomy between Continuants and Occurrents. Continuants are entities that persist, endure, or continue to exist through time while maintaining their identities, whereas occurrents are entities that unfold them in time (or are instantaneous boundaries thereof). An entity is anything that exists or has existed or will exist, and has a definition and entities are linked together in relations. There are two types of relations: Is-a, expressing subclass links between all classes in the central backbone taxonomy, and other or binary relations which exist in the PS domain. Fig. 2 provides an overview of the BFO2 Is-a hierarchy.

3.1. First alignment

The first alignment between PS-CAST categories and BFO2 classes involved three steps; first, we performed an analysis of the definitions and elucidations proposed in BFO 2.0.

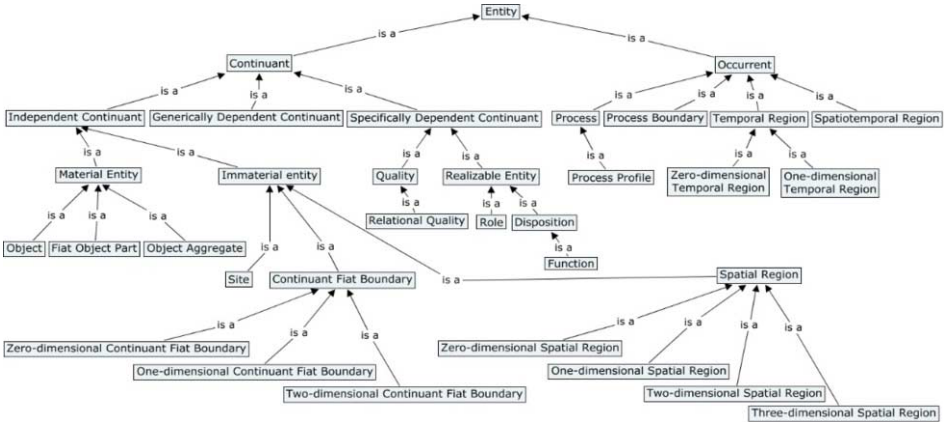


Figure 2: The BFO2 *Is-a* hierarchy overview

Then a search for alignments between PS-CAST and these classes was performed. The proposed alignments were reviewed by the experts. Whenever class definitions were unclear or inaccurate, or a relation was questionable, new proposals were made.

3.2. Second alignment

The second alignment concerns the terminology value sets of the class *Incident* (named *Incident Types* proposed by ICPS [6]), and the hospital acquired complications from ICD 10 CHADx [15, 16], and ICD 11[17]. The ontology definition of incident proposed by Ceusters et al. [18]: ”*an incident (a perdurant) that occurred to a patient during the past, that is documented in a database of adverse events and that is an expectation of some future happening that can be prevented’* shows that it is a complex entity which corresponds very often to a occurrent event or process as well as to an independent continuant as physical object or to dependent continuant as a wrong drug prescription by a physician.

Arapinis [19] argues that inherent polysemy arises only when there are dependence relations involved, such as between components, and that such complex objects are mereological sums of their aspect components. Fig. 3 and 4 show the BFO2 classes concerned by these terminologies. They clearly separate what is an event which exists in a single instant of time and when the realization is dependent of specific entity as healthcare.

4. Results

The final MIMPS was approved in 2015 on one hand as the basis for comparison between national reporting systems and international vigilance systems. The assessment of this goal has not yet be finalised.

The use of the MIMPS in four countries having not any PS risk reporting has started in 2016 and is ongoing.

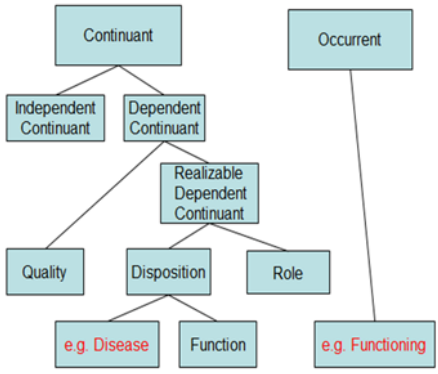


Figure 3: BFO 2.0 taxonomy

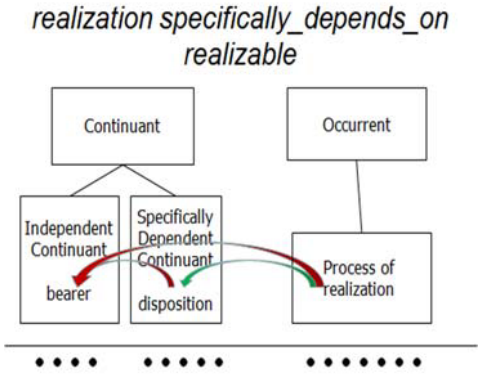


Figure 4: BFO 2.0 extract for the incidents types three categories of entities

The MIMPS is as following:

1. *Patient description* denotes the person who is a beneficiary of healthcare and a direct or indirect particular of the PS incident described only by age and gender.
2. *Time* refers to date and time when the PS incident occurred. A patient incident is a process in the sense of BFO.
3. *Location* refers to the physical environment (three-dimensional spatial region) in which a PS incident occurs.
4. *Cause* refers to an independent continuant such as an object, a substance, or a person, identified as the main cause when analysing the PS incident.
5. *Contributing factors* refer to an independent continuant such as an object, a substance, or a person, identified as having participated in the origin or development of a PS incident or to have increased the risk of a PS incident.
6. *Mitigating factors* refer to an independent continuant identified as (i) preventing a PS incident, or (ii) participating in the moderation of the progression of an PS incident, or (iii) participating in a process that reduces the risk of a PS incident.
7. *Incident type* denotes classes of PS incidents of a common nature, grouped by shared, agreed features in a defined representation of the world.
8. *Incident outcomes* refer to all impacts upon a patient or an organization wholly or partially attributable to a PS incident. Ontologically, this category may encompass processes in which the patient or the organization participates, as well as independent continuants that are part of the patient or the organization, as well as qualities or dispositions (dependent continuants), which inhere in the patient or the organization.
9. *Resulting actions* refers to all actions taken as consequence of a PS incident.
10. *Reporter's role* refers to the role in the organization of the person who collects and writes information about the PS incident and not his/her personal identity.

The terminologies for *Incident Types* which is a complex entity are polysemic. Nevertheless we can distinguish three main BFO 2.0 categories of entities for the types of incidents described in ICPS list [6] and only two in CHADx hospital acquired disorders based on ICD 10 AM [15, 16], or in ICD11 beta chapter 23 [17]:

- bfo:*Occurrent* (incident unrelated to any healthcare intervention)
    - ICPS Patient accidents; Behaviour.
    - CHADx: Accidental injuries.
    - ICD11 beta Chapter 23: "External causes of morbidity and mortality: Unintentional causes"
- bfo:*Occurrent* and **PrecededBy** some *HealthCareIntervention* (incident happening in relation with an health care intervention)
    - ICPS Health care associated infection, Problems in the management of Clinical process, Medication, Blood/Blood products, Nutrition, Medical fluids, Medical devices.
    - CHADx: Post-procedural complications; Adverse Drug Event and codes Not Present On Admission (NPOA)
    - ICD11 beta Chapter 23 External causes of morbidity and mortality: causes of healthcare related harm or injury

Note that we use the relation of temporal relatedness **PrecededBy**, which does not express causation. The reason is that although much of those events are of causal nature, the proof of this is often impossible.

- bfo:*IndependentContinuant*: incident related to a permanent entity as a structure or organization)
    - ICPS Infrastructure/building /Fixtures;
    - CHADx: none
    - ICD 11 beta none
- bfo:*GenericallyDependentContinuant*: Information entities:
    - ICPS Documentation; Human resource management (plan)
    - CHADx: none
    - ICD 11 beta none

5. Discussion

- The MIMPS can be criticized as being a structured information model not allowing for instance a real Root Cause analysis, which needs information that is more detailed. This limitation can be addressed by allowing for the Incident class a free text description of the instance.
- The different existing terminologies of types of incidents (ICPS, CHADx and ICD11 beta) were challenging to align with BFO 2, because the definitions of types and subtypes are not always aligned with the taxonomic principle of Is-a hierarchies, e.g. to decide whether an infection is related to some healthcare intervention or an event not related to an healthcare intervention.
- The final goal assessment of this initiative is one to create a reference data base on PS risk management in all the countries and two to make possible comparison of PS incidents and their different attributes across countries to support learning in less developed countries from more advanced ones in PS risk management.

6. Conclusion

The aim of this work was to facilitate the sharing of data between the different interpretations of the domain of Patient Safety endorsed by WHO in 2009 [6]. To this end, we applied semantic approaches, *viz.* Categorial Structure [10] and Upper Level Ontology BFO 2.0 [14]. Adding constraints, e.g., providing more rigour on the underlying organization and on formal definitions facilitated the semantic interoperability between different reporting systems in a country and across countries.

The MIMPS and the terminology of PS incidents, aligned with an upper level ontology provide a simple tool, especially for developing countries, which lack generic national reporting systems and/or not use the different international vigilance templates.

It is presently tested in four countries having no national PS or specific vigilance reporting systems.

As MIMPS and the PS incident terminology are based on a bottom up approach from the most developed national or international reporting systems for PS risk management they can be used to develop a comparative data base providing learning facilities for PS incident identification, correcting and preventing actions and enhancing permanent control of the safety of patients particularly in less developed countries.

The MIMPS can be seen as the framework to Quality and Security of Care integrated management.

Acknowledgements

PS-CAST project has been contracted by WHO (Registration [14]635-0, Order 200094768, Reg. File H15 APW 221) as part of 2IMPS development. The views expressed are those of the authors only.

References

[1] Shiloach, M., Frencher Jr, S. K., Steeger, J. E., Rowell, K. S., Bartzokis, K., Tomeh, M. G.,... & Hall, B. L. (2010). Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. Journal of the American College of Surgeons, 210(1), 6-16.

[2] Kahn, M. G., Raebel, M. A., Glanz, J. M., Riedlinger, K., & Steiner, J. F. (2012). A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. Medical care, 50.

[3] Blake, R., & Mangiameli, P. (2011). The effects and interactions of data quality and problem complexity on classification. Journal of Data and Information Quality (JDIQ), 2(2), 8.

[4] Benson, T. (2010). Principles of health interoperability HL7 and SNOMED (p. 263). London: Springer.

[5] Larizgoitia, I., Bouesseau, M. C., & Kelley, E. (2013). WHO efforts to promote reporting of adverse events and global learning. Journal of Public Health Research, 2(3), e29.

[6] Runciman, W., Hibbert, P., Thomson, R., Van Der Schaaf, T., Sherman, H., Lewalle, P. (2009). Towards an International Classification for Patient Safety: key concepts and terms. International Journal for Quality in Health Care, 21(1), 18-26.

[7] Schulz, S., Karlsson, D., Daniel, C., Cools, H., & Lovis, C. Is the International Classification for Patient Safety a classification? Stud Health Technol Inform. 2009; 150, p. 502 506).

[8] Ceusters, W., Capolupo, M., De Moor, G., Devlies, J., & Smith, B. (2010). An evolutionary approach to realism-based adverse event representations. Methods of Information in Medicine, 50(1), 62-73.

[9] Souvignet, J., Bousquet, C., Lewalle, P., Trombert-Paviot, B., & Rodrigues, J. M. (2011). Modeling patient safety incidents knowledge with the categorial structure method. In AMIA Annual Symposium Proceedings (Vol. 2011, p. 1300).

[10] Rodrigues JM, Kumar A, Bousquet C, Trombert B. Standards and biomedical terminologies: the CEN TC 251 and ISO TC 215 categorial structures. A step towards increased interoperability. Stud Health Technol Inform. 2008;136:857-62.
[11] Souvignet J, Rodrigues JM. Toward a patient safety upper level ontology. Stud Health Technol Inform. 05/2015; 210:160-4.
[12] Rodrigues JM, Schulz S, Souvignet J. Integrated care: an Information Model for Patient Safety and Vigilance Reporting Systems. Stud Health Technol Inform. 05/2015; 210:434-8.
[13] Arp, R, Smith, B., & Spear, A. (2015). Building Ontologies with Basic formal ontology. ISBN 978-0-262-52781-1. MIT Press.
[14] Basic Formal Ontology 2.0.specifications and user guide at https://bfo.googlecode.com. Last access 18/01/2016.
[15] Jackson TJ, Michel JL, Roberts R, Jorm C, Wakefield JA. Classification of Hospital Acquired Diagnoses for use with routine hospital data. Medical Journal of Australia (2009) 191(10):544-548.
[16] Trentino KM, Swain SG, Burrows SA, Sprivulis PC and Daly FFS. Measuring the incidence of hospital-acquired complications and their effect on length of stay using CHADx. Med J Aust 2013; 199 (8): 543-547.
[17] ICD11 chapter 23 at: http://apps.who.int/classifications/icd11/browse/l-m/en#/Last access 18/01/2016;
[18] Ceusters W et al. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3103706/ Last access 15/02/2016
[19] Arapinis A, Referring to institutional entities: Semantic and ontological perspectives, Journal of Applied Ontology 8 (2013), 31–57.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-85

Automatic Analysis of Critical Incident Reports: Requirements and Use Cases

Kerstin DENECKE [a, 1]
[a] *Bern University of Applied Sciences*

Abstract. Increasingly, critical incident reports are used as a means to increase patient safety and quality of care. The entire potential of these sources of experiential knowledge remains often unconsidered since retrieval and analysis is difficult and time-consuming, and the reporting systems often do not provide support for these tasks. The objective of this paper is to identify potential use cases for automatic methods that analyse critical incident reports. In more detail, we will describe how faceted search could offer an intuitive retrieval of critical incident reports and how text mining could support in analysing relations among events. To realise an automated analysis, natural language processing needs to be applied. Therefore, we analyse the language of critical incident reports and derive requirements towards automatic processing methods. We learned that there is a huge potential for an automatic analysis of incident reports, but there are still challenges to be solved.

Keywords. Data mining, Critical incidents reporting, Natural language processing

1. Introduction

Incident reporting systems have been a key tool to improve safety and enhance organizational learning from incidents in a range of high-risk organizations (commercial aviation, rail industry, and others). Their objective of the system is to enable users, e.g. health care professionals working for a hospital, to report in an anonymous manner critical events that occurred in their working environment. Incident reporting has been instituted in healthcare systems in many countries for some time now, e.g. in Switzerland in 1997 [1], but not in all healthcare systems it is obligatory to report critical incidents in healthcare. It has been shown that those anecdotal reports bear important information on limitations of systems and processes. For example, incident reports provide information on measures that helped in avoiding serious harm. From the quality management perspective, it is interesting to learn about the things that worked well in an incidental situation.

An analysis of the reports is difficult and currently done manually. For this purpose, incidents are collected and analysed to work out countermeasures. Often through manual analysis the total of each incident class is determined, occurrence factors and time periods are studied to identify the causes of accidents. Given the anonymous nature of the free textual messages, it is difficult to analyse roots and causes of incidents. So far, the information on critical incidents is collected, but only few hospitals are analysing the data and draw conclusions with respect to quality measures and adaptions of processes or structures. However, we can learn from frequent occurrences of messages on similar problems and from unusual constellations. At local level, i.e. within a hospital the reports

[1] Corresponding author: Kerstin Denecke, Bern University of Aöölied Sceinces, Höheweg 80, CH - 2502 Biel. Email: kerstin.denecke@bfh.ch

can uncover problems or limitations in structure or workflows of a hospital. In a regional context, general problems could be determined where guidelines could help to address the challenges. Analysing beyond institutional borders is realised by external organizations such as the Stiftung Patientensicherheit (foundation for patient safety) in Switzerland or the German Medical Association in Germany.

As described, the current analysis of the reports is done manually by the quality management department in hospitals or physicians working in organizations for patient safety that run regional critical incident databases. This is clearly necessary, when going into depth and analysing origins of critical incidents. However, studying correlations and dependencies or getting a quantitative overview on frequently occurring incidents requires considering many reports at once.

Given the unstructured nature of critical incident reports, we hypothesize that natural language processing (NLP) methods could support in analysing and processing critical incident reports. NLP of clinical narratives has sparked increasing attentions in recent years, resulting in effective algorithms for named entity recognition and relation extraction methods [2]. NLP techniques are required to make content of unstructured text machine-processable. They typically involve tokenization and sentence splitting, stop-list filtering, stemming, lemmatization, part of speech tagging, chunking and shallow or deep parsing and named entity recognition [3]. Named-entity recognition (NER) aims at identifying within a collection of text all of the instances of a name for a specific type of thing [3] (e.g., mentions of diseases or symptoms, but also persons and locations). For these basic tasks open source software is available, for example NLP tools provided by the Stanford language processing group (http://nlp.stanford.edu/software/). Domain-specific named entity recognition is enabled through tools such as MetaMap (provided by the National Library of Medicine, http://metamap.nlm.nih.gov/ [4]) or CTakes [5]. These tools map natural language text to concepts of an underlying ontology such as the UMLS. Both tools were successfully implemented and tested on clinical documents and biomedical literature. Due to the flexibility in language usage, the same meaning can be expressed in different ways, e.g. through a noun, its synonym, an abbreviation etc. Through mapping of terms to concepts of a terminology, texts can be represented semantically and become interpretable for computer algorithms.

However, it is still unclear whether these methods are already able to process incident reports or which other requirements need to be addressed. In order to develop methods for such analysis, we need to learn more about the linguistic and semantic characteristics of these reports. In this paper, we collect characteristics by means of a manual analysis of critical incident reports written in German. Further, we describe two use cases derived from discussions with physicians showing the potentials of an automatic analysis. Besides this, the main contribution of the paper is to provide a vision for an automatic analysis of critical incident reports for the purpose of using the content for quality management.

2. Material and Methods

2.1. Material

With an increasing demand for the prevention of medical accidents, critical incidents are collected in online platforms. One of those platforms is CIRS medical (http://www.cirsmedical.de) which is a reporting and learning system of the German

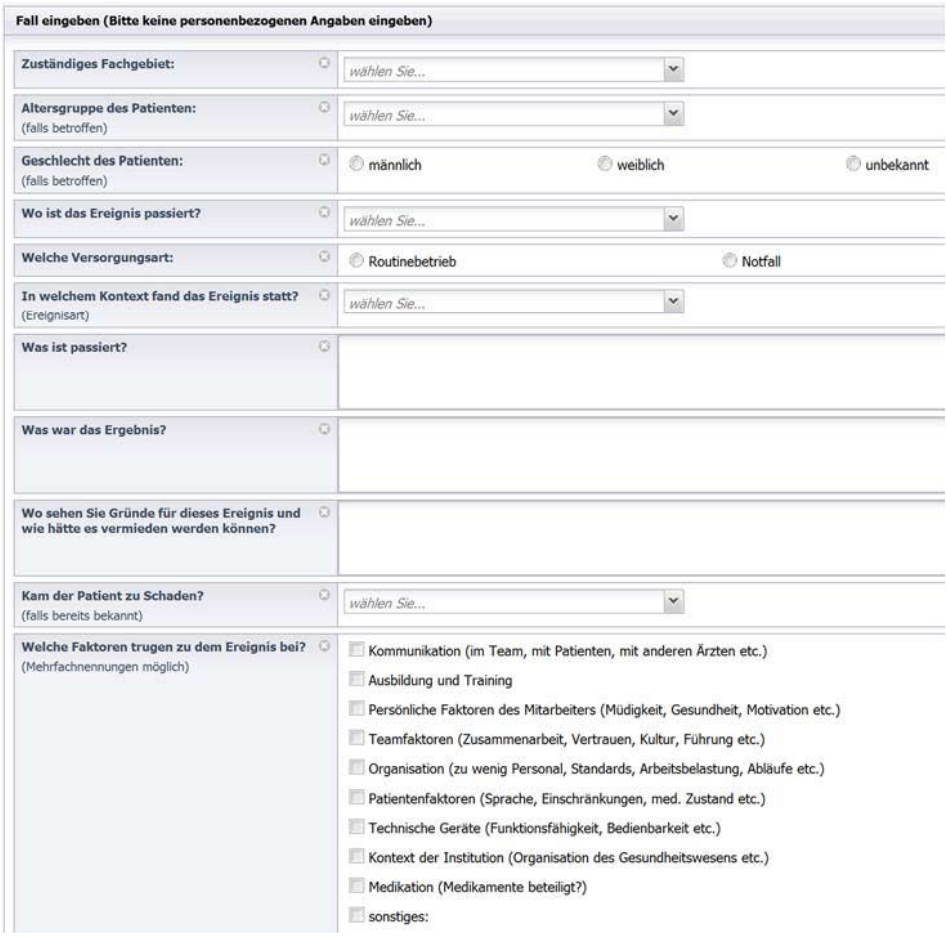


Figure 1: Entry Form of CIRS medical (*http://www.cirsmedical.de*)

association of physicians (Bundesärztekammer). The platform is accessible for everyone via Internet. Staff working in the healthcare domain can report all events associated with patient security and patient safety including errors, almost-damages, critical incidents or undesired events. A form (see Figure 1) allows to report the critical incident in a semi-structured way. Some fields are queried by multiple choices; others require free textual entries. The reports may not contain any data that allow to draw conclusions on the involved persons or institutions.

For our analysis, we collected reports from the CIRS medical website. At the time of data collection, the database contained 4703 reports (08.01.2016). For a detailed analysis, we randomly selected 100 reports from that database. No other filter criteria where applied.

2.2. Methods

The author of this paper performed a manual analysis of language peculiarities and content of the text in the free text field “What happened”, “What was good” and “What

was not good". These text fields were analysed considering the questions: 1) Which characteristics and peculiarities with respect to syntax can be determined in the reports? and 2) How can the reports be described semantically? Are there certain semantic categories into which the reports can be categorised?

From our previous work on information extraction from clinical texts [6], we are aware of the peculiarities of clinical documents. Thus, we identified differences and similarities between both text types. Further, we interviewed two physicians, both responsible for analysing critical incident reports. One of them is working for the Stiftung Patientensicherheit in Switzerland, responsible for analysing the incoming reports and for selecting relevant ones for providing recommendations for the future. The other person is responsible for the incident reporting at the Inselspital Bern. We collected their experiences with critical incident reporting systems, asked for the processes and the need for analysis support. The observations from the linguistic and semantic analysis were discussed with the physicians and the results extended by their experiences.

3. Results

In this section, we summarise the results of the analysis of critical incident reports and describe the collected use cases.

3.1. Linguistic Characteristics

From the sentence structure, it can be recognised, that there is a broad spectrum ranging from short phrases to complex sentences or enumerations of phrases. Information is summarised in a few words, resulting in compound words. As already known from other clinical texts, abbreviations are used for locations (e.g. "iv" for "intravenous") or for procedures or technical devices (e.g. "ECMO" which is a kind of life support machine). Clinical texts are known to be error prone since they are written under time pressure [6]. In contrast, the analysed incident reports did not contain many writing errors. Since often processes are described, temporal and other number expressions are exploited. Table 1 summarises the characteristics and provides examples from the reports. Differences to

Table 1. Linguistic Characteristics of Critical Incident Reports

Characteristic	Example
Broad spectrum in sentence complexity: Complex sentences, short phrases	„Der Kreislauf verschlechterte sich rapide, es konnte der Katecholaminschenkel schnell auf eine peripher venöse Kanüle gewechselt werden, welche zufälligerweise noch nicht, wie geplant, entfernt worden war."
	„Ungenügende Fixierung des ZVK" (insufficient fixation of central venous catheter)
Abbreviations	ECLS/vaECMO
Number expressions	Zugriffszeit ca. 2 min. (access time 2 min)
	nach ca. 15 Sekunden (after 15 seconds)
	SpO2 von 100% (SpO2 of 100%)
Subjective vs. Factual information	Die größte Gefahr der Verletzung besteht beim…(The largest danger of injury occurs during…)
Compound words	Zustandsverschlechterung (State worsening)

Table 2. Categories of Critical Incidents with Examples

Category	Example
Technical events or events related to equipment	*Oxygenator is burning*
Administrative events	mistaken identity, ambiguous abbreviation as diagnosis, language barrier, mixing of laboratory probes
Hygiene-related events	Report on a case of Methicillin-resistant Staphylococcus aureus (MRSA) / hospital infection
Workflow-related issues	application of wrong drug, allergy information was not transmitted to all involved persons
Events related to transport and positioning of patient	Patient is falling from operating table

clinical documents are mainly related to the semantics and content described. They are summarised in the next section.

3.2. Semantic Characteristics

Semantically, the texts are rich of information. While specific terms referring to diagnoses are seldom explicitly mentioned, the texts contain descriptions of symptoms, procedures (e.g. *Intubationsnarkose (intubation anesthesia), nachbeatmet (ventilated), Ambulanz-OP (ambulant surgery)),* or names of pharmaceutical agents, narcotics etc. In contrast to clinical documents such as discharge summaries, various material, either technical or non-technical are mentioned including *monitor, emergency rucksack, blood pressure cuff* etc. Additionally, the locations of events occurrences are described such as *laboratory, emergency department, anaesthetic recovery room, intensive care unit* and others.

Different persons involved in the event are presented in the reports by their role. Besides the patient, e.g. *staff working on the various wards, ward physician, anaesthetist, emergency doctor, person responsible for blood transfusion, child, nurse, director* and others are mentioned in the reports. All these pieces of information together result in detailed descriptions of critical incident events in their timely order. Sometimes, it is only mentioned what happened without any further explanation. There are also non-medical events described (such as mistaken identity of patients) that resulted in clinical events. More specifically, we can distinguish five categories of events. They are listed with examples in Table 2.

In particular in the section asking for good or bad things, we can find subjective information and factual information or even speculations. Consider the sentence „Es hätte zu einer Entzündung, Thrombophlebitis und einer Fibronisierung des Fremdkörpers kommen können“ (An inflammation, thrombophlebitis and fibronisation of the foreign body could have occurred). It is not yet confirmed that these issues happened to the patient, but the author of the report is mentioning them already.

3.3. Automatic Analysis of Critical Incident Reports: Use Cases

Through interviews with physicians, we identified several use cases where an automatic analysis of incident reports is required: 1) Retrieval of and access to reports and 2) automatic assessment and analysis of incident reports. They are described in more detail in the following.

3.3.1. Faceted Search for Retrieval of Critical Incident Reports

Situation: A physician wants to know, whether a problem that he recognised happened before and which countermeasures had been taken.

Problem: The database of critical incident reports is difficult to query. If at all, only predefined categories allow to filter the reports. A free textual or even semantic search for similar cases is not yet implemented. The free text fields are not searchable.

Use Cases: There are two possible use cases. On the one hand, an improved retrieval of relevant reports is necessary. One option would be that for a given critical incident report relevant cases are to be retrieved. The query would be the report and the result would comprise a set of similar reports.

On the other hand, the entire database of reports could be actively searched, e.g. for reports of a certain category. Consider an incident that occurred, and a quality manager or physician who would like to know which measures had been taken when the event occurred before. A structured and semantically enriched critical incident data base could provide answers by returning all relevant cases with retrieval results structured by the taken measure.

Technical Solution: To enable these use cases, in particular the content of the free text fields needs to be semantically analysed and mapped to concepts of an ontology. This would result in a normalised representation of the text. This representation could form the basis for enabling a faceted search [7]. Faceted search, or faceted navigation is a technique that allows to assess information using a faceted classification system. The technique enables users to explore a collection of information by applying multiple filters. Each information element is assigned to multiple explicit dimensions, called facets. For the first use case, the retrieved records would be made browsable by their facets. In the second use case, the retrieved reports would be accessible by the measure taken.

3.3.2. Critical Incident Analytics

Situation: The quality management of a hospital is interested in learning about frequently occurring critical incidents and wants to analyse correlations for developing in-house guidelines or improved processes. They want to learn about limitations in process or structure quality. Further, they would like to know how often which type of incident occurs.

Problem: The web-based systems for reporting critical incidents not yet provide methods for an automatic assessment and aggregation of data. A challenge is the unstructured text in the reports.

Use Case: For a given problem or event statistics have to be created, how often the event occurred etc. An automatic analysis could determine event-cause chains that allow to get insights into processes causing the critical events.

Technical Solution: Using text and data mining methods, and semantic technologies the incident reports could be automatically analysed, semantically enriched with categories. Relations such as those between an event and its cause could be determined and prepared for user assessment. The texts could be automatically assigned to semantic categories and text could be normalised by indexing (assigning concepts of an ontology to normalise the content). The categorization could be realised according to type of reported problem (e.g. hygiene, complication during treatment), location (e.g. intensive care unit, laboratory), or severity. Methods for detecting events, causes and measures in unstructured text are required for this purpose.

4. Discussion and Conclusions

Critical incident reports are still an unused resource of knowledge on quality issues in the health care sector in general or in a hospital in particular. While it is recognised that the reports could support in identifying risks and limitations, there is no technical support available to analyse the content systematically or in retrieving relevant information. Currently, the reports are manually analysed which results in an almost non-use of the data. A semantic analysis of the data could enable multiple use cases. Yamamoto et al. performed a linguistic analysis for incident reports in English. They extracted characteristic words using natural language processing and they evaluated the degree of similarities between incident documents [8]. Lee et al. [9] used data mining to identify critical factors in patient falls using a web based reporting system. Using artificial neural network analysis they developed a predictive model and identified several critical factors. However, to the best of our knowledge there is no system available that supports the use cases described in this work.

As a first step towards developing technologies that enable the introduced use cases, in this paper, we collected the linguistic peculiarities of critical incident reports: They make several demands on the technologies which partially differ from those of clinical documents. In contrast to natural language processing methods developed specifically to process clinical documents, methods for analysing critical incident reports require:

- Methods for identifying persons and locations,
- Methods for separating factual from experiential or hypothetical information,
- Methods for categorizing incident reports semantically and of events according to severity,
- Methods for detecting events and their relations,
- Methods for analysing time and number expressions.

Additionally, methods for semantic structuring and analysis are necessary. As a next step, we will study which existing tools and methods can be used to support these tasks. For identifying persons and locations named entity recognition tools such as those provided by the Stanford NLP group could be used. Some first research has been done to separate factual from experiential or hypothetical information from clinical documents [10] and for extracting time information (e.g. Heideltime [11]).

Further, the use of ontologies and formal concepts of a domain is necessary for adapting inference functionalities to various situations and application scenarios, but also to make unstructured text automatically processable within the context of reasoning. For mapping to an ontology, we need to study, whether incident reports could be indexed of classified using the International Classification of Patient Safety (ICPS), a classification system for incident cases [12]. This classification system could form the basis for a faceted search and semantic annotation of the reports. Methods for an automatic mapping to this classification system are still missing. First attempts regarding automatic classification of incident reports using machine learning algorithms have been made by Ong et al. [13]. But, they used only two categories and not the ICPS. In summary, as next steps, methods for realising the text analysis tasks mentioned before need to be developed or existing methods adapted. An incident report analysis tool need to combine multiple methods for natural language processing.

Our analysis was done in a qualitative manner. The results could be confirmed by a quantitative assessment of the reports with respect to word classes and frequent occurring patterns. However, for collecting requirements towards the development of NLP methods, our analysis provide already comprehensive results. We used the data from CIRS medical for our analysis. It might be that these messages are selected manually by the hosts of the system to show only representative reports. This might be an explanation why we could not identify writing errors as expected.

Given the results from our study in this paper, we recognize potentials and use cases for text and data mining of critical incident reports. An improved access to experiential knowledge has the potential to improve patient safety, and quality of care which in turn are benefits that would help in increasing physician's motivation of reporting such events.

References

[1] S. Staender, J. Daviers, B.Helmreich, B. Sexton, M. Kaufmann: The anaesthesia critical incident reporting system: an experience based database. Intern J Med Inform 47 (1997): 87-90

[2] C. Friedman, T.C. Rindflesch, M. Corn: Natural language processing: state of the art and prospects for significant progress, A workshop sponsored by the national library of medicine. J Biomed Inform 2013;46(5):765–73

[3] A. Cohen and W. Hersh: A survey of current work in biomedical text mining. Brief Bioinform, 6(1):57–71, Jan. 2005

[4] A. R. Aronson: Effective mapping of biomedical text to the umls metathesaurus: The metamap program. In Proceedings of the AMIA 2001, 17-21

[5] G.K. Savova, J.J. Masanz, et al.: Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. J Am Med Inform Assoc. 2010;17(5):507-13

[6] K. Denecke: Semantic Structuring of and Information Extraction from Medical Documents using the UMLS. Methods of Information in Medicine, 2008; 5(47), 425-34

[7] B. Wei, J. Liu, Q. Zheng, W. Zhang, X. Fu, B. Feng: A Survey of faceted search. Journal of Web Engineering, Vol. 12, No. 1 & 2 (2013), 41-64

[8] K. Fujita, A. Akiyama, K. Park, E.N. Yamaguchi, H. Furukawa. Linguistic Analysis of Large-Scale Medical Incident Reports for Patient Safety. Studies in Health Technology and Informatics. 2012, Volume 180: Quality of Life through Quality of Information, 250 –254

[9] T.T. Lee, C.Y. Liu, Y.H. Kuo, M.E. Mills, J.G. Fong, C. Hung.: Application of data mining to the identification of critical factors in patient falls using a web-based reporting system. Int J Med Inform. 2011 Feb;80(2):141-50

[10] K. Denecke, Y. Deng: Sentiment analysis in medical settings: New opportunities and challenges. Artif Intell Med. 2015 May;64(1):17-27

[11] J. Strötgen, M. Gertz: Multilingual and cross-domain temporal tagging. Language Resources and Evaluation 2012;1–30

[12] http://www.who.int/patientsafety/implementation/taxonomy/en (last access: 13.01.2016)

[13] M.S. Ong, F. Magrabi, E. Coiera: Automated categorisation of clinical incident reports using statistical text classification. Qual Saf Health Care. 2010 Dec;19(6):e55

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-93

SEMCARE: Multilingual Semantic Search in Semi-Structured Clinical Data

Pablo LÓPEZ-GARCÍA[a,1], Markus KREUZTHALER[a], Stefan SCHULZ[a], Daniel SCHERR[b], Philipp DAUMKE[c], Kornél MARKÓ[c], Jan A. KORS[d], Erik M. van MULLIGEN[d], Xinkai WANG[e], Hanney GONNA[e], Elijah BEHR[e], Ángel HONRADO[f]
[a]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria; [b]Division of Cardiology Department of Medicine, Medical University of Graz, Austria; [c]Averbis GmbH, Freiburg, Germany; [d]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands; [e]St. George's University of London, United Kingdom; [f]SYNAPSE Research Management Partners S.L.

Abstract. The vast amount of clinical data in electronic health records constitutes a great potential for secondary use. However, most of this content consists of unstructured or semi-structured texts, which is difficult to process. Several challenges are still pending: medical language idiosyncrasies in different natural languages, and the large variety of medical terminology systems. In this paper we present SEMCARE, a European initiative designed to minimize these problems by providing a multi-lingual platform (English, German, and Dutch) that allows users to express complex queries and obtain relevant search results from clinical texts. SEMCARE is based on a selection of adapted biomedical terminologies, together with Apache UIMA and Apache Solr as open source state-of-the-art natural language pipeline and indexing technologies. SEMCARE has been deployed and is currently being tested at three medical institutions in the UK, Austria, and the Netherlands, showing promising results in a cardiology use case.

Keywords. Information Storage and Retrieval; Health Records, Personal; Data Mining

1. Introduction

Clinical data in electronic health records (EHRs) have a promising potential in many areas of healthcare [1]: to monitor and improve healthcare delivery [2], to identify disease mechanisms [3], to enhance drug safety [4], and to facilitate patient recruitment for clinical trials [5]. For example, only 18% of clinical trials in Europe and 7% in the US meet their patient enrolment quotas on time, causing delays in bringing new drugs to market. Exploiting patient-level data can also optimize clinical research in several ways, e.g. by enabling the definition of appropriate study designs or ensuring that inclusion/exclusion criteria map to an existing patient population [6].

Cardiovascular disease, in particular ischemic heart disease, is the leading cause of death in the developed world [22]. Almost half of all ischemic heart disease related deaths occur suddenly and are due to abnormal heart rhythms caused by a diseased heart muscle. The primary biomarker used to identify patients at high risk of sudden death is

[1] Corresponding Author: Pablo LÓPEZ-GARCÍA, Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Auenbruggerplatz 2, 8036 Graz, Austria, E-Mail: pablo.lopez@medunigraz.at

the ejection fraction, which is an imaging-based measure of the strength of cardiac contraction. This value may be defined in a quantitative manner expressed as a percentage value, or a qualitative manner described as normal, mild, moderate, or severe impairment. Other important biomarkers for risk stratification include ECG measures such as presence of cardiac rhythm abnormalities, QRS and QT duration, symptoms of breathlessness and transient loss of consciousness. A relevant blood based marker is creatinine, a surrogate for kidney function. A prototypical example of a combination of clinical criteria of interest would be: *"Patients older than 60 years with an ejection fraction less than 35 and a QRS duration ranging between 120 and 130 who are taking Heparin or Dalteparin"*.

Most use cases, such as the one above, have a similar goal in common: to identify patient cohorts based on existing data and clinical criteria like age, gender, diagnosis, signs, symptoms, and lab results. There are two main difficulties for achieving this goal: Firstly, processing the data: large parts of patient-level data are still scattered in heterogeneous resources and stored in unstructured or semi-structured form as free text [7]. Furthermore, at each institution these texts can be available in a different natural language, additionally complicated by medical language idiosyncrasies (e.g., ambiguous terms, acronyms, compounds, derivations, spelling variants, spelling errors, jargon, telegram style), and be mixed with quantitative data (e.g., lab results, drug dosages, dates/times) [8]. A second difficulty is the establishing of the clinical criteria: the system needs to enable users to precisely express the selection criteria, which requires optimised user interfaces.

In this paper we present SEMCARE [9], a recently-finished EU-funded project involving three medical institutions, aimed at helping to solve these problems by providing a multi-lingual platform (English, German, and Dutch) that allows users to express complex queries and get relevant search results from unstructured data in form of free texts.

2. Methods

In order to meet the defined use cases also targeting different languages to which the search scenario is applied, the SEMCARE architecture is constituted by the following components: **Extract-Transform-Load (ETL), Data Integration:** this component transforms heterogeneous data sources (structure and content) into standardized input for further processing (see "Text Mining" component). Data sources might consist of plain text, PDF, Microsoft Word documents, or HL7 messages. **Data Storage:** a structured copy of the data is stored in an i2b2 star schema [10], an efficient and well-established relational clinical data model that allows third party applications to be used for data analysis and visualization. **Anonymization:** a de-identification service can be applied to ensure patient and carer privacy. This component was not used in our case, as the project required that all clinical texts were previously anonymized. **Terminology Management:** this component allows users to browse and administrate (add, edit, and delete concepts and terms) the biomedical terminologies used by the text mining component. **Text Mining:** this component constitutes the core of the semantic middleware. Its main functionality is to provide extended structured information (annotations) using the biomedical terminologies available in the system, and open source Natural Language Processing (NLP) (Apache UIMA [11, 25, 26]) and indexing (Apache Solr [12])

technologies. **Search and Analytics:** a semantically enriched web-based front end for semantic search and analysis of the processed data.

All NLP pipeline components were embedded within UIMA to exploit a flexible and adaptable configuration framework and constitute of the following main components: a tokenizer (Lucene Snowball) and a sentence detector (rule based); Word de-compounding based on morphosemantic analysis [23]; Concept mapping via a customised version of the UIMA Concept Mapper annotator acting on the terminologies explained above; Negation detection based on NegEx [24], as well as a dictionary-based abbreviation and acronym resolver.

3. Results

SEMCARE was installed in three hospitals: St. George's Hospital (London, UK), Erasmus Medical Center (Rotterdam, Netherlands), and the University Hospital of Graz (Austria). Each installation was deployed in isolation from the clinical information system. With approval by the local ethics committees, documents in English, Dutch, and German were extracted from the operational hospital EHR systems, and only limited access was given to authorized staff. All documents were anonymized before being accessed by SEMCARE. The semi-structured discharge letters were transferred to the SEMCARE search server (Solr) in a batch process, obtaining an enhanced index exploiting an NLP pipeline via UIMA. Standard biomedical terminologies were used (SNOMED CT [13] and ICD10 [14] for English; ICD10-GM [15], LOINC [16], and ABDAMED [17] for German).

In addition, German and Dutch terminologies were extended using semi-automatic translation techniques. These terminologies enable semantic search by mapping synonymous terms to concepts, which are related by hierarchical links.

Figure 1 shows the graphical user interface of SEMCARE providing the services to end users, with its three main constituting parts: A **Query builder** (top) allows users to express complex queries graphically in an accessible manner. The listing of **Search**

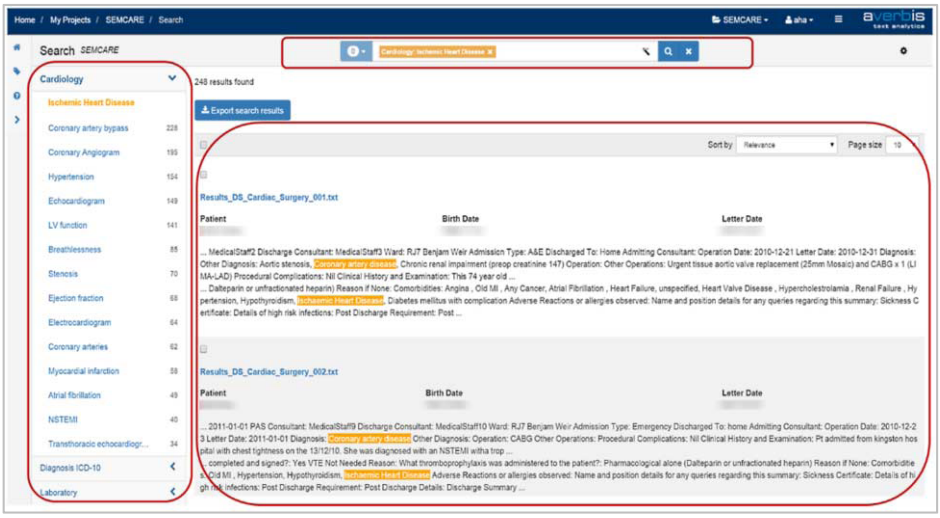


Figure 1. Graphical user interface.

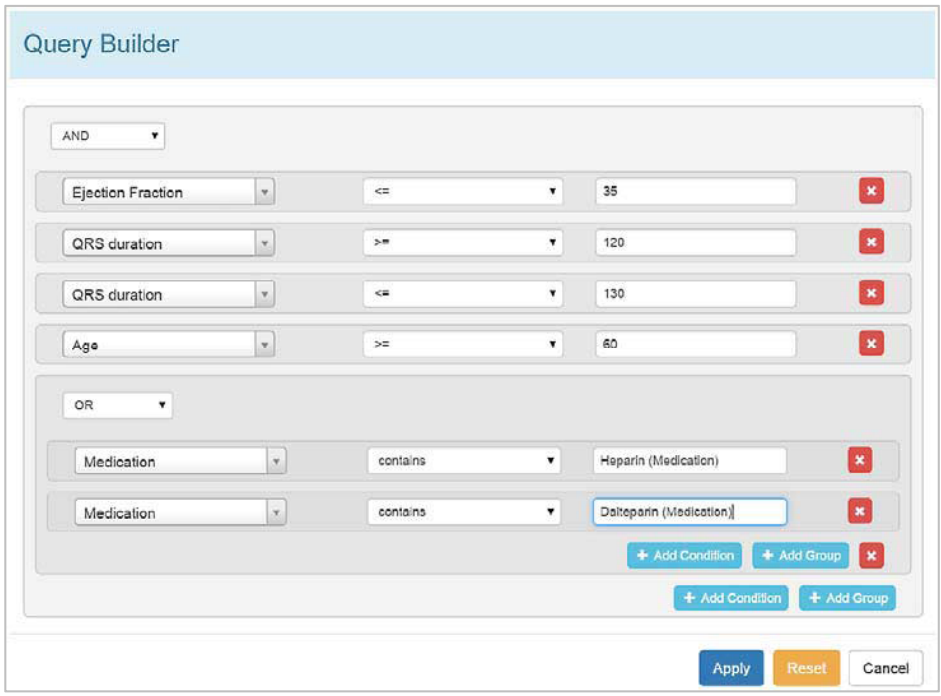


Figure 2. Query builder.

results (centre) shows all the (unstructured) clinical notes matching the criteria from the query. **Facets** (left) permit a quick overview, grouping, and filtering of results based on relevant semantic axes [27] like Disorders, Drugs or Lab, according to the scope of the source terminologies and their subdivisions.

Figure 2 shows the input window when the query builder is clicked, and how to express the cardiology use case presented in the introduction: *"Patients older than 60 years with an ejection fraction less than 35 and a QRS duration ranging between 120 and 130 who are taking Heparin or Dalteparin"*.

Figure 3 shows a search result matching the criteria, and Figure 4 shows the result conveniently summarized for the end user in a structured way.

4. Discussion and Related Work

Most existing clinical data are stored in unstructured form, mainly in EHRs. The difficulties in analysing and capturing semantics from free texts have been studied since

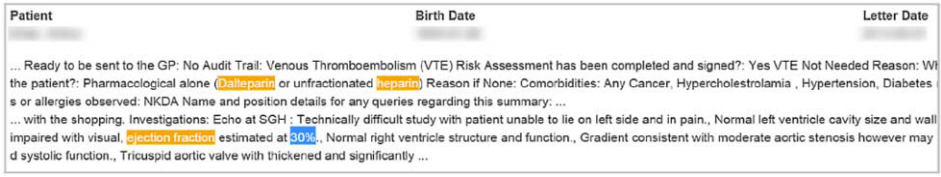


Figure 3. A specific search result matching the input query. Entities and values are marked in the text.

Ejection Fraction	30
QRS duration	125
Diagnosis ICD-10	Oedema Essential (primary) hypertension Diabetes mellitus Living alone Pain Pure hypercholesterolaemia Diabetes mellitus with complications
Cardiology	Ischemic Heart Disease Aortic stenosis Breathlessness Coronary Angiogram Echocardiogram Hypertension Ejection fraction QRS duration

Figure 4. Structured summary of the search result shown in the previous figure.

long ago in the field of computational linguistics, and several well-established conferences are devoted to the matter (e.g. ACL, EACL, SIGIR). However, clinical NLP is still a relatively small sub-area, especially when compared to other biomedical text-mining areas that focus on literature abstracts instead of EHR content. One of the main reasons for the early stage of clinical NLP research is the difficulty of accessing and sharing real patient data due to privacy concerns, which itself leads to a lack of gold standards or references for evaluation.

This is gradually changing within the English speaking community, particularly in the United States, as scientific challenges have been established to foster clinical NLP developments, like the i2b2 [10], SemEval [18], or TREC [19] medical records track challenges. However, the situation in other languages is still very preliminary due to the lack of a joined effort to foster research for processing clinical narratives.

To improve the situation in Europe and languages other than English, SEMCARE adds to previous EU-funded projects, like EHR4CR [20] or MANTRA [21].

5. Conclusion and Future Work

The vast amount of existing clinical data in EHR systems promises great potential. However, several challenges are still to be solved for data to be useful in real applications. A major obstacle is given by the fact that important content is only available as unstructured texts; with clinical texts being especially difficult to process, due to language idiosyncrasies in different natural languages, as well as the limited coverage of domain-specific terminology resources for all languages other than English.

The SEMCARE system addresses these difficulties by providing a multi-lingual platform for performing queries on unstructured medical data. It was deployed and is being tested at three medical institutions in Europe, with documents in English, German, and Dutch. The core technologies in SEMCARE are a solid basis of adapted biomedical terminologies, and a state-of-the-art NLP pipeline.

Project participants are now testing SEMCARE at their institutions and have just finished a more complete evaluation that includes NLP performance, identifying high-

risk individuals for cardiology use cases. A user satisfaction survey has shown that the system excels by very good response times and enables users to get a quick insight into retrospective patient cohorts, which can be further exploited for e.g. hypothesis generation, rare disease detection or high-risk patient profiling. SEMCARE's semantic retrieval approach, powered by terminologies and enriched by synonyms that reflect clinicians' language preferences, guaranteed an efficient exploration of large document spaces with good retrieval quality.

As future work, we plan to fine tune the platform, expand the use of SEMCARE to other domains, and publish the evaluation results in detail.

Acknowledgments

The SEMCARE platform was developed under the European Union's Seventh Framework Program for research, technological development, and demonstration activities (grant no. 611388).

References

[1] P.B. Jensen, L.J Jensen, and S. Brunak. Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics* **13**(6) (2012), 395–405.
[2] P.J. Embi et al. Development of an electronic health record-based Clinical Trial Alert system to enhance recruitment at the point of care. AMIA Annual Symposium Proceedings. Vol. 2005. American Medical Informatics Association, 2005.
[3] I.S. Kohane. Using electronic health records to drive discovery in disease genomics. *Nature Reviews Genetics* **12**(6) (2011), 417–428.
[4] M.J. Schuemie, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. *Medical Care* **50**(10) (2012), 890–897.
[5] T. van Staa et al. Pragmatic randomised trials using routine electronic health records: putting them to the test. *BMJ* **344**:e55 (2012).
[6] J. Powell and B. Iain. Electronic health records should support clinical research. *Journal of Medical Internet Research* **7***(1):*e4, (2005).
[7] S.M. Meystre, et al. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb Med Inform* **35** (2008), 128–44.
[8] W.W. Chapman et al. Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. *JAMIA* **18**(5) (2011), 540–543.
[9] SEMCARE, http://semcare.eu, last access: 10.02.2016.
[10] I2B2, https://www.i2b2.org/, last access: 10.02.2016.
[11] Apache UIMA, https://uima.apache.org/, last access: 10.02.2016.
[12] Apache Solr, http://lucene.apache.org/solr/, last access: 10.02.2016.
[13] SNOMED CT, http://www.ihtsdo.org/, last access: 10.02.2016.
[14] ICD10, http://www.who.int/classifications/icd/en/, last access: 10.02.2016.
[15] ICD10-GM, https://www.dimdi.de/static/en/klassi/icd-10-gm/, last access: 10.02.2016.
[16] LOINC, https://loinc.org/international/german, last access: 10.02.2016.
[17] ABDAMED, http://www.wuv-gmbh.de/abdata-pharma-daten-service/datenangebot/abdamed/, last access: 10.02.2016.
[18] SEMEVAL, http://alt.qcri.org/semeval2016/, last access: 10.02.2016.
[19] TREC, http://trec.nist.gov/, last access: 10.02.2106.
[20] EHR4CR, http://www.ehr4cr.eu/, last access: 10.02.2106.
[21] MANTRA, https://sites.google.com/site/mantraeu/, last access: 10.02.2106.
[22] T.A. Gaziano et al. (2010). Growing epidemic of coronary heart disease in low-and middle-income countries. Current problems in cardiology, **35**(2), 72-115.
[23] Markó, K., Schulz, S., & Hahn, U. (2005). MorphoSaurus Design and Evaluation of an Interlingua-based, Cross-language Document Retrieval Engine for the Medical Domain. *Methods Archive*, **44**(4), 537-545.

[24] Chapman, W. W., Bridewell, W., Hanbury, P., Cooper, G. F., & Buchanan, B. G. (2001). A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of biomedical informatics*, **34**(5), 301-310.
[25] Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. Journal of the American Medical Informatics Association, **17**(5), 507-513.
[26] http://ohnlp.sourceforge.net/MedKATp/, last access: 10.03.2016
[27] Natarajan, K., Stein, D., Jain, S., & Elhadad, N. (2010). An analysis of clinical queries in an electronic health record search utility. International journal of medical informatics, **79**(7), 515-522.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-100

Bridging the Gap between HL7 CDA and HL7 FHIR: A JSON Based Mapping

Christoph RINNER[a,1] and Georg DUFTSCHMID[a]
[a]*Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna*

Abstract. The Austrian electronic health record (EHR) system ELGA went live in December 2016. It is a document oriented EHR system and is based on the HL7 Clinical Document Architecture (CDA). The HL7 Fast Healthcare Interoperability Resources (FHIR) is a relatively new standard that combines the advantages of HL7 messages and CDA Documents. In order to offer easier access to information stored in ELGA we present a method based on adapted FHIR resources to map CDA documents to FHIR resources. A proof-of-concept tool using Java, the open-source FHIR framework HAPI-FHIR and publicly available FHIR servers was created to evaluate the presented mapping. In contrast to other approaches the close resemblance of the mapping file to the FHIR specification allows existing FHIR infrastructure to be reused. In order to reduce information overload and facilitate the access to CDA documents, FHIR could offer a standardized way to query CDA data on a fine granular base in Austria.

Keywords. Electronic Health Records, HL7 CDA, HL7 FHIR.

1. Introduction

In December 2015 the Austrian electronic health record system - Elektronische Gesundheitsakte (ELGA) - went live in two Austrian test regions. ELGA is designed as a distributed system to exchange medical documents with central components for identification of patients and health care providers and authorization management [1]. The Health Level Seven International (HL7) Clinical Document Architecture (CDA) [2] is used to exchange and persist medical information (i.e. ELGA adopts a document centric view). Currently four different document types (i.e. "Physician's Discharge Summary", "Nursing Discharge Summary", "Laboratory Report" and "Diagnostic Imaging Report") are available in the ELGA system. In July 2016 the first "e-Medication report" including prescription, dispensing, medication summary and pharmaceutical advice will be available in ELGA. By summer 2016 ELGA will be operational in all Austrian provinces.

The current architecture to exchange medical information in ELGA is based on the *Integrating the Healthcare Enterprise* (IHE) *Cross-Enterprise Document Sharing* (XDS) profile [3] and CDA. This architecture was first proposed in a feasibility study in 2006 [4]. IHE XDS is a distributed architecture and is used in different regions and countries worldwide [5]. Despite the widespread use of IHE XDS and its practical significance, IHE XDS is very much restricted in its search capabilities due to the following three reasons: First, IHE XDS only allows coarse filtering of medical documents via a limited

[1] Corresponding Author: Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, BT88, Floor 4, Spitalgasse 23, 1090 Vienna, Austria. E-Mail: christoph.rinner@meduniwien.ac.at

set of document metadata (e.g., patient, author, date, type of document) and only complete documents are returned as the result of a query [6]. Second, IHE XDS uses the Simple Object Access Protocol (SOAP), which works well in server environments, but due to its complexity is rarely used in the mobile world. This was no problem a decade ago, but the IT landscape has since changed significantly and the use of mobile internet access overtook the fixed internet access in 2014 [7]. Third, CDA documents are predestined to replace paper-based documentation adapting a similar process as paper based approaches to persist and sign documents in the digital world. Yet the experience has shown that CDA is difficult to implement due to the complexity in its specification in combination with the additional layer of constraints from specific CDA implementations (e.g. ELGA document types).

HL7 Fast Healthcare Interoperability Resources (FHIR) [8] is a relatively new HL7 initiative that combines the advantages of the HL7 v2/v3 messages and CDA documents. It is based on the Representational State Transfer (REST) architectural style [9], which was specifically designed for thin clients like web browsers as well as fast and easy implementation. FHIR uses modular components called Resources to exchange, query, load, persist or delete health information on a granular level. FHIR resources are serialized using the Extensible Markup Language (XML) or the Java Script Object Notation (JSON), converters between the two formats exits.

The easy implementation of FHIR in combination with good documentation and open source reference implementations available has led to a wide adoption of FHIR[10, 11]. FHIR is also used as a connector to existing health data infrastructures like i2b2 [12] or openMRS [13]. To allow mobile devices or other resource constrained systems to access IHE XDS health information, IHE created the *Mobile access to Health Documents* (MHD) profile. MHD is currently on hold and the profile will be updated to reflect the changes made in new FHIR standard once FHIR reaches a stable point [14].

The goal of the present work was to allow easier access to health data stored in ELGA CDA documents. We present a mechanism based on adapted FHIR resources to map and transform ELGA CDA documents to FHIR resources. This can enable various stakeholders like health care providers, patients, app developers and researchers to access ELGA compliant data more easily. The approach was tested using ELGA e-Medication documents and a proof-of-concept Java implementation.

2. Methods

To formalize the mapping of CDA documents to FHIR resources in a computer processable format we used JSON. In order to simplify the initial setup of the mapping file, the JSON representation of the resulting FHIR resources was used as boilerplate for the mapping, i.e. the mapping file matches the JSON representation of the resulting FHIR resources without any data but with minor additional elements. In the mapping file we added metadata about the author, document types, to which the mapping files can be applied to and versions of the used standards. XPATH expressions were used to directly pinpoint a corresponding element in the source CDA document to an attribute in a FHIR resource. Additional mapping information to transform CDA specific values to FHIR values (e.g. gender, dosage instructions, etc.) were also available using references to existing ConceptMap resources. A mapping file was created for each document type separately. Figure 1 shows an example of the JSON mapping to transform the Austrian e-Medication Medication List to a patient and the various Medication FHIR resources.

```
1. {  "CDA2FHIR_oid" : "1.2.40.0.34.11.8.3",
2.     "CDA2FHIR_label" : "Document-level template of a Medication List
           document within the Austrian e-Medication projectrn",
3.     "FHIR_version" : "ca.uhn.fhir.model.dstu2.resource",
4.     ...
5.     "resources" :
6.     [...{
7.       "cda-path" : "/ClinicalDocument/recordTarget",
8.       "resourceType" : "Patient",
9.       "text" : {...},
10.      "gender" : {
11.        "cda-value" :
               "./patientRole/patient/administrativeGenderCode/@code",
12.        "cda-mapping" : {
                "http://spark.furore.com/fhir/ConceptMap/
                 v3-administrative-gender" },
13.     },
14.     ...
15.   },{
16.     "cda-path" :
            "//entry[//manufacturedProduct[manufacturedMaterial/
            templateId/@root='1.2.40.0.34.11.2.3.4']]",
17.     "resourceType" : "Medication",
18.     "code" : {
19.      "coding" : [{
20.        "system" : "system test",
21.        "code" : {
22.         "cda-value" : "//manufacturedMaterial/code/@code"
23.        },
24.        "display" : {"cda-value":
              "//manufacturedMaterial/code/@displayName"}
25.   ... }]}
```

Figure 1: Example JSON mapping of Austrian e-Medication Medication List. Elements added for mapping are printed in bold.

In the metadata part of the JSON mapping the unique ID (line 1) and a human readable description (line 2) matching the CDA document type described in an implementation guide are referenced. Additionally the FHIR version used (line 3) in the mapping as well as information about the author and the version of the mapping file and general comments are located there.

Since a single CDA document can be decomposed into various FHIR resources, the metadata part includes an Object "resources" (line 5) containing an Array with the specific mappings for each FHIR resource that represents a target of the mapped document type (the mapping for the first resource starts on line 7, for the second on line 19).

Each mapping section of a FHIR resource includes an object "cda-path" (line 7 and line 16) with a String containing the XPATH to the specific XML subtree in the CDA document corresponding to the FHIR resource. A "cda-path" always indicates that the XPATH returns XML nodes. A specific value is mapped using an object "cda-value" (e.g. line 11 or 22) and the corresponding XPATH relative to the XPATH of the corresponding resource. In our example the XPATH in line 11 is applied to the XML nodes returned from XPATH in line 7. If the returned value needs transformation, a "cda-mapping" JSON object (see line 12) can be added. Besides mapped values with the object "cda-value" also fixed values "String", "Boolean", "Number" can be added (e.g. line 8) or prefixes and suffixes for the values (e.g 'urn:oid') can be specified. The date and

timestamp conversions between CDA and FHIR is hardcoded in the parser and no mapping is required.

In the JSON mapping the resources have to be in chronological order. If a patient is referenced for example in the Medication resource, the Patient resource has to be defined before the Medication resource in the mapping file.

3. Results

In order to validate the proposed method to transform CDA documents to FHIR resources using a JSON mapping we created a proof-of-concept implementation using Java, the HAPI-FHIR library [15] and Java reflections. HAPI-FHIR is an open-source implementation of the FHIR specification in Java. Our prototype transforms ELGA e-Medication CDA documents into FHIR resources using the proposed JSON mapping.

The JSON mapping file is parsed using JSON.simple [16], a simple Java toolkit for JSON. FHIR classes from the HAPI-FHIR library are initiated based on the names in the mapping file using Java reflections. To set the fields of the instantiated FHIR classes the suitable setter functions are searched (i.e. *OBJECT.getMethods()*) and also called using reflections. If objects in the JSON mapping do not correspond to the FHIR specification, they are part of the mapping and are handled as described above.

The ID of a FHIR resource is the unique URL created by a FHIR server once the resource is sent to the server. Unless a created FHIR resource is immediately sent to the server, the URL of the newly created FHIR resource is not known. To ensure valid states during the transformation of CDA documents FHIR resources are only uploaded to a server once the whole document has been transformed successfully. Hence a temporary ID for each FHIR resource is created and used to cross-reference the various resources created form the same CDA document.

The resulting Java FHIR objects are serialized as FHIR bundle in JSON format. Using the HAPI-FHIR validator the resulting bundle is validated. We did not implement a FHIR server ourselves, since various open-source implementations of FHIR compliant servers are available online [17].

4. Discussion

In this work we presented a proof-of-concept mapping mechanism from HL7 CDA to FHIR loosely coupled to FHIR resources. FHIR profiles have a similar aim as mappings; they constrain FHIR resources whereas mappings describe how to initialize FHIR resources. For our approach to be applied in a productive setting, FHIR profiles should be more closely incorporated and used by the mapping mechanism. We are currently working on a FHIR profile for the ELGA e-Medication. Based on this profile we aim to analyze how the JSON mapping and future ELGA profiles can complement each other (e.g. validate transformation result with profile). By using a JSON mapping our work represents an initial design proposal to incorporate mappings natively into FHIR. We plan to further close the gap between our mapping and the FHIR specification. Our insights gained from the bottom-up approach may then be used as a starting point for the development of a FHIR resource for the mapping.

Our approach was tested using FHIR DSTU 2 and ELGA e-Medication implementation guide version 2.06.1. The mapping was specifically fine-tuned to these

specifications. If the FHIR specification changes or the ELGA implementation guides are updated, our mapping has to be updated as well. Using the CDA2FHIR_oid and the FHIR_version (see figure 1, line 1 and 3 respectively) the eligible mapping files need to be selected.

Using Java reflections it was easy to reuse the HAPI-FHIR and create a parser. This ease of creation using reflections comes with the price of lower performance compared to other FHIR parsers. In our use case the performance was no problem. Our initial versions of the JSON mapping did not correspond as closely to the FHIR standard and hence our own parser was developed. For future implementations the adaptation of existing FHIR parsers should be considered.

Various approaches to transform HL7 v2 Messages or CDA documents to FHIR resources exist. In [18, 19] HL7 v2 messages are converted into FHIR resources. In the first approach the created Resources are immediately uploaded to the server and no temporary IDs are used. In the second approach a message bundle with temporary IDs similar to our approach is used to serialize the result. The uploading of the bundle to a FHIR server is done by a proprietary FHIR driver developed using HAPI-FHIR. The driver checks if the resource with the temporary ID already exists on the server. If it exists and was authored by the same source, the resource is purged and updated with the new resource. If the resource does not yet exist, a new resource is created. In contrast to [19], we do not check if the resource already exists, we currently only upload the resource.

As part of the Argonaut Project [20] all parts of the CCDA are mapped to FHIR resources and presented as CCDA profile for FHIR. We used their mapping as a starting point to create our own mapping.

In [21] the "CDA to FHIR Transformer" is presented that allows transforming of CDA documents to FHIR resources and vice versa. Unfortunately implementation details were not published. In [22] the structural definitions of a FHIR resource are described using Archetypes. Archetypes enable the authors to use their Archetype based tool to map CDA documents described by Archetypes to FHIR resources, which results in an XQuery containing all the mappings and transformation instructions. In [23] CDA documents are transformed into FHIR resources using XSLT scripts.

Using JSON to describe the mapping between FHIR and legacy data offers the following benefits compared to the other approaches mentioned earlier: XSLT and XQuery are well established standards and are predestined to be used for transformations of XML documents. Yet, whereas XSLT and XQuery are specific implementations of a transformation, our JSON mapping does not predetermine how the actual transformation is performed. The JSON mapping can be used as a starting point to automatically create the needed XSLT script or, as in our case, it could be used from within a programming language like JAVA to perform the transformation.

For legacy reasons CDA documents will stay the preferred way to archive ELGA health data in Austria. In order to reduce information overload and facilitate the access to CDA documents the FHIR API could offer a standardized way to query data at a fine granular level not excluding mobile devices. Using FHIR to access CDA documents opens a plethora of opportunities to use existing tools in the CDA world and could help to increase the value of ELGA in Austria.

Our current mapping file covers the Composition, Patient, Practitioner, Medication, MedicationDispense and the MedicationOrder resources. We are currently refining the JSON mapping to be itself a valid FHIR instance and are creating mappings for the Organization, Location, Encounter and Condition resources also needed to fully cover

the ELGA e-Medication document type. Further mappings for the other ELGA document types are planned.

Acknowledgements

We want to thank Victoria Moser and Anar Osmanbayli for their initial idea to use JSON as a mapping language in their bachelor thesis and Michael Kohler for the inspiring discussions concerning implementation details.

References

[1] Herbek S, Eisl HA, Hurch M, Schator A, Sabutsch S, Rauchegger G, et al. The Electronic Health Record in Austria: a strong network between health care and patients. Eur Surg. 2012;44(3):155-63.
[2] Dolin RH, Alschuler L, Boyer S, Beebe C, Beilen FM, Biron PV, et al. HL7 Clinical Document Architecture, Release 2. Journal of the American Medical Informatics Association. 2006;13(1):30-9.
[3] Integrating the Healthcare Enterprise (IHE). IHE IT Infrastructure (ITI) Technical Framework Revision 9.0, 2012 Mai 2013 [21.1.2016]. Available from: http://www.ihe.net/technical_framework/index.cfm#IT.
[4] IBM Österreich GmbH. Machbarkeitsstudie ELGA. Bundesministerium für Gesundheit und Frauen Familie und Jugend, 2007, Mai 2013.
[5] Boone KW. Where in the World is XDS Mai 2013. Available from: http://motorcycleguy.blogspot.co.at/2010/01/where-in-world-is-xds.html.
[6] Duftschmid G, Rinner C, Kohler M, Huebner-Bloder G, Saboor S, Ammenwerth E. The EHR-ARCHE project: satisfying clinical information needs in a Shared Electronic Health Record system based on IHE XDS and Archetypes. International journal of medical informatics. 2013;82(12):1195-207.
[7] comScore. The U.S. Mobile App Report 2014 [21.1.2016]. Available from: http://www.comscore.com/USMobileAppReport.
[8] Health Level Seven International. Introducing HL7 FHIR 2012 [21.1.2016]. Available from: http://www.hl7.org/implement/standards/fhir/summary.html.
[9] Fielding RT. Architectural styles and the design of network-based software architectures: University of California, Irvine; 2000.
[10] Cerner Corporation. Cerner Launches Developer Experience For SMART on FHIR Applications 2016 [2.3.2016]. Available from: http://www.cerner.com/Cerner_Launches_Developer_Experience_for_SMART_on_FHIR_Applications/.
[11] Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. Journal of the American Medical Informatics Association : JAMIA. 2016.
[12] Wagholikar K. Read only SMART-FHIR façade for i2b2, AID WG session at the HL7 Working Group Meeting held in Atlanta 2015 [21.1.2016]. Available from: http://www.hl7.org/documentcenter/public/wg/java/20151004_Kavi_i2b2-fhir-smart-hl7-10-15.pdf.
[13] Kasthurirathne SN, Mamlin B, Kumara H, Grieve G, Biondich P. Enabling Better Interoperability for HealthCare: Lessons in Developing a Standards Based Application Programing Interface for Electronic Medical Record Systems. Journal of medical systems. 2015;39(11):182.
[14] Integrating the Healthcare Enterprise. Mobile access to Health Documents (MHD) 2014 [21.1.2016]. Available from: http://wiki.ihe.net/index.php?title=Mobile_access_to_Health_Documents_(MHD).
[15] Agnew J. HAPI-FHIR library 1.2 2016 [21.1.2016]. Available from: http://jamesagnew.github.io/hapi-fhir/.
[16] Fang Y. JSON.simple - A simple Java toolkit for JSON 2014 [21.1.2016]. Available from: https://code.google.com/archive/p/json-simple/.
[17] Health Level Seven International. Publicly Available FHIR Servers for testing 2016 [20.1.2016]. Available from: http://wiki.hl7.org/index.php?title=Publicly_Available_FHIR_Servers_for_testing.
[18] Hay D. Mapping HL7 Version 2 to FHIR Messages 2014 [21.1.2016]. Available from: http://fhirblog.com/2014/10/05/mapping-hl7-version-2-to-fhir-messages/.
[19] Heckmann S, Paschke K. Cloverleaf on FHIR: HL7 V2 meets the Message Bundle, Paris, FHIR Hackathon #9, Track 2 2015 [21.1.2016]. Available from:

https://www.hl7.org/documentcenter/public_temp_1ACCEDC7-1C23-BA17-0CD561661A2C29A6/wg/java/20150510_FHIR_Hackathon_AID_Paris.pdf.

[20] Health Level Seven International. Argonaut Project 2015 [21.1.2016]. Available from: http://argonautwiki.hl7.org/.

[21] Lantana Consulting Group. CDA to FHIR Transformer 2014 [21.1.2016]. Available from: https://dev.lantanagroup.com/resources/free-tools/.

[22] Bosca D, Moner D, Maldonado JA, Robles M. Combining Archetypes with Fast Health Interoperability Resources in Future-proof Health Information Systems. Studies in health technology and informatics. 2015;210:180-4.

[23] Smits M, Kramer E, Harthoorn M, Cornet R. A comparison of two Detailed Clinical Model representations: FHIR and CDA. European Journal for Biomedical Informatics. 2015;11(2):en7–en17.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-107

Accessing Patient Information for Probabilistic Patient Models Using Existing Standards

Jan GAEBEL[a,1], Mario A. CYPKO[a] and Heinz U. LEMKE[b]
[a]*Innovation Center Computer Assisted Surgery, University Leipzig, Leipzig, Germany*
[b]*International Foundation for Computer Assisted Radiology and Surgery, Küssaberg, Germany*

Abstract. Clinical decision support systems (CDSS) are developed to facilitate physicians' decision making, particularly for complex, oncological diseases. Access to relevant patient specific information from electronic health records (EHR) is limited to the structure and transmission formats in the respective hospital information system. We propose a system-architecture for a standardized access to patient specific information for a CDSS for laryngeal cancer. Following the idea of a CDSS using Bayesian Networks, we developed an architecture concept applying clinical standards. We recommend the application of Arden Syntax for the definition and processing of needed medical knowledge and clinical information, as well as the use of HL7 FHIR to identify the relevant data elements in an EHR to increase the interoperability the CDSS.

Keywords. Clinical Decision Support Systems, Standardization, Electronic Health Records, Hospital Information Systems.

1. Introduction

1.1. Decision support for complex diseases

Electronic health records comprise of a lot of heterogeneous patient information. This information is gathered from multiple different sources and exists in different formats. Assessing and comprehending all relevant patient information is a difficult and time-consuming process. It is particularly demanding for complex diseases, especially in oncology. Tumor diseases require interdisciplinary collaboration of clinical experts which makes information gathering even more difficult. In specialized tumor boards, patient cases are discussed by these experts. There are systems being developed to enhance the information access of the physicians, e.g. the clinically integrated Oncoflow system developed by Meier et al. [1].

In the course of personalized medicine specific approaches aim to map patient information to digital patient models [2]. A patient model represents the characteristics and specificities of one patient with a certain disease. In clinical decision support systems (CDSS) digital patient models are used to process and interpret the individual case. The goal is to provide intelligent support for physicians and assistance of cognitive processes in the mind of the user. The support opportunities range from diagnosing a patient to

[1] Corresponding Author: Jan Gaebel, Innovation Center Computer Assisted Surgery, University Leipzig, Semmelweisstraße 14, 04103 Leipzig, Germany, E-Mail: jan.gaebel@medizin.uni-leipzig.de

suggesting possible therapy options with the aim of finding the most suitable treatment [3].

1.2. Probabilistic patient models

Lemke et al. propose clinical therapy decision support based on Multi Entity Bayesian Networks (MEBN) [*4*]. With a MEBN model specific therapy decisions with all relevant variables can be described. The relations of nodes in a MEBN, representing medical information entities (IEs), are characterized by conditional probability distributions. They represent the magnitude of influence of one IE on another. MEBNs integrate first order logic with Bayesian probability theory. By setting patient specific information (e.g. clinical and pathological examinations), a patient specific Bayesian Network (PSBN) is generated. The PSBN then infers the probability of occurrence for all unobserved variables in the model (e.g. therapy options, complications and quality of life). This approach is implemented with an exemplary model for laryngeal cancer by Cypko et al. [5]. The model consists of over 1000 information entities and more than 1300 dependencies.

1.3. Restricted information access

The information access of CDSSs is restricted by the communication capabilities by the respective hospital information system (HIS). Although communication via HL7 messages is widely implemented by system vendors, methods to apply the information to additional CDSSs are not considered [6]. There are no efficient approaches to utilize patient specific information for probabilistic patient models. Isolated solutions within one hospital, e.g. adapting existing interfaces of a certain HIS to the CDSS, are unrewarding. The same CDSS would not be usable in a different institution.

In July of 2015 the European Commission decided on 27 IHE-Profiles to be referenced in public procurements. This current development has “potential to increase interoperability of eHealth services and applications to the benefit of patients and medical community” [7]. Following the idea of IHE profiles for other clinical use cases, the information access for digital patient models needs to be revised.

2. Methods

HL7 standards provide various methods and tools aiming to interoperability when dealing with patient information. The HL7 messaging standard is internationally accepted for transferring medical and administrative data between application systems. However, the increasing complexity of the standard made it difficult to fully implement [8]. Furthermore, given its main purpose of data exchange, it is not destined for the inclusion of interpretation or conditional evaluation of the information. Other standards of the HL7 international standards organization (e.g. Arden Syntax and FHIR) may suit the purpose of CDSS integration better.

2.1. Arden Syntax

Arden Syntax is a programming language specifically designed to represent and share procedural medical knowledge. In Arden's Medical Logic Modules (MLMs) single medical decisions are implemented to be used in a broader clinical context [9]. MLMs are optimized to be used in decision support systems. With a variety of tools and methods, the clinical information can be processed and enhanced, always considering the specific patient context. This processed information can be used by CDSSs. Arden's principle of fuzziness is a tool for representing medical vagueness [10]. It makes the description of certain medical conditions more accurate.

MLMs can be called by applications but can also place calls and queries back to HIS, e.g. to request data. These calls, however, are not standardized in any way. In so called *curly brace statements* a specific string is transferred to a host system which needs to interpret it correctly. Then, a certain action is triggered, e.g. a data base query. Depending of the implementation, this string may be interpreted differently in a different institution. This circumstance limits the capabilities of exchanging MLMs between different information systems or institutions respectively.

2.2. FHIR

HL7 FHIR (*Fast Healthcare Interoperability Resources*) is a framework currently in development. It aims to combine features from other HL7 specifications, like HL7 versions 2 and 3 and Clinical Document Architecture (CDA). It targets web technologies as a way of transferring information to be applicable in modern application systems, e.g. mobile health applications [11]. FHIR allows the use of RESTful services to communicate and transfer data in the standardized data formats XML and JSON.

The core component of FHIR is called a *resource*. Resources define and identify relevant data elements within an electronic health record (EHR). A resource represents generalized medical concepts, e.g. observation, body site or order entry. Combining different resources allows constituting any clinical use case as needed. Thereby it can solve the issue of ambiguous data element identification in Arden's MLMs.

2.3. System development

Developing CDSSs, we managed to apply Arden Syntax to adapt and process clinical information in our system. MLMs define decision support rules, representing clinical knowledge for an exemplary use case from ENT oncology, the laryngeal cancer model. For example, we implemented MLMs to evaluate the cancer cell infiltration of the larynx using data from different clinical examinations. The information is aggregated and adjusted, so it can be used by our CDSS. MLMs allow quantifying certain information using other data processing methods, e.g. natural language processing, and adding elements of uncertainty as needed. However, when querying the data from the data source (e.g. the database of the patient data management system), our implementation is fixed to the characteristics of this particular system. Exchanging our implementation with another institution would require intense adjustments concerning the data access.

3. Results

As the result of our investigations, FHIR in combination with Arden Syntax is a well-suited technology to integrate CDSSs into existing information system environments. We propose a combination of MLMs and FHIR resources to be able to access information accurately within a CDSS. This approach extends the suggestions of Kimura and Ishihara in [12]. In the following, our system architecture is presented.

We based the CDSS on the laryngeal cancer disease model [5]. Figure 1 illustrates the architecture and communication paths. Our suggested system architecture follows the Medical Information and Model Management System (MIMMS) architecture, proposed by Lemke et al. in [13]. It separates the tools and engines for modeling and model processing from the actual data source and presentation destination. By itself it makes interoperability possible. The probabilistic disease model, based on Bayesian Networks *(a) [see Fig. 1]*, contains all relevant information entities that are needed to evaluate a patient's case and calculate possible therapy options. In MLMs *(d)*, we defined the clinical knowledge that the network considers. For each node we created a MLM that delivers the node specific information. All relevant clinical rules and processing steps are defined. The user selects the patient context *(1)* and respective model *(2)* over a clinical user interface *(g)*. To be able to calculate the PSBN *(f)* the needed patient information is requested via curly brace statements in the appropriate MLM. To enable interoperability and reusability the curly brace statement calls a FHIR resource. Appropriate FHIR templates are stored and filled with references to the desired information when called. References, like patient ID or a SNOMED code for clinical findings, are handed over through the curly brace statements. This allows describing and identifying the needed data elements. These FHIR resources are sent *(4)* to the patient database *(e)* to request the desired information. The corresponding response *(4)* and *(5)* contains the information in the FHIR format, as well. Any information system that integrates HL7s specifications should be able to receive the requests and correctly deliver the information. The information is then assessed by the MLM that requested the data.

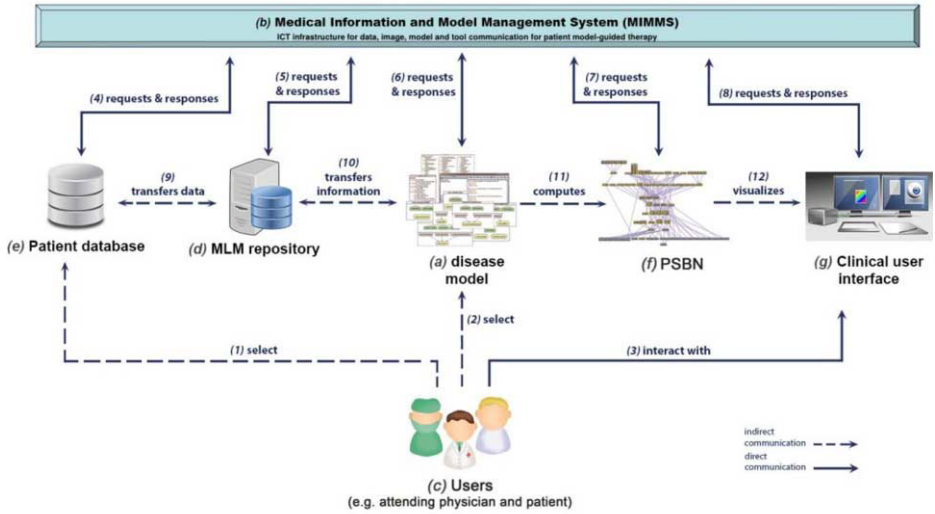


Figure 1. System architecture for CDSS based on Bayesian Networks and access to patient specific data.

Depending on the context that is implemented in the MLM, required calculations are performed on the data, e.g. aggregation of several examinations, quantification of alphanumeric data or fuzzification of uncertainty in findings. This information is handed over to the disease model *(9)*, *(10)* complying with the MIMMS architecture. It is applied to the respective node of the Bayesian Network. Depending on the implementation these nodes expect either numeric values, probability values or even Boolean values, which are provided by the MLMs. The patient specific model is computed *(11)* and the results are transferred to the clinical user interface *(12)* where it is finally presented to the user.

Applying modern communication technologies, e.g. REST, allows our system to be used between different application systems, operating systems and possibly different devices, like mobile tablet computers, as well. Applying standards to the individual elements of the system, like an Arden engine and the way it requests information, increases the interoperability and reusability even more. In addition, separating the definition of clinical knowledge and decision support rules from the data identification and -transfer, facilitates updating and maintaining the CDSS. For example, Medical experts in collaboration with technicians can add new medical knowledge and reasoning while technicians or computer scientists revise the data access.

4. Discussion

Our proposed CDSS combines different standardization frameworks and strategies. With Arden Syntax MLMs we define the medical knowledge and relations to evaluate a patient's case and provide the necessary processing for the calculation of PSBNs. With FHIR resources the needed data elements are defined to precisely identify patient specific information. We aim to enabling the interoperability of clinical data and the reusability of CDSSs in different institutions and hospital information system environments respectively.

MLMs have the advantage that single medical conditions and decisions can be described. Several MLMs, either in a nested architecture or called concurrently, can express very complex relations. The complexity is predetermined by the level of detail in the Bayesian Network. Assessing and processing complex diseases, e.g. laryngeal cancer, implies complex digital models. Our proposed architecture concept combines several technical methodologies. This initially increases the complexity of the disease modeling and system development process. Maintenance efforts increase as well. New medical knowledge, e.g. new diagnostic or therapeutic procedures, requires an adaptation of the fundamental Bayesian Network. For every alteration the corresponding MLMs and FHIR resources also need to be adjusted or new ones need to be added. Yet, to be able to describe a complex disease and assess it in a patient specific context, different elaborate technologies, like Bayesian Networks in combination with Arden Syntax and FHIR, must be applied.

Where Arden Syntax is a well-established standard, used in many decision support systems from different institutions, FHIR is still in development. Its predecessors, e.g. HL7 version 2 and 3, are widely used. FHIR will combine the best features to extend the application possibilities of clinical data communication. On the other hand, relying on standards always depends on software vendors to apply the standards as well. Current developments, like the recognition of IHE profiles on a European level, show that in the future the integration of different information- and decision support systems will be improved.

We will extend our implemented CDSS with effective methods of connecting MLMs with FHIR resources. The goal is to create an added value to routinely recorded patient data by making it available in CDSSs. Thereby we hope to facilitate the physicians' work and increase the patients' safety by finding the best suitable treatments.

5. References

[1] Meier J, Boehm A, Kielhorn A, Dietz A, Bohn S, Neumuth T. Design and evaluation of a multimedia electronic patient record "oncoflow" with clinical workflow assistance for head and neck tumor therapy. Int J Comput Assist Radiol Surg, 2014, ISSN: 1861-6429.

[2] Chan IS, Ginsburg GS. Personalized Medicine: Progress and Promise. Annual Review of Genomics and Human Genetics. 2011;12(1):217–44.

[3] Yu, PP. Knowledge Bases, Clinical Decision Support Systems, and Rapid Learning in Oncology. Journal of Oncology Practice, 24. Februar 2015, JOP.2014.000620. doi:10.1200/JOP.2014.000620.

[4] Lemke HU, Berliner L (2008): Modellgestützte Therapie, patientenspezifisches Modell und modellbasierte Evidenz (Model based therapy, patient specific model and model based evidence). In: Niederlag W, Lemke HU, Meixensberger J, Baumann M (Editors): Modellgestützte Therapie, Health Academy. Dresden, Germany, Vol 13, pp 13-24.

[5] Cypko MA, Stoehr M, Denecke K, Dietz A, Lemke HU. User interaction with MEBNs for large patient-specific treatment decision models with an example for laryngeal cancer. Int J CARS. 2014;10(1).

[6] Lemke HU. Medical device integration in the OR, Interoperability Standards and Issues relating to International Approval Procedures. Health Management Journal. 2015; 15(1):66-73.

[7] Commission Decision (EU) 2015/1302 of 28 July 2015 on the identification of 'Integrating the Healthcare Enterprise' profiles for referencing in public procurement.

[8] Braunstein, M. Practitioner's Guide to Health Informatics. Springer Publishing Company, Incorporated, 2015. ISBN 3319176617.

[9] Samwald M, Fehre K, de Bruin J, Adlassnig K-P. The Arden Syntax standard for clinical decision support: experiences and directions. J Biomed Inform. 2012 Aug;45(4):711–8.

[10] Vetterlein T, Mandl H, Adlassnig K-P. Fuzzy Arden Syntax: A fuzzy programming language for medicine. Artif Intell Med. 2010 May;49(1):1–10.

[11] HL7 Fast Healthcare Interoperability Resources Specification 'FHIR™', Release 1. HL7 International. 2014-02-02. Retrieved 2016-01-26.

[12] Kimura E, Ishihara K. Internal domain-specific language based on Arden Syntax and FHIR. Stud Health Technol Inform. 2015;216:955. PubMed PMID: 26262257.

[13] Lemke, HU, Cypko MA, Berliner L. 5. Der virtuelle Patient im Rahmen der Therapie-planung am Beispiel des Larynxkarzinoms. In: Der virtuelle Patient. De Gruyter, 2014. ISBN 9783110334296.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-113

Developing an Approach for Aggregating Information From a EHR into a Fully Structured Patient Summary

Patrick MANGESIUS[a,1], Samrend SABOOR[b] and Thomas SCHABETSBERGER[a]
[a]*ITH icoserve technology for healthcare GmbH, Innsbruck, Austria*
[b]*Siemens Healthcare GmbH, Erlangen, Germany*

Abstract. Background: Providing healthcare professionals with adequate access to well-filled electronic patient records and health-related information contributes to an improvement of the treatment process. Compared to conventional Electronic Medical Records these EHR systems commonly contain more medical artifacts due to their cross institutional, multipurpose use cases. Physicians and health professionals are therefore concerned about information overflow. Objectives: Goal of this paper is to elaborate new concepts for the automated aggregation of a fully-structured patient summary document based on information extracted from documents which are published in large-scale EHRs. Methods: The first step was the conduction of semi-structured group interviews with experts and customers. This was followed by a qualitative literature analysis. Consequently technical and medical standards in the field of interoperability were screened. Results: The result of this paper is the elaboration of an architectural approach to integrate an automatic patient summary creation into well established workflows of large ehealth projects based on standard IHE XDS infrastructures, taking the Austrian ELGA as an example..

Keywords. Clinical Data analysis, Electronic Health Records, Patient Summary.

1. Introduction

eHealth is often referred to as an essential prerequisite for more efficient, cooperative healthcare [1]. Providing medical practitioners and other healthcare professionals with adequate access to well-filled electronic patient records and health-related information contributes to an improvement of the treatment process and, ultimately, the quality of patient care. In order to achieve this goal, countries and healthcare organizations introduce generic eHealth infrastructures for Electronic Health Records (EHR) [2]. The requirements and quality aspects of these infrastructures have already been examined in various approaches [3]. One of the most recent examples of such infrastructures is the Austrian "Elektronische Gesundheitsakte" (ELGA) which officially went online in December 2015. It is a nation-wide virtual EHR that interconnects all hospitals (grouped according the different federal states). Such an EHR, however, is expected to contain much more documents and health-related artefacts than conventional EMRs (Electronic Medical Records). This raises concerns of information overflow amongst the physicians and health professionals [4, 5]. It is therefore crucial to reduce and aggregate the amount of information that is returned by queries of the EHR. The healthcare professionals

[1] Corresponding Author: Patrick Mangesius, ITH icoserve technologies for healthcare GmbH, Innrain 98, 6020 Innsbruck, Austria. E-Mail: patrick.mangesius@ith-icoserve.com.

should be only presented with those details that are absolutely necessary for answering their needs in respect to their information context (e.g. emergency or normal treatment situation) and role (e.g. physician, nurse, radiologist). Ideally, this aggregation is conducted by automated extraction mechanisms that are well integrated into the EHR infrastructure (i.e. use existing interfaces and standards) and provides the possibility to flexibly define which pieces of information are of relevance for answering a specific information need. A patient summary is a standardized set of basic medical data that includes the most important clinical facts required to ensure safe and secure healthcare. epSOS defined a special form for electronic health records in European context [7].
In a first stage the compilation of a patient summary document would be already sufficient. Of course, these mechanisms must be able to deal with the amount of information available in national EHRs. This paper describes the concept and architecture on a technical level, how it fits the Austrian EHR ELGA and explains how it was developed.

2. Methods

The authors have been actively involved in the development of the concepts for the Austrian EHR ELGA. They have implemented ELGA-compliant software products that are in productive use. Based on this domain knowledge, a combination of qualitative research methods was used to elaborate requirements for a system that automatically extracts relevant information out of an EHR and aggregates these into a fully-structured patient summary:

In the first step, semi-structured group interviews with experts and customers were conducted. The aim was basically to confirm the necessity of information aggregation in EHRs which is also stated in recent publications such as Duftschmid et al [9]. The second aim of these interviews was to gain an understanding of the extent or scope of the extraction and aggregation.

In the second step, a qualitative literature analysis was conducted. It focused on the prerequisites of generic patient summaries in combination with EHRs. The analysis was conducted on publications registered in PubMed and Google Scholar and included only papers published 2005 and later. The analysis was carried out in iterative steps which used the following combinations of keywords:

- First iteration: “patient summary“ AND “IHE”. This combination resulted in five hits
- Second iteration: “patient summary” AND “EHR”. This combination resulted in 41 hits
- Third iteration: (“IHE“ AND “search”) AND “content”. This resulted in four hits

The review stopped when the amount of redundant results increased and a content-wise saturation became obvious. In the end, three papers were identified as relevant.

As a last step, technical and medical standards in the field of interoperability were screened. The main focus was on health record frameworks defined by IHE [6], such as IHE-XDS (Cross-Enterprise Document Sharing). Furthermore, ELGA and its usage of the previously mentioned standards was analyzed using [8].

3. Results

The resulting requirements from the interviews and the literature analysis, amongst others, mainly expressed the need for:

- Content based searching to avoid information overload by manually screening the documents contained in the whole EHR
- Content based search capable of answering queries based on fine grained information items, such as keywords or codes
- Summarization functionality of patient condition, history, hot spots and diseases based on the available information in connected (sub) systems.

Last but not least, a main finding was that the document indexes of current EHRs which are based on the IHE profile XDS are by far not able to provide the necessary content for a fully-structured patient summary document. Thus, an index which comprises more document-related details is required as a crucial part of an automated aggregation architecture. Of course, the respective indexing services for populating the index with the required information are needs as well. Therefore, two new actors are introduced into the workflow of publishing document information:

- Content Indexer
- Content Index

Together these two services are intended to integrate into existing architectures and workflows and work possible with standardized interfaces as far as. In contrast to the approach described in the EHR-ARCHE project [9], the proposed architecture does not focus on pairing the indexing actors with the consuming parts (i.e. the XDS Consumer). Instead, the indexing actors are integrated into the content producing workflows. In large eHealth projects like the ELGA, all accesses to the document infrastructure within the ELGA context must be associated with a responsible natural person and not by automated means. This is ensured by an active patient contact [8].

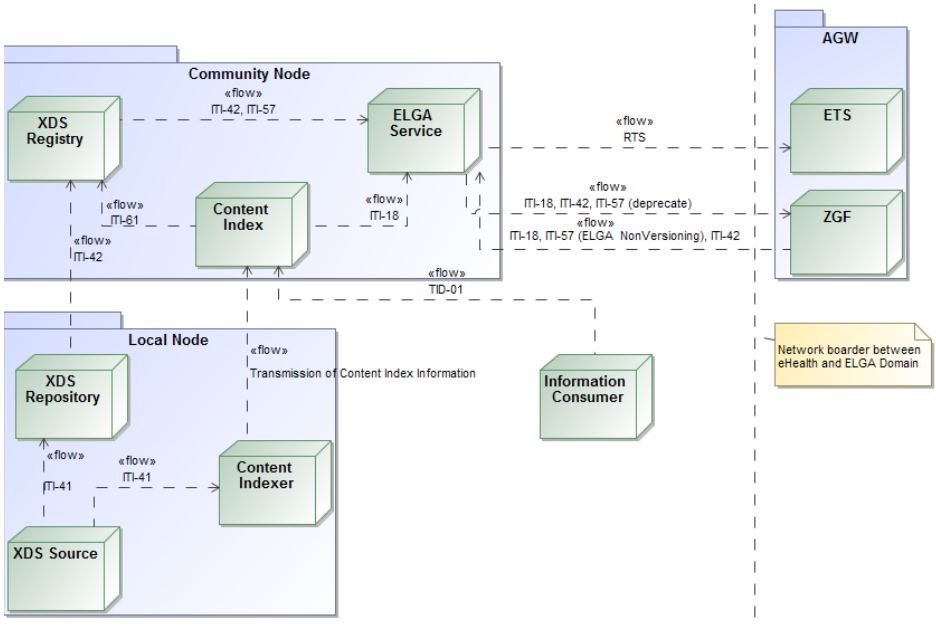


Figure 1. The indexing components integrated into an eHealth domain using ELGA Variant C [8].

An example how the services fit into an ELGA architectural variant C according to [8] deployment is outlined in **Figure 1**.

3.1. Content Indexer

This service is responsible for parsing and processing documents. As an asynchronous process it extracts all needed information. There are already a lot of publications available on the topic of what and how information can be gathered and analyzed [9, 10]. The *Content Indexer* is intended to be in the scope of the healthcare provider's infrastructure. This ensures that the documents never leave the local provider. The service is capable of retrieving IHE ITI-41 (Provide and Register Document Set) compliant requests [6]. This ensures that it seamlessly integrates into an IHE compliant source infrastructure. Using the IHE XDR (Cross-enterprise Document Reliable Interchange) [6] profile, the client system (which represents the author of the document) is also capable of deciding which documents should be used for indexing and which not. After the process of analyzing and extracting data from the document, the information is sent to the *Content Index*. This transaction must contain the following information:

- Identification information of the document (e.g. Document Unique ID, the ID by which the document is uniquely identified within the IHE Affinity domain and therefore is known to the XDS Registry)
- Index Information: information that is needed to fulfill search requests for arbitrary content of the documents. These could be proprietary to already existing industry established full-text search engines, such as Apache Lucene [11]. Those search engines commonly provide patterns and architectures for remote indexing and processing.
- Extracted patient summary specific information. The document is processed on site at the hospital. The central component of the indexing system, the Content Index is not intended to get in contact with the original document. Therefore the information must be preprocessed on site and the results of this operation are sent to the central index. The extraction process has already been discussed in [9].

3.2. Content Index

In contrast to the *Content Index*er, the index is located in the community node. The service must be provided by the operator of the affinity domain, since it is a central component like the XDS Registry or PIXPDQ Manager. The Service has two main functionalities:

- Storing the index information for responding to search requests containing parts of the full-text of the documents. For search requests no already existing standard could be identified. Therefore a new transaction needs to be defined, which will be referred as TID-01. This transaction should be an extended ITI-18 query. These stored query requests defined by IHE should be extended with a new Stored Query UUID [6]. This will identify a content based search request, the parameters of which are transported in the slot of the message.
- Assembling and registering an on-demand Document [6] to the XDS Registry. This document contains the patient summary. It will be registered to the XDS affinity domain and can be viewed by any consumer. This is done by executing an IHE ITI-61 transaction from the Content Index services to the XDS Registry.

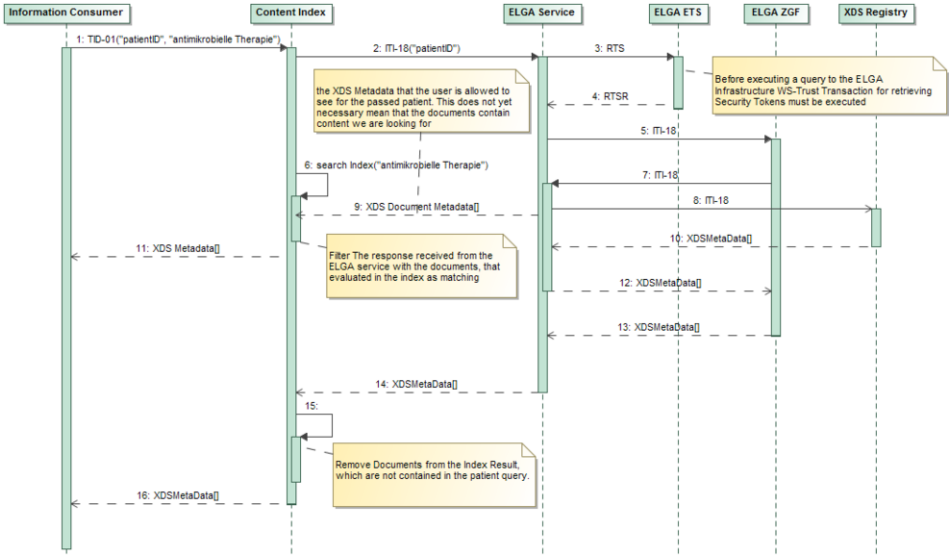


Figure 2. Workflow sequences of querying the full-text search index in an ELGA Variant C scenario. ELGA ETS and ELGA ZGF are explained in [8].

Figure 2 outlines the integration of the workflow in an ELGA Variant C query process.

4. Discussion

The objective of this paper was to introduce a new concept for the automated aggregation of a fully-structured patient summary document based on information extracted from documents which are published in large-scale EHRs (such as the Austrian ELGA). A main prerequisite here is that such a system seamlessly integrates into existing state-of-the-art EHRs. The international de facto standard for information exchange between EHRs is the IHE XDS profile along with related IHE profiles. Thus, such an aggregation architecture must be compliant with the respective IHE profiles.

In order to clarify whether a patient summary is actually needed in productive EHR-environments and to elaborate the basic requirements for an aggregation architecture, semi-structured group interviews with experts and customers were conducted. They clearly confirmed the necessity of such a system. The subsequent qualitative literature analyses resulted only in three relevant publications although the queries were conducted using broad search terms. An explanation for this circumstance might be that XDS-based EHRs just started to be established in the last couple of years (e.g., the XDS profile was first introduced around 2007). However, further effort needs to be invested into extending the analysis.

The developed approach successfully integrates into workflows and processes in large eHealth projects, as shown on the example of ELGA. Capturing data in its registration process allows the extraction of information without executing automated transactions that do not involve human actors (which is always regarded suspiciously and hard to explain to the patient). Since all document indexes are associated with

document uniqueIDs [6] it is possible to integrate into current scenarios and access control systems. In case of ELGA, the developed architecture can interact with the workflow in a way that patient consents are respected. No information is sent to the user that is not aligned with the access rights.

However, despite these benefits it also has some severe short comings that must be addressed in future research. One of the most important drawbacks is that currently the full-text search and the information contained in the patient summary only represents the information from one affinity domain. There is no concept yet on how to achieve a cross community full-text search. Such a cross-community-spanning query requires a concept for a hierarchical index (or alternatively, for doing queries across multiple affinity domains). This is one of the most crucial future works on this topics. Cross affinity domain communication is a core concept of many eHealth projects and, therefore, there is the need to answer the question on how to create a patient summary containing all the needed data. Therefore, one of the major next steps will be to establish a concept of developing cross community full text searches.

Up to now, only the concept for the extraction and aggregation architecture exists. In order to be able to assess the practical needs and usability of it, a proof of concept is required.

Last but not least, auditing requirements for such a system must also be elaborated in order to reflect data security and privacy. IHE systems are required to audit operations that are executed using the IHE ATNA profile [6]. This also needs to be applied for the *Content Indexer* and *Content Index*.

References

[1] Stroetmann K, Jones T, Dobrew A, Stroetmann V: The economic benefits of implemented eHealth solutions at ten European sites. Bonn (Germany): eHealth-impact 2006.

[2] Fayn J, Ghedira C, Telisson D, Atoui H, Placide J, Simon-Chautemps L, Chevalier P, Rubel P: Towards new integrated information and communication infrastructures in e-health. Examples from cardiology. In: 2003. IEEE: 113-116.

[3] Hoerbst A, Schabetsberger T, Hackl W, Ammenwerth E: Requirements regarding quality certification of Electronic Health Records. Stud Health Technol Inform 2009, 150:384-388.

[4] accenture: Connected Health: The Drive to Integrated Healthcare Delivery. In. http://www.himss.eu/sites/default/files/Accenture-Connected-Health-Global-Report-FinalWeb.pdf: accenture; 2012.

[5] O'malley AS, Grossman JM, Cohen GR, Kemper NM, Pham HH: Are electronic medical records helpful for care coordination? Experiences of physician practices. Journal of general internal medicine 2010, 25(3):177-185.

[6] (IHE) IHE: Infrastructure Technical Framework. In: Cross-Enterprise Clinical Documents Sharing (XDS), by the IHE ITI Technical Committee, Version. vol. 12; 2015.

[7] (Swedish Association of Local Authorities and Regions) epSOS: Electronic patient smart open services. Available at http://www.epsos.eu. Last accessed on March 02, 2016.

[8] GmbH E: ELGA Gesamtarchitektur, 2.20 edn: ELGA GmbH; 2015.

[9] Duftschmid G, Rinner C, Kohler M, Huebner-Bloder G, Saboor S, Ammenwerth E: The EHR-ARCHE project: Satisfying clinical information needs in a Shared Electronic Health Record System based on IHE XDS and Archetypes. International journal of medical informatics 2013, 82(12):1195-1207.

[10] Krummenacher R, Simperl E, Cerizza D, Della Valle E, Nixon LJ, Foxvog D: Enabling the European patient summary through triplespaces. Computer methods and programs in biomedicine 2009, 95(2):S33-S43.

[11] McCandless M, Hatcher E, Gospodnetic O: Lucene in Action: Covers Apache Lucene 3.0: Manning Publications Co.; 2010.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-119

Impact of Internet on Cytology Information Management

Ljerka LUIĆ[a,1] and Livia MOLNAR[a]
[a] *Karlovac University of Applied Sciences, Croatia*

Abstract. Internet technologies and services impose global information standards in the sphere of healthcare as a whole, which are then implied and applied in the domain of cytology laboratories. Web-based operations form a significant operating segment of any contemporary cytology laboratory as they enable operations by the use of technology, which is usually free of the restrictions imposed by the traditional way of business (geographic area and narrow localisation of activities). In their operations, almost all healthcare organisations currently create and use electronic data anddocuments , which can originate both inside and outside the organisation. An enormous amount of information thus used and exchanged may be processed timely and in a high-quality way only by integrated information systems, given three basic safety requirements: data confidentiality, integrity and availability. In the Republic of Croatia, integration of private and public healthcare information systems has been ongoing for several years but the private healthcare does not yet operate as an integrated system. Instead, each office operates using its own separate information system, i.e. database. This paper elaborates the argument that the sample private cytology laboratory possesses an IT system that meets current market and stakeholder needs of the healthcare sector in Croatia, given that private doctors' offices/polyclinics use IT technologies in their operations but make only partial use of Internet capacities in the segment of communication with their business associates and patients, implying the need to continue the research on a statistically relevant sample of EU countries.

Keywords: Internet, Cytology Laboratory, Management Information Systems

1. Introduction

The Internet currently stands for a global multimedia distributed information system encompassing almost the entire world and enabling not only access to spatially distributed multimedia content, but also generation of specific business/medical content tailored to specific user queries in an interactive operating regime, such as requests by patients and/or doctors' offices/polyclinics for cytology test results. Accordingly, this paper aims to present the impact of Internet technologies on the management of business and health information in a cytology laboratory, given the increasing importance of Internet worldwide, to which Croatia is no exception. This paper intends to clarify the functioning of the IT system of a private cytology laboratory in Croatia observed in terms of studying the impacts of informatization, and how well the observed laboratory has adapted to using the Internet to meet its stakeholders' needs.

The paper focuses on exploring the implications of the Internet on the implementation of an integrated IT system in private healthcare institutions in Croatia, as seen in the practice of medical laboratories, on the example of the three currently

[1] Corresponding Author: Ljerka Luic, Karlovac University of Applied Sciences, Trg J.J. Strossmayera 9, 47000 Karlovac, Croatia. Email: ljerka.luic@vuka.hr

operating private cytology laboratories, in which 90% of the throughput consists of pap-smear analysis (gynaecological cytology). This is owing to the fact that private gynaecology clinics in Croatia by far outnumber cytology laboratories and increasingly seek collaboration with them, mainly to ensure operational transparency, but also to allow for faster turnaround time compared to results obtained from a state-owned laboratory. In light of these facts, it is clear that there is an outspoken demand for fast turnaround time. However, the speed of processing should not impair the quality of analysis and of medical data processing. In the sample laboratory studied for this paper, all specimens are analysed personally by medical experts – cytologists, who perform the analysis in person by use of microscopy as well as the most recent technologies and make their conclusions based on their own knowledge and experience as well as team consultations. The performance standard of the laboratory requires that all received specimens be processed within a maximum of 2 work days, which fulfils the market requirements in terms of turnaround. Improvement of turnaround time and the quality of results/medical data are the main reasons why the management of the observed private laboratory decided to design and implement an integrated business-medical information system, with an imperative to meet the laboratory's operating standards as well as requirements of both corporate and individual clients.

2. Methods

Since our research topic, implications of Internet on the cytology information system, required us to draw generalizations (applicable in most private laboratories) from a specific case (model private cytology laboratory), as well as to apply general reasoning to specific cases, the paper relies on both inductive and deductive methods, but with a greater focus on inductive ones. Although the methods of analysis and synthesis work in different directions, they intersect and complement each other, which is why we have used both. In examining the observed implications of Internet we used the method of comparison and differentiation, which resulted in the description of the identified respondent groups. As a method of consolidating the data collected during analysis, synthesis has been used in the stage of summarizing the results, where we aimed to provide an appropriate graphic representation of the findings (shown below). A questionnaire has been used as a quantifying instrument for the research, which may be used as a control sample in further study. We divided a total of 14 questions into two groups, by the following two topics: (1) Harmonization of the business and medical information system, and (2) Use of Internet technologies in daily operations.

3. Results

This research was conducted in 15 doctors' offices/polyclinics out of a total of 40 with which the sample laboratory collaborates. This satisfies the criterion of Yin's[2] reflection of 'literal replication' in which all the cases are theoretically the same, which also fulfills the main aim of preliminary research.

The respondents were asked to circle the job they perform in the doctor's office/polyclinic. If they performed several jobs, they were required to circle the job

[2] Yin, R. K., *Case Study Research: Design and Methods*, Newbury Park: Sage Publications, 1989

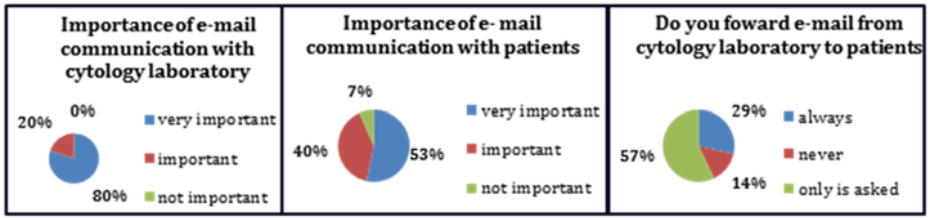
which they think they perform most of the time. Out of all the respondents that completed the survey, 46% were physicians, 27% managers, 20% nurses and 7% administrative staff. Selected results are shown in Table 1.

- *When asked about the purpose of the information system*: 60% of the respondents answered that the purpose of the information system was monitoring the information flow, 73.3% of the respondents think that the purpose is connecting those who directly participate in the business process; 80% think that the purpose also includes connecting with clients and colleagues. In addition, 53.3% of the respondents think that the purpose of the information system is faster and better decision-making; while 46.7% of the respondents answered that the information system serves to meet only the basic business requirements.
- *When asked about the importance of information flow*: 66.7% of the respondents think that information flow is important for improving the operational efficiency of a doctor's office/polyclinic; 26.7% think it brings an increase in profitability; 86.7% think it contributes to customer satisfaction; while 6.7% of the respondents think that the information flow has no significant impact; 6.7% think that it is just an additional inconvenience.
- *When asked about the impact of the usage of information and communications technologies on the doctor's office/polyclinic operations:* 26.7% of the respondents answered that they consider them useful but not indispensable; 73.3% think that their work depends on them; 46% think that information technologies might influence the practice in the future; while 26.7% think that their usage is crucial for future business operations. Only 6.7% of the respondents consider information technologies a distraction and as something that slows down the work process.
- *When asked about computer skills and the importance of the Internet:* 40% of respondents answered that they have average computer skills; 33% of them think their computer skills are very good; 20% that they are excellent; and 7% that they are below average. When asked whether they use a computer in their everyday work, all the respondents gave an affirmative answer.
- *When asked about the importance of client communication via the Internet with their colleagues and patients*: 66.7% of respondents answered that communication via the Internet is very important to them, while 33.3% of them do not attach any importance to communicating with colleagues via the Internet. 60% of the respondents consider communicating with clients via the Internet extremely important, while 26.7% think it is important, and 13.3% think that it is not important at all.
- *When asked about using the Internet to receive medical reports from the cytology laboratory via email, 80% of the respondents said they consider it* very important, and 20% that they consider it unimportant. 53% of the respondents consider *using the Internet in forwarding medical reports via email* very important for the patients; 40% of them consider it important; and only 7% unimportant.
- *Importance of reports received from the laboratory in paper form via regular mail*: 46.7% of the respondents consider them to be more important than the reports received via email. This means that 53.3% of the respondents still

consider the printed medical report received 'on paper' to be more important than the one received via email.

- *Importance of reports sent to the patient in paper form via regular mail:* 46.2% of the respondents consider them to be more important to the patients and only 23.1% of the respondents consider the report received via email to be more important to the patients primarily because they can access the data faster.
- When asked how often they forwarded reports to the patient via email, upon receiving them from the laboratory in that form, only 29% of the respondents said they almost always forward the finished report; 57% only forward them if the patients request it themselves; and 14% never forward the report using the Internet (via email).

Table 1: Using the Internet to receive medical reports from the cytology lab via email and forwarding medical reports to patients via email.



4. Discussion

Analysis of the data obtained in the research of "implications of Internet on the cytology information system", clearly shows that the researched polyclinics/doctors' offices all have a basic form of healthcare information system. In practice, however, these systems differ greatly in terms of scope, usage, and integration via the Internet (particularly in the form of email). The reasons for this can be found in the economic maturity of a particular polyclinic/doctor's office and in the strategic direction of their management concerning the application of Internet in everyday business. Furthermore, it is evident that the healthcare information system of public institutions in the Republic of Croatia is compliant with the healthcare information systems of the most developed European countries, and we can see the areas for its further improvement in the private sector.

4.1. The usage of Internet in the sample cytology laboratory in Croatia

Since the cytology laboratory usually works with polyclinics and doctors' offices, and not directly with patients, finished reports are sent to email addresses of the doctors' offices with which the laboratory collaborates. In the case that the patient personally comes to the laboratory, the report is sent via email to the patient and to his or her home address via regular mail in paper form. With regard to archiving reports, it is important to mention that all laboratories in the Republic of Croatia have a legal obligation to keep the reports in paper form for 10 years. This includes microscope slides and the accompanying form, as well as a copy of the report. Colleagues often need a transcript of a report.

The report is then retrieved from the archives and an official transcript is created and signed by the cytologist. Likewise, reports need to be retrieved from the archives if the sample needs to be reanalyzed for any reason. The archives are also available in electronic form in a structured database. All patient data for every patient who was ever processed in the laboratory can be found in that database. The relevant report can be found by entering any single point of patient data. It is also possible to search by doctor's office ID number, patient name and surname, date, or laboratory notebook number. The laboratory information system provides access to a large amount of medical and personal patient information easily and quickly, and it can be accessed from any place that has interned access.

4.2. Usage of Internet in cytology laboratories in EU countries

By using the inductive method of reasoning in the analysis of the results of research conducted in the sample cytology laboratory in Croatia, which we chose as our case study, it can be concluded that the healthcare sector in Croatia follows the trends of Internet usage in its operations, but it lags behind in using the Internet for communicating with patients who use laboratory services.

The continuation of research on the "implications of Internet on the cytology information system" is justified within the context of contribution to scientific findings, such as finding algorithms and variables which would allow us to describe more precisely the changes within the constructs of the simulation model, as well as the interactions of their elements: laboratory - polyclinic - doctor's office - patient, and for which a scientific study should be conducted on a large statistically significant sample of respondents that would include all stakeholders.

Further research on the "implications of Internet on the cytology information system" is also justified in the context of applying the mentioned model of integrated information systems in cytology laboratories in EU countries. It is beyond doubt that the effectiveness of the usage of the Internet technologies in the healthcare business operations as a whole lies in the synergistic communication of its nucleus – the patient as an end user of laboratory services. Undertaking further research in laboratories throughout Europe is also justified in the domain of management in healthcare institutions, where it can be seen in the context of rethinking management of polyclinics/doctors' offices through administrative use of Internet.

References

[1] R. K. Yin, *Case Study Research: Design and Methods*, Newbury Park: Sage Publications, 1989.
[2] G. Surján, R. Engelbecht, P. McNair, *Health Data in the Information Society,* IOS Press Amsterdam, 2002.
[3] E. Dreizen, M.S. Lima, R.B. Kitagawa, F.J. Corronos, O. Spinelli, Development of a database for cervical database reports, *Informatics in Primary Care* **19** (2011), 119-124.
[4] V. Radisic Biljak, I. Ozvald, A. Radeljak, K. Majedic, B. Lasic, Z. Sifar, M. Vucic Lovencic, Z. Flegar-Mestric, Validation of a laboratory and hospital information system in a medical laboratory accredited according to ISO 15189, *Biochemia medica* **1** (2012), 86-91

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-124

Requirements for Workflow-Based EHR Systems – Results of a Qualitative Study

Marco SCHWEITZER [a,1], Nelia LASIERRA [a] and Alexander HOERBST[a]
[a] *eHealth Research and Innovation Unit,*
UMIT - University for Health Sciences, Medical Informatics and Technology

Abstract. Background: Today's high quality healthcare delivery strongly relies on efficient electronic health records (EHR). These EHR systems or in general healthcare IT-systems are usually developed in a static manner according to a given workflow. Hence, they are not flexible enough to enable access to EHR data and to execute individual actions within a consultation. Objectives: This paper reports on requirements pointed by experts in the domain of diabetes mellitus to design a system for supporting dynamic workflows to serve personalization within a medical activity. Methods: Requirements were collected by means of expert interviews. These interviews completed a conducted triangulation approach, aimed to gather requirements for workflow-based EHR interactions. The data from the interviews was analyzed through a qualitative approach resulting in a set of requirements enhancing EHR functionality from the user's perspective. Results: Requirements were classified according to four different categorizations: (1) process-related requirements, (2) information needs, (3) required functions, (4) non-functional requirements. Conclusion: Workflow related requirements were identified which should be considered when developing and deploying EHR systems.

Keywords. expert opinions, electronic health records, diabetes mellitus, semantics.

1. Introduction

The electronic health record (EHR) is seen as a major factor for reaching high structure-, process- and outcome quality in healthcare. One of its major advantages is given through the possibility of sharing a lifelong interoperable patient documentation across different institutions. Today, a major challenge of EHRs is related to data access: In order to avoid information overload, it is necessary to facilitate the retrieval of precise and particular required data. During a patient examination, the main goal of data retrieval for a physician is to get all the required information, displayed in an appropriate manner for creating a further treatment decision. Therefore, interfacing the EHR in a usable and functional manner is necessary. Available works in the literature have elaborated on how to offer the needed information [1,2]. However, none of them considered the further processing of the retrieved data as part of the physicians' duties.

In the context of a particular clinical situation, information logistics should go along with the intended workflow a physician, nurse or any other healthcare professional is following. Thus, any IT-system should adapt to the workflow and proactively support the physician without hindering medical work.

Available EHR systems or in general healthcare IT-systems offer user interfaces tailored to a static sequence of activities and thus are not flexible enough to fit individual,

[1] Corresponding Author: Marco Schweitzer, University for Health Sciences, Medical Informatics and Technology (UMIT), Eduard-Wallnoefer-Zentrum 1, 6060 Hall i.Tirol, E-Mail: marco.schweitzer@umit.at

dynamic workflows. From a physician's perspective, functions for accessing the EHR are limited to basic query retrieve actions without any particular medical workflow relation. The project OntoHealth[2] tries to foster EHR utilization by executing functional flexible, individual workflows over a semantic service-based platform. As a first step within the design of such a system, we investigated on explicit requirements for IT-systems to fit healthcare professionals' workflows. To gather necessary functional and non-functional requirements we followed a triangulation approach that comprised a systematic literature review, observations in the clinical settings and expert interviews. The literature review resulted in a first classification of workflow building blocks [3] and the subsequent observations revealed process-related information about what tasks are executed considering EHRs for different situations [4]. This paper completes the above-mentioned requirement gathering tasks with expert interviews and reports on requirements pointed by physicians working with diabetes patients to design a system for supporting dynamic workflows for EHR activities.

Interviews with experts were planned in order to validate and enrich the already collected functional requirements for the design of a functional flexible, workflow-based and user-centered EHR utilization. We aimed also to gain insights regarding non-functional requirements (NFP) for the selection of services according to the user's needs. Based on these goals, we formulated the following research questions: (a) Do the results from the interviews correspond to the information we have identified in the literature review (information needs) and the observations (required functions, process-related requirements)?, (b) What additional information needs, required functions and process-related requirements are needed for valuable EHR utilization?, (c) What non-functional requirements do physicians consider mandatory when accessing proper services?

These research questions lead us to derive the following five topics, which we used as the core of our interview guideline:

- Examination process: What tasks are executed during routine consultation?
- Information needs: What Information is needed for different diabetes situations?
- Currently used clinical documentation system: What are the strengths and weaknesses of the person's currently utilized IT-system?
- Future changes: How to support clinical physicians' patient contacts with future challenges? What functions will be needed?
- Non-functional requirements: What non-functional requirements are necessary according to physicians for the service selection?

2. Methods

Physicians with different medical expertise relating to diabetes care were invited to participate in our study. The complete interview process comprised four steps: (1) interview preparation, (2) participant recruitment, (3) interview conduction and (4) qualitative analysis.

The mind-map depicted in Figure 1 contains all the topics and sub-topics derived from the research questions. This mind-map was used to guide the process of the qualitative part of the interview and its implementation was done by means of open questions.

[2] www.ontohealth.org

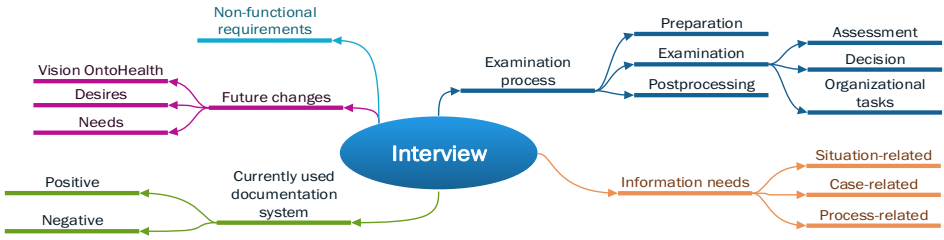


Figure 1. Mind-map showing the different topics for the interviews.

For each of the five topics within the guideline we formulated between two and eight questions. For example, for the topic "current used clinical documentation system" the following questions were asked: (1) What are important functions in the system? (2) What are the strengths of the system? (3) What weaknesses would you consider? (4) Are there any missing functions you would like to have? (5) Can you name any time-consuming tasks when using this software?

In addition, two quantitative questions with a five-point Likert scale were asked about how familiar the interviewed person was with IT in everyday usage and how familiar the person was with IT used in everyday work. Furthermore, information about the corresponding institution, gender, role and work experience after education end were surveyed.

Our interview study was divided in two stages. The first stage dealt with the evaluation of the design of the interview itself. For this pre-interview stage, we invited one physician with expertise in diabetes care. The goal of this initial task was to assure the correct design of the interview in terms of time, tasks and coverage of topics to answer the research questions of the study.

Theoretical sampling was used to select physicians, who got recruited through telephone or email. Telephone calls conveyed a very short introduction about the project and if the person indicated any interest, we switched to further email contact. Emails included an information document with a general description of the OntoHealth project and our intention regarding the need to gather physicians' requirements. Also, this document included the main topics of the interview in order to make the physicians familiar with the content of the conversation. The number of participants was determined by using the method of theoretical sampling according to the grounded theory approach [5]. All the interviews were held personally at the physicians' place, either in the hospital or office. The interviews were audio-recorded to facilitate further analysis.

After the talks, interview-recordings were manually transcribed to text documents. The content was transcribed literally without focusing on verbal expressions such as breaks or mood. Using the qualitative approach from Mayring et al. [6] we used MaxQDA[3], a software tool for qualitative analysis which facilitates to create and assign categories to the text content. Main attention was given to statements concerning information items, required functions, process-related requirements and non-function requirements according to EHR workflows. The initial selection and categorization was done deductively and in a very granular way. Once half of the content was annotated, all the topics were revised to inductively group the granular structure to a more generalized categorization before analyzing the remaining content.

[3] www.maxqda.de

3. Results

The evaluation of the initial test interview led to some modifications. For example, we had initially planned to use the resulting models from our literature review and observations as a starting point for the interviews. However, we found out during the test stage that the explanation of the general models was highly time consuming and that this approach could led to an influenced opinion on the interviewed person. Thus, it was decided to ask the questions independently of previous results and compare and combine the results afterwards. Furthermore, the test interview enabled us to verify the understandability of our questions as well as the correctness of the used vocabularies.

Once the interview design was validated, we sent the recruitment call to 15 physicians from Tyrol, Austria, of whom five male and two female physicians agreed to participate in the interviews. Participants were diabetes specialists, working in ambulatory metabolism wards in public hospitals (n=4) and physicians from private offices (n=3). The latter group comprised two diabetes specialists and one general physicians. All the selected physicians are engaged in the domain of diabetes mellitus. Mean work experience after education end of all participants was 19,7 years (ranging from 10 to 32 years). Four participants indicated that they are familiar[4] with IT in everyday usage and in everyday work.

All the interviews were conducted in mid-2015 by one researcher. The mean duration of an interview was 44 minutes. All the talks were held in German. Although transcripts were written in German, the code-system was directly created in English. Analyzed results lead to a total of 442 assigned codes categorized through four main categories: (1) process-related requirements (N[5]=146), (2) information needs during diabetes examination (N=157), (3) required functions (N=113) and (4) non-functional requirements (N=26). As shown in Figure 2, the discussed topics in the interview led to extract different results.

After a short introduction about the OntoHealth project and the interview process, we asked the physicians about typical workflows for diabetes consultations and the comprising tasks they were executing.

Three different situations for diabetes treatment were identified: (1) first consultation (NP[6]=6), (2) routine follow up consultation (NP=7) and (3) emergency (NP=3). The tasks followed in a diabetes routine consultation as well as in a follow up

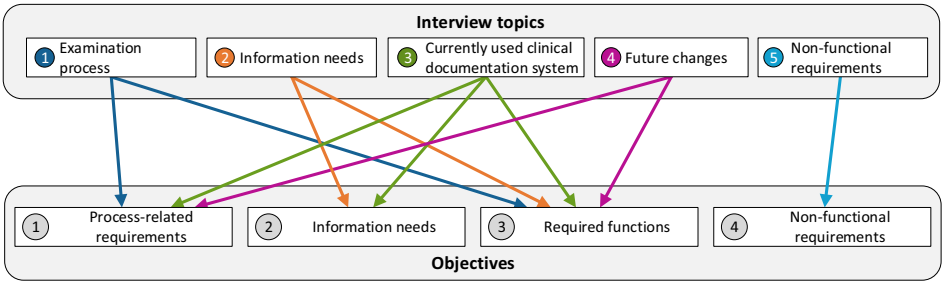


Figure 2. Arrows show relation between discussed topics and results topics. E.g. the interview topic "examination process" led to results for "process-related requirements" and "required functions".

[4] Equals a rating of 2 on the 5-point Likert scale

[5] Number of assigned Codes within that category – multiple assignments possible

[6] Number of interviewed persons giving that answer

consultation correspond from a process-related perspective. The main difference between these two workflows relates to the different information needs. One Physician claimed (Dr. B): *"If it's the first time for a patient at my office, I'm doing a more detailed informative conversation, create a comprehensive laboratory checkup and read and scan documents the patients brought."*

According to the analysis of the different interviews, a physician-patient consultation can be grouped in the following parts: (1) Patient evaluation: inquiry (assessment talk, measuring vital signs, laboratory checkup, physical examination) and documentation (patient history, documents, self-documentation); (2) Therapy decision: discussion/decision for therapy, prescription; (3) Organizational management: referral, appointment planning.

Nowadays, the information that can be requested is highly dependent on each specific IT-system and physicians often need to adhere to those rigid structures. Physicians claimed to miss a seamless patient documentation. Due to the lack of clinical information from other institutions, some tests are done several times (NP=3). On the other side, sometimes too much information is present in the patient chart: Two physicians stated that they are missing filtering functions for laboratory results, as they get displayed too much information not needed for diabetes treatment.

Integration of other reports often leads to interpretation limitations, as one physician (Dr. F) said: *"It's sometimes a problem to interpret the results of other institution's reports without further efforts. For example, if a patient brings a report, important values, like HbA1c, could have a percentage unit while my system uses mmol/mol. I first would need to convert the values in order to compare them with my results. Further it's often hard to find the interesting facts, like pathologic results, as all reports use a different style."*

One part of a consultation is the patient elucidation, i.e. to give information for the patient about his/her test results and explain the further therapy. For a diabetes routine checkup this information comprises: change of therapy with receipt and/or diabetes pump adjustments, present laboratory results, and organizational information, e.g. the next routine checkup or a planned referral to a specialist. All interviewed physicians told that the patient gets this information through printouts of the test results (e.g. laboratory) and/or a general report of the consultation.

In total the interviews led to the identification of 64 information items (including parent categorizations), of which 12 new items were added to the existing categorization of data elements. All the results were properly included in the existing categorization of information needs from previous literature analysis and observations. The main categories of identified information items were administrative data (demographic data, recent medical activities) and medical data (vital parameter, self-documentation, clinical documents, clinical problems and general patient information). A full list of all results can be obtained from the authors.

Documentation handling differs between hospitals and offices as well as among offices. While all the selected physicians document electronically, some physicians document at the time of the patient contact (NP=4), others take paper-based notes during the consultation and transfer to IT later (NP=3), which is usual in private offices. Asking more about this documentation behavior revealed complicated and non-intuitive use of the software, missing knowledge of necessary functions or non-adherence to the workflow. One (Dr. E) said: "*Even though I could document immediately in the system, it's too complicated and I like to do the assessment in the conventional way, using a preprinted sheet of paper. My secretary is then adding this information in the electronic*

record system. […] During the patient contact I take my general notes on a sheet of paper as well. At the end of the day I add all those notes in the IT-system for each patient".

The required data needs to be processed through functions. The interviews revealed 15 main functions and 14 sub functions, physicians need to use when accessing diabetes related information. All functions were successfully aligned with the previous gathering results about functions. The most stated functions in the interviews comprised: (1) Filter functions: Extract information of interest in an individual, case-specific manner (N=11, NP=5); (2) Use of templates: predefined set of all necessary items/functions for a situation/patient (N=10, NP=5); (3) Flexible workflow management (N=9, NP=6); (4) Data integration: avoid multiple documentation (N=7, NP=3); (5) Shared workflow exchange: physicians/organizations know planned and executed tasks according to a complete disease workflow. (N=5, NP=4); (6) Asynchronous documentation: document time-independently (N=2, NP=2); (7) User groups for workflow rights: e.g. hospitals don't want to give the right for allowing the physicians to fully design their own workflows. (N=2, NP=2); (8) Offline access to information (N=1, NP=1).

We asked the physicians to imagine our planned OntoHealth system, where components could be flexible selected according to the users' needs with restriction to quality parameters. We asked them, if they would embrace the feature of selecting a proper component, needed for the fulfilling of a defined goal, by specifying its functional and quality needs. Whenever needed, the interviewer gave an example of such quality parameters like "costs" or "performance": So if the physician wants to retrieve prescribed medication from the EHR, there could be several services available for fulfilling that goal. The main difference between those services would relate to quality parameters, e.g. one service is really fast in executing, but is tied to higher costs, the other one is slower but cheaper.

Answers of physicians were mixed. Some physicians (NP=3, all related hospital doctors) claimed that they would not admit to choose quality parameters for their needed components, e.g. Dr. B: *"I just want to use the system. I don't have time to define any quality parameters for the proper use of the system. Of course in common it makes sense to allow the flexibility in the system, but selection of quality parameters should be part of a higher entity, e.g. the management"*. For physicians working in a hospital, selecting functions and/or related quality parameters should not be part of their responsibility. However, all of the physicians owning their own practices/offices stated, that creating own workflows would be valuable and the proper selection of suitable services through non-functional parameters would lead to an efficient way of getting the needed functionality for accessing the right data.

Mentioned non-functional parameters from the interviewed physicians were: Usability (NP=5), Efficiency (NP=4), Security/privacy (NP=4), Performance (NP=3) and Availability (NP=3). Note that some of the requirements such as usability relates to a criterion related to the interface design rather than the service requirements.

4. Discussion

The interviews about the process workflow were limited to direct physician patient contacts (e.g. nursing consultation or administrative workflows were not discussed). The reason for this is that EHRs are used most frequently in this situation [7] and the results could be aligned with our previous results from the observations and literature analysis. The current number of participants reduced generalizability, however, the interviews

delivered enough details on the needed requirements to generalize and validate the complete results from the triangulation approach. The fact that only one researcher performed the qualitative content analysis could entail a limitation of the study.

The results from this study provide evidence for the demand and need of experts (in this case in diabetes) for more flexible case-specific IT-Systems. Future systems should provide individual and case-specific selection of required information as well as functions for processing this information, e.g. at the beginning of a diabetes routine consultation a physician usually needs a compact but complete patient overview comprising, among others, a list of currently prescribed medication with information about the related drug, regime, prescribing physician and reason for the prescription. Several steps for processing the data could lead to the desired goal: e.g. retrieve patient's documents, find medications, present result. Containers and templates are needed, because same sub-processes could be re-used, e.g. "medication overview" as part of the general patient assessment and when the physician is editing medication later in the workflow.

The interviews also revealed the differences in workflows between physicians: Whenever needed, the physician should have the choice for defining the precise information content (precise lab parameters, etc.) within the task definition of a certain workflow. A structure of granular building blocks representing the processing of information items is needed to compose user-centered workflows. The main physician's need during a consultation is to get the exact required information without unwanted details and the proper presentation of this information on the screen [8].

According to the interviews, physicians found that it is a good idea to move from a static data-driven approach to workflow-based EHR processing. However, some physicians, in particular those working in hospitals, were not satisfied to define the workflows themselves, as they find, IT-systems should act the same way for all physicians within the same hospital or at least within the same ward. On the other hand, physicians from private offices would embrace the feature of choosing functions themselves, as they do not need to adhere to common structures.

Our interviews did not reveal significant additional information needs to our prior results [3,4] but revealed more information about needed functions and workflow-related requirements which had not been identified through literature analysis and observations. Process-related information about how to access clinical data (information needs) for different situations could be obtained, for which two perspectives can be differentiated: (1) Disease-specific information needs, i.e. information is defined according to the disease known factors of having an influence on diagnosis and therapy. In our case, the information depends on diabetes related factors. And (2) Case-specific information needs, i.e. all patients are different. There are secondary disorders, allergies and other factors that have influence on the diagnosis and therapy. According to these two cases, a workflow-based EHR system should cover disease- and case-specific structures. Stated functions in the interviews relate to results from literature [1,9], enhancing the granularity structure of needed functions.

Non-functional requirements for EHRs can be found in the literature [9], as well as information about the generic selection of generic Web Services [10]. Our work augmented the state of the art with combined NFPs for the selection of services for the particular case of improving user-related EHR workflows. Results show, that hospital physicians do not want to deal much with individual, flexible service selection through defining proper NFPs, as higher authorities should take those decisions (e.g. management). Physicians of private offices show higher interest, as they prefer high

flexibility to foster their EHR usage. The reason for this could be different responsibility: While physicians in hospitals merely act as the consumer of health IT, physicians practicing in their own offices are responsible for the complete healthcare system they utilize. According to the analysis of the persons' experience with used IT, we could not find any relation between gender, role, years in business and IT-experience. IT-experience in work tend to be slightly worse than compared to everyday IT-experience (all results had same rating or worse in contrast to the related everyday experience).

Current IT-systems in healthcare do not exploit the potential of structure data. This leads to some drawbacks such as information overload when interfacing the EHR. There is an urgent need of managing healthcare data with syntactical, semantical and process interoperability. Static software is too inflexible to manage different users' needs and the different and often changing medical situations. This leads to unstructured data, information gaps and unsupported workflows. Service oriented architectures such as the one designed in the OntoHealth project [3] can offer a good solution to enable access to EHR data fulfilling individual user goals.

Acknowledgments

This work has been supported by the Austrian Science Foundation (FWF), project number P 25895-B24.

References

[1] A.H. Krist, J.W. Beasley, J.C. Crosson, D.C. Kibbe, M.S. Klinkman, C.U. Lehmann, et al., Electronic health record functionality needed to better support primary care., *J. Am. Med. Inform. Assoc.* **21**(5) (2014), 1–8.

[2] T.T. Van Vleck, D.M. Stein, P.D. Stetson, S.B. Johnson, Assessing data relevance for automated generation of a clinical summary., *AMIA Annu. Symp. Proc.* **2007** (2007) , 761–5.

[3] M. Schweitzer, N. Lasierra, S. Oberbichler, I. Toma, A. Fensel, A. Hoerbst, Structuring Clinical Workflows for Diabetes Care-An Overview of the OntoHealth Approach., *Appl Clin Inf.* **2**(5) (2014) , 512–26.

[4] M. Schweitzer, N. Lasierra, A. Hoerbst, Observing health professionals' workflow patterns for diabetes care - First steps towards an ontology for EHR services., *Stud. Health Technol. Inform.* **210**(2015) , 25–29.

[5] J. Corbin, A. Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 4rd ed, SAGE Publications, Los Angeles, 2014.

[6] P. Mayring, Qualitative Content Analysis [28 paragraphs], *Forum Qual. Soc. Res.* **1**(2) (2000), http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204, last access: 26.02.2016.

[7] L. Chen, U. Guo, S.J. Peterson, Racing Against the Clock : Internal Medicine Residents ' Time Spent On Electronic Health Records, *J Grad Med Educ.* **8**(1) (2016), 39–44.

[8] S.H. Walsh, The clinician's perspective on electronic health records and how they can affect patient care., *BMJ* **328**(7449) (2004) , 1184–7.

[9] A. Hoerbst, E. Ammenwerth, Electronic health records: A systematic review on quality requirements, *Methods Inf. Med.* **49**(4) (2010) , 320–336.

[10] I. Toma, D. Foxvog, Non-functional properties in Web services, *WSMO Deliv.* (2006) , 1–57.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-132

AAL Service Performance Measurement Cube – Key Criteria for AAL New Service Development

Johannes KRIEGEL[a,1], Klemens AUINGER[a], Luise RECKWITZ[b], Stephanie SCHMITT-RÜTH[c], Susanne WEISSENBERGER[a] and Linda TUTTLE-WEIDINGER[a]
[a] *Fakultät Gesundheit und Soziales, Fachhochschule Oberösterreich, Linz/Austria*
[b] *Fakultät für Mathematik und Informatik, Georg-August Universität Göttingen/Germany*
[c] *Fraunhofer-Arbeitsgruppe für Supply Chain Services, Nürnberg/Germany*

Abstract. The living environments of senior citizens are gaining in complexity with regard to health, mobility, information, support and behaviour. The development of Ambient Assisted Living (AAL) services in order to reduce this complexity is becoming increasingly important. The question is: What relevant criteria support the development, measurement and evaluation of business models of hybrid AAL services which have to be considered in an appropriate Performance Measurement Set? Within the EU funded research project DALIA (Assistant for Daily Life Activities at Home) a Service Performance Measurement Criteria (SPMC) Set has been developed and described. With the help of literature review and expert interviews relevant performance criteria were identified and described in the context of Analytic Hierarchy Process (AHP). In conjunction with an AAL business models scanning, a set of performance measurement criteria could be created. Discussion: The development and application of a specific AAL SPMC Set offers the possibility in a targeted and conceptual way advance the development of marketable AAL services. Here it will be important to integrate with software support the SPMC Set in the service development process of future marketable AAL applications. With the application of an adjusted AAL Service Performance Measurement Cube, the conceptual development of marketable AAL services can be maintained and relevant decisions can be supported.

Keywords: ambient assisted living, new service development, performance measurement, criteria set, business model

1. Transparency, Assistance, Automation and Empowerment through AAL Applications

1.1. Service Quality as an main Objective in Health Care and Social Care

The social and health care services in developed industrial nations are currently dominated by historical structures and processes as well as by changing objectives and earnings estimates. They are fast approaching resource limits that force the actors involved to make decisions regarding utilization, distribution and financing [1]. Further information asymmetries and information intransparencies as well as the increasing need for qualified health professionals to growing problems in the health care developed

[1] Corresponding Author: : Johannes Kriegel, University of Applied Sciences Upper Austria, Garnisonstraße 21, 4020 Linz/Austria, E-Mail: johannes.kriegel@fh-linz.at

industrialized countries [2]. The aim therefore must be to adapt the entire continuum of care and support for older people in the physical, psychological and social dimensions of current and future requirements. One approach to greater transparency, assistance, automation and empowerment of supportive services for the elderly is the development and the use of AAL applications. In recent years the AAL area has been subject to constant change and constant development [3]. Considering the recent research projects and solutions in the development of AAL services, it is clarified that only a limited proportion has reached the stage of marketable realization and the commercial use or inclusion in the performance catalog of service providers [4]. The analysis illustrates the variety of causes for problems in the development of hybrid AAL services. Here, six different dimensions can be identified (environment, technology, customer, organization, methods, measurement) for consideration as part of a conceptual AAL service development [5].

Based on the identified challenges in the AAL service development, it is necessary to strive for improved achievement and outcome measurement by means of appropriate structures and processes as part of new service development [6]. In that regard, in the last years the application of a business model across all sectors [7] (or in the service sector a service model) has been developed [8]. A service business model is a business model that serves as a model description or blueprint of a hybrid service [9]. In addition to the resources needed and expected value creation, the necessary structures were described, analysed and strategically aligned. Here, the various external factors and requirements are described in addition to intense business-related perspective within the framework of a service business model. [10]. The aim is to enable an optimal fit between offered supply and realizable demand for quality, service, flexibility and cost [9]. One approach to visualize the different relevant structural dimensions is the service business model according to Osterwalder/Pigneur [11]. The different building blocks and their further consideration are product- or service-related information which supports a targeted development solution- and customer-oriented AAL service. As part of the social and health services for the specific AAL sector it makes sense to expand the original Osterwalder/Pigneur model with the dimensions of privacy, ethics and emotions [5].

The established methods and tools used in product development (e.g., blueprinting, stages of model development, push strategy) lead to insufficient results due to the special requirements of services [12]. Previous experiences in the field of AAL service development have shown that the successful introduction and establishment of new and innovative AAL services require a systematic development and design of services using appropriate methods and procedures (e.g., SWOT analysis, expert and user survey). Here a targeted and conceptual model enables a timely and economic approach to the development and establishment of marketable AAL service [13]. Particularly, the customer and system-related requirements have to be taken into account, which comprise the formative dimensions problem / idea, ethics / culture, organization / service, technology, economics, law, emotion and prototype [14].

1.2. Performance Measurement as a Challenge

It becomes clear that the current challenges in the social and health services will let a window of opportunity open up services for AAL solutions and AAL services. [15]. However, these require a request and demand design of conceivable AAL services [16]. The recent developments and research activities as well as the (non) establishment of related results suggest that in the past the focus was on technology-driven product

development (push perspective) [17]. In the future, we must also include more the customer or user perspective (pull perspective) in developments [18]. Furthermore, it is necessary to edit the respective requirements, interim results and objectives in detail by means of appropriate criteria and to present in a transparent and timely manner through meaningful indicators [17]. This leads to the following question: How should a suitable set of relevant criteria and indicators to support the development, measurement and evaluation of business models of hybrid AAL services be designed?

2. Methods

2.1. Analytic Hierarchy Process for Developing a SPMC Set

For the systematic and structured identification, prioritization and analysis of a suitable set of key criteria to support the development, measurement and assessment of business models of hybrid e-health services, the Analytic Hierarchy Process (AHP) provides a solution [19]. First, the overarching objective has been identified and defined. The derived operative objective, a Service Performance Measurement Criteria set for the AAL service development was formulated in the second step. The third stage of AHP based approach includes on the one hand the collection of possible influencing factors by literature search in relevant national and international databases (e.g., ScienceDirect College Edtition, SpringerLink, PubMed, Emerald collections, Thieme Connect) using targeted buzz words (e.g., business model, success factors, AAL , e-Health, customer satisfaction, etc.). In a first step, 148 criteria were identified. From these were 45 selected regarding the number of entries and on the basis of expert assessments. Secondly, the collected factors were analysed with respect to their significance and effect relationships by twelve experts and players from Austria (including service providers, users, developers, scientists) in the field of e-Health and AAL. The experts were chosen for their qualifications, work experience, focus on innovation and institutional affiliation. Furthermore, the meanings and descriptions of the 45 different criteria collected were connected with the research results. In the next two steps, the synthesis and evaluation of priorities and a review of the consistency of the evaluations were conducted in experts` discussions. This resulted in assigning 36 criteria to the 12 dimensions of the business model grid [20]. The seventh step involves the visualization of dependencies using an AAL Service Development Loom [20] and the outline of a Service Performance Measurement Cube for linking structure (business model grid) [21], process (service development process) [22] and results (Service Performance Measurement Criteria Set) [5]. The eighth step involves the identification and formulation of appropriate alternatives and operational measures.

2.2. Service Performance Measurement Criteria Set Development

The specific development of a Service Performance Measurement Criteria Set (SPMC-set) under the AHP was carried out in three iterative steps. In a first stage, the current situation (e.g. challenges) of AAL service development was identified and analysed. Based on these results, the identification and analysis of the target situation (e.g. customer requirements) were carried out. The third stage of development of a SPMC set includes the validation of the process and outcome measurement of the SPMC set based on exemplary AAL development projects, which are ongoing.

3. Results

3.1. Description of relevant Business Model Grid Dimensions

The complexity and lack of transparent services markets and emerging AAL services market requires a structured and conceptual approach in the development of hybrid AAL services [23]. Towards this end a business model grid provides a solution. A business model grid (also Business Model Canvas) is a structured template for the development of new or the documentation of existing business models. Through visiualization both the existing solutions as well as the development of future or alternative combinations and business models are possible. With regard to the peculiarities of the AAL market, the extension of the classic business model canvas by Osterwalder/Pigneur model (2011) to include the dimensions of privacy, ethics and emotions is presented [20]. The practical application of the grid has shown that the creative development process and the visualization of the results can be supported by further questions. Table 1 gives an overview of the twelve different dimensions of an AAL business model and the dimension-related issues.

In the context of a specific AAL service development, these questions should be discussed and answered decidedly. For example, in the business model dimension customer segments the question: Which customers (segments) are addressed by the proposed service? In relation to a possible AAL service arise a multitude of different customer groups (e.g. residents, patients, relatives, people with disabilities in the domestic environment, housing association, cardiac rehabilitation patients, health insurance, mobile care, rehabilitation facilities, assisted living, pharmaceutical companies, Workplace health promotion and so on.) that need to be investigated and analysed hereinafter in objectives, willingness to pay, utility. First from the collection of different information and linking the various dimensions of business model grid arise alternative applications / use and business cases and concrete operational AAL scenarios.

Table 1: Dimension of an AAL business model

Dimension		Deskription	Dimension		Deskription
1	Customer Segments	Customers have different needs and expectations with respect to the product/service and the provision of services . The segmentation in very homogeneous groups of customers enables targeted customer processing. The question is: Which customers (segments) are expected to be addressed by the service?	7	Value Propostion	The value proposition of a product is the core of the provider-customer relationship and aims usually at the solution or satisfaction of a customer´s problem or need. Value proposition promises an explicit benefit for the customer. The question is: Why should the customer make use of the service?
2	Revenue Sources	The sources of income are the inflow of income which the service provider/supplier receives and ensure success in the long term. The questions is: What types of services are paid to what extent and to what amount by the customer or cost unit?	8	Cost Structure	The cost structure includes and describes the total costs resulting from the offer as well as the provision of a service , arise. The question is: What are the relevant costs associated with the business model?
3	Comm-unication and Sale Channels	In connection with the provision of services at the point of service, there is an exchanges of information, data, products and resources between providers and customers through appropriate channels. The questions is: Via which channels do the exchange and communication of hardware/software, money, information, etc. takes place?	9	Key Ressources	Key resources include the most important capital goods which are required for the offer and provision of a service. The question is: Which resources (labor resources, labor, information, disposition) are needed for the provision of the service?
4	Customer Relationships	Customer relationships enable the reciprocal interaction between customers and suppliers. The question is: What kinds of customer relationships and customer contact are expected or offered ?	10	Key Activities	Key activities include the most important processes and sub-activities that are required as part of providing the service. The question is: Which processes and (sub-) activities are required or provided by the provision of the service?
5	Ethic	Ethics comprise society orientated and informal values (moral) of human activity and their evaluation. The question is: What are the ethical criteria which need be taken into account within the business model?	11	Privacy	Privacy includes securing the privacy of the customer (individuals or groups). In particular, in an increasingly automated and digitized world, personal data and information should be protected against the unauthorized access by third parties. The question is: Which data protection frame is connected to the business model and necessary players?
6	Emotions	Emotions of the customer are individual and subjective, which are based inter alia on different psychological experiences, priorities, social behavior and reactions. They influence buying behavior significant. The question is: Which desires and emotions services should provide?	12	Key Partner	Key partners are actors and networks that are required for the development and provision of the service. The question is: What external partners can be/are required or used concerning the provision of services?

3.2. Collection of relevant Performance Measurement Criteria

In addition to monitoring, developing and analysing the various dimensions of possible hybrid AAL services, it is necessary for their description, measuring and evaluation to identify, collect and analyse appropriate and meaningful criteria. As part of the conducted literature reviews, expert interviews (n = 12) and two creative workshops, 45 different performance measurement criteria were identified and described (see Table 2 and 3). In addition, naturally there are other criteria as well as monetary and non-monetary and qualitative and quantitative indicators available. Depending on their purpose they are conceivable and make sense. However, a final collection and presentation of criteria relevant for decision making purposes is beyond the scope of this work.

4. Discussion

4.1. Development of a Service Performance Measurement Cube

The added value of innovation and the benefits of a new service for the stakeholders concerned are significantly influenced by the associated which objectives, structures, processes and results. It is the combination of these different dimensions of quality of services may result in establishment on the market and sustainable service success. It is therefore necessary to describe transparently, to check and analyse the intermediate results during the complex and collaborative AAL new service development [6, 69, 70]. Here, the responsible players have to rely on an appropriate criteria and performance measurement system. A multi-dimensional and qualitative performance measurement system (PMS), serves as a control and documentation tool, which comprises a selected number of interrelated decision criteria and key figures. The aim is, through the use of appropriate instruments (e.g., EDV, cockpit, dashboard, reporting) to provide target-oriented, condensed, up-to-date, decision-relevant information. A winning pin end Performance Measurement by Service Performance Measurement Criteria Set is to see solutions in conjunction with the respective strategic objectives in the context of a targeted and conceptual product and service development of AAL and align them. It is

Table 2: Overview of relevant service performance measurement criteria of an AAL business model

Key Performance Criteria	*Description*
Level of Attention	Extent of perception by the customer regarding a product and/or vendor. Attraction and concentration count as indicators of the intensity and duration of attention. [24]
Break-Even-Analysis	Analysis of the breakeven point (break-even point), in which proceeds and costs of a product are equal. The profit margin is identical to the fixed costs. [25]
Protection of Privacy	Personal and health-related consumer data enable targeted and customized advertising and marketing opportunities. Therefore, these are always at the forefront by vendors and developers. It is important to ensure appropriate security as well on the supplier side. [26]
Privacy Policy	In conjunction with the digitization and the associated products and business models, the question arises about the privacy of customers in the processing of personal data . For this purpose, appropriate frameworks and guidelines must be observed(z.B. 95/46/EG). [27]
Data Security	A certain degree of information security refers to the characteristics products which store and process information. It is essential to ensure confidentiality, availability, integrity and protection against hazards or threats, prevention of economic damages, and minimizing risks. [28]
Distribution Network	Network and channels or processes by which the distribution and the sale of products between producers, suppliers and distributors through to the (end) customer take place. [29]

Table 3: Overview of relevant service performance measurement criteria of an AAL business model

Key Performance Criteria	*Describtion*
Personal Initiative	Personal initiative includes one's own initiative and the first steps to a responsible decision and action. [30]
Research and Developement (R&D) Cost Component	The use of resources of a company/consortium influencing the innovation success. Through external R&D-activities and a targeted management of research and development costs, better innovation performance is possible. [31]
R&D-Networking	For dealing with time-, content- and market-related challenges in the context of product development, it is important to develop appropriate solutions like simultaneous technological and market development, dynamic development of hybrid service bundles by means of networking by independent businesses. [32]
Horizontal and Vertical Integration	By grouping homogeneous players and functions (as well as upstream and downstream production levels) under a uniform management enables economies of scale and scope as well as optimizing the supply chain and networks with regard to common innovation successes. [33]
ICT-Infrastructure	Use of technology and structure in the field of information and communication. Factors such as distribution, interfaces, capacity or speed play an important role. [34]
Degree of Innovation	The level of innovation is a function that describe the cost-benefit ratio of the new product and previous solutions (state of the art). [35]
Interdisciplinarity	The use of different perspectives, approaches and methods from various disciplines with the aim of enabling future and successful innovations. [36]
Share of Investment / Return on Investment (ROI)	Profitability, the ratio of a profitability measure to the capital investment of an accounting period, measures the efficiency of product development and enables a comparison in competition. [37]
Purchasing Power	Monetary amount remaining per business entity after all fixed payment obligations were made, and which is available for potential buying decisions. [38]
Communication	The development of innovations dependents on the transfer and the exchange of information, knowledge, awareness or experience . It is important to develop and establish appropriate channels, methods and skills. [39]
Cooperation	Today the development of innovation requires targeted and purposeful interaction of two or more actors to achieve a common goal in a coordinated division of labor and resource allocation. [40]
Coordination	Consortium formation and systematic classification with regard to innovation and product development requires the control of players and projects and necessary strategically oriented activities. [41]
Debtors and Creditor Analysis	Overview and analysis of creditors, which means the suppliers and the debtors, the customer / company / products. [42]
Customer-/ User Integration	Customer involvement in entrepreneurial innovation projects aims at a reinforced market-based-view and is seen as an important factor for successful service development. The customer perspective is involved as a co-designer. [43]
Customer Acquisition	Status, planning and implementing targeted actions of customer acquisition. [44]
Customer Acceptance	Customer acceptance is the positive acceptance decision by the customer and includes, inter alia, aspects of the benefit assessment, the usability, the expectation of conformity, the costs and the context- and network effects. [45]
Customer Needs	The subjectively perceived lack of a business entity with the desire to eliminate these by demand and consumption. [46]
Customer Loyalty	Status, planning and implementing targeted actions of customer loyalty. [47]
Customer Authorisation	Customer or Consumer Empowerment aims to strengthen the customer or ownership and participation by the customer (regarding decisions, codetermination, participation in the development- and value creation process). [48]
Customer Profitability	Customer profitability, based on selected customer segments and groups, sets the acquisition costs in relation to revenues. [49]
Customer Satisfaction	The difference between customer expectations and satisfaction of needs is evidence of improvement and product innovation. [50]
Mass Suitability	Property of a product that makes this attractive and desirable for a larger number of customers / clients. [51]
Employee Qualification	The mix of qualifications of the employees determines the team- and personal working capacity , which is composed of professional and social skills, and should be controlled through strategic human resource development. [52]
Mobility	A sustainable mobility offers customers and stakeholders the movement in physical , geographic, social or virtual spaces. [53]
Modularity	The modular design of complete systems, forms standardized individual components with defined interfaces. The modular product design allows a higher system complexity and promotes product innovation. [54]
New Media	New media describe current and time-related new media technologies (e.g. Internet, blog, social networks, word of mouth) for information transmission and communication current or potential customers / clients. [55]
User Rating	The benefit assessment is a subjective assessment by the customer, whereby the benefits of a product are assessed in relation to the resources that has to spend the customer. [56]
Terms of Use	In conjunction with the general terms and conditions (e.g. directive 93/13 / EEC) it is necessary to protect customers from unfair contract terms. [57]
Patents	The development of products and thus of proprietary claims offer the opportunity to differentiate themselves from competitors and represents an increase in the value of the innovation or the product . Therefore, the number of patent applications is growing at a fast pace. Companies unable to process the patent information, are losing competitiveness. [58]
Personnel Costs	Personnel expense includes the costs incurred through the use of workers in the development and production process. [59]
Privacy	Privacy is that non-public area and need where, a human being undisturbed by others, lives out his human right to free development of personality. [60]
Product Quality	The product quality is composed of the dimensions of structure, process and outcome. This can be considered neutral, which means the sum of all characteristics of a product, as well as subjectively, which means the goodness of all the properties of a product. [61]
Self-determination	Self-determination includes the degree, the ability and the cognitive ability of customers to act in accordance with their own will, and to decide on their actions, behavior, and their body freely. [62]
Security Potential	Assistance systems and activity monitoring for the elderly are gaining importance and create a sense of security. [63]
Simulation	Simulation includes the analysis of complex systems. By means of an abstraction of the simulated system, different scenarios of product development and –variation are passed through, concerning structure, function and behavior by using specific parameters. [64]
Strategic Orientation	Medium- to long term planned objectives and approach of a company in terms of innovation, market and use of resources also includes the control of implementation incl. performance measurement. [65]
Time-to-Market	The lead time includes the time from product development to product placement on the market. In this period costs for the product arise, but it still generates no revenue. [66]
Use Case Building	The development and description of applications is the collection of visions and expectations for the technology and the user context. [67]
Willingness to Pay	Monetary amount paid for a product which a buyers or customers is willing to pay out of their income. [68]

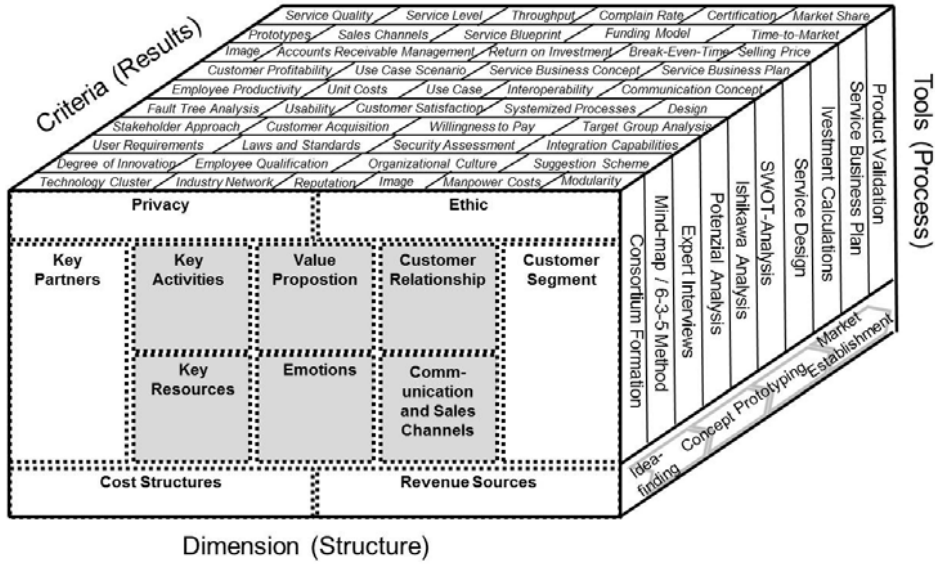


Figure 1: Service performance measurement cube for AAL service development

advisable to focus on selected key performance criteria. This results in a focused concentration on essential compliance levels and progress as well as on important goals and critical success factors in the context of AAL service development.

4.2. Evaluating the Spectrum and Validity of the SPMC Set

For the stakeholders, the development and application of an individual AAL Service Performance Measurement Cube creates the chance to drive forward the development of marketable AAL services in a targeted and conceptual way and to support relevant decisions. Therefore the service performance measurement cube can be used as a management tool in the AAL new service development process. For the future development of the SPMC set and their application in the context of the development of new services and service management of existing AAL applications, different challenges and research needs have to be identified. On the one hand it is necessary to evaluate the validity and reliability of the SPMC with further research and development projects [71]. Secondly, it is necessary to integrate the large number of different criteria and indicators by software support (for example, by means of e-Performance Measurement Cubes) practical and application-oriented in the new service development process of future marketable AAL applications [72].

References

[1] K. Böhm, A. Schmid, R. Götze, C. Landwehr, H. Rothgang, Five types of OECD healthcare systems: Empirical results of a deductive classification, Health Policy 113 (2013), 258-269
[2] L. Hajen, H. Paetow, H. Schumacher H, Gesundheitsökonomie, Kohlhammer, Stuttgart, 2010
[3] M. Gersch, J. Liesenfeld J, AAL- und E-Health-Geschäftsmodelle, Springer, Wiesbaden, 2012

[4] R. Luzsa, S. Schmitt-Rüth, F. Danzinger, Alterssensible Marktforschung - Ältere Kunden verstehen - Methoden, Ansätze & Lösungen. Fraunhofer, Stuttgart, 2014
[5] J. Kriegel, S. Schmitt-Rüth, K. Auinger, Performance Measurement für die Entwicklung von hybriden eHealth/AAL Dienstleistungen, in: e-Health 2015 – Informationstechnologien und Telematik im Gesundheitswesen, VDMJ, Solingen, 2014. pp. 223-229
[6] K. Spitalewsky, J. Rochon, M. Ganzinger, P. Knaup, Potential and Requirements of IT for Ambient Assisted Living Technologies, Methods Inf Med 52 (2013), 231-238
[7] C. Zott, R. Amit, L. Massa, The Business Model: Recent Developments and Future Research, Journal of Management 37 (2011), 1019-1042
[8] J.C. Aurich, C. Mannweiler, E. Schweitzer, How to design and offer services successfully, CIRP Journal of Manufacturing Science and Technology 2 (2010), 136-143
[9] S.M. Shafer, H.J. Smith, J.C. Linder, The power of business models, Business Horizons 48 (2005), 199-207
[10] B. Edvardsson, B. Enquist, R. Johnston, Cocreating Customer Value Through Hyperreality in the Prepurchase Service Experience, Journal of Service Research 8 (2008), 149-161
[11] Osterwalder, Y. Pigneur, Business Model Generation, Campus, Frankfurt/M., 2011
[12] H.U. Buhl, B. Heinrich, M. Henneberger, A. Krammer, Service Science, Wirtschaftsinformatik 50 (2008), 60-65
[13] P. Rashidi, A. Mihailidis, A survey on ambient-assisted living tools for older adults, IEEE journal of biomedical and health informatics 17 (2013), 579-590
[14] J. Kriegel, S. Schmitt-Rüth, B. Güntert, P. Mallory, New Service Development in German and Austrian Health Care – Bringing e-Health Services into the market, International Journal of Healthcare Management 6 (2013), 77-86
[15] R.C. Leventhal, Aging consumers and their effects on the marketplace, Journal of Consumer Marketing 4 (1997), 276-281
[16] B. Eberhardt et al., Better health and ambient assisted living (AAL) from a global, regional and local economic perspective, Int. J. Behavioural and Healthcare Research 2 (2010), 172-191
[17] C. Jaschinski, S.B. Allouch, An Extended View on Benefits and Barriers of Ambient Assisted Living Solutions, International journal on advances in life sciences 7 (2015), 40-53
[18] Brem, K.I. Voigt, Integration of market pull and technology push in the corporate front end and innovation management - Insights from the German software industry. Technovation 5 (2009), 351-367
[19] T.L. Saaty, How to make decision: The Analytic Hierarchy Process, European Journal of Operational Research 48(1990), 9-26
[20] J. Kriegel, K. Auinger, AAL Service Development Loom – From the Idea to a Marketable Business Model, in: eHealth2015 - Health Informatics meets eHealth. Proceedings of eHealth2015, Stud Health Tech Inform, Vienna, 2015. pp. 125-133
[21] J. Kriegel, S. Schmitt-Rüth, T. Ortner, Entwicklung von eHealth- und AAL-Geschäftsmodellen am Beispiel der Forschungsprojekte PIN und TMC in Oberösterreich, in: e-Health 2014 – Informationstechnologien und Telematik im Gesundheitswesen, Solingen, VDMJ, 2013. pp. 314-321
[22] J. Kriegel, S. Schmitt-Rüth, Hybride Dienstleistungen im e-Health-Bereich, ÖKZ 51 (2010), 35-36
[23] J.M. Leimeister, C. Glauner, Hybride Produkte – Einordnung und Herausforderungen für die Wirtschaftsinformatik, Wirtschaftsinformatik 50 (2008), 248-25
[24] A.F. Payne, K. Storbacka, P. Frow, Managing the co-creation of value, Journal of the Academy of Marketing Science 36 (2008), 83-96
[25] P.J. Williamson, Cost Innovation: Preparing for a 'Value-for-Money' Revolution, Long Range Planning 43 (2010), 343-353
[26] J. Wirtz, M.O. Lwin, J.D. Williams, Causes and consequences of consumer online privacy concern, International Journal of Service Industry Management 18 (2007), 326-348
[27] S. Spiekermann, A. Acquisti, R. Böhme, K.L. Hui, The challenges of personal data markets and privacy, Electronic Markets 25 (2015), 161-167
[28] R. Piplode, U.K. Singh, An Overview and Study of Security Issues & Challenges in Cloud Computing, Journal of Advanced Research in Computer Science and Software Engineering 2 (2012), 115-120
[29] M. Restuccia, U. de Brentani, R. Legoux, J.F. Ouellet, Product Life-Cycle Management and Distributor Contribution to New Product Development. Journal of Product Innovation Management (2015), doi: 10.1111/jpim.12261
[30] S. Nambisan, Designing Virtual Customer Environments for New Product Development: Toward a Theory, Academy of Management Review 27 (2002), 392-413

[31] L. Berchicci, Towards an open R&D system: Internal R&D investment, external knowledge acquisition and innovative performance, Research Policy 42 (2013), 117-127
[32] C.C. Snow, D. Fjeldstad, C. Lettl, R.E. Miles, Organizing Continuous Product Development and Commercialization: The Collaborative Community of Firms Model, Journal of Product Innovation Management 28 (2011), 3-16
[33] E. Bendoly, A. Bharadwaj, S. Bharadwaj, Complementary Drivers of New Product Development Performance: Cross-Functional Coordination, Information System Capability, and Intelligence Quality, Production and Operations Management 21 (2012), 653-667
[34] J. Vilaseca-Requena, J. Torrent-Sellens, A.I. Jiménez-Zarco, ICT use in marketing as innovation success factor: Enhancing cooperation in new product development processes, European Journal of Innovation Management 10 (2007), 268-288
[35] R.J. Calantone, K. Chan, A.S. Cui, Decomposing Product Innovativeness and Its Effects on New Product Success, Journal of Product Innovation Management 23 (2006), 408-421
[36] E.M. Olson, O.C. Walker, R.W. Ruekerf, J.M. Bonnerd, Patterns of cooperation during new product development among marketing, operations and R&D: Implications for project performance, Journal of Product Innovation Management 18 (2001), 258-271
[37] H. Ernst, Success Factors of New Product Development: A Review of the Empirical Literature, International Journal of Management Reviews 4 (2002), 1-40
[38] N.M. Puccinelli, R.C. Goodstein, D. Grewal, R. Price, P. Raghubir, D. Stewart, Customer Experience Management in Retailing: Understanding the Buying Process, Journal of Retailing 85 (2009), 15-30
[39] R.T.A.J. Leenders, J.M.L. van Engelen, J. Kratzer, Virtuality, communication, and new product team creativity: a social network perspective, Journal of Engineering and Technology Management 20 (2003), 69-92
[40] Z. Emden, R.J. Calantone, C. Droge, Collaborating for New Product Development: Selecting the Partner with Maximum Potential to Create Value, Journal of Product Innovation Management 23 (2006), 330-341
[41] K.J. Petersen, R.B. Handfield, G.L. Ragatz, Supplier integration into new product development: coordinating product, process and supply chain design, Journal of Operations Management 23 (2005), 371-388
[42] S. Bougheas, S. Mateut, P. Mizen, Corporate trade credit and inventories: New evidence of a trade-off from accounts payable and receivable, Journal of Banking & Finance 33 (2009), 300-307
[43] P. Carbonell, A.I. Rodríguez-Escudero, D. Pujari, Customer involvement in new service development: an examination of antecedents and outcomes, Journal of Product Innovation Management 26 (2009), 536-550
[44] Rindfleisch, C. Moorman, The Acquisition and Utilization of Information in New Product Alliances: A Strength-of-Ties Perspective, Journal of Marketing 65 (2001), 1-18
[45] F. Langerak, E.J. Hultink, H.S.J. Robben, The impact of market orientation, product advantage, and launch proficiency on new product performance and organizational performance, Journal of Product Innovation Management 21 (2004), 79-94
[46] Herrmann, F. Huber, C. Braunstein, Market-driven product and service design: Bridging the gap between customer needs, quality management, and customer satisfaction, International Journal of Production Economics 66 (2000), 77-96
[47] W.D. Hoyer, R. Chandy, M. Dorotic, M. Krafft, S.S. Singh, Consumer Cocreation in New Product Development, Journal of Service Research 13 (2010), 283-296
[48] C. Fuchs, M. Schreier, Customer Empowerment in New Product Development, Journal of Product Innovation Management 28 (2011), 17-32
[49] R.G. Cooper, E.J. Kleinschmidt, Benchmarking the Firm's Critical Success Factors in New Product Development, Journal of Product Innovation Management 12 (1995), 374-391
[50] N.K. Lankton, E.V. Wilson, Factors Influencing Expectations of e-Health Services within a Direct-Effects Model of User Satisfaction, e-Service Journal 5 (2007), 85-111
[51] R. Duray, Mass customization origins: mass or custom manufacturing? International Journal of Operations & Production Management 22 (2002), 314-328
[52] H. Melton, M.D. Hartline, Employee Collaboration, Learning Orientation, and New Service Development Performance, Journal of Service Research 16 (2013), 67-81
[53] M. Alessandro, E. Gianluca, S. Giustina, P. Giuseppina, Sustainable mobility: An integrative framework and its application for new service design, International Journal of Technology Management & Sustainable Development 11 (2012); 31-49
[54] A.K.W. Lau, R.C.M. Yam, E. Tang, The Impact of Product Modularity on New Product Performance: Mediation by Product Innovativeness, Journal of Product Innovation Management 28 (2011), 270-284

[55] B. Lis, C. Neßler, Digitale Mundpropaganda, Wirtschaftsinformatik 56 (2014), 73-75
[56] S. Sandström, B. Edvardsson, P. Kristensson, P. Magnusson, Value in use through service experience, Managing Service Quality 18 (2008), 112-126
[57] D. Svantesson, R. Clarke, Privacy and consumer risks in cloud computing, Computer Law & Security Review 26 (2010), 391-397
[58] A.J.C. Trappey, C.V. Trappey, C.Y. Wu, C.W. Lin, A patent quality analysis for innovative technology and product development, Advanced Engineering Informatics 26 (2012), 26-34
[59] D.H. Henard, M.A. McFadyen, Resource Dedication and New Product Performance: A Resource-Based View, Journal of Product Innovation Management 29 (2012), 193-204
[60] R.H. Weber, Internet of things: Privacy issues revisited, Computer Law & Security Review 31 (2015), 618-627
[61] R.C. McNally, M.B. Akdeniz, R.J. Calantone, New Product Development Processes and New Product Profitability: Exploring the Mediating Role of Speed to Market and Product Quality, Journal of Product Innovation Management 28 (2011), 63-77
[62] Dholakia UM (2006) How Customer Self-Determination Influences Relational Marketing Outcomes: Evidence from Longitudinal Field Studies. Journal of Marketing Research 43(1):109-120
[63] M. Chan, E. Campo, D. Estève, J.Y. Fourniols, Smart homes - Current features and future perspectives, Maturitas 64 (2009), 90-97
[64] S.L. Chan, W.H. Ip, A dynamic decision support system to predict the value of customer for new product development, Decision Support Systems 52 (2011), 178-188
[65] C.K. Prahalad, V. Ramaswamy, The New Frontier of Experience Innovation, MIT Sloan management review 44 (2003), 12-18
[66] J. Chen, R.R. Reilly, G.S. Lynn, New Product Development Speed: Too Much of a Good Thing? Journal of Product Innovation Management 29 (2012), 288-303
[67] S. De Rouck, A. Jacobs, M. Leys, A methodology for shifting the focus of e-health support design onto user needs: A case in the homecare field, International Journal of Medical Informatics 77 (2008), 589-601
[68] K.M. Miller, R. Hofstetter, H. Krohmer, Z.J. Zhang, How Should Consumers' Willingness to Pay Be Measured? An Empirical Comparison of State-of-the-Art Approaches, Journal of Marketing Research 48 (2011), 172-184
[69] G. Van Den Broek, F. Cavallo, C. Wehrmann, AALIANCE Ambient Assisted Living Roadmap, Amsterdam, IOS, 2010
[70] S. Gronum, J. Steen, M.L. Verreynne, Business model design and innovation: Unlocking the performance benefits of innovation, Australian Journal of Management (2015), doi: 10.1177/0312896215587315
[71] H.S. Kim, Y.G. Kim, A CRM performance measurement framework: Its development process and application, Industrial Marketing Management 38 (2009), 477-489
[72] P. Cocca, M. Alberti, A framework to assess performance measurement systems in SMEs, International Journal of Productivity and Performance Management 59 (2010), 186-200

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-142

Requirements for Evidence-Based Templates in Electronic Case Report Forms

Marco SCHWEITZER [a,1], and Stefan OBERBICHLER[a]
[a] *eHealth Research and Innovation Unit,*
UMIT - University for Health Sciences, Medical Informatics and Technology

Abstract. Background: Electronic case report forms (eCRF) play a key role in medical studies and medical data registries (MDR) for clinical research. The creation of suitable eCRFs is a challenging and yet a very individual task. Objectives: We plan to create evidence-based templates for eCRFs which are aligned with existing studies or MDRs. Methods: In this paper we investigated existing standards for eCRFs, defined uses cases and derived requirements needed to identify and annotate data items and pertinent information within study protocols, literature or patient cases. Results: In order to establish evidence-based eCRF templates based on annotated text documents, a standard-based, hierarchical structure with linkage to existing data repositories needs to be modeled. Standards like ISO/IEC 11179 provide a necessary base, which needs to be extended with proper linking functions. Conclusion: Linking evidence-based sources with eCRFs allows creating templates, which could be used to define eCRFs for new clinical studies or even compare studies among them. As a next step, the derived requirements from this paper will be used to establish an ontology-based structure for annotating existing text-documents with eCRF data elements.

Keywords. electronic case report forms, standards common data elements, evidence-based design

1. Introduction

1.1. Background

Electronic Data Capture (EDC) is widely used in clinical or epidemiologic research. EDC systems support data integration, reduce cost and can decrease the time needed to develop or verify new medication or therapies [1]. Such data capture and documentation systems used in medical research focusing on collecting health specific patient related data for a specific purpose are referred to as Medical Data Registries (MDR). MDRs play a significant role in medical research [2]. Please note, in this paper we use the term MDR for Medical Data Registry, which should not be confused with metadata registries.

A key facilitator for MDRs is data capturing through electronic case report forms (eCRFs). Such forms are used to collect relevant information according to a certain study goal in a structured manner. The main goals of eCRFs are to assist and facilitate clinical data capturing and standardized data storage [3]. This enables re-use of captured data as well as sharing the data between research institutions with necessary syntactic and semantic interoperability.

[1] Corresponding Author: Marco Schweitzer, University for Health Sciences, Medical Informatics and Technology (UMIT), Eduard-Wallnoefer-Zentrum 1, 6060 Hall i.Tirol, E-Mail: marco.schweitzer@umit.at

Usually eCRFs are designed from scratch to answer newly raised research questions, although there may already exist forms for data capturing within the field of interest. Several initiatives have been started to address these issues [4,5]. Re-using defined sets of eCRFs and its related data elements can help to decrease complexity, costs and workload. Additionally, by using already existing eCRFs comparability between different studies or MDRs can be ensured.

1.2. Motivation

Designing new eCRFs considers the study conducting team to thoroughly choose the data elements, which are necessary to answer the intended research question. Therefor data should be collected unambiguously and in sufficient detail and at the same time avoid redundancy and unwanted details [6]. Additionally, existing designs of eCRF e.g. published in scientific literature may provide suitable information.

eDiFy[2] (evidence-based Design of Electronic Case Report Forms) is a project funded by the Tyrolean Science Fund (TWF) and intended to address the outlined problems. The main intention is to find ways to annotate and link eCRFs as well as their data elements with existing statements published in scientific literature, medical cases, data element repositories or any other suitable source. By using such annotated and linked data, templates for evidence-based eCRFs can be derived. Using those templates can increase comparability between clinical studies as well as decrease the needed effort to implement clinical trials or MDRs.

1.3. Objectives

eDiFy aims to facilitate the challenging task of recreating eCRFs for new studies by establishing a link between existing text-based documents such as scientific literature or patient related medical cases associated to the research question and suitable common data elements (CDEs), e.g. used in medical studies focusing on related topics or defined in CDE repositories. To address these issues the structure and existing work in the design of eCRFs and data elements has to be evaluated. Additional functional requirements should be identified for evidence-based eCRFs. The standards and requirements have to be matched to derive an abstracted design approach.

Therefore, the goals of this paper are to:

1. Analyze existing eCRF design approaches and identify data element types and eCRF standards.
2. Identify requirements to build templates for an evidence-based design of eCRFs.
3. Correlate and map the defined standards and requirements to gain insights on how such templates can be assembled.

2. Methods

The collection and validation of requirements is a main part of software development. Thus, the concept of Requirements Engineering (RE) defines the tasks of systematically gathering requirements, which comprise 1) definition, 2) documentation, 3) maintenance and 4) validation [7]. We elicited requirements through analysis of existing eCRF

[2] https://ehealth.umit.at/index.php/edify

standards (system requirements) and the definition of proper use cases (user requirements). The combination of those with respect to the proposed project goal revealed necessary functional requirements for the system.

A literature review was conducted in order to identify and analyze existing standards for the design and use of eCRFs and data elements. The literature review started with analyzing the paper and references from Richesson and Nadkarni [8], which gives an overview of existing eCRF standards. A subsequent search on google scholar with search terms “eCRF standards” and “data element standards” leads to additional papers. All research results were screened by abstract, aiming at suitable eCRF or data element standards. The full texts of remaining papers were analyzed through qualitative content analysis [9], resulting in a descriptive list of available standards as well as their characteristics, advantages and disadvantages of the most established results. Whenever needed, websites of the standard developing organizations were also included as an information source. According to the idea of the eDiFy project, proper use cases were modeled using the method of use case diagrams from the Unified Modelling Language (UML) [10], which depicts the interaction of the system and its related actors in a graphical and structured way.

3. Results

3.1. Standards and classification

After analyzing the initial paper and a general google scholar search, we found 13 papers. Those papers got screened by abstract and resulted in a 5 paper set, which got further thoroughly analyzed trough qualitative content analysis using MaxQDA[3]. The resulted categories were inductively created during text analysis and finally comprised the identified eCRF and data element standards with related aims, structures, future challenges, barriers, interrelations and examples of application.

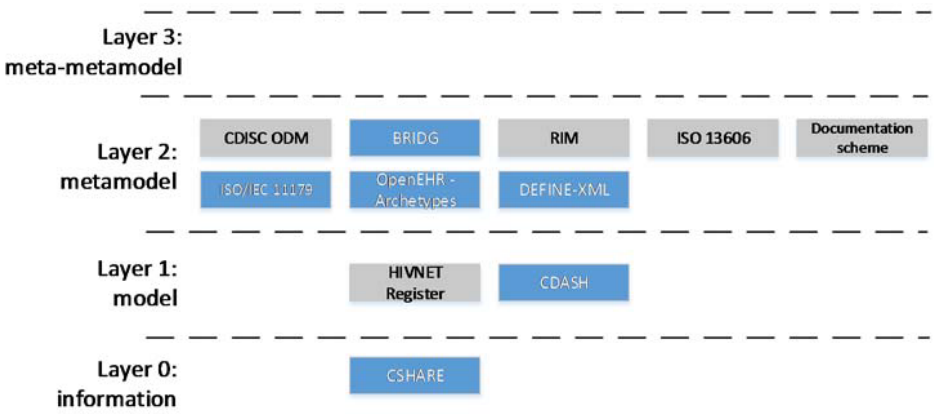


Figure 1. Existing standards for managing eCRFs and data elements grouped according to the levels of MOF. Figure is adopted from [12] (grey) without vocabularies and got extended by additional found standards (blue).

[3] http://www.maxqda.de/

Our investigations revealed that currently no universal CRF standard exists [8]. However, the review revealed 10 standards used in healthcare suitable for the creation and management of eCRFs and data elements which can be grouped into 4 levels according to the Meta Object Facility (MOF) specification (see figure 1). This categorization got extended with additional found standards (blue boxes). The next section gives a short overview of the most suitable standards for eDiFy and describes advantages and disadvantages of them. All the standards, which are part of the metamodel layer, are the most appropriate ones for eDiFy as they describe a common conceptualization of eCRF data.

3.2. ISO/IEC 11179 - Metadata Registries

The ISO/IEC 11179 standard defines a schema for describing metadata in MDRs. This metadata specifies data to describe, explain and locate an information resource (data about data) [11], which is necessary for eCRFs in the sense of having a shared conceptual data model describing the related clinical information. Those data elements can then be used as data types (e.g. "person's body weight in kg") for describing a certain clinical information item of a medical trial ("body weight of person ID123: 75kg").

The aim of the standard is to provide interoperability for data exchange through a common understanding of metadata. The core of the standard is the *data element*, which defines properties like designation, identification, representation, classification and a list of valid attributes. As described in the *Data Description Package* of the standard, a data element is compounded of a *data element concept* (conceptual level, e.g. "person's smoking status") and a *value domain* (representational level, e.g. "Every-day smoker|Some-day smoker|Former smoker"). On the other side, the *Concepts Package* defines the constitution of concepts through controlled vocabularies and its relations (e.g. "smoking status" is part of "tobacco addiction", which is then part of "Substance Abuse and Addiction").

ISO/IEC 11179 is used in several MDR projects, e.g. from Cancer Data Standards Registry and Repository (caDSR) [13] or the Australian Metadata Online Registry (METeOR) [14]. It has some limitations, e.g. relation between a data element and its explanation text is modeled one to one, and thus no multi-language explanation texts are possible. More detailed information about ISO/IEC 11179 information can be found in the standard's specification [15].

3.3. CDISC ODM

The Operational Data Model (ODM), developed from Clinical Data Interchange Standards Consortium (CDISC), is a XML-based standard for metadata models and data interchange models [16]. ODM facilitates the exchange of clinical protocol data and manages all data about a complete clinical study (events, forms, item definitions, facts about clinical subjects, audit log entries, administrative information, reference data, versioning support and interfaces for custom vendor extensions). The ODM format is used as an implementation for CDASH (Clinical Data Acquisition Standards Harmonization), which describes a collection of recommended data fields for 18 domains used in clinical trials (e.g. demographics, adverse drug events). The advantage of ODM is to have one standard that covers the documentation and management of all parts within a clinical study. However, ODM also has some disadvantages, e.g. in ODM semantically identical data elements may have different names across different systems and thus

cannot be mapped without further processing. An advanced format of ODM was published under DEFINE-XML (Case Report Tabulations Data Definitions in an XML format) [17], which is also used as the mandatory format for the Food and drug association (FDA) for reporting study data.

3.4. openEHR archetypes

The openEHR[4] foundation develops standards for establishing electronic health records (EHR). The main idea is to have a dual model approach that means to describe clinical data with an information model (generic) and a conceptual model of archetypes (domain-specific). By utilizing this method clinical knowledge (e.g. "blood pressure" which describes exemplary attributes for type: systolic/diastolic, unit: mm[HG], measured entity: arm/leg) is used to model the related information (e.g. "arm measured blood pressure of patient X is 120/80 mm[Hg] on 19th feb. 2015 at 18:15"). Archetypes can be used to model eCRFs and related data elements as described in [18]. In general, the modular approach of archetypes allows their use for the direct implementation of data elements within eCRFs, though due to the common relation to EHRs, openEHR archetypes are preferred to gather clinical information from EHRs for clinical trials [19].

3.5. Use Cases

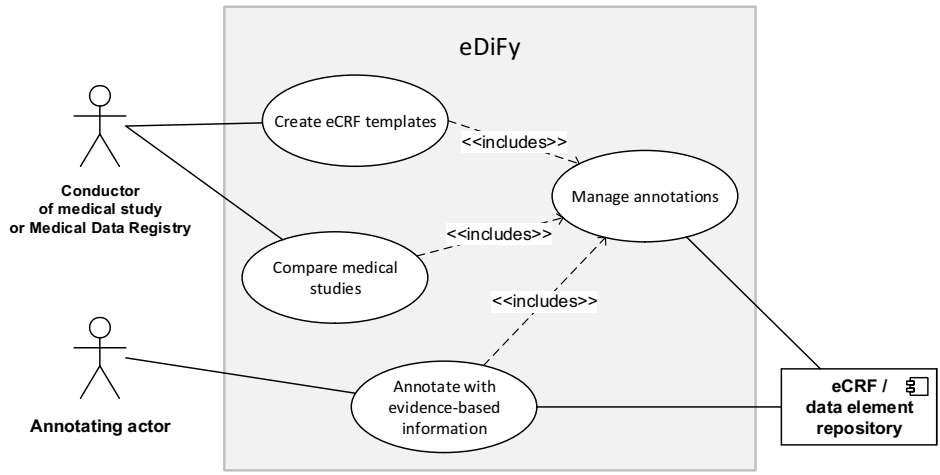


Figure 2. Use Case diagram for the proposed eDiFy setting.

Taking into account the proposed eDiFy system, suitable use cases can be defined as depicted in figure 2, which describe the interactions the system should support. A conductor of a medical study or MDR should be able to create a new eCRF built from evidence-based templates. The templates are composed of semantically related items, groups, sections and forms related to annotations of protocols, literature documents, patient related medical cases and others. An annotating person should be able to annotate those local managed text documents with external, shared CDEs from approved sources.

[4] http://www.openehr.org/

If necessary new data elements can be created which are stored in suitable repositories. The fully established system can be further used to compare medical studies or literature according to its referenced data elements.

3.6. Requirements

Using the results of the literature analysis and the use cases, proper requirements could be derived for a system that is able to create and manage evidence-based eCRF templates based on semantic annotations. The requirements encompass the following:

R1: Linking knowledge representation with existing repositories: All elements of suitable eCRF creation/utilization should be linked to CDEs from shared, accessible repositories.

R2: Standards conformity: Adhering to existing standards is necessary to provide syntactically and semantically interoperability needed for sharing the content (eCRF, data elements, etc.) between different parties.

R3: Define a proper structure for eCRFs and its items: A metamodel should define the structural composition of eCRFs and related templates. This structure should define how certain data elements are structured into groups of semantically depended entities.

R4: Utilize a semantic enabled structure: In order to facilitate a flexible and versatile use, all components of eCRFs or templates should be nested in a suitable structure. Enriching this structure with semantics enables to generate new knowledge out of existing conditions.

R5: Link eCRF content with evidence based information: Allow to annotate arbitrary content of textual resources with related data elements. Existing study protocols, literature documents, patient related medical cases or eCRFs should be assigned to one or more CDE.

eDiFy provides a suitable implementation of this approach complying to the above-mentioned requirements and the elaborated use cases.

3.7. Correlation and mapping of standards and requirements

In this section we formulate our conceptual ideas, based on the general requirements and described standards.

The main idea is to establish a system that is able to annotate text documents with references to CDEs and, by using this information, to create evidence-based eCRF templates. Therefore, annotating capabilities as well as the link to a shared CDE repository are necessary. CDISC ODM and ISO/IEC 11179 fit with the need to build a structural and conceptual base for such a system (R1, R2[5]). This base needs to be extended by an easy to use text annotation functionality which establishes relations between content of text documents and existing CDEs (R3, R5).

Accessible data element repositories are available online: National Institutes of Health (NIH) CDE is an initiative, which provides an integrated, shared repository of CDEs and eCRFs for different research domains and diseases [5] (R1). However, currently it is not possible to automatically search and process those repositories and compose data elements needed in a certain eCRF.

[5] Relation to requirements

For the creation and usage of evidence-based eCRF templates, semantically related elements need to be provided (R4). Thus, a pertinent hierarchical structure is needed, e.g. like the Richesson and Nadkarni [8], who divide elements of eCRF in four structural nested parts: form, section, group and item (R3).

4. Discussion

This paper gives an overview of most important standards for eCRFs, related data elements and provides use cases and derived requirements for evidence-based eCRF templates. Further it states conceptual ideas on how to use the resulting requirements in the project eDiFy.

The literature review revealed eCRF related standards. The main attention was given to get an overview and a general understanding of the available standards with suitability for the aimed goals rather than providing a complete and detailed elaboration. The alignment of existing standards with described use cases proved to be a good choice of deriving requirements during the requirements elicitation phase, as the standards define a proven structure while the use cases model the processes needed to reach the project goal. Matching the characteristics of the standard with the outlined processes wherever possible ensures an interoperable approach.

According to the results, all investigated standards seem to cover the necessary structural requirements. However, we decided to use ISO/IEC 11179 for the eDiFy project as it offers high flexibility in modification by using a combination of conceptual and representation models. The standard is frequently used for managing clinical trial data elements in MDRs but also often adapted to satisfy specific organizational requirements [12]. According to the proposed problem in this paper, the ISO standard needs to be extended with semantic annotation capabilities to fit the needs. Further work will rely on existing efforts to extend the standard with additional semantic relations, like the work from Takis and Lange [20].

Requirements were derived from proper use cases and standards characteristics. It is necessary to state that those requirements are technical related, focus on the current situation and do not represent general requirements for the creation or design of eCRFs. To the best of the authors' knowledge there do not exist requirements for comparable systems. Comparing eDiFy with other projects, papers with partly similar content could be found, though they never match with the complete idea of eDiFy: There are projects aiming at the annotation possibilities of text documents, though they do not attempt to create eCRF or templates [21]. On the other side, papers could be identified that emphasize on the need of sharing eCRFs and CDEs, but they don't include any relation to evidence-based literature [22].

At present most eCRFs are not available to the scientific community. Due to this lack of transparency, standardization is hindered. The goal of designing evidence-based eCRFs has proven to be yet a very complex task. As an ongoing work to address that complexity, the requirements derived in this paper are used to develop an ontology for eDiFy as a next step. This approach should then allow to assemble annotation elements of text documents with existing eCRF categories and items in a very dynamic and flexible manner.

References

[1] J. Welker, Implementation of electronic data capture systems: Barriers and solutions, *Contemp. Clin. Trials* **3**(28) (2007), 329–336.

[2] B.C. Drolet, K.B. Johnson, Categorizing the world of registries., *J. Biomed. Inform.* **6**(41) (2008), 1009–20.

[3] M. Nahm, J. Shepherd, A. Buzenberg, R. Rostami, A. Corcoran, J. Mccall, et al., Design and implementation of an institutional case report form library, *Clin. Trials* **8**(8) (2011), 94–102.

[4] M. Dugas, K.-H. Jöckel, T. Friede, O. Gefeller, M. Kieser, M. Marschollek, et al., Memorandum "Open Metadata": Open Access to Documentation Forms and Item Catalogs in Healthcare, *Methods Inf. Med.* **4**(54) (2015), 376–378.

[5] U.S. National Library of Medicine, NIH Common Data Element (CDE) Resource Portal. https://www.nlm.nih.gov/cde/, last access 15.01.2016.

[6] S. Bellary, B. Krishnankutty, M.S. Latha, Basics of case report form designing in clinical research., *Perspect. Clin. Res.* **4**(5) (2014), 159–66.

[7] IEEE, Systems and software engineering -- Life cycle processes -- Requirements engineering, *ISO/IEC/IEEE 29148* (2011), 1–94.

[8] R.L. Richesson, P. Nadkarni, Data standards for clinical research data collection forms: current status and challenges., *J. Am. Med. Inform. Assoc.* **3**(18) (2011), 341–346.

[9] P. Mayring, *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, 8th ed., Beltz Verlag, Weinheim und Basel, 2000.

[10] Object Management Group (OMG), Unified Modeling Language™ (UML) v2.5, http://www.omg.org/spec/UML/2.5, last access: 15.01.2016.

[11] NISO Press, Understanding Metadata, (2004), 15. http://www.niso.org/publications/press/UnderstandingMetadata.pdf, last access: 03.02.2016.

[12] S.M.N. Ngouongo, M. Löbe, J. Stausberg, The ISO/IEC 11179 norm for metadata registries: Does it cover healthcare standards in empirical research?, *J. Biomed. Inform.* **2**(46) (2013), 318–327.

[13] D.B. Warzel, C. Andonaydis, B. McCurry, R. Chilukuri, S. Ishmukhamedov, P. Covitz, Common data element (CDE) management and deployment in clinical trials., *AMIA Annu. Symp. Proc.* **2003** (2003), 1048.

[14] Australian Government, Australian Metadata Online Registry (METeOR), http://meteor.aihw.gov.au/, last access 15.01.2016.

[15] ISO/IEC JTC1 SC32 WG2, ISO/IEC 11179 Information Technology -- Metadata Registries (MDR) - Part 1: Framework Ed 3, http://metadata-standards.org/11179/, last access: 13.02.2016.

[16] V. Huser, C. Sastry, M. Breymaier, A. Idriss, J.J. Cimino, Standardizing data exchange for clinical research protocols and case report forms : An assessment of the suitability of the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM), *J. Biomed. Inform.* (57) (2015), 88–99.

[17] CDISC, DEFINE-XML, http://www.cdisc.org/define-xml, last access: 28.02.2016.

[18] C.H. Lin, Y.C. Fann, D.M. Liou, An exploratory study using an openEHR 2-level modeling approach to represent common data elements, *J. Am. Med. Informatics Assoc.* **155** (2016).

[19] C. Rinner, M. Kohler, G. Hübner-Bloder, S. Saboor, E. Ammenwerth, G. Duftschmid, Creating ISO/EN 13606 archetypes based on clinical information needs, *Stud. Health Technol. Inform.* (165) (2011), 43–48.

[20] Y.R. Park, Y.J. Yoon, H.H. Kim, J.H. Kim, Establishing semantic interoperability of biomedical metadata registries using extended semantic relationships, *Stud. Health Technol. Inform.* **1-2**(192) (2013), 618–621.

[21] J. Takis, C. Lange, Crowdsourced Semantic Annotation of Scientific Publications and Tabular Data in PDF, *Proceedings of the 11th International Conference on Semantic Systems* (2015), 1–8.

[22] K. Stone, Comparative effectiveness research in neurology: healthcare reform will increase the focus on finding the most effective--and affordable--treatment., *Ann. Neurol.* **1**(68) (2010), A10–A11.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-150

Identifying eHealth Opportunities to Support Medication Adherence – Findings of a Focus Group Study

Yulong GU[a, b, 1], John KENNELY[b], Jim WARREN[b], Amelia Been AHN[b], Matire HARWOOD[b], Pat NEUWELT[b]
[a]*Stockton University, Galloway, New Jersey, United States of America*
[b]*University of Auckland, Auckland, New Zealand*

Abstract. Background: The burden of cardiovascular disease (CVD) among New Zealand (NZ) indigenous people (Māori) is well recognized. A major challenge to CVD risk management is to improve adherence to long-term medications. Objectives: To elicit patients' and providers' perspectives on how to support Māori with high CVD risk and low medication adherence to achieve better adherence. Methods: Analysis of electronic health records (EHR) of four NZ general practices identified medication adherence status of Māori patients with high CVD risk (≥15%, 5-year). A random sample of these patients participated in focus group discussions on barriers to long-term medication adherence. Their primary care providers also participated in separate focus groups on the same topic. Results: A range of factors are identified influencing adherence behaviour, including patient's medication knowledge, patient-doctor communication effectiveness and cost. Conclusion: Analysis of barriers to medication adherence in primary care suggests opportunities for health information technology to improve adherence, including patient education, decision support, clinician training and self-service facilities.

Keywords. Medication adherence, cardiovascular diseases, electronic health records, health information technology.

1. Introduction

New Zealand (NZ) indigenous people (Māori) are at high risk for cardiovascular disease (CVD), with 31.9 CVD hospitalizations per 1000 Māori (35+ years) and 2.9‰ dying from CVD annually [1]. NZ Māori suffer higher mortality and morbidity rates from CVD, with the mortality from coronary heart disease among Māori three to four times higher than in the general population [2, 3]. The NZ health system has a commendable focus on screening people for CVD risk, with guidelines recommending risk assessment for Māori, Pacific and Indo-Asian people 10 years earlier than other low-risk population groups (e.g., 35+ for Māori men and 45+ for Māori women) [4]. The challenge remains to improve the management of people with elevated risk, including populations at highest risk such as Māori. One key barrier is long-term adherence to evidence-based medication, as adherence to CVD medications is approaching only 50% among lower socio-economic groups [5]. Along with other risk factors, low medication adherence is likely to have contributed to the high CVD rates in Māori and Pacific people in NZ [6]. Among high-risk patients, those who adhered to medications had lower blood pressure, total-to-

[1] Corresponding Author: Yulong Gu, School of Health Sciences, Stockton University, 101 Vera King Farris Drive, Galloway, NJ 08205, USA. E-Mail: yulong.gu@stockton.edu.

HDL cholesterol ratio and HbA1c than non-adherers [7]. A 2013 systematic review highlighted association between adherence and improved outcomes as well as reduced costs [8]. Poor adherence has been recognized as a critical issue for public health in terms of both quality of life and health economics [9]. A 2011 estimate of the economic impact of medication non-adherence in the United States (US) accounted for $290 billion per year in avoidable medical spending or 13% of total health care expenditures [10].

In previous studies, we developed and tested a structured primary care model to improve CVD medication adherence, predominantly in NZ Pacific populations [11-13]. Our experience suggested that patient counselling and education delivered in a general practice nurse led model can improve CVD medication adherence and physiological measures in high-needs population such as Pacific peoples [11, 12]. Health services must reconfigure themselves in order to support an active, voluntary and collaborative establishment of treatment goals between patients and providers. The findings demonstrated effectiveness, but improvement in adherence rates was only achieved with substantial additional nursing time and cost [11]. Building on this experience, we conducted a qualitative study with focus group interviews to understand the perspectives of Māori patients and their primary care providers regarding what is (not) working and what might work in primary care to achieve better adherence to CVD medication among Māori. We aimed to inform the development of appropriate, effective and sustainable interventions to improve CVD risk management among Māori.

2. Methods

Participants were recruited into either patient or provider focus groups. Potential patient participants were identified by analysis of the electronic health records (EHR) data of four NZ general practices (two in a metropolitan region and two in rural regions). Extraction and analysis methods of EHR data (including demographics, CVD risk screening results, and CVD medication prescription records) were based on previous studies [6]. Analysis of the EHR data extracted in 2014 for this study identified medication adherence status of all adult Māori patients (20+ years) with high CVD event risk (≥15%, 5-year risk) who were enrolled and funded at the participating practices. Using EHR prescription records on three classes of CVD medication (antihypertensive, cholesterol and oral diabetic medication), we calculated medication possession ratio and defined adherence as 1) having at least 80% of days in last fifteen months covered by the prescribed CVD medication(s), and 2) having at least one prescription written in the past six months for the classes of drugs the patient was on. A random sample of patients with high CVD risk who had self-identified as Māori was drawn and invited to participate in a focus group session, including those with high CVD medication adherence status and those with low adherence (e.g. four adherers and four non-adherers per practice). The patients' families ('whānau') were encouraged to participate, as recommended in best practice guidelines for undertaking research and developing medicine interventions with Māori [14]. Four patient focus groups (one per practice) were led by the research team. Patient-perceived barriers to adherence were discussed in depth and suggestions on how to support patients were sought from the patients and families. The researcher introduced some potential intervention options (based on the literature and previous research by the authors) in each new session as topic prompts, to test face-value and acceptance. These included process improvement, additional protocols (e.g. our experience of nursing support [12, 13]) and technology options such as automatically-generated text message

reminders, and smart bottle / dispensing (including blister pack). The goal of discussion was to explore combination and tailoring of options most acceptable to patients.

In the provider focus groups (one per practice) a summary of EHR analysis findings was reported by a researcher, including aggregated adherence levels in the practice and CVD risk assessment status. This analysis was discussed to identify opportunities for clinical care quality improvement in the practice, particularly addressing adherence. Provider feedback was gathered on enablers, challenges, and suggestions on how to improve adherence among their Māori patients. Patient and provider group findings up to that point, including suggested protocols, interventions and technologies were discussed in each new provider group to assess the perceived feasibility of implementing such interventions. Constant comparison and triangulation of literature, patient perspectives, and provider perceptions across different practices were undertaken to identify interventions that might be acceptable to a wide group of Māori patients and general practice staff, and might be achievable within current health workforce constraints; i.e. promoting a transformation of 'usual care' that can be delivered by the same people using no more than the same amount of time as current practice.

We adopted an iterative participatory action research approach to bring together action and reflection, theory and practice, in the pursuit of practical solutions to issues of concern to individual persons and their communities [15]. The goal was to identify key components of feasible interventions for medication adherence promotion among Māori. Each practice served as an iteration. Data collected from their participants were compared by JK and YG with that collected from other practices and with literature. Then findings were synthesized and topics brought into the next practice's discussions. General inductive analysis, constant comparison and data triangulation were used throughout data analysis to identify key themes of the discussions.

3. Results and Discussion

3.1. Medication adherence gaps among high-risk Māori

The analysis included 3,414 Māori adults, with 16% (N=549) having ≥15% CVD risk. Among the high-risk Māori, only 59% had ≥80% CVD medication adherence (Table 1). Adherence related issues were discussed with practice staff in provider focus groups. A random sample of high-CVD-risk Māori were invited to participate in patient focus groups and an average of six patient/family participants attended each group.

Table 1. CVD medication adherence status of funded adult Māori patients with identified high CVD risk

Adherence status	N	%
≥80% adherence to all CVD medications prescribed	322	59%
<80% adherence to ≥1 CVD medication prescribed	127	23%
No antihypertensive, cholesterol or oral diabetic medication prescribed in last two years	100	18%
Total:	**549**	**100%**

3.2. Key factors influencing long-term medication adherence

A range of factors were reported by participating patients and family as influencing their long-term medication adherence behaviour, including patients' knowledge about medication, side effects, doctor-patient communication effectiveness, costs and value of medication and of clinic visits, and forgetfulness. These factors were repeatedly

mentioned across practices as having either positive or negative impact on adherence-related beliefs and behaviors. These factors appear interrelated and were also identified by participating practice staff. For instance, the level of patients' knowledge about their medication(s), including understanding of effects and potential side effects, is associated with how effective doctor-patient communications are perceived to be by both sides.

Doctor-patient communication requires doctors' ability to convey information and knowledge, as well as their active listening skills. Finding an effective and convenient medication for an individual often takes more than one appointment. It is a continuous and productive interaction process between provider and patient. This interaction depends on patients being informed and activated, and staff being prepared and proactive [16]. The process involves patient education and a 'trial' with the drug class, dosage, effects, to determine whether it disrupts the patient's work / life, and how it fits into her/his daily schedule. As related by a patient, "these pills that make you smell when you pass water. It's caused by these pills, and they [doctors] know it. And I said I found out what's causing that stink, it's these pills of yours. So when I've taken it I go, chuck those out and take the pink ones for the blood thinners. And the other ones, out it goes. … Because we keep each other informed they [doctors] are going to change those pills."

The medication trial process can inform clinicians and remind them to involve the patient in shared decision-making and to take patient's circumstances into consideration when choosing or changing medications. For instance, with patients who work night shifts, adherence is difficult and shared decision-making with providers to create suitable solutions is desired, "like a tailor made system for a person that works different hours," as suggested by one patient. Without close monitoring and open communication in the trial process, patient safety may be at risk and those patients who experience side effects or inconvenience may not take the medication as prescribed, if at all. As suggested by a patient, "I really believe that when they change, putting you on blockers to lower your blood pressure or anything like that, I believe that it should be monitored for at least a week, or maybe 2 weeks. ... Because it's so easy and it happens a lot, and the next minute you're wondering why you get all light headed when you stand up, and your bloody blood pressure's too low." Another patient mentioned, "I have high blood pressure as well, and that one [diuretic] that makes you wee, oh I don't want that then. Because when you're working you don't want to be running to the loo all the time. Anyway, I stopped that." Interestingly, the patient went on describing the following experience, "I checked my blood pressure too with the little cuff and it went up to 166 [mmHg] over 90 [mmHg] I think. So I took that little pill for 2 or 3 days, came back, checked it again, 130 [mmHg]."

Participants, both patients and providers, agreed that effective communication between doctors and patients plays a crucial role in promoting adherence and delivering high quality care. A patient related, "it's having the ability to talk to the health professionals on a one on one and face to face [base], and voice your concerns. Working with them I think has been a big step forward." A challenge for patients is to keep up their understanding when a medication is changed. Confusion of a drug's generic name versus brand name also requires provider attention. One approach to improve communication may be to involve family, as mentioned by a patient, "Well there's whānau, my mother-in-law, [wife's name] talks for her. She [mother-in-law] wants to sit there and talk for herself, but [wife's name] has to talk for her because she's not telling the doctor the truth. You know, she's always hiding things." In addition to effective communication, the relationship is also influenced by factors such as trust, respect, manner, culture, and language barriers. As mentioned by several patients, "if you're not treated right they're not going to come back." "I think it starts with those reception girls.

They have to be welcoming. The last thing you want is to walk through that door and see a grumpy face." "They actually need training in people skills and cultural awareness."

The cost and cost/value perception issue emerged as another key factor that affects patient behavior. Although the NZ government subsidizes long-term CVD medications, there is a dispensing fee ($5NZ per prescribed medication) payable by patients at community pharmacies. Given the number of medications high-risk patients are taking to manage their CVD risk, often alongside medications for other comorbidities, several participants felt that their medications "cost too much." Further barriers are the costs (≥$15 per visit) and time commitment ("an hour's drive" to the practice in rural area, plus taking time off work / other life priorities) associated with travelling to see a doctor for repeat prescriptions. All the costs relevant to maintain a personal medication supply may hinder not only good medication adherence but also access to care in general.

A related issue raised by patients was the value of medication adherence, particularly in the absence of symptoms. Research has identified that patients are more likely to take medication when symptoms are present [17]. A participant related, "I take my pills when my body says it's time to take pills now. As soon as I get the little boom, boom, boom in the back of the head, oh heck, my blood's gone thick again, I'd better start taking them. … over some years I've been taking my pills that way." Patient-perceived value of medication appeared to be associated with their (dis)beliefs in the efficacy of modern medication. This issue was often discussed in contrast to 'traditional' therapies. A patient related, "like some of you I was brought up the same way, Māori medicine came first before anything else. And you know, they were freely available. You've got the cobwebs up the wall, and you peel a potato if you have earache ..." Another patient mentioned, "Because of that affordability they'd rather use the old method, the old ways, the bush, the sea, the ground, the land." If patients didn't see value in doctor-recommended medication, they were unlikely to adhere to it. From the provider perspective, adherence was of great concern, as one clinician related "I get the feeling that, well for a certain percentage of the patients, they see the doctor, you can hand them the medication; as soon as they walk out of the building, half of it would end up in the rubbish bin anyway."

Patients also reported difficulty in remembering to take medication. "I definitely have problems remembering … did I take it or didn't I?" Another patient talked about a calendar system to keep track of medication "like a medication sheet." Many believed that blister packs, which also mark the date and time, help manage medication by prompting to take the right medications at the right time. Remembering to pick up prescription from doctors and medications from pharmacists could be difficult with the fast pace of life as well. Providers mentioned some services available that assist patients in obtaining medications, such as postal delivery from pharmacy to patients at rural locations and, occasionally, mail order prescriptions without face-to-face consultation.

3.3. Potential technology opportunities to support medication adherence

Previous research has suggested that the more a patient knows about their long-term conditions and their medications, the better equipped they are in managing their health. This requires patient understanding of what a prescribed therapy (e.g., medication) is for and how it works. As suggested by a patient, "I think another thing that helps … is if the doctors explain a little bit more what the pills do." However, it is a challenging task to develop adequate health knowledge/literacy within limited doctor-patient interactions. One patient stated, "I don't really listen when he [doctor] tells me what they [my medications] are for. And that's half the trouble, if I don't understand what I'm taking

and I don't know what they do, it's that mindset of oh it's okay, I can forget about taking it for today. And then I might miss a couple of days, and then I wonder why I get sick."

Targeting the information needs of patients is an opportunity for health information technologies (IT) development, e.g., to support clinicians in the patient education task, via a checklist to guide the conversation, and pointers to reliable sources of Internet information (e.g., sharing bookmarks to certified or accredited resources). They could also include decision support feature for patients and providers to use during and between appointments. A patient seeking medications information may ask their health providers, their personal contacts, or search on the Internet. The use of the Internet for patient health information has become regular, with the majority of US doctors reporting experience with a patient bringing Internet information to a visit [18]. As stated by one patient in this study, "Look it up, do a search on the Internet." However, the quality of Internet information remains of concern. To support both patients and providers, many governments have invested in providing quality Internet health information services, e.g., the NZ catalog on medicines and medical devices (http://www.medsafe.govt.nz/).

It is also important to help patients understand the consequences of non-adherence. One patient said what many others expressed, "We need to take it and if we don't we'll get sick, so we take it." This demonstrated the process of associating perceived threat with the efficacy of recommended therapy to lower risks, as suggested in the health belief model. At one rural practice, staff described a successful education program (based on Lorig's model [19]) that they ran for CVD prevention using pictures, games, goal setting exercises, and outreach nurses conducting family group classes. A staff member related, "a lot of our Māori people up here are illiterate in some form as well, so to give them pamphlets just full of writing, or handouts full of writing … They'll never read it and they can't read it, and they won't understand it." The key approach was "to try and explain the same thing using all different senses, then it gave everyone the opportunity, a chance to learn the same thing." IT may play a key role in sharing experience from successful programs such as this with wider practitioner community to encourage quicker program adoption and evaluation to build up evidence base.

Further, the technology environment of population groups with high health needs must be accounted for when designing patient-interacting systems. As put by one patient, "We live in the far north, we haven't got emails." As per a systematic review, although patient-oriented health IT can increase medication adherence, patient factors must be considered to ensure acceptance. Such factors include socio-demographic characteristics, health- and treatment-related variables, and prior experience or exposure to computer / health technology [20]. In this context, mobile phone text messaging may be a viable option to enhance patient-provider communication. Opportunity exists to develop features in the EHR to identify medication adherence gaps so as to support providers in patient management, and to trigger automatic text messages to patient as reminders for picking up prescriptions. Provider experience also suggested that reminder systems worked; for example, ringing up a patient about an overdue prescription. However, such actions take staff time and were considered unlikely to be sustainable within the current health system and workforce constraints, particularly with the shortage of Māori nursing and health workers. Automatic telephone services allowing patients to request repeat prescriptions were suggested by providers and patients, saving time when it is safe to do so. It was also suggested that kiosk technology, if installed in convenient locations and capable of data sharing, e.g., with PHRs, may assist self-monitoring, e.g., of biometrics such as blood pressure, and thus promote adherence to therapy. Information sharing of self-monitoring data with clinicians may also improve care quality and patient safety.

Blister packs were seen by participating patients as useful to manage complex medication regimen. Research trials on combination drug (polypill) also suggest advantage in promoting adherence [21]. Potentially, IT solutions such as smart bottle may serve the same purpose; and such technology may record better data on patient behavior to support further improvement. Table 2 summarizes the key aims of potential IT solutions to improve medication adherence by addressing key influencing factors.

Table 2. Potential health IT to address adherence factors

Key target area for potential health IT	Adherence factors
To utilize quality Internet information services; to share education materials, experience, and information on consequences of non-adherence	Improving patient knowledge
To identify medication adherence gaps; to support patient self-monitoring; to deliver effective interventions to wider community; to facilitate provider training and knowledge sharing	Supporting patient-provider communication
To simplify prescription refill procedures; to support convenient data collection and sharing in patient monitoring	Reducing costs
To identify patient needs; to support data collection regarding patient behavior	Tackling forgetfulness

3.4. Implications for eHealth practitioners

State-of-art health IT solutions, especially in knowledge management, clinical decision support, telehealth, patient health records (PHR), and mobile health, may help address adherence-related issues. The key challenges in this target area include:

- Provision of effective patient education on medication, including side-effects and impact of non-adherence. We call for developing locally-appropriate multimedia education materials, including materials to assist clinician-led education and consult (e.g., checklist-based) and self-help materials catering for patients with various levels of health literacy and deliverable via tablet computer (e.g., at clinics), kiosk (e.g., at community centers), and Internet.
- Decision support integrated with EHR to identify medication adherence gaps (e.g., based on our algorithm), and subsequently to prompt patients to act (e.g., via system-generated text messages) and providers to intervene (e.g., alerts).
- Clinician training with a focus on cultural sensitivity and competency. We recommend utilizing computer-based simulation technologies in this task.
- Patient support, in terms of self-monitoring, data management, and communication with providers. Kiosk, PHR (interoperable with EHR), and telehealth are promising technologies to reduce cost and empower patients.

In summary, there are opportunities for health IT to provide information and communication support to population groups with high health need, e.g., Māori in NZ.

Acknowledgements

This study was supported by the University of Auckland School of Population Health Performance-Based Research Fund (PBRF) Distribution Seeding Grants. We also thank Ms Tai Boyce, the participating patients, their whānau, and the practice staff. This research was conducted under ethics protocol 14/NTA/145 approved by the Northern A Health and Disability Ethics Committee.

References

[1] Ministry of Health. Ngā mana hauora tūtohu: Health status indicators - Cardiovascular disease Wellington: Ministry of Health; 2015 [updated Oct 07 2015; cited 2015 Dec 17]. Available from: http://www.health.govt.nz/our-work/populations/maori-health/tatau-kahukura-maori-health-statistics/nga-mana-hauora-tutohu-health-status-indicators/cardiovascular-disease

[2] Thornley C, Tobias M, Bonne M. Priorities for Mäori and Pacific Health: Evidence from epidemiology Wellington: Ministry of Health; 2001 [cited 2013 Nov 11]. Available from: http://www.moh.govt.nz/notebook/nbbooks.nsf/0/A31842D91480064FCC256A55007A980A/$file/PrioritiesForMaoriandPacificHealth.pdf

[3] Ministry of Health. Diabetes and Cardiovascular Disease Quality Improvement Plan Wellington: Ministry of Health; 2007 [updated 2007; cited 2012 Dec-26]. Available from: http://www.moh.govt.nz/notebook/nbbooks.nsf/0/D5D08CAB1256B1F4CC2573FB006A1CBE/$file/diabetes-cardio-quality-improvement-plan-feb08-v2.pdf

[4] New Zealand Gudelines Group. New Zealand primary care handbook 2012 Wellington: New Zealand Guidelines Group; 2012 [cited 2015 Mar 12]; 3rd. Available from: http://www.health.govt.nz/system/files/documents/publications/nz-primary-care-handbook-2012.pdf

[5] Mabotuwana T, Warren J, Harrison J, Kenealy T. What can primary care prescribing data tell us about individual adherence to long-term medication?-comparison to pharmacy dispensing data. Pharmacoepidemiology and drug safety. 2009;18(10):956-64.

[6] Gu Y, Warren J, Kennelly J, Neuwelt P, Harwood M. Cardiovascular disease risk management for Maori in New Zealand general practice. J Prim Health Care. 2014;6(4):286-94.

[7] Gu Y, Warren J, Walker N, Kennelly J. Gender differences in cardiovascular disease risk management for Pacific Islanders in primary care. Quality in primary care. 2013;21(5):275-85.

[8] Bitton A, Choudhry NK, Matlin OS, Swanton K, Shrank WH. The impact of medication adherence on coronary artery disease costs and outcomes: a systematic review. The American journal of medicine. 2013;126(4):357 e7- e27.

[9] World Health Organization. Adherence to long-term therapies: evidence for action Geneva, Switzerland: World Health Organization; 2003. Available from: http://www.who.int/chp/knowledge/publications/adherence_report/en/

[10] New England Healthcare Institute. Thinking Outside the Pillbox: A System-wide Approach to Improving Patient Medication Adherence for Chronic Disease: New England Healthcare Institute; 2009 [updated August 12 2009; cited 2014 May 16]. Available from: http://adhereforhealth.org/wp-content/uploads/pdf/ThinkingOutsidethePillbox_Report.pdf

[11] Warren J, Kennelly J, Warren D, Elley CR, Chang Wai K, Manukia M, et al. Using the General Practice EMR for Improving Blood Pressure Medication Adherence. Studies in Health Technology and Informatics. 2012;178:228-34.

[12] Warren J, Gaikwad R, Mabotuwana T, Kennelly J, Kenealy T. Utilising practice management system data for quality improvement in use of blood pressure lowering medications in general practice. N Z Med J. 2008;121(1285):53-62.

[13] Warren J, Gu Y, Kennelly J, editors. The Caring Does Matter (CDM) Initiative: To Improve Cardiovascular Medication Adherence in Pacific People. Proceedings of Health Informatics New Zealand 11th Annual Conference and Exhibition; 2012 Nov 8-9; Rotorua.

[14] Health Research Council of New Zealand. Guidelines for Researchers on Health Research Involving Māori Auckland: Health Research Council of New Zealand; 2010 [cited 2016 Jan 25]. Available from: http://www.hrc.govt.nz/news-and-publications/publications/guidelines-researchers-health-research-involving-m%C4%81ori

[15] Reason P, Bradbury H, editors. Handbook of action research. 1 ed. London: Sage; 2001.

[16] Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q. 1996;74(4):511-44.

[17] Marshall IJ, Wolfe CDA, McKevitt C. Lay perspectives on hypertension and drug adherence: systematic review of qualitative research. BMJ. 2012;345.

[18] Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the Internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. J Med Internet Res. 2003;5(3):e17.

[19] Lorig KR, Holman HR, Sobel DS, Laurent DD, Gonzalez VM, M. M. Living a healthy life with chronic conditions. Palo Alto, CA: Bull Publishing Company; 1993.

[20] Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc. 2009;16(4):550-60.

[21] de Cates AN, Farr MR, Wright N, Jarvis MC, Rees K, Ebrahim S, et al. Fixed-dose combination therapy for the prevention of cardiovascular disease. Cochrane Database Syst Rev. 2014;4:CD009868.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-158

Specific Requirements of Physiotherapists on the Practical Use of Software in the Therapeutical Process

Karin MESSER-MISAK[1, a], Rudolf EGGER[b]
[a]*Fachhochschule Joanneum, Institut of eHealth*
[b]*Karl-Franzens-Universität Graz, Institute of Educational Sciences*

Abstract: The current healthcare system requires more effective management. New media and technology are supposed to support the demands of the current healthcare system. By the example of physiotherapy, the primary objective of this study was to define the specific requirements of therapists on the practical use of software which cover the administration, documentation and evaluation of the entire therapy process, including a database with pictures/videos about exercises which can be adapted individually by the therapists. Another objective was to show what conditions for a successful implementation of advanced applications during the entire treatment process have to be fulfilled. The approach of mixed-methods designs was chosen. In the first part a two-stage qualitative study was carried out, followed by a quantitative survey. The results show that the use of the software regarding the therapy-related part is dependent on how adaptable the software is to the special needs of the therapists, that the whole treatment process is mapped on the software and that an additional training during the professional practice must be implemented in order to deploy the use of the software successfully in the therapeutic process.

Keywords: Physical Therapy Software, Software Tools for Therapy, Requirements, Physiotherapy, Competence

1. Introduction and Background

The education of health professionals in Austria has recently been moved to a tertiary level. With this move, an increased use of new technologies is expected in this professional area [1]. The new opportunities provided by media and technology are to accompany and strengthen this improvement process. The digitalisation of entire healthcare systems has triggered complex developments which are not always predictable. Therefore physiotherapists do not only need knowledge about the possibilities of existing technologies and advanced software solutions but also the ability to apply them to specific situations. It is not sufficient to deal selectively with new software applications, rather a holistic consideration of the physiotherapeutic process is required to enable them to use the technologies for best achievement of the goals set [2].

After the switchover during which the professional training was moved to a tertiary level according to the Bologna process in 2006, a further step towards the modernization and deployment of technologies and software solutions especially in the therapeutic environment is expected [3,4]. The gap between the available advanced software

[1] Corresponding Author: Mag. Karin Messer-Misak, FH Joanneum Gesellschaft mbH, Eggenberger Allee 11, 8020 Graz, E-Mail: karin.messer-misak@fh-joanneum.at

solutions and their actual use in therapeutic practice is getting bigger with the progressing digitalisation of society.

The interactive application of media and their resources provides new opportunities of designing professional scenarios. Software in therapeutic practice is mainly used to streamline administration procedures. The multifaceted options for interacting with the individual optimization of efficient establishment of findings or the definition of therapeutic targets and strategies as well as the access to a comprehensive databases is not yet considered in the therapeutic process[5]. The aim was to solve a practical problem, namely how to improve the practical use of therapy relevant software in a therapeutic environment. The increase in knowledge is gained primarily from practice-relevant results. The results were obtained directly in the research field. The objective of this paper is to show to what extent and in which phase of the therapy the use of advanced software solutions is considered to be helpful and which requirements must be met by the therapists to integrate them in their professional field of action as well as in the treatment process to gain the maximum benefit from it.

2. Methods

In order to analyse the current state of research and the theoretical background as well as possible, a systematic literature review was initially carried out.

In a first step, specific terms and relevant keywords were determined, and a systematic search within scientific databases (ERIC, SpringerLink, ScienceDirect College Edition, Pubmed, Thieme-connect, Medline), academic publishers and online bookstores was done. The following search criteria have been chosen: physioth* and competen*, physioth* and IT, physioth* and software*, physioth* and computer, physioth* and competen* and technic*. The search criteria were restricted to only peer reviewed articles and were limited to the years 2010-2015. In total more than 22,250 hits were recorded, of which 15 studies were selected for their relevance and reviewed.

For the study, the research strategy of mixed-methods [6] was chosen. The work consists of a two-stage qualitative and quantitative study, carried out consecutively and referring to each other.

2.1. Methods Qualitative Research

For the qualitative study, a semi-structured qualitative interview guideline for the expert interviews (n=5) and the interviews with the therapists (n=8) was designed. The interviews were conducted in the period February - April 2015. All interviews were been set up as One-on-One Interviews, one hour was allocated. The interview guide contained 18 questions, which were divided into further sub-questions regarding the factors influencing the practical use of software and the competence that is needed to operate with software during the therapeutic process. In addition to the ideal conditions for the use the interviewees were asked about the advantages and drawbacks of these options. Interviews were conducted with persons with relevant qualifications and practical experience in physiotherapy. Various therapy organizations were asked to name interested volunteers (n = 8). It was possible to interview these volunteers at their place of work. The experts (n = 5) were asked in person for an interview because of their specific expertise and their functions in political or management positions, which are decisive for physiotherapy in Austria.

The requirements of typical action scenarios in which therapy-relevant software solutions are used were analyzed according to the physiotherapeutic process. The four phases of the process are as follows:

- Problem recognition
- Therapy planning
- Implementation and realisation
- Comprehensive activities

This approach was chosen because the physiotherapeutic process follows the logically structured process of treatment as laid down in the International Classification of Functioning, Disability and Health (ICF) as well as in the current model of health issued by the World Health Organization (WHO) [7]. As software solutions are already used before the first treatment to inform the client about the range of therapies, the so-called Preliminary phase is additionally described. The interviews were transcribed and a software for qualitative data analysis (MAXQDA) was used. The text material was screened and coded. In a second step, the material was examined for similarities and patterns of action. Codes were created to which the text material was assigned. In a third step the material was selectively encoded and interpreted.

2.2. Methods Quantitative Research

In a subsequent quantitative study in the period March – April 2015 the results of the qualitative part of the study were considered. Using a standardized questionnaire, which previously was subjected to a pretest, the filtered aspects identified through the qualitative research were tested by a larger sample for their relevance (n = 306, this part of the questionnaire was fully answered by 179 participants). It was sent to approximately 1,000 therapists via a link in an email, with a 58.5 per cent response rate. The mail addresses were exported from the public database of the “Bundesverband der PhysiotherapeutInnen Österreichs”. The form was also included in the Facebook page of the Association of PhysioAustria.

The free online survey application "EHR LimeSurvey" was used. The therapists were to complete the questionnaire independently online. Caching the results (after a password was set) and subsequent completion were possible. Upon completion of the survey, the raw data were exported and an analysis was performed by using the statistical software R version 3.1.3 Core Team (2015) and the package psych Revelle (2015). The graphics were created with the package ggplot 2 (Wickham 2009).

The questionnaire comprised 3 topical areas. The first part of the questionnaire includes questions about the age, gender, the duration of their professional practice and the medical discipline they are associated with. The second part referred to the competences needed to use software in connection with the therapeutical process. The specific requirements on the practical use of software in the therapeutical process made up the third part of the questionnaire.

3. Results

3.1. Results of the qualitative findings

A total of 225 codings out of thirteen documents and two groups of documents (interviews with experts and therapists) associated with the topic “requirements on software solutions used during the physiotherapeutic process” were recorded. The leading questions were in which phases of the therapy is software in use, which requirements are needed and how the integration in the treatment process can be guaranteed.

The results gained from the expert interviews and the statements of therapists, referring to the physiotherapeutic process, are as follows:

- Prephase: Since 2005 YouTube has offered a video portal that allows users to watch video clips for free, rate and upload them. About 133,000 results (as of 01.06.2016) appear, when the term "excer* physiotherapy*" is entered. All surveyed therapists consider the application of the recommended exercises - performed without instructions – to be a health risk for the clients. The therapists support health literacy in general, but only recommend the use of quality-assured exercises and information. Exercises with therapeutic indications should only be performed under therapeutic supervision.
- Problem recognition: In this phase, the medical history concentrates on the health complaints, the social environment, the history and development of the disease and the effects of the disease on the client at the level of "activities" and the level of "participation". Therapists claim that a smooth exchange of data between institutions should be ensured, and that the software should provide the possibility to guide through a standardized diagnostic assessment which can be adapted by the therapists.
- Therapy planning: In this phase the therapists determine in cooperation with the client which treatment goals are to be achieved in a realistic time-span and what kind of measures have to be set in order to reach the goal within a planned timeframe. Software solutions offer an optimal way to access databases/videos/images/models, by means of which the therapists can explain the effect of therapeutic measures. The recording of the current situation in the form of an image or video (for example, analysis of stance and gait) is desired for the presentation and documentation. Based on the findings and the objectives to be achieved, the software should display individually selectable exercise suggestions.
- Stage of implementation: The selected activities are carried out at this stage. Software solutions are considered to be helpful, if there are possibilities to assemble individually tailored exercises and add individual hints and suggestions for the clients. The individual exercises can be printed for the client or the client has the possibility to obtain an "app". If the client wants to access their digital practice records, special attention must be placed on data protection. Reminders and (game based) feedback functions while taking advantage of an app are encouraged by the client. The information to the therapists about the performance of the exercise, the number and frequency as well as the individual feedback of the clients are deemed helpful. The information stored should be

retrievable for later analysis as well as for measuring and assuring the quality of results.

- Phase Cross-cutting activities: Relief and support in all administrative activities, such as planning appointments and billing as well as documentation and an interdisciplinary exchange of information between the persons involved in the treatment, are vital. Data should be available for future evaluation purposes to contribute to improved data representation in the context of evidence-based therapy, and to evaluate the quality of treatment results.

One particular requirement of therapists is that software solutions should be available as a complete solution. This emerges from the fact that the conventional administration of client data also covers the entire therapeutic process. This was confirmed by one of the interviewees as follows:

> *„It would be particularly useful for us if the entire treatment process were mapped on the software".*

Out of 13 interviewees, 12 had already gained experience with software solutions in practice. Only one therapist has experience with a "complete solution" and also uses it for the entire therapeutic process. However he pointed out that there is still a need to improve the usability and he explicitly referred to the insufficient quality of the images, photos and exercises in this software solution.

In addition, fears exist concerning data protection and the question of who assumes liability in case of damage caused by improper use of software solutions. The question of how additional expenses will be paid for the enrollment of the client or the cost of a software solution has also arisen. Therapists also fear that the "Face-to-face-therapy sessions" could be reduced through the use of teletherapy. It is also unclear whether the supervised entities are paid for by the clients or by the health insurance provider.

Ultimately it was pointed out that the available market software solutions are not sufficiently mature and their use would currently be connected to considerable extra efforts.

3.2. Results of the quantitative findings

The sample size does not allow any generalization to all Austrian physiotherapists because rather younger people with little work experience participated in the survey. However it can serve as an indicator of the future use of technologies.

Based on the literature review and the results of the qualitative survey typical plot scenarios emerged in which software solutions are used. These were listed and the participants were asked to what extent the use of software in therapeutic issues were perceived as helpful (see Figure 1):

The question "In what situations is the use of software solutions helpful?" aimed at the situation-dependent use of software solutions. Different sub-items (shown in figure 1) can be grouped thematically in two categories:

- Six questions relating to the therapeutic process on problem identification
- Six questions relating to the planning and implementation phase of therapy

To assess the internal consistency of the results across those items Cronbach´s α (alpha) was used. It is assumed that the average correlation of the set of items (on problem identification and the planning and implementation phase of therapy) is the average

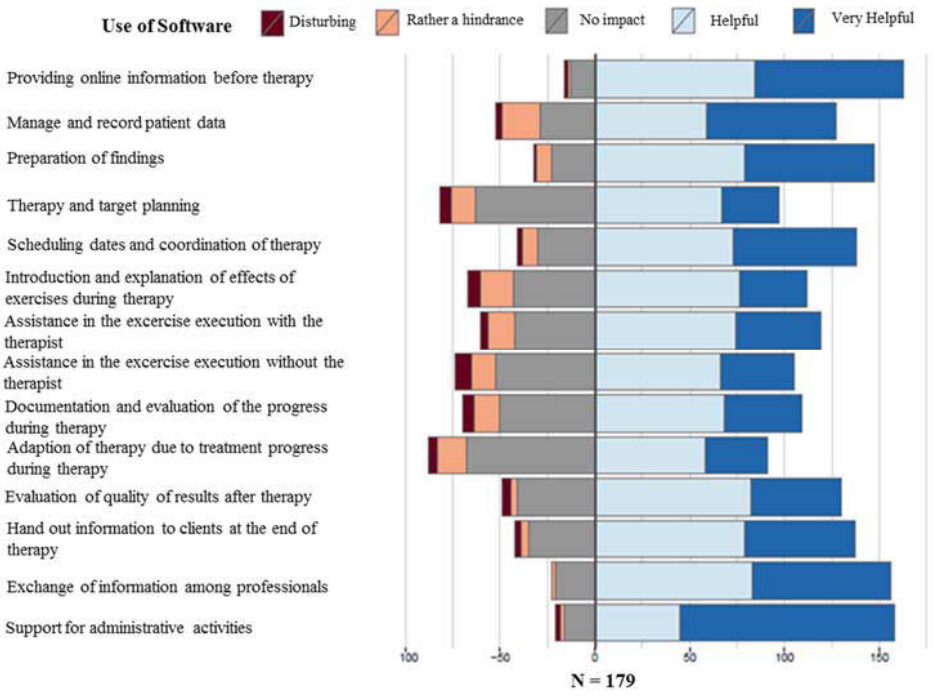


Figure 1. Scenarios – Software Solutions helpful.

correlation of all items that pertain to this construct. Using ordinal α is to verify the conformity of the answers given for the categories "phase of problem identification" and "phase of the planning and implementation". All considered coincidence measurements yielded values above 0.79 which suggests a high degree of accordance (see Table 1).

One of the issue adressed in the questionaire was in which phases of the physiotherapeutical process software with therapeutical relevant aspects is used and for what kind of specific actions. It turned out that the use is only selective. So far more than 73% of surveyed physiotherapists have managed their patient data digitally. More than half of the therapists (57.87%) create their findings supported by digital technologies, or record them digitally after the anamnesis for the purpose of documenting. In therapy planning only 46.06% of respondents use digital tools. The tools for individual exercise compilation for clients is used by 50.39% of respondents. For statistical purposes and evaluations with respect to the treatment outcome, only 22% of respondents use software solutions. A small percentage of therapists (14.2%) either work occasionally (5.9%) or slightly more frequently (8.3%) with software tools involving the clients at the same time.

Table 1. Degree of Accordance

Questions regarding	**ordinal α**
Problem identification	0.79
Planning and implementati	0.87

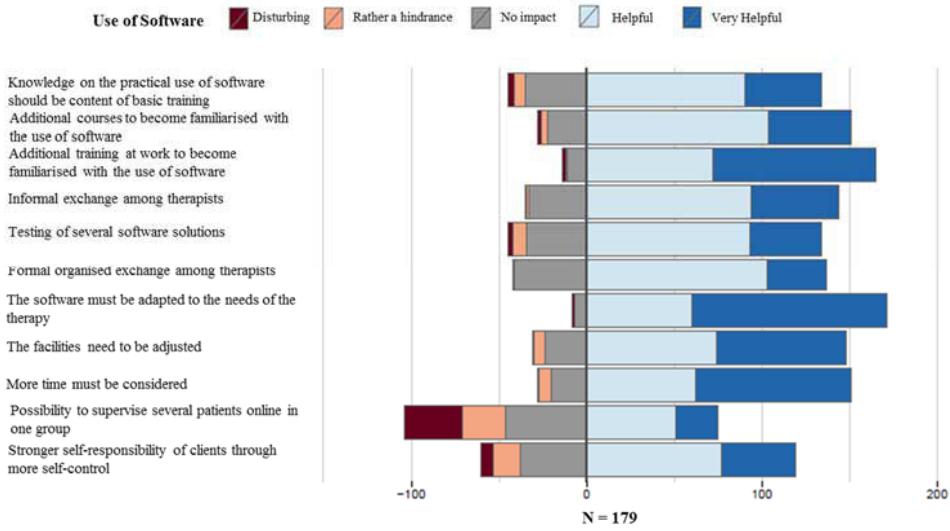


Figure 2. Requirements for the Use of Software.

Regarding the requirements that therapists need to effectively deploy software solutions in the therapeutic process (see Figure 2), it was emphasized that the technology must cover the specific demands of the field of physiotherapy. The software has to be easily adaptable (62.01 % very helpful and 33.52 % helpful). Additionally the therapists require training at work on how to use the technologies (51.96 % very helpful and 40.22 % helpful). From an organisational point of view it is relevant that the extra-time that might be needed for additional instructions and explanations to the clients is included in the treatment times (49.72 % very helpful and 34.64 % helpful). If additional space is required, it should be made available accordingly (82,68 % very useful and helpful).

Providing information via the Physiotherapist's Association, an informal exchange between therapists, the testing of several software solutions and transfer of basic knowledge about the use of software solutions should already be taught during the training period. This is also perceived as helpful.

4. Conclusion

The detailed evaluations show that no holistic, continuous use and integration of software technologies is part of the therapeutic process. The therapists only selectively apply some of the available possibilities and individual functions. So they do not fully benefit from them as support for the entire, continuous therapeutic process.

The existing and ongoing research refers predominantly to the technology acceptance of clients and/or therapists [8-10] and shows a number of factors that influence their decision about how and when therapists will use technology in general. On the other hand, there are studies [11] about the competences therapists need in general within the physiotherapeutic process. However the use of therapy relevant software tools during this process is not considered. This study takes both into account, the aspects of technology acceptance as well as the entire workflow in the therapeutical process. To

fulfill the specific requirements of physiotherapists on the practical use of software is the first step to use of new work equipment in the physiotherapeutic process.

The next stage in this study will be to include the findings about the competencies therapists need to understand und use software for the therapeutic process to gain the maximum benefit from the integration of technology in their professional field of action as well as in the treatment process.

The adaption of the necessary skills and incorporation of the use of software that opens up new areas of application, or new technologies in general, in the learning contents will need to be considered in the near future. However comprehensive formal knowledge acquisition does not automatically lead to the applicability and implementation of knowledge in practical professional life [12]. Many requirements relate directly to the demands of working and life challenges [13] and have to be adapted to the actual situations [14,15]. Individuals as well as healthcare institutions have to be responsive, participate actively in opportunities and possibilities in their ongoing development, as circumstances and technologies change rapidly.

References

[1] Rosenbach, M., Schmergal, C., Schmundt, H., Der gläserne Patient: Digitalisierung, Total vermessen Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Bd. 50(2015), 10-18.

[2] Rychen, D.S., et.al., Definition and Selection of Competencies: Theoretical and Conceptual Foundations: Executive Summary, URL: https://www.oecd.org/pisa/35693281.pdf, last access: 10.3.2016.

[3] Waegemann, C.P., Patienten- und Arztkommunikation mit der Elektronischen Patientenakte, in: Personalisierte Medizin & Informationstechnologie, Niederlag, Lemke, Rienhoff, Dresden, 2010. pp. 58-69.

[4] EPatient RSD GmbH (2015), ´4. Nationaler ePatient Survey 2015: Die Internetnutzung von Patienten und deren Auswirkungen auf ihre Therapie und den Gesundheitsmarkt. Die jährliche Befragung unter 50 Mio. Gesundheits-Surfern im deutschsprachigen Internet´. URL: https://medonline.at/2015/e-patient-survey-2015/, last access: 8.2.2016.

[5] Burgsteiner, H., Messer-Misak, K., Spieler, B., Gödl-Purrer, B., Salchinger B., Development and Evaluation of a web-based Application for Digital Findings and Documentation in Physiotherapy Education, in: Proceedings of the eHealth 2015, Ammenwerth, Hörbst, Hayn, Schreyer, Wien, 2015. pp. 182-189.

[6] Döring, N. & Bortz, J., Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften, Springer-Lehrbuch, 5., vollst. Überarb., akt. u. erw. Aufl. 2016, Springer Verlag Berlin, 2016. pp. 184-185.

[7] ÖBIG – Österreichisches Bundesinstitut für Gesundheitswesen im Auftrag des Bundesministeriums für soziale Sicherheit und Generationen (2004a), ´Berufsprofil der/des diplomierten Physiotherapeutin/Physiotherapeuten: Projekt MTD Qualitätssicherung. URL: https://www.physioaustria.at/system/files/general/berufsprofil.pdf, last access: 8.2.2016.

[8] Kairy, D, et.al., The Patient´s Perspective of in-Home Telerehabilitation Physiotherapie Services Following Total Knee Arthroplasty, Int. J. Environ. Res. Public Health 2013, 10(9) 3998-4011; doi:10.3390/ijerph10093998.

[9] Fraiwan, M.A., Khasawneh, N., Malkawi, A., et.al., Therapy central: On the development of computer games for physiotherapy, in: Innovations in Information Technology (IIT), URL: http://dx.doi.org/10.1109/Innovations.2013.6544388, IEEE, 2013. pp. 24-29.

[10] Feistritzer-Groebl, P., Computerunterstützter Therapieansatz zur sportwissenschaftlichen und physiotherapeutischen Skoliosebehandlung, Karl-Franzens-Universität Graz, 2012.

[11] Becker, V., Schlüsselkompetenzen Deutscher Physiotherapeuten, in: ´Erfolgreich in die Zukunft´, Heyse, V. & Giger, M., medhochzwei Verlag, Heidelberg, 2014. pp. 289-319.

[12] Heyse, V., Aufbruch in die Zukunft. Erfolgreiche Entwicklungen von Schlüsselkompetenzen in Schulen und Hochschulen: Aktuelle persönliche Erfahrungen aus Deutschland, Österreich und der Schweiz, Waxmann Verlag Münster, 2014. p. 13.

[13] Heyse, V., Schircks, A.D., Kompetenzprofile in der Humanmedizin: Konzepte und Instrumente für die Ausrichtung von Aus- und Weiterbildung auf Schlüsselkompetenzen, Vol. Bd. 8, Waxmann, Münster, 2012. pp. 18-19.
[14] Heyse, V., Kompetenzen erkennen, bilanzieren und entwickeln, Waxmann, Münster, 2004. p. 8.
[15] Hof, C., (Wie) lassen sich soziale Kompetenzen bewerten?, in: ´Kompetenzentwicklung in der beruflichen Bildung´, U. Clement & R. Arnold, Schriften der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE), Leske und Budrich, Opladen, 2002. p. 155.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-167

KATIS: An eHealth System for Complementary Medicine

Venus OGULTARHAN[a,1], Alban SHOSHI[a], Rico MAGNUCKI[a], Benjamin KORMEIER[a] and Ralf HOFESTÄDT[a]
[a] *Bioinformatics/Medical Informatics Department, Bielefeld University, Bielefeld, Germany*

Abstract. Background: Much of the information on the complementary medicine is spread across literature and the internet. However, various literature and web resources provide information just of one specialist field. In addition, these resources do not allow users to search for suitable therapies based on patient-specific indications. Objectives: Aggregating knowledge about complementary medicine into one database makes the search more efficient. Methods: Data integration is a promising method for providing well-based knowledge. Therefore, integrative methods were used to create the database ALTMEDA, which includes complementary and drug-related data. Results: Based on this comprehensive database ALTMEDA, the new eHealth system KATIS and the corresponding app ALMEKO for the mobile usage were implemented. Conclusion: KATIS is a web-based system for complementary medicine. KATIS provides knowledge about ten different specialist fields, which enables users not only to look up a particular complementary therapy, but also to find suitable therapies for indications more efficiently. [http://www.komplementäre-medizin.de]

Keywords. Information systems, complementary therapies, decision support techniques

1. Introduction

More and more patients use complementary and alternative medicine for treatment of indications. The term summarizes several specialist fields, which do not belong to the conventional scientifically-based medicine. Herbalism, homeopathy or monastic medicine of the complementary medicine can even so supplement the conventional medicine in a useful manner. In particular, patients with mild or chronic diseases trust in complementary therapies. Using the internet, this data is available and easy to find. Various web resources for complementary medicine, such as the portal "aponet" [1], provide a large amount of relevant information for patients. Unfortunately, such portals are always limited to one specialist field of complementary medicine. In addition, these portals do not offer an indication check, which enables users the ability to identify appropriate therapies based on specific symptoms and diagnoses. For this purpose, we have developed a new eHealth system named KATIS, which provides fused data of complementary medicine and a tool for checking indications. The indication check allows users a specific and individual therapy selection. Furthermore, the corresponding app "ALMEKO" was implemented for the mobile usage.

[1] Corresponding Author: Venus Ogultarhan, Bioinformatics/Medical Informatics Department, Bielefeld University, P.O. Box 10 01 31, D-33615 Bielefeld, Germany, E-Mail: venus.ogultarhan@uni-bielefeld.de

2. Methods

KATIS is a web-based information system for complementary medicine. The user-friendly application is designed to assist patients and health care professionals during the process of searching for alternative therapies efficiently. The core of KATIS is represented by the database "ALTMEDA", which provides data of ten different specialist fields. This includes herbalism, Bach flower, homeopathy, food, healing stones, Maria Treben, Hildegard von Bingen, essential oils, Schüssler salts, nutritional supplements associated with drugs, synonyms and fields of application. Due to this comprehensive knowledge, KATIS provides an intuitively designed graphical user interface (GUI) with "so-called" auto-complete functionality. This feature is used to suggest possible remedies or indications to users, which guarantee search results.

2.1. Data integration

Several parsers were developed to realize the automatic integration of relevant data [2]. These parsers integrate the information from sources as shown in figure 2. The application logic of the parsers was implemented in PHP [3], which provides a large number of suitable functions. The functionalities were encapsulated in order to increase the maintainability of the parsers. Furthermore, it provides a better overview of the information gathering process and the processing of data. Therefore, each parser consists of minimum two files, one for data collection and processing and one for handling database communications.

First of all, the parser filters the incoming text of the source by separating it into elements. The elements such as headings, which serve for collection and separation, were defined due to manual check of the source. The required data are extracted and stored based on the strict separated text.

A common structuring of the collected data grants the transformation into a unified character encoding by formatting, encoding and removing unnecessary spaces. During all of these steps, the parsers use pre-designed functions of PHP. Another important step in the data integration process was to distinguish symptoms from diagnoses. Codes of ICD-10, an International Classification of Diseases, were used to distinguish symptoms from diagnoses by searching the database "GELBE LISTE" [4] for potential ICD-10 codes that match indications within ALTMEDA. If a code exists, the matching indication is a diagnosis. Otherwise, it is a symptom. This additional information was also stored in ALTMEDA. The ICD-10 codes were obtained from the database "GELBE LISTE", which is a drug reference with more than 30.000 drugs representing relevant drug data.

At the end of the integration process, database entries were generated to set up the specific relationships between the entities. A single connection relation merges therapies on symptoms and diagnoses and vice versa. Thereby, the connection relation assigns the identification numbers of therapies to those of diagnosis and symptoms.

2.2. Data fusion

The integrated data were mapped to the databases "Priscus-Liste" [5], "ABDAMED" [6] and "CYP" in order to provide the basis for implementing an interaction check and to enhance the data quality of the new database ALTMEDA.

"Priscus-Liste" is a collection of 83 drugs from 18 drug classes, which were classified as potentially inappropriate for elderly people.

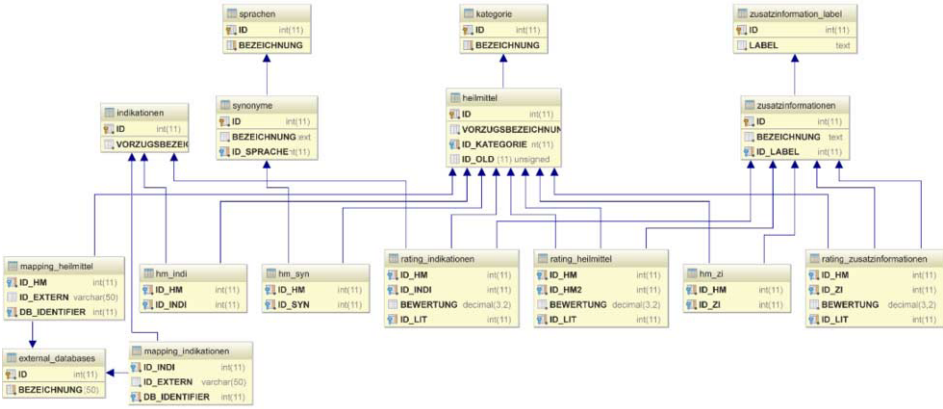


Figure 1. Database structure of ALTMEDA

ABDAMED is a commercial pharmaceutical database, which contains approved drug-related data, such as active ingredients, excipients, risks, indications, contraindications and adverse effects.

The additional constructed database “CYP” includes information on cytochrome P450 enzymes, which plays an important role in the degradation of drugs. For instance, if a drug induces a CYP that is also active in another drug’s metabolism, the dosage of the first drug must be enhanced to achieve a therapeutic effect [7].

As summarized in figure 1, the database structure with 116 relations and 17 remedy categories was reduced to 16 relations in total by optimizing the structure. The base of the new database consists of the relations “heilmittel”, “indikationen”, “zusatzinformationen” and “synonyme”.

While the relation “indikationen” does not require further classifications, the remaining three relations were associated to additional relations that were used to identify the data on unique key-value pairs. The relation “heilmittel” combines the relations “indikationen”, “zusatzinformationen” and “synonyme”. The implementation of ratings was realized on the relations with the prefix “rating”. Therefore, ratings of medications, additional information and indications can be mapped.

The number of 17 remedy categories in the raw data database was reduced in ALTMEDA to a total number of 10 by optimizing (Figure 2). The categories of medicinal plants and herbs were merged based on the knowledge that herbs belongs biologically to plants. A similar procedure was performed with the nutrients. Minerals, trace elements and vitamins were merged under the preamble nutrients. The differentiation was realized by additional information in the relation “zusatzinformationen”. The remedies of Maria Treben and Hildegard von Bingen were also integrated into the categories medicinal plants and healing stones, which represents their remedies. In order to provide a central point for synonyms, they were summarized in a separate relation. The relation “synonyme” includes English, German, Latin, and Chinese names. In the same way, a central point was created for the additional information. The relation “zusatzinformationen” contains all the additional information on the remedies. This fusion provides three major advantages as opposed to the raw data structure: all information of a remedy can be detected independently of category. Secondly, the storage requirement is reduced drastically, because no redundant entries are stored. Furthermore, it is easy to add more remedies with associated data.

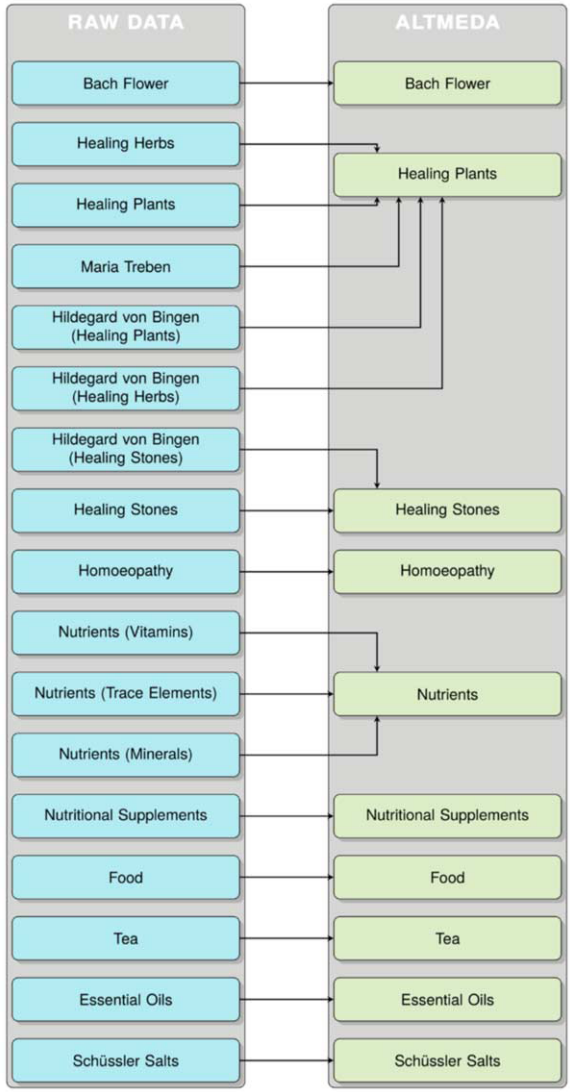


Figure 2. The raw data of the sources are shown on the left hand side, which were mapped to remedy specialist fields of ALTMEDA on the right hand side.

The fusion of "Priscus-Liste", "CYP" and ABDAMED was realized via a relation, which connects the primary key of both databases to refer to each other. The reference was obtained by comparing the strings. Based on the data from the database ALTMEDA, the contents of other databases were iterated and searched for matches. In case of a match, both primary keys were inserted into the relation tables, which consists of three attributes. The attributes describe the primary key of the record provided by ALTMEDA, the primary key of the record provided by the foreign database and an identifier of the source.

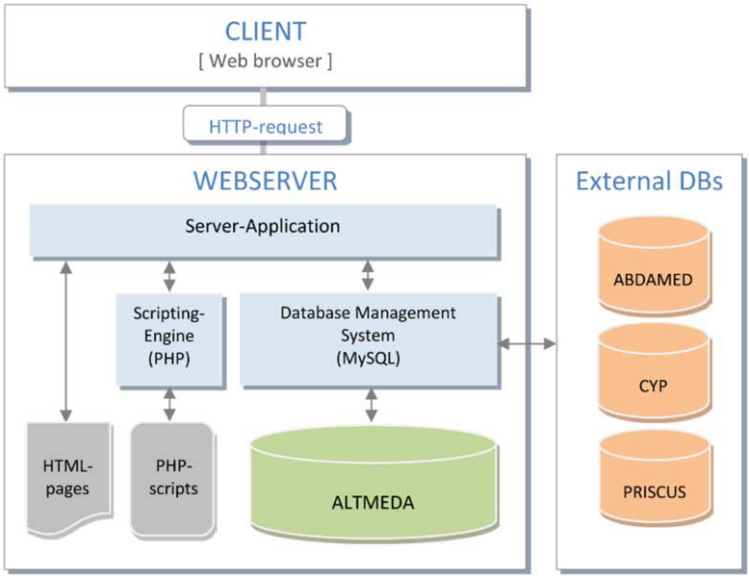


Figure 3. Client/Server-architecture of the information system KATIS.

2.3. Implementation

KATIS was developed as a user-friendly information system for complementary medicine. It is a platform-independent eHealth system and accessible via the internet, which is based on a common client/server-architecture. The use of a client/server-architecture guarantees an advantageous separation and decoupling of aspects of graphical user interface (GUI), of actual application logic and data management. The application logic was implemented in PHP, HTML [8] and CSS [9]. Animations and so-called autocomplete features of the jQuery library [10] were used to increase the usability by navigating and entering indications. The chosen development strategy does not only allow a high number of users, it provides also a good scalability and maintenance. All user inputs were sent to a web server via an asynchronous HTTP-request. On the server side, PHP scripts generate dynamically HTML pages with the results of the database query and response the results back to the client (Figure 3). The most important requirement for the design of the GUI was the usability, which includes the effectiveness, efficiency and satisfaction. A contribution to the usability makes the simple search form and menu. The used control mechanism in the search form allows just correct entries. Menu is a simple list of categories. Due to ergonomics, the main menu was structured wide instead of deep. Before launching the system, an in-house evaluation was performed.

3. Results

Comprehensive data of ten specialist fields are stored in ALTMEDA. Around 2.700 remedies and more than 4000 synonyms are listed and ordered by their categories. With 9.100 indications and 2.500 additional information ALTMEDA provides much information about complementary and alternative therapies. Conditional on the well-

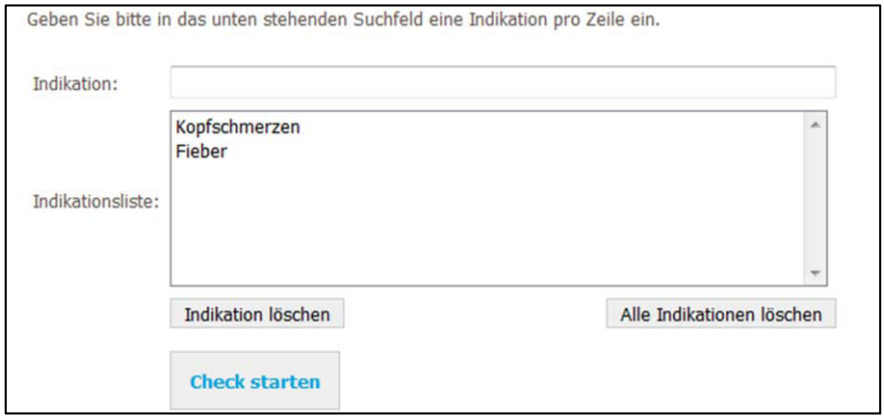


Figure 4. The indications "Kopfschmerzen" (headache) and "Fieber" (fever), as user-specific input in the indication check, where taken

based structure ALTMEDA provides information for an indication and interaction check. Based on the comprehensive resource ALTMEDA, the web-based information system KATIS was implemented. KATIS allows an efficient search for therapies based on ten different specialist fields. The main feature of KATIS is the indication check. Based on the user-specific indications (Figure 4) the system analyzes the knowledge with regard to useful therapies (Figure 5).

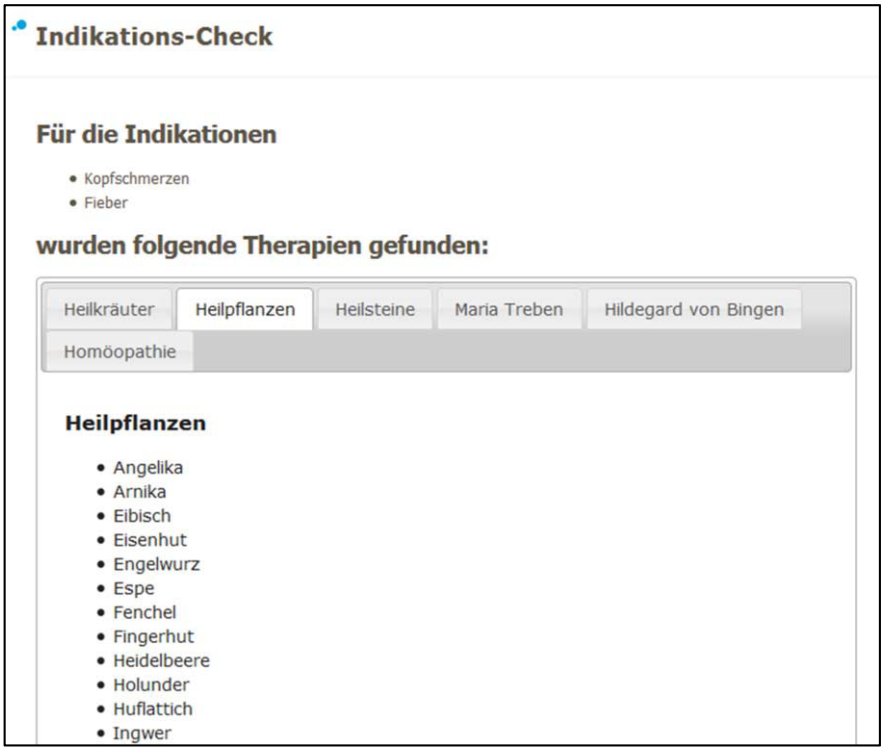


Figure 5. Results of the indication check to the indications "Kopfschmerzen" (headache) and "Fieber" (fever) by the category healing plants (Heilpflanzen)

Furthermore, the Android app ALMEKO for mobile use was also realized based on ALTMEDA, which is available at Google Play Store.

4. Discussion

The new database ALTMEDA represents the integration of data from heterogeneous databases and information systems of complementary medicine. Based on ALTMEDA we implemented the information system KATIS, which combines the knowledge of the complementary medicine and an indication check. The indication check allows user-specific identification of different therapies. Currently, the system is designed for the German market and available only in German. Due to developed data structure, it is possible to internationalize KATIS and make it available in several languages. The integration of the drug database ABDAMED allows more precise structuring of the data with regard to the remedies. Identification and assignment of specific active agents to each remedy are planned to be realized in future projects. Finally, this will allow filtering by using drugs with similar active agents as well as delivering better results by mapping to other databases with drug interactions. That will be a powerful extension of KATIS to identify side effects and drug interactions of remedies.

Availability and requirements

KATIS is available at http://www.komplementäre-medizin.de. To access all features of the system, the browser versions such as Internet Explorer 9+, Mozilla Firefox 23+ and Google Chrome 29+ should first be installed.
ALMEKO is available at Google Play Store (https://play.google.com/store/apps/details?id=de.icancode.almeko&hl=de). To access all features of the app, the Android version 4.1+ should first be installed.

References

[1] ABDA - Bundesvereinigung Deutscher Apothekerverbände, Aponet, www.aponet.de/heilpflanzenlexikon.html, 28.01.2016
[2] A. Shoshi, T. Hoppe, B. Kormeier, V. Ogultarhan, R. Hofestädt, GraphSAW: A web-based system for graphical analysis of drug interactions and side effects using pharmaceutical and molecular data, BMC Medical Informatics and Decision Making (2015),
[3] The PHP Group, PHP: Hypertext Preprocessor, www.php.net, 28.01.2016
[4] Medizinische Medien Informations GmbH MMI, Gelbe Liste Pharmaindex – Arzneimittelinformationssystem, www.mmi.de/gelbe-liste-pharmindex, 28.01.2016
[5] S. Holt, S. Schmiedl, P. Thürmann: Potentially Inappropriate Medications in the Elderly: The PRISCUS List. In: Dtsch Arztebl Int 107 (2010), August, Nr. 31-32, S. 543-551
[6] Werbe- und Vertriebsgesellschaft Deutscher Apotheker mbH, ABDAMED – WuV, http://www.wuv-gmbh.de/abdata-pharma-daten-service/datenangebot/abdamed/, 28.01.2016
[7] Lu,C., Hatsis,P., Berg,C., Lee,F.W. and Balani,S.K. (2008), Prediction of pharmacokinetic drug-drug interactions using human hepatocyte suspension in plasma and cytochrome P450 phenotypic data. II. In vitro-in vivo correlation with ketoconazole. Drug Metab. Dispos., 36, 1255–1260.
[8] World Wide Web Consortium (W3C), W3C HTML, https://www.w3.org/html/, 28.01.2016
[9] World Wide Web Consortium (W3C), Cascading Style Sheets, https://www.w3.org/Style/CSS/, 28.01.2016
[10] jQuery Team, jQuery, http://jquery.com/, 28.01.2016

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-174

Electronic Certification of Death in Slovenia – System Considerations and Development Opportunities

Dalibor STANIMIROVIC[1,a]
[a]*National Institute of Public Health, Slovenia*

Abstract. Background: Accurate and consistent death certification facilitates morbidity and mortality surveillance, and consequently supports evidence-informed health policies. Objectives: The paper initially explores the current death certification practice in Slovenia, and identifies related deficiencies and system inconsistencies. Finally, the paper outlines a conceptualization of ICT-based model of death certification including renovation of business processes and organizational changes. Methods: The research is based on focus group methodology. Structured discussions were conducted with 29 experts from cross-sectional areas related to death certification. Results: Research results imply that effective ICT-based transformation of the existing death certification model should involve a redefinition of functions and relationships between the main actors, as well as a reconfiguration of the technological, organizational, and regulatory elements in the field. Conclusion: The paper provides an insight into the complexities of the death certification and may provide the groundwork for ICT-based transformation of the death certification model in Slovenia.

Keywords. Public Health, Mortality, Vital statistics, Cause of Death, Data Collection, Information Management. Focus Groups.

1. Introduction

Death certification, in various forms and with different purposes, has been one of the essential measures for the monitoring of personal and legal existence since the medieval history [1]. As such, it is still one of the most typical instruments, which allows the analysis of different aspects and causes of death in the modern public health era. Certification of death has been recurrently highlighted as a foundation for monitoring mortality patterns and documenting the leading causes of death, with the results being used to inform health policies and improve prevention strategies [2, 3]. Mortality statistics including the causes of death are essential data to monitor population health, undertake epidemiological studies, and international comparisons [4, 5]. Accordingly, they are widely used and often provide a major source of data for comparing health characteristics between different countries [6, 7]. Clearly in order for anyone to make use of this kind of information, the cause of death must be determined by a qualified person, and reported in a systematic, accurate and consistent way [5, 8]. The only way to obtain good-quality mortality statistics is to have deaths certified by a medically trained and experienced doctor [9]. In order to facilitate high quality death certification and to standardize certification practices among various countries, the United Nations

[1] Corresponding Author: Dalibor Stanimirovic, National Institute of Public Health, Trubarjeva 2, 1000 Ljubljana, Slovenia, dalibor.stanimirovic@nijz.si

(UN) and World Health Organization (WHO) periodically develop protocols and guidelines for the management of death certification procedures [10]. To ensure that doctors are able to competently certify deaths in accordance with WHO guidelines and standards (International Statistical Classification of Diseases and Related Health Problems – ICD) [11], they must receive intricate training in death certification [12-14].

Throughout the European Union (EU), as in most other parts of the world, the completion of a death certificate is a mandatory requirement of the doctor or other qualified individual. However, the accuracy of death certification has often been inadequate and fragmented even in EU, despite the fact that EU member states are generally thought to have the most complete and accurate data records of all regions [4], [15]. The situation is yet more serious in some developing countries, whereas death registration may have a number of flaws, including not only miss-diagnosis of causes of death, but also underreporting [2, 16]. Factors contributing to deficient registration of death and direct sources of fault have been generally identified [17, 18]: diagnostic errors, late registration, missing information, coding errors, unavailability of medical records, and difficulties in ascertaining causal sequence of events leading to death. In order to overcome these deficiencies, European Commission has repeatedly expressed the importance of high-quality mortality data and methods that improve the international comparability of mortality statistics [19]. In the last years, electronic certification of death (eDeath certification) is becoming an issue of great interest concerning public health surveillance in all EU member states [20]. Due to objective problems referred to above and other circumstances related to the access to the deceased's medical record, speed of data transmission, the confidentiality of personal data, faster execution of business processes, data archiving and standardization of death certification practices, there are a number of national initiatives in the field, which support the establishment of ICT-based solution for the certification of death [12, 21].

Slovenia does not have a comprehensive national health information system including the electronic death certification solutions. The current death certification practice in Slovenia, somewhat unchanged over the last decades, prevents the cross-functional integration of processes and considerably inhibits the utilization of the stakeholders' organizational capabilities [22]. The aim of this paper is to examine the existing situation concerning the death certification model in Slovenia, and explore the possibilities regarding its transformation through innovative and comprehensive informatization, which has become a general trend in the Slovenian health sector, particularly strong within the recent efforts to implement various eHealth solutions. Informatization in the case of death certification implies the construction of the specialized web-based information system (including application, infrastructure and mobile devices, and access to the deceased's medical documentation), reengineering and streamlining of the business process, and the redefinition of the business model in the field. Accordingly, this paper primarily focuses on the following interrelated research objectives:

- An overview and analysis of the death certification practice in Slovenia and identification of related deficiencies and system inconsistencies.
- Conceptualization of ICT-based model of death certification including renovation of business processes and organizational changes.

2. Methods

This paper employs a focus group methodology to investigate the research questions. Controlled and structured focus groups were used as the main data collection technique during the entire research process. A selection of the research method was adapted to the particularities of the research problem [23] 24]; given the fact that quantitative empirical research would not yield credible results. Namely, the death certification system itself is exceptionally complex, because a number of institutions and professional groups are involved in the intertwined and often deficiently defined processes. These agencies include doctors and/or coroners, government bodies, the police, judicial system, the local authorities, public health institutes, and the national statistics agencies. It should be stressed, that any errors and irregularities in the death certification process are due to objective reasons (burial, cremation) very difficult or almost impossible to detect.

2.1. Sample

Selection of the focus groups participants was based primarily on their expertise and experience, which was supposed to ensure credibility of their views and facilitate constructive participation in the study. A non-random stratified sampling approach was used to ensure a representative sample of the health care experts that satisfy the required conditions. The final sample size comprised 29 experts from the cross-sectional areas strongly related to different aspects of death certification. Participants were affiliated with different institutions: Institute of Forensic Medicine (2 participants), National Institute of Public Health (7 participants), Ministry of Health (1 participant), Ministry of the Interior (5 participants), coroners form different health care providers (12 participants) and ICT companies (2 participants). The participants were qualified health care professionals (general practitioners and specialists), ICT experts from the health care and private sector (consultants and analysts), health care managers (managers of public health care institutions), experts in public health (analysts and statisticians), and senior government officials. Quotas of experts in each area were determined after reaching saturation point. The sample was relatively equally distributed in terms of gender, as 45 % of the sample was male (13 participants), and 55 % was female (16 participants). The participants were aged between 40–60 years.

2.2. Data Collection and Analysis

The final goals of the focus group sessions were revised with the participants in line with their comments and suggestions, which helped to resolve some conceptual ambiguities. Focus group sessions were conducted in the period from January 2014 to December 2015. They lasted approximately 90 to 120 minutes and were held at the premises of National Institute of Public Health. All participants were explained the purpose and objectives of the study in order to clarify the last details and potential uncertainties pertaining to their assignments. All focus group participants were provided anonymity and assured confidentiality of the information obtained. Discussions and responses of the focus group participants were recorded in writing.

Conceptualization of the ICT-based death certification model was conducted in collaboration with the participating experts, who took a constructive part throughout the study. The role of the participating experts was twofold. First, they had to participate in the analysis of the existing death certification practice, and second, they had to provide

their vision of the transformation, and propose a conceptual design of the ICT-based death certification model.

3. Results

The occurrence of death generates a very complicated course of action that triggers activities in different institutional subsystems. Accordingly, death certification concept can be observed and analyzed from different perspectives, which are often complex and intertwined with each other.

3.1. Legal and administrative issues

Certification of death is legally regulated from three perspectives; medical, administrative and general perspective. Despite relatively extensive legislation in the field, several inconsistencies concerning the non-compliance were revealed:

- Non-compliance with the existing legislation in the field (especially the Rules on the conditions and method for the performance of post mortem examination service (OG RS 56/93 and 15/08)) [25]
- Specific authorizations and licenses to the coroners have not been issued
- Divergence between legislation and practice can be frequently detected
- The normative acts do not foresee the possibilities for eDeath certification
- Evaluation of doctors' workload has not been adequately defined

3.2. Education and training issues

Coronary service in Slovenia is exclusively performed by the medical doctors. During formal education on medical faculties, the medicine students learn the basics of forensic medicine. Afterwards, in the course of compulsory residency, the doctors gain knowledge of data collecting procedures and essentials of death certification. Therefore, despite the presumable lack of specific knowledge, contents of education and training programs for post mortem examination service have not been defined, and consequently, education and training of the doctors has not been implemented.

3.3. Organizational issues

All organizational aspects of the post mortem examination service are defined in the Rules on the conditions and method for the performance of post mortem examination service (OG RS 56/93 and 15/08) [25]. These Rules regulate the responsibilities, organizational structure, functions of the coroners, necessary documentation, autopsy and burial, and other technical issues relating to the post mortem examination service. Although rather clear and strict provisions, there are several discrepancies in practice:

- A network of post mortem examination services has not been established
- Official post mortem examination services are available only in large medical centers (services in rural areas are mainly carried out by family doctors)
- Inconsistent practices (between municipalities and large medical centers)

3.4. Informational issues

Existing database on the deceased includes verified data on all circumstances and causes of death, both physiological and external. The collecting of data is a manual process, which has not changed a lot over the years. These data are additionally supplemented with socio-economic variables using secondary sources from the Statistical Office of the Republic of Slovenia (SORS). Mortality data is collected from Administrative units, which regularly send in "Notifications on death" including annexed form "Medical death certificate and report on the cause of death". This form is completed in the field by the coroner and used for determining the underlying cause of death. Death certification consists of rather complex sequence of events and is carried out by using paper-based forms. Insight into the current data and information process uncovered significant inconsistencies and shortcomings:

- The slow data and information flow between the stakeholders and the absence of fast and effective control of the data entered (paper-based forms)
- Rights of access to information are vaguely defined (unauthorized access)
- Certain amount of data collected is doubled (different forms – the same data)
- Inaccessibility of data from the medical records of the deceased
- Inadequate management of related documentation

The issues revealed are manifested in several limitations of the current death certification practice and have a significant impact on the business model in the field. Death certification system is consequently not optimally organized and structured; including paper-based manual data entries during the process, inadequate legal regulation and non-compliance, unresolved issues concerning the licensing, status and jurisdiction of the coroners, and unsecured and untimely information flows between the main stakeholders.

3.5. Conceptualization of ICT-based model of death certification

Deriving from the research findings, we have been striving for the conceptualization of an ICT-based death certification model. The proposed conceptual solution addresses the identified issues directly through the ICT-based solution for eDeath certification, and indirectly through the promotion of guidelines for the national implementation of eDeath certification, including the regulatory amendments, organizational and educational changes, and business process reengineering. Hence, conceptual proposal for eDeath certification exceeds the current incapacities concerning the data/informational aspect and ensures better speed of data flows, protection of personal data, access to deceased's medical documentation, higher quality and control of data, adequate documentation and archiving of data, and unification of national databases and registries. The conceptual solution for eDeath certification is based on more transparent data flow; sequence of events is presented in Table 1. The graphical sketch of the conceptual solution displays the main actions within the eDeath certification process and the main actors including their inherent health care / administrative functions (Figure 1). Conceptual solution suggests the redefinition of the functions and relationships of and between the main actors within the death certification field, and the reconfiguration of the technological, organizational, and regulatory elements according to declared public health objectives.

At the same time, the conceptual solution includes and promotes reengineered, optimized and streamlined business processes and considerable organizational changes

Table 1. Sequence of events within the eDeath certification process

Sequence	Event
1.	Occurrence of death;
2.	Post mortem examination – CERTIFIED CORONER with professional card/authorization;
3.	Online notification on death to CPR (administrative part of death certification);
4.	Access/insight to all medical documentation (electronic health record), consultations with treating/personal doctor, etc.;
5.	Decision: autopsy YES/autopsy NO;
6.	Autopsy NO – recording all necessary diagnoses for "Medical death certificate and report on the cause of death" form;
7.	"Medical death certificate and report on the cause of death" form "waits" for the confirmation of final identification of the deceased by CPR;
8.	After confirmation "Medical death certificate and report on the cause of death" form is handled by the National Institute of Public Health;
9.	Autopsy YES – "Medical death certificate and report on the cause of death" form "waits" for the confirmation of final identification of the deceased by CPR and for autopsy diagnoses;
10.	After obtaining autopsy report, the coroner finalizes filling in "Medical death certificate and report on the cause of death" form;
11.	Finalized "Medical death certificate and report on the cause of death" form is sent to the National Institute of Public Health.

(certification of coroners, education and training, network of coronary services, etc.), dependent on the preliminary amendments of the regulatory framework.

The proposed solution assumes the complete transformation of the existing business model, including the translation of existing paper based forms into the web-based application. Use of the web-based application for eDeath certification presumes that all transactions of the certified coroners will be signed with a qualified digital certificate, thus providing a higher level of data protection throughout the entire process of data

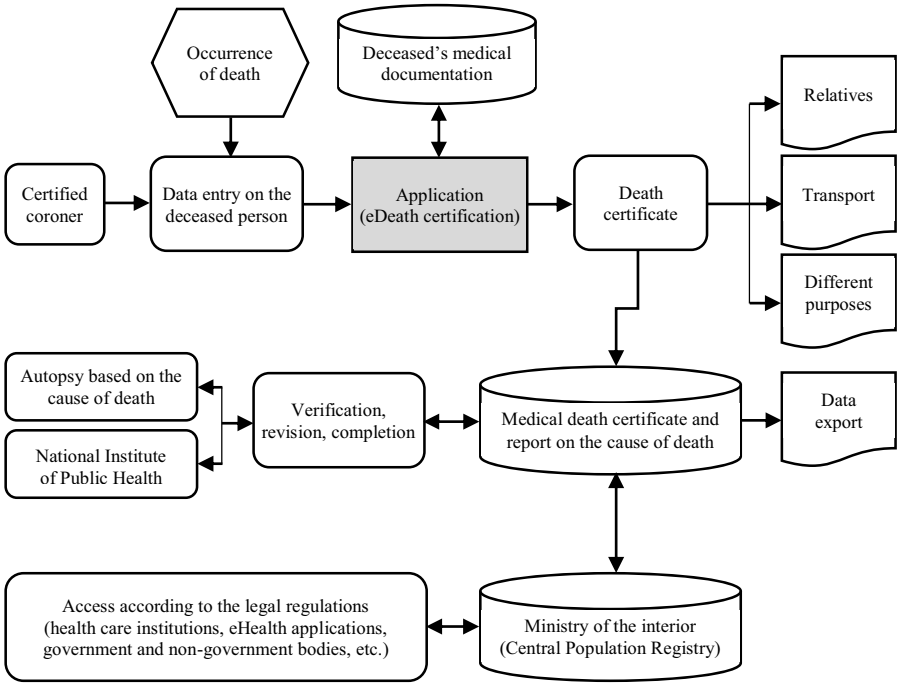


Figure 1. Conceptual design of ICT-based model of death certification.

transmission. In the first phase, while establishing the identification of the deceased, the process of data transmission will be unified, in the later phases the process of data transmission will bifurcate into two distinctive parts, namely health care part and administrative part. This functionality will prevent unauthorized access of administrative workers to personal health data and provide opportunities to health professionals to complement the data on the deceased based on the final conclusions and lab results. The presented ICT-based solution for eDeath certification has already been developed, whereas the referred systemic measures and institutional actions are still in the initial stage. Currently, talks are under way with major decision-makers and other stakeholders to support the preparation of the national guidelines in the field, and promote gradual introduction of eDeath certification. Obviously, prior to potential national implementation of eDeath certification, political and professional consensus on the highest level would have to be reached regarding the several open issues stressed in the study, especially amendments of the regulatory provisions, professionalization and certification of the coronary services, and related expenses.

4. Discussion

The presented research findings imply that effective transformation of the death certification practice should consider the wide array of factors mainly from the health care and administrative ecosystems. Besides the large number of interconnected actors who play an active role within the death certification process, necessary modifications should ponder the importance of the concept of public health, and the rather distinctive and restrictive characteristics of the health care system environment.

The main methodological dilemmas and limitations of the study relate to the somewhat arbitrary identification of the potential effects of informatisation, and especially the following conceptualization of the transformed death certification model, which was, due to obvious reasons, hypothesized without prior pilot testing and practical validation in the health care environment. Accordingly, the issues related to the conceptualization of eDeath certification model raise some important questions of principle, while the research outcomes may therefore be arguable. These issues should be properly resolved in further research and successive experiments aimed at comprehensive transformation of the obsolete death certification model.

The research presented does not seek to impose a ‘one-size-fits-all’ solution for the myriad of problems related to development and implementation of eDeath certification. However, the presented study provides a valuable insight into the complex dynamics of the existing death certification practice in Slovenia and may provide the groundwork for further developments in this area. Notwithstanding the intricacies revealed, transformation of the death certification model in Slovenia, and possibly elsewhere, certainly represents a development opportunity which can, subject to proficient coordination with other ecosystem factors, ensure better utilization of public health resources and provide tangible public health benefits.

Acknowledgement

The project e-Certification of Causes of Death was funded by the European Commission - Eurostat, under the Grant Agreement Number 07154.2013.002-2013.629.

References

[1] F. Sim, M. McKee, *Issues in Public Health (2 ed.)*. Open University Press, Oxford, 2011.
[2] A.M. Sibai, Mortality certification and cause-of-death reporting in developing countries, *Bulletin of the World Health Organization* **82**(2)(2004), 83–83.
[3] E.H. Burger, P. Groenewald, A. Rossouw, D. Bradshaw, Medical certification of death in South Africa-moving forward, *SAMJ: South African Medical Journal* **105**(1)(2015), 27–30.
[4] C.D. Mathers, D. Ma Fat, M. Inoue, C. Rao, A.D. Lopez, Counting the dead and what they died from: an assessment of the global status of cause of death data, *Bulletin of the World Health Organization* **83**(33)(2005), 171–177.
[5] J. Cohen, J. Bilsen, G. Miccinesi, R. Löfmark, J. Addington-Hall, S. Kaasa, M. Norup, G. van der Wal, L. Deliens, Using death certificate data to study place of death in 9 European countries: opportunities and weaknesses, *BMC Public Health* **7** (2007), 283–283.
[6] European Commission – EC, *Quality and comparability improvement of European causes of death statistics*, DG Sanco, Eurostat EDC DGV/F3 SOC 98 20108 – INSERM SC8/Cepidc, Brussels, 2001.
[7] N.M.A. Ali, R.R. Hamadeh, Improving the Accuracy of Death Certification among Secondary Care Physicians, *Bahrain Medical Bulletin* **35**(2)(2013), 1–6.
[8] M. D'Amico, E. Agozzino, A. Biagino, A. Simonetti, P. Marinelli, Ill -defined and multiple causes on death certificates – A study of misclassification in mortality statistics, *European Journal of Epidemiology* **15**(2)(1999), 141–148.
[9] T.H. Lu, Using ACME (Automatic Classification of Medical Entry) software to monitor and improve the quality of cause of death statistics, *Journal of Epidemiology and Community Health* **5**(6)(2003), 470–471.
[10] World Health Organization – WHO, *Strengthening civil registration and vital statistics for births, deaths and causes of death*, *Resource kit*, World Health Organization, Geneva, 2013.
[11] World Health Organization – WHO, *International Statistical Classification of Diseases and Related Health Problems – ICD, 10th Revision*, World Health Organization, Geneva, 1992.
[12] E.G. Brooks, K.D. Reed, Principles and Pitfalls: a Guide to Death Certification, *Clin Med Res* **13**(2)(2015), 74–82.
[13] N. Crowcroft, A. Majeed, Improving the certification of death and the usefulness of routine mortality statistics, *Clinical Medicine* **1**(2)(2001), 122–125.
[14] G. Maudsley, L. Williams, Death certification — a sad state of affairs, *Journal of Public Health Medicine* **16** (1994), 370–381.
[15] B. Swift, K. West, Death certification: an audit of practice entering the 21st century, *Journal of Clinical Pathology* **55**(4)(2002), 275–279.
[16] K. Jain, D.V. Bala, K. Trivedi, H. Chandwani, Situational analysis of Medical Certification of Cause of Death (MCCD) scheme in Municipal Corporation of Ahmedabad, *Indian Journal of Forensic and Community Medicine* **2**(2)(2015), 95–99.
[17] S.K. Dash, B.K. Behera, S. Patro, Accuracy in certification of cause of death in a tertiary care hospital – A retrospective analysis, *Journal of Forensic and Legal Medicine* **24** (2014), 33–36.
[18] M. Wall, J. Huang, J. Oswald, D. McCullen, Factors associated with reporting multiple causes of Death, *BMC Medical Research Methodology* **5**(1)(2005), 1–4.
[19] E. Jougla, F. Rossolin, A. Niyosenga, J. Chappert, L. Johansson, G. Pavillon, *Comparability and quality improvement in European causes of death statistics*, Eurostat, Luxembourg, 2001.
[20] European Commission – Eurostat, *Report of the Task Force on Satellite Lists for Causes of Deaths (COD)*, Luxembourg, 2-3 December, 2013.
[21] D. Lefeuvre, G. Pavillon, A. Aouba, A. Lamarche-Vadel, A. Fouillet, E. Jougla, G. Rey, Quality comparison of electronic versus paper death certificates in France, 2010, *Population health metrics* **12**(1)(2014), 1–9.
[22] D. Stanimirovic, Framework for Information and Communication Technology Induced Transformation of the Healthcare Business Model in Slovenia, *Journal of Global Information Technology Management* **18**(1)(2015), 29–47.
[23] M. Patton, *Qualitative evaluation and research methods*, Sage Publications, Thousand Oaks, 1990.
[24] R. Yin, *Case study research: design and methods (4th ed.)*, Sage Publications, Thousand Oaks, 2009.
[25] Ministry of Health, *Rules on the conditions and method for the performance of post mortem examination service (OG RS 56/93 and 15/2008),* Government of the Republic of Slovenia, Ljubljana.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-182

A Process Model for IT Migrations in the Context of a Hospital Merger – Results from an Austrian Case Study

Katharina STEININGER[a,1], Birgit KEMPINGER[a],
Stefan SCHIFFER[b], Gustav POMBERGER[b]
[a]*Department of Business Informatics – Information Engineering*
[b]*Department of Business Informatics – Software Engineering*
Johannes Kepler University Linz

Abstract. In 2016, a new university hospital merged from three former independent Austrian hospitals started its operation. This paper presents a process model developed to coordinate the IT migration after the merger, using five phases to meet the requirements of the specific setting. A methodological mix of interviews, surveys and workshops was applied during the IT migration process. High stakeholder participation and a transparent methodical approach led to a broad agreement on success factors, migration objectives, and evaluation results. Thus, acceptance for the finally selected migration scenario was very high among employees, which is known to be crucial for the success of migration projects.

Keywords. Hospital merger, IT migration, hospital information system, migration scenario.

1. Introduction and Background

1.1. Theoretical Background

In general, several conditions can cause the need for software or information system migration projects throughout various industries and enterprises. Most important are the obsolescence of a technology, the pressure of users to modernize the information infrastructure, or the need to build a single coherent information system after a company merger [1]. Mergers are usually dominated by legal and financial analyses and negotiations, while strategic analysis and planning of holistic organizational integrations are often neglected. Nevertheless, it is shown that substantial holistic analytical and planning activities are important for the long-term success of a merged institution [2]. In fact, studies have shown that among other organizational issues, a poor definition of the new corporate information system and its infrastructure requirements as well as a reluctance to determine the objectives and conditions of the integration process in advance are a significant cause for poor ex-poste performance [3]. Moreover, the IT integration itself is found to be critical for the merger success and improvements in merger planning can often be achieved by including IT staff in pre-merger planning activities [4]. These findings are also applicable in hospital merger settings, where the

[1] Corresponding Author: Katharina Steininger, Johannes Kepler University Linz, Altenbergerstr. 69, 4040 Linz, E-Mail: katharina.steininger@jku.at.

migration of hospital information systems (HIS) and consequently the electronic health record is crucial for the success of the future hospital operation [5].
In fact, only a uniform HIS can ensure that the new organization can exploit the full potential of its high level of IT support and process automation. Previous research shows that HIS provide major benefits in patient care, such as improvements in reporting, organizing and locating clinical information [6] as well as clinical decision support [7], coordination and management of patient care [8], and patient safety [9]. These benefits can only be achieved if the hospital has a coherent and universal HIS in place. In addition, standardized, valid and comprehensive clinical data in hospital information systems are a crucial source for clinical and health outcome research [10].
Even though the benefits of IT integration after a merger are transparent, little is known on how to set up the migration process in a hospital setting in order to ensure an effective IT integration of the HIS. Studies show, that implementing a general process model following the phases of (i) Planning, (ii) Implementation/Integration, and (iii) Review/Evaluation has been effective to structure post-merger IT integration [3].

1.2. Project Background

As a political response to the immanent shortage of medical doctors in Europe in general and in Austria in particular, the Austrian national and Upper Austrian state government decided on the founding of a fourth national school of medicine. This school of medicine was established as part of a newly founded medical faculty at the existing Johannes Kepler University in Linz (JKU) in October 2014.

In Austria, a school of medicine needs to cooperate with a university hospital, which is characterized as a hospital that serves as a whole or in part as research and teaching institution for a medical university or medical faculty. Since Austrian law provides that medical schools and universities can be associated to only one corresponding university hospital [11], state and local governments agreed on merging three formerly independent state and community hospitals belonging to two individual hospital organizations to form a sufficiently large and diverse hospital institution in order to satisfy the requirements of the university in terms of research and teaching resources. Therefore, the Linz General Hospital (AKh), the Women's and Children's State Hospital (LFKK) and the Psychiatric and Neurologic State Hospital Wagner Jauregg (LNK-WJ) were merged to form the Kepler University Hospital (KUK). Table 1 presents several indicators to demonstrate the project size of the merger. When starting its official operation in January 2016, the KUK was Austria's second largest hospital (1,825 beds) with the Vienna General Hospital being the largest (1,990 beds) regarding the indicator "number of beds" [12].

Table 1. Project size of the KUK hospital merger (* in million Euro)

Indicator	AKh	LFKK	LNK-WJ	total
Legal ownership	City of Linz	State of Upper Austria	State of Upper Austria	
Number of beds	886	270	669	1,825
Stationary stays	60,000	19,566	17,894	97,440
Outpatients	275,000	21,107	18,917	315,024
Total spending*	230.6	106.5	143	480.1
Staff	2,838	1,142	1,838	5,817

In addition to the challenges arising from the merger of the different legal and cultural environments of the three hospitals as well as the requirement of adding teaching

and research tasks in many departments that were until then oriented mainly to clinical care, the hospital management had to face the fact that two completely different Hospital Information Systems (HIS) were in place in the institutions. Therefore, the migration of the existing HIS systems was a central challenge within the merger project.

In the present case, a committee consisting of members from local politics, the university, the three hospitals, and legal as well as healthcare experts was established to accompany the project in a strategic and holistic way. The objective of the merger was to create a new hospital, so in terms of the strategic goals from an IT perspective, this objective was adopted in a very early state. Therefore, the main objective of the IT migration was defined as follows: *The KUK will offer an highly effective IT support for all users working in clinical care, administration, research, and teaching, and this highly effective IT support will be realized through the merger and standardization of information and communication systems from the three former hospitals.*

Since time and cost constraints of the merger and IT migration project were accordingly tight, an accurate and specially adapted process model for the migration of the HIS of the KUK was needed. Existing studies on process models for IT migration projects are often focusing rather abstractly on high-level process phases without explaining details on specific methodologies or milestones [3]. Others tend to concentrate on detailed technical procedures [13]. Therefore, the research goal was to develop and apply a customized but still generic process model for IT migrations after hospital mergers, including the selection process of an appropriate future HIS.

2. Process Model and Applied Methods

Evidence shows that in terms of the migration process, several factors are critical in order to achieve a successful migration outcome. Beside thorough IS planning, positive support by the management, and high-quality communication to end-users, also high level of end-user involvement in strategic planning during the process is crucial [14].

In order to address these findings as well as the specific demands and needs of the present situation within the three hospitals, a customized process model was developed by the authors specifically for the current IT migration. When designing the process model, it was essential to describe the phases and methods in a generic way so that the model can be applied to other migration projects in similar settings. The model consists of five phases:

- Phase 1: Definition of migration objectives, identification of basic conditions
- Phase 2: Evaluation of the IT infrastructure in the former individual hospitals
- Phase 3: Development of migration scenarios, selection of one scenario
- Phase 4: Design of a project plan, installing of operational teams
- Phase 5: Implementation of the IT migration

The migration and research project started in January 2014. At the time the present paper was written, phase 1-3 were successfully completed. Furthermore, phase 4 and 5 focus on standardized project planning and realization procedures. Therefore, in the following the paper focuses on the presentation of applied methods, findings and lessons learned from the first three phases.

2.1. Phase 1: Definition of migration objectives, identification of basic conditions

The main objective of phase 1 was to set the basic framework for the following migration process. With regards to the findings of Robbins and Stylianou [14] concerning critical factors identified for successful migrations, end-user involvement as well as end-user communication and support by the hospital management were thoroughly pursued in order to defining overall IT objectives for the migration process as well as identifying basic migration conditions. Starting point for the investigations was the identification of HIS key factors for the success of IT support along clinical care as well as clinical research processes.

To reach the objectives above, phase 1 was structured into seven process steps applying the methods interview, survey, and world café as shown in Figure 1.

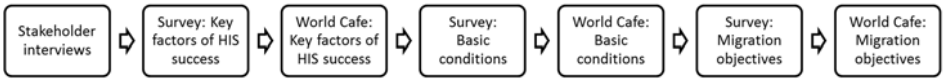


Figure 1. Process Steps of Phase 1.

Initially, stakeholder from various professional groups within the three hospitals were identified and divided into the following stakeholder groups: (i) clinical care and hospital management, (ii) research and teaching, and (iii) IT provider. Then, a total of 35 structured interviews were conducted with members of all three stakeholder groups. The aim of the interviews was to get a broad understanding of expectations and concerns regarding the IT migration process, of perceived key factors influencing the success of the future HIS, and of subjective IT migration objectives. In total, more than 50 hours of interviews were conducted, resulting in over 280 pages of transcripts.

A content analysis was conducted to identify a consolidated list of perceived key factors for the IT success in the new hospital, basic conditions for the IT migration as well as potential IT migration objectives. These findings form the basis for the subsequent surveys and workshops.

In three rounds each starting with a survey followed by a World Cafe workshop 37 members of all stakeholder groups were invited to evaluate, rate, and comment on the findings of the initial interviews. User participation was high (survey response rate between 89.9 and 94.6 percent) and the working atmosphere during the workshops was very constructive with management and clinical care personnel working in partnership of equals when agreeing on the main key factors, conditions and objectives.

In phase 1, a total of 52 key factors for HIS success, 35 basic conditions for the IT migration process and 22 IT migration objectives were identified, defined, rated and commented.

2.2. Phase 2: Evaluation of the IT infrastructure in the former individual hospitals

The aim of phase 2 was to determine the gap between the IT migration objectives defined in phase 1 and the current IT status within the three hospitals. The results of phase 2 were an important input for the development of detailed migration scenarios and subsequently a project plan. The analyses of the IT landscapes and IT infrastructure within the three hospitals were conducted by independent healthcare analytics as well as healthcare IT consultant companies and accompanied by the research team. First, an internationally

accepted score measuring the maturity of the existing HIS in all three hospitals was determined [15]. Second, more detailed analyses of the functionalities and technical abilities of both HIS were performed. The areas analyzed and categorized were based on an IT capability model developed by the consulting company referring to the following domains: (i) strategic alignment of IT, (ii) IT governance, (iii) architecture management, (iv) solution development, (v) service management and operation, (vi) IT security, (vii) human resource and knowledge management.

Finally, clinical and administrative processes commonly supported by a HIS were analyzed, including administration, planning, care support, as well as research and teaching. Finally, the degree to which each IT migration objective defined in phase 1 is already fulfilled in each hospital was determined. In total, 49 interviews with members of various professional groups (IT decision maker, IT provider and IT staff, doctors, and nurses) within the three hospitals were conducted between November 2014 and February 2015.

2.3. Phase 3: Development of migration scenarios, selection of one scenario

Phase 3 was designed to derive various migration scenarios based on the findings of phase 1 and 2, evaluate the scenarios in terms of defined criteria and select the best suitable migration scenario as a starting point for the development of a project plan.

A migration strategy team of 13 people from the following stakeholder groups was formed to complete the tasks planned for phase 3: (i) two strategic IT representatives from both former hospital organizations serving as interim migration managers, (ii) two representatives from medical informatics, (iii) two representatives from the IT provider each servicing one of the two existing hospital systems, (iv) two external IT migration consultants, (v) one representative from the university rectorate, (vi) one external IT manager from a university hospital in Germany, (vii) three members of the research team. The group was responsible for gathering information from the hospitals and the HIS market, involving potential HIS vendors, defining the conditions for the migration scenarios, evaluating the scenarios and finally presenting the findings to the hospital management so they could make a decision on the future HIS as well as the migration process. Due to financial, legal and time restrictions it soon became clear that only one of the two already implemented HIS would be suitable to be extended to the whole university hospital, therefore only two HIS vendors were involved in phase 3.

The development, evaluation and rating of migration scenarios is an extensive and complex task that requires in depth focus and a multi-level approach by a critical amount of strategic key stakeholder and experts that are difficult to be brought together during a time-consuming migration project. Thus, the research team designed a five days intensive workshop bringing together the migration strategy team and up to 20 additional specialists. The first three days were used to further develop and specify two migration scenarios that had been prepared by the HIS vendors, each based on the enrollment and enhancement of one of the two existing HIS for the whole university hospital. On the last two days the members of the migration strategy team developed evaluation criteria for the scenarios using the card brainstorming technique [16]. The initial amount of more than 80 criteria was reduced by removing redundancies and clustering evaluation topics. The remaining 30 criteria were then categorized into three dimensions: HIS evaluation, vendor evaluation and scenario evaluation. Members of the migration strategy team as well as nine until then not involved clinical care representatives from the three hospitals were asked to evaluate the two scenarios with regards to the criteria. Each rater received

90 virtual points that could be assigned to a criteria-scenario-combination being completely free of leaving several criteria unrated. Finally, the external experts, the interim migration manager and the hospital IT provider each conducted SWOT (strengths, weaknesses, opportunities, threats) analyses that were combined to create a transparent picture of the advantages and disadvantages of the two migration scenarios for the hospital management.

3. Results

The main objective in phase 1 was to set the fundamental framework to all following IT migration phases. For this purpose, 35 structured interviews with stakeholders from clinical care and hospital management, research and teaching staff, and IT provider were conducted and analyzed. The content analyses led to 35 basic conditions (e.g. "*Data protection and data security have at least to be sustained at its current level"* or "*Time and financial resources for training have to be provided"*) as well as to 52 key factors of IT success, which were categorized (human/process/system) and evaluated using two dimensions (importance and urgency) on a 1-5 scale (1="I do not agree at all"; 5="I completely agree"; see Table 2).

Table 2. Top 10 key factors sorted by importance (n=34, values are mean values)

Key Factor of IT Success	Category	Importance	Urgency
Sufficient financial resources	process	4,90	4,67
Qualified IT-personal	human	4,81	4,65
Reliability of the IT-systems	system	4,79	4,5
Clearly defined IT-strategy	human	4,79	4,43
Sufficient personal resources	human	4,74	4,63
Trained staff	human	4,73	4,42
Uniform system for patient administration	process	4,67	4,50
Process orientation in IT	process	4,66	4,18
Human success factor	human	4,64	4,36
Maximum risk reduction	process	4,63	4,53

Interview content analyses summarizing and classifying the statements along with further categorization of key factors led to the identification of 22 IT migration objectives. These IT migration objectives were presented to the stakeholders in order to ask for their consent (see Table 3 for selected migration objectives ranked by degree of consent, labeled with date of achievement).

Table 3. Selected key factors and IT migration objectives (*n=33)

Consent (n=34)	IT Migration Objectives	To achieve
100.00	High data security	anytime
100.00*	High reliability in patient related IT systems	anytime
100.00	Informed and IT-trained staff	until 01.01.16
100.00	Realization of IT synergy effects	after 01.01.16
100.00	Uniform HIS	after 01.01.16
97.06	Open IT infrastructure	until 01.01.16
91.18	Investment in sustainable technologies	after 01.01.16
88.24	Use of mobile technologies	after 01.01.16
87.88*	Paperless hospital	after 01.01.16
85.29	Outstanding IT systems for research, teaching and care	after 01.01.16

To determine the gap between the 22 IT migration objectives identified in phase 1 and the current maturity of the IT infrastructure of the three hospitals, a holistic IT evaluation was conducted in phase 2. One aspect of this evaluation was the determination of an HIS maturity level (scale from 0 to 7, with 7 being the highest level [15]). Reaching maturity levels exceeding European average (3.0 in 2015 [17]). For the next level, full clinical decision support and closed loop medication would be needed. Detailed IT evaluations as well as the degree to which the IT migration objectives were already achieved, showed that all three hospitals were at a similar maturity level, providing a solid IT basis for the University hospital.

Table 4. Selection of evaluation criteria sorted by according category (scenario, vendor, HIS)

Evaluation Criteria	Scenario	Vendor	HIS
Controllability of risks	*		
Consideration of organizational and personnel restraints	*		
Customer orientation		*	
Innovative power		*	
Professional and personnel competence			*
Degree of coverage of clinical care processes			*
Enhancing opportunities for cooperation with university hospitals			*
Sustainable investment in terms of technology			*
IT support of research & teaching tasks			*

Based on the results from phase 1 and 2, two IT migration scenarios were developed and evaluated by a migration strategy team formed in phase 3 using a set of 30 criteria to evaluate the scenarios. Each member of the migration strategy team received 90 virtual points that could be split up and assigned to any criteria-scenario combination. The criteria were developed during an intensive workshop and applied to the migration scenarios anonymously, resulting in a strong preference for one migration scenario (62% vs. 38%, see Table 4 for a selection of evaluation criteria). The same evaluation procedure was conducted with nine clinical care representatives from the three hospitals, leading to nearly identical results. Interestingly, criteria from the category HIS had the strongest influence on the results.

Finally, selected participants of the intensive workshop conducted SWOT analyses evaluating the scenarios in detail resulting in various assessments. Naturally, IT manager and IT provider rather supported the HIS already in place in their own institution, while external experts had a clear primary preference for the scenario that was already preferred in the criteria-based evaluation. The decision on the migration scenario made by the hospital management was based on the findings presented above.

4. Discussion

Because of tight time and cost constraints within the presented migration project, a customized process model was developed and presented in this paper. Three out of five phases had been successfully completed by the time the paper was written.

One of the main principles applied in the present model was the high degree of user integration. In phase 1, various stakeholders were interviewed and surveyed in order to identify key success factors, basic conditions and IT migration objectives. In phase 2, evaluations were conducted interviewing a large amount of process owners and system experts within the institutions. All findings were discussed intensively with

representatives from various professional groups in all three hospitals, which lead to a broad acceptance throughout all phases and finally the final decision on the migration scenario. This acceptance was necessary in order to avoid a perceived situation of winners and losers because of an imminent system change. The authors claim that this acceptance could not have been easily achieved without following a transparent and involving process model like presented. As Ahmad et al. [18] show, the appropriate involvement of various stakeholders is not only important in innovation processes but also in decision making processes and conclude that the involvement of stakeholders at an early stage can lead to decisions compatible with structural and cultural contexts. Involving stakeholders throughout the whole migration process is crucial for a sustainable success of migration projects. Not less than five key factors assigned to the category *human* are found among the top 10 most important key factors (Table 2). In comparison, only one key factor assigned to the category *system* can be found in the top 10. These findings also correspond to Bowns et al. [19], who relate most problems in the implementation of information management to human rather than technical factors. Therefore, it is crucial to address and involve stakeholders in IT migration projects in an early, decision making phase.

The extensive willingness of more than 40 stakeholders to participate in the project made it possible to create a good and productive working environment under the guidance of the research group implementing the present process model. Working together in such a heterogeneous group was not easily achieved because of more than 40 stakeholders from different professional groups, different hierarchic levels and different organization with even divergent interests from time to time.

Equally challenging like creating a productive and goal oriented atmosphere was the realization of the five days intensive workshop in phase 3 in order to develop and evaluate two IT migration scenarios. In addition to the migration strategy team representatives of the two HIS vendors were present and willing to work together in an unexpectedly constructive and respectful manner. It can be assumed that the setup of the intensive workshop was indispensable for successfully concluding phase 3 and for preparing a solid scenario decision for the hospital management accepted and supported by stakeholders.

One of the biggest problems was the fact that the new organization had not been legally founded at the time main decisions had to be made. The absence of decision-making competencies in single hospitals as well as no defined future IT provider led to temporary problems with the motivation among the staff and consequently to delays in milestone achievement. Further difficulties were caused by uncertainties concerning available financial and personnel resources. Having these conditions undetermined, technical decisions can only be based on assumptions and tend to be unnecessarily restrictive. Therefore, a competent project structure with transparent responsibilities and clearly defined budgets for various investments is highly recommended.

In addition, the establishment of a carefully designed and organized change management process is considered crucial in order to reduce the risks of frustration, resistance or refusal among employees. An established change management process accompanied by a proactive information policy can keep performance loss low [20].

References

[1] F. Fleurey, E. Breton, B. Baudry, A. Nicolas, J.M. Jézéquel, Model-Driven Engineering for Software Migration in a Large Industrial Context, Model Driven Engineering Languages and Systems 4735 (2007), 482–497.
[2] M. Alaranta, S. Henningsson, An approach to analyzing and planning post-merger IS integration: Insights from two field studies, Information Systems Frontiers 10(3) (2008), 307–319.
[3] P. McKiernan, Y. Merali, Integrating information systems after a merger, Long Range Planning 28(4) (1995), 4–62.
[4] A.C. Stylianou, C.J. Jeffries, S.S. Robbins, Corporate mergers and the problems of IS integration, Information & Management 31(4) (1996) 203-213.
[5] J. Øvretveit, T. Scott, T.G. Rundall, S.M. Shortell, M. Brommels, Improving quality through effective implementation of information technology in healthcare, International Journal for Quality in Health Care, 19(5) (2007), 259-266.
[6] S.Y. Park, S.Y. Lee, Y. Chen, The effects of EMR deployment on doctors' work practices: A qualitative study in the emergency department of a teaching hospital, International Journal of Medical Informatics 81(3) (2012), 204–217.
[7] D. Gans, J. Kralewski, T. Hammons, B. Dowd, Medical Groups' Adoption Of Electronic Health Records And Information Systems, Health Affairs 24(5) (2005), 1323–1333.
[8] R.H. Miller, I. Sim, Physicians' Use Of Electronic Medical Records: Barriers And Solutions, Health Affairs 23(2) (2004), 116–126.
[9] D.W. Bates, A.A. Gawande, Improving Safety with Information Technology, New England Journal of Medicine 348(25) (2003), 2526–2534.
[10] B.B. Dean, J. Lam, J.L. Natoli, Q. Butler, D. Aguilar, R.J. Nordyke, Review: Use of Electronic Medical Records for Health Outcomes Research: A Literature Review, Medical Care Research and Review 66(6) (2009), 611–638.
[11] Austrian National Council, Federal Act on the Organization of Universities and their Studies 2002 (UG 2002), https://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2015_I_131/BGBLA_2015_I_131.pdf, last access 7.1.2016.
[12] Vienna Hospital Association, Number of beds classified according to type, www.akhwien.at/default.aspx?pid=871, last access: 28.12.2015.
[13] W. Teppe, The ARNO Project: Challenges and Experiences in a Large-scale Industrial Software Migration Project, European Conference on Software Maintenance and Reengineering (2009), 149-158.
[14] S.S. Robbins, A.C. Stylianou, Post-merger systems integration: the impact on IS capabilities, Information Management 36(4) (1999), 205–212.
[15] HIMSS Europe, HIMSS Analytics, www.himssanalytics.eu/emram, last access: 12.1.2016.
[16] E. Beck, M. Obrist, R. Bernhaupt, M. Tscheligi, Instant card technique: how and why to apply in user-centered design, Proc. of the Tenth Anniversary Conference on Participatory Design (2008), 162–165.
[17] HIMSS Europe, European hospitals EMRAM maturity overview, HIMSS CIO Summit, Valencia (2015).
[18] R. Ahmad, Y. Kyratsis, A. Holmes, When the user is not the chooser: learning from stakeholder involvement in technology adoption decisions in infection control, Journal of Hospital Infection 81(3) (2012), 163–168.
[19] I.R. Bowns, G. Rotherham, S. Paisley, Factors associated with success in the implementation of information management and technology in the NHS, Health Informatics Journal 5(3) (1999), 136–145.
[20] N. Lorenzi, Organizational ISSUES=change, International Journal of Medical Informatics 69(2–3) (2003), 197–203.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-191

Animated Scatterplot – Analysis of Time-Oriented Data of Diabetes Patients

Margit POHL[a,1], Harald ENDL[a] and Ulrich FELS[a]
[a]*Vienna University of Technology, Austria*

Abstract. Animated scatterplot visualizations promise to be useful tools for analyzing trends in time-oriented data. We evaluated such a visualization for the analysis of time-oriented medical data. We found 10 medical professionals to test the software which visualizes clinical diabetes patient cohorts. To analyze the usability of the software, the methods "Thinking Aloud" and structured interviews were used. Results show that animated scatterplots do support medical professionals in their daily work, in contrast to more negative results of scientific research in the past.

Keywords: medical data, animation, time, information visualization, time-oriented data, evaluation

1. Introduction

One of the major problems for physicians treating patients with chronic diseases is that in such cases a large amount of data accumulates over the years (e.g. blood tests, blood pressure etc.). Especially older patients may also have multiple diseases at the same time. This places a large burden on the physicians who have to study all these tests and reports. Information visualization is a possibility to ease this burden and to provide medical doctors with an overview of these data at a glance. It is especially suitable to represent heterogeneous data from various sources which come in different formats, e.g. data from blood tests, medication and body mass index.

A program using animated scatterplots is a possibility to represent such data. The main goal of animations is to show time-oriented data in a "natural form" – time is supposed to represent time. In contrast to timeline representations, the x-axis of the visualization can be used to show a second variable. This enables the user to see the temporal development of the relationship between two variables (e.g. body mass index and blood sugar). The temporal development is shown as the movement of the dots (which represent the patients) on the screen. In the literature, there is a controversial discussion about animation. Some authors argue that animation can be very confusing, and that small multiples is the better method to support exploration [8]. Tversky et al [9] are also sceptical about the usage of animation, although she argues that interaction can help to alleviate the problems with animations. Recent research indicates that some forms of interaction can help to get insights with animations. Kriglstein et al [5] conducted a literature review and argue that the reviewed studies indicate that VCR-like control elements (see Figure 2) can help users to get insights more easily. It is especially valuable if the users are able to control the speed of the animation.

[1] Corresponding Author: Margit Pohl, Vienna University of Technology, Argentinierstrasse 8, 2nd floor, Vienna, Austria. Email: margit@igw.tuwien.ac.at

The goal of the study was to test the usability and utility of the Animated Scatterplot software for the presentation of time-oriented medical data. The results indicate that the program had some usability problems which were solved in a later version. Nevertheless, the users were able to use the program fairly well after a trial time. We also investigated more general issues, especially the question whether animation can be beneficial for the work of physicians and whether such a program can be learned easily.

1.1. Related Work

There is a controversial discussion in scientific literature concerning animation. Bartram [1] discusses potential advantages of the use of animation. She describes animation as an additional and useful display dimension to represent large amounts of multivariate dynamic data. Griffin et al. [4] describe the additional advantage of using animation in recognizing trends regarding timing (i.e., frame rate).

Nowell et al. [6] see a potential disadvantage in the perceptual effect of change blindness, where significant changes are not recognized. In Tversky et al. [9] static and animated representations were compared. They state that static and animated cases were not comparable (e.g., representations differ or include less or more information). They expressed three main concerns on animation: (i) developments may be hard to detect, (ii) developments may be perceived as a sequence of frames, and (iii) there is no or less interactivity. They recommend that users should be able to change orientation of parts or the whole animation and are able to control the speed, view, review, start, stop, zoom in and out.

In the literature there is no clear view on animation in visualization. Therefore further research is necessary to explore animation in visualization.

1.2. Description of the Animated Scatterplot Program

The basic feature of the Animated Scatterplot program (see Figure 1) is a scatterplot for two variables. Users can select variables from two combo boxes next to the respective axes (e.g. blood sugar level versus body mass index). Development over time is represented by the movements of the dots of this scatterplot diagram. One dot represents one patient. Additional variables (e.g. sex of the patients) can be represented by the form or the color of the dots. A VCR-like panel allows the users to control the animation, but it is also possible to use a timeslider and choose the speed of the presentation manually. The program offers several different possibilities of interaction. It is, for example, possible to zoom into the data within the scatterplot diagram. It is also possible to filter the data (e.g. to show only patients with values above a certain threshold). Users can also get detailed information of individual patients (by hovering over the dot of this specific patient). There are two possibilities to show the “age” of data, that is the temporal distance from the last measurement of variables: transparency and traces. In transparency mode, marks fade out more and more. This makes patients with current data clearly stand out. In trace mode, all traces of the movement of dots stay visible. Thus, at the end of the animation, complete patient histories can be seen (see Figure 3). Rind, et al [7] provide a more extensive description of the software.

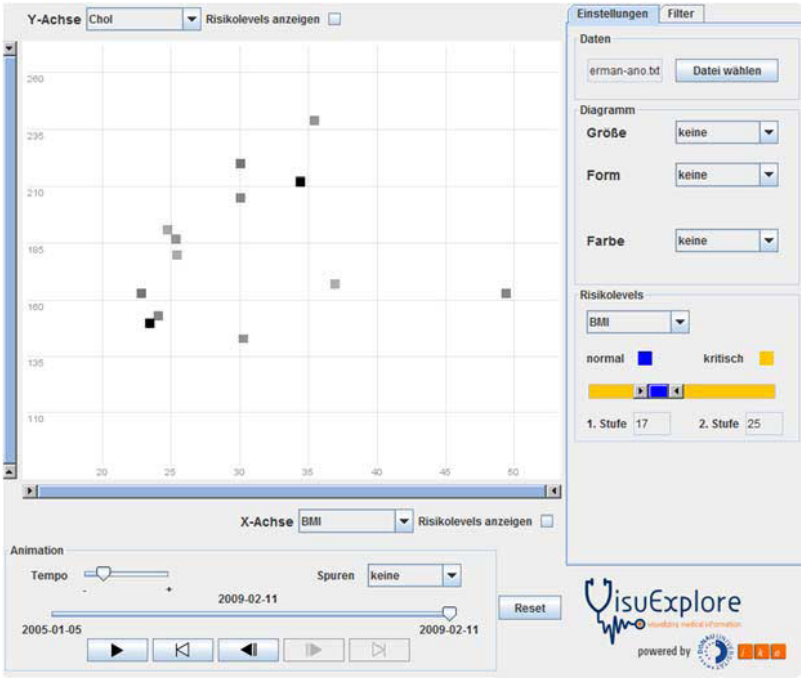


Figure 1. Animated Scatterplot (main screen)

2. Methods

The goal of the study was to test the usability of the Animated Scatterplot program for the presentation of time-oriented medical data (specifically diabetes data). The participants of the study were 10 medical doctors. This makes the study especially valuable. It is well known that it is difficult to get real experts, especially physicians, for evaluation studies in information visualization. Therefore, visualizations are often tested with students which makes the results in many cases less valid.

The methods which were used were content analysis based on screen capture of the interactions of the users and interviews which were conducted after the experiment. For the content analysis, two different taxonomies were used for categorization purposes, the taxonomy developed by Yi, et al [10] and the categorization scheme developed by Forsell and Johansson [3]. The categorization scheme developed by Forsell and Johansson is especially relevant for this paper. Their categorization scheme is based on other well-known heuristics. They derived empirically the most important usability heuristics for information visualization: information coding (mapping), minimal actions, flexibility (number of possible ways to achieve a goal), orientation and help, spatial organization (e.g. distribution of elements on the screen), consistency, recognition rather than recall, prompting (all means to support the users to find alternative ways of doing things), remove the extraneous (distraction through unnecessary information), data set reduction. Even if these categories are mainly meant for heuristic evaluation, we think that they can also form a valuable framework for the interpretation of user actions working with information visualizations.

The participants of the study first got a brief introduction into the system. Then they had to solve 4 tasks which were developed together with medical experts and which reflected realistic activities in medical diagnosis. Task 1 was designed to let the participants familiarize themselves with the software, tasks 2 and 3 had concrete questions to answer, and task 4 allowed the participants to freely explore and interact with the software.

Below, we want to show task 3 as an example for the tasks used in this investigation:

Parameters: x-axis: NBZ (fasting glucose level)
y-axis: RR diast (mmHG) (diastolic blood pressure)

Task description: Limit the data set to NBZ < 100; RR diast < 80. Choose a setting that gives a good overview of the trends of the patients. Which patients shows a favorable trend? What is the general trend of the group? Explore at will. Describe your findings.

The interactions of the participants were logged with a screen capture program. The participants were asked to think aloud while they solved the tasks. This was also recorded with the screen capture program. After the participants had solved the 4 tasks, an interview was conducted to identify the subjective attitude of the participants concerning the program. The duration of the whole experiment was approximately 60 to 90 minutes.

According to Boren and Ramey [2] the method of "Thinking Aloud" can provide immediate insights into the reasoning processes accompanying the solution of cognitive tasks. While study methods that are applied after the interaction with the system can also yield interesting results, there is a tendency of the users to re-interpret their activities retrospectively. In thinking aloud it is possible to avoid this source of misinterpretation. Therefore, this method is often used by usability researchers to get a more detailed view of the users' interaction processes with the system under investigation.

3. Results

3.1. Screen Capture and Thinking Aloud

The screen capture showed that all the participants could solve all the tasks with the animated scatterplot program. The solution process for task 4 showed that the participants also tried out new parameters and new features on their own in task 4 which asked them to freely explore the data. 8 of the 10 participants used new parameters in this task, only 2 used parameters from previous tasks to study them in more detail.

The largest amount of usability errors according to the classification of Forsell and Johannsson (2010) was in the areas of spatial organization and consistency. It is clear that the spatial organization of the system was novel for the participants and therefore, they had difficulties with the visualization. Physicians, in general, tend to use visualization like line charts and bar charts and avoid more novel forms of representation. Therefore, at first they tended to interpret the x-axis as the time axis because they were used to this from static line graphs. At the beginning, they also had problems to understand the dynamic nature of the animation. For some patients no data existed for the first time periods. Participants were confused that dots appeared for these patients all of a sudden while they worked with the visualization. It must be mentioned, however,

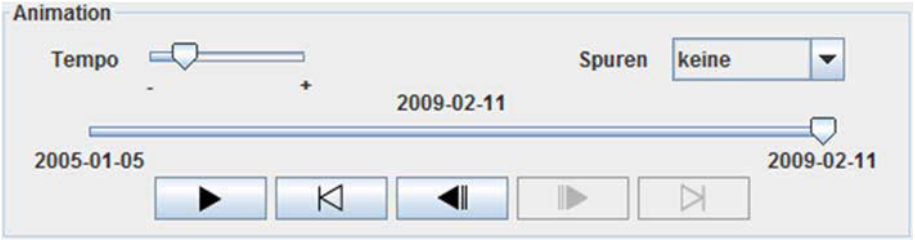


Figure 2. VCR-type of control and time slider

that the participants got used to these novel features fairly quickly and understood them at the end of the testing period.

Consistency was a problem for the users because the animated scatter plot program does not resemble standard Microsoft application programs. The other categories identified by Forsell and Johansson [3] only played a minor role as sources of usability errors.

The participants appreciated the possibility of the program to show groups of patients and several points in time in the same visualization. The animation was new and interesting, but it took some time to adapt to it. None of the participants used the VCR-type of control of the animation, but only the time slider (see Figure 2). They commented that the speed of the animation was too slow. They dragged the time slider manually again and again (all in all the participants used the time slider approximately 600 times). They argued that this helped them to understand the temporal development better. This conforms to the assumption of Tversky et al [9] who posits that animation can be beneficial if interaction possibilities are provided.

3.2. Interviews

The main questions which were asked in the interviews addressed general problems of the program, advantages of the program, possible application areas of the program, interaction possibilities, and visualization and animation.

Advantages of the program: Participants mentioned that the main advantage of the program was the possibility to see the data of all patients at the same time and to get an overview of the data.

Possible Application areas: Participants said that a possible application area could be for monitoring the effects of new medications. Another possible application was the education of healthcare professionals.

Interaction possibilities: Participants mentioned that they would appreciate the possibility to highlight single or a small group of patients which was not possible in their version of the program. They also enjoyed the extensive possibilities to filter data. This was seen as very beneficial. Even the combination of several criteria for filtering was intuitive and easy to understand. Zooming into the data was only appreciated by 3 out of 10 participants. This might be due to the design of this functionality.

Visualization and animation: Participants mentioned that at first they interpreted the x-axis as time, but they quickly got used to that. Participants had different opinions concerning the number of patients (=dots) they wanted to see on the screen. Some preferred more patients to be able to see trends more clearly, and others found the large number of dots too confusing. Participants preferred to manually operate the time slider instead of the VCR-type of control. The development of the patients (that is, the

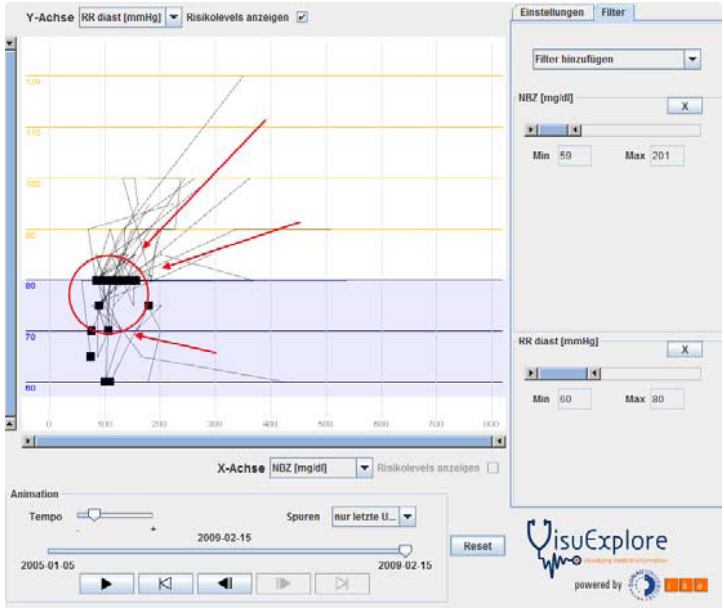


Figure 3. Trend recognition (traces)

movement of the dots) during the animation can be shown as a static picture using traces. These traces show the movement of the dots on the screen as static lines. Traces can be used during the animation to support the recognition of a group trend (see Figure 3). Participants appreciated this possibility to show trends, although some mentioned that traces get cluttered when too many patients and time steps were being shown.

One of the most serious usability problems was the design of the risk levels. The "normal range" was highlighted, not the "at risk" range, and although the physicians know what the normal range of the different medical parameters is, they focused on the highlighted area and interpreted it as risk level (see Figure 2). Apparently, the affordance of the highlighted area implies that it shows the risk level, not the normal range.

Strategy: All participants mentioned that they first studied the whole group of patients and then significant single patients – either outliers or patients who are typical for a certain group. They start with the whole group and then reduce the number of patients they want to investigate.

4. Discussion

In this paper we describe a usability and utility evaluation of an animated scatterplot software for medical data. Although the evaluation addressed one specific system, some general conclusions can still be drawn. On the one hand, this study is a contribution to the discussion concerning the value of animations in information visualization. On the other hand, our results can also contribute to the discussion about how to deploy novel systems.

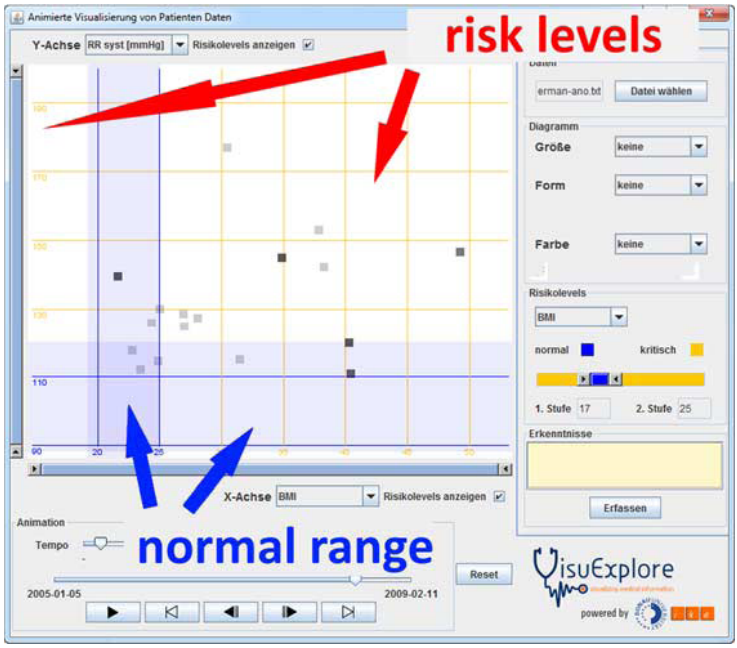


Figure 4. Risk levels

The animated scatterplot system was at first confusing for the physicians who participated in the study. We could observe, however, that they quickly learned how to use the system. They could solve all the tasks presented to them fairly easily. We also found that interaction plays an important role in this context. The participants did not use the VCR-like control but the time-slider which allowed them much more control over the speed of the animation and provided the possibility to stop the animation anytime they wanted to. These results indicate that animations can be used to help experts to explore data and get insights, contrary to Robertson's et al [8] sceptical attitude. We would like to point out that it is important to design the visualization appropriately, so that it is intuitive to use. The study also indicates that systems that are easy to learn can help in the implementation process. In general, physicians are sceptical about novel visualization because their time constraints are very severe. Our study shows that this can be overcome by a system that is easy to learn.

We also identified some usability problems as, e.g., the design of the risk level. Based on this study the system was redesigned to enable physicians to work with it more easily.

There are also some limitations concerning this study. It can be argued that the situation in which the Animated Scatterplot system was tested was not realistic. In addition, it seems plausible to assume that users will behave differently if they have used a system for a longer period of time and are more acquainted with it. Possible future work would be to test the system over a longer period of time, so that the participants of the study get a more realistic picture of the system. In such a situation, users are better able to assess the advantages and disadvantages of the system and how it can be used most efficiently to support the work of physicians.

Acknowledgements

This work is conducted in the context of the CVAST (Centre for Visual Analytics Science and Technology) project. It is funded by the Austrian Federal Ministry of Science, Research, and Economy in the exceptional Laura Bassi Centres of Excellence initiative, project number 822746. We want to thank B. Neubauer for the development of the Animated Scatterplot System.

References

[1] Bartram, L., Perceptual and interpretative properties of motion for information visualization. In: Proc. Workshop New Paradigms inInformation Visualization and Manipulation, 1997, pp.3-7.

[2] Boren, T.M. &Ramey, J., Thinking Aloud: Reconciling theory and practice. IEEE transactions on professional communication, September 2000, 43 (3), pp.261-278.

[3] Forsell, C. & Johansson, J., An heuristic set for evaluation in information visualization. In: Proceedings of the International Conference on Advanced Visual Interfaces 2010, pp.199-206.

[4] Griffin, A.L., MacEachren, A.M., Hardisty, F., Steiner, E., Li, B., A comparison of animated with static small multiple maps for visually identifying space-time clusters. Annals of the Association of American Geographers 96(4), 2006, pp.740-753.

[5] Kriglstein, S., Pohl, M. & Smuc, M., Pep Up Your Time Machine: Recommendations for the Design of Information Visualizations of Time-Dependent Data. In: W. Huang, ed. 2014. *Handbook of Human Centric Visualization*. Springer New York, 2014, pp.203-225.

[6] Nowell, L., Hetzler, E., Tanasse, T., Change blindness in information visualization: A case study. In: Proc. IEEE Symp. Information Visualization, 2001, pp.15-22.

[7] Rind, A., Aigner, W., Miksch, S., Wiltner, S., Pohl, M., Drexler, F., Neubauer, B. & Suchy, N.,Visually Exploring Multivariate Trends in Patient Cohorts using Animated Scatter Plots. M.M. Robertson (ed.), Ergonomics and Health Aspects of Work with Computers, In: Proceedings of the International Conference held as part of HCI International 2011, LNCS 6779, pp.139-148, Heidelberg (DE-BW): Springer. doi:10.1007/978-3-642-21716-6_15.

[8] Robertson, G., Fernandez, R., Fisher, D., Lee, B. & Stasko, J., Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 2008, pp.1325-1332.

[9] Tversky, B., Morrison, J.B., Betrancourt, M., Animation: can it facilitate? Int. J.Human-Computer Studies 57, 2002, pp.247-262.

[10] Yi, J.S., Kang, Y.A., Stasko, J.T. & Jacko, J.A., Toward a deeper understanding of the role of interaction in information visualization. IEEE Transactions on visualization and computer graphics, 13(6), 2007, pp.1224-1231.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-199

How Is My Field Evolving? – Network Based Analysis of Biomedical Scientific Discourse

Cord SPRECKELSEN[a,1], Klaus KABINO[a]
[a]*Department of Medical Informatics, RWTH Aachen University*

Abstract. *Background*: Gaining overview on a biomedical field of research becomes challenging due to the increasing amount of publications. Computer-supported discourse analysis based on bibliometric data could help scientists to identify focal points and trends. *Objectives*: The project aimed at implementing an automatic processing pipeline starting from a PubMed query and leading to a suitable visualization of the thematic evolution of a domain fostering meaningful interpretation. *Methods*: The four-step processing pipeline includes bibliographic data acquisition, preprocessing, network-based analysis of co-occurring keywords (degree and betweenness centrality), and its visualization by heatmap diagrams. *Results*: Applying the implemented workflow we analyzed the field of clinical decision support systems based on 5.094 PubMed results yielding a total of 174.663 inter-keyword links. The resulting heatmap shows e.g. increased relevance of electronic health record and (patients') age. *Conclusion*: Network-based discourse analysis can be implemented using an efficient processing pipeline and may add valuable insight on the thematic evolution of scientific domains.

Keywords. Bibliometrics, Health Information Management, Abstracting and Indexing as Topic

1. Introduction

Facing the overwhelming and ever increasing number of publications in the biomedical sciences in general it becomes difficult to gain an overview on the scientific discussion of a field of interest. Network-based bibliometric analysis supported by an effective visualization of the results may foster the identification and interpretation of focal points and trends within the scientific discourse. Discourse is used here as: "a mode of organizing knowledge, ideas, or experiences that is rooted in language and its concrete contexts" [1]. Discourse analysis then investigates documented language (e.g. text corpora) in order to reveal the evolving conceptual (and institutional) patterns of a given field of interest.

Bibliographic databases such as PubMed are an invaluable resource for literature research and retrieval of relevant information. Some bibliometric databases support time-related analyses: e.g. PubMed's "Result per year"-chart shows the development of publication activity. Third party tools (e.g. GoPubmed) further analyze PubMed results e.g. by retrieving all included keywords and offering interactive modification of the search focus, by geo-localizing the affiliations, and by creating collaboration diagrams based on co-authorship.

[1] Corresponding Author: First Author, Department of Medical Informatics, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany, E-Mail: CSpreckelsen@mi.rwth-aachen.de

While these basic approaches provide useful insight, they, nonetheless, need to be complemented in order to produce a more substantial view on the past development and future trends of a given field based on bibliometric discourse analysis. More complex bibliometric analysis alone is not enough: It needs to be complemented by suitable methods of visualizing the results to comprehensively and efficiently inform the user and to prepare meaningful interpretation.

1.1. Aim

The project reported here aimed at implementing an automatic processing pipeline starting from a PubMed query and leading to a graphic representations of the development of the field in question. The approach should 1) identify the relevance of topics/concepts based on bibliometric measures, 2) visualize temporal evolution of topic relevance, and 3) identify similarly evolving topics (i.e. thematic clusters defined by similarly evolving relevance). Assuming a renaissance of artificial intelligence (AI) in the field of medical decision making, it is promising to apply the approach to the field of clinical decision support systems in order to 1) investigate present trends and 2) show feasibility.

1.2. Related work

Bibliometric methods and text-mining algorithms have been applied to the analysis of biomedical literature extensively in order to aggregate content, evaluate scientific impact/relevance, and identify domain experts or research trends: Co-authoring analyses, co-occurrence w.r.t citations, keywords, and journals have been exploited statistically (see e.g. [2][I3]). To identify topics discussed from bibliographic and/or full-text data latent semantic indexing (LSI) [4], Dirichlet allocations (LDA) [5], author topic modeling (ATM) [6], and dynamic topic modeling (DTM) [7] have been adopted. These techniques are based on term statistics, treating topics as hidden (latent) variables [8]. Li et al. used ATM and DTM in order to inform smoking prevention strategies [9]. Westgate et al. used LDA in order to identify research gaps in conservation biology [10]. Neshati et al. [11] used topic modelling techniques in order to identify experts in an articles' author set.

As a complementary approach, network analytics is used in order to go beyond co-occurrence analysis and qualify the relevance of topics. In the past, network analysis has been applied successfully to various fields from the analysis of social networks to trade economics - recently including applications to medical fields e.g. the analysis of genes (metabolic) pathways [12], and interacting brain regions [13].

In the context of bibliometric analysis Hsu and Kao adopted network analysis to document triage, i.e. to filter relevant articles [14]. Furthermore, scientific collaboration has been analyzed based on co-authorship network analysis [15][16]. Recently, Ineichen and Christen analyzed co-occurring terms extracted from articles on deep brain stimulation using network centrality measures [17]. Nonetheless, the potential of network analysis for bibliometric investigations in the biomedical field has to be further exploited.

Dedicated efforts address the visualization of text outside the biomedical domain ranging from network diagrams based on graph layout algorithms showing e.g. co-citation networks to sophisticated river plots representing the topic flow of time-varying

text data [18]. The ISOVIS Text Visualization Browser (http://textvis.lnu.se/) presents an inspiring showcase of text visualization.

2. Methods

The project reported here adopts network-theoretic measures (namely "degree centrality" and "betweenness centrality") to discourse analysis in medical fields. The network to be analyzed is derived from co-occurring keywords. The results are visualized by heatmaps. Figure 1 shows the processing pipeline for network-based bibliometric discourse analysis.

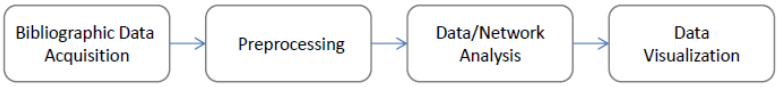


Figure 1. Processing pipeline for network-based bibliometric discourse analysis

2.1. Processing pipeline

Automatic data acquisition starts from bibliographic databases (in our case PubMed) by using web-service based application programming interfaces. The National Center for Biotechnology Information (NCBI) offers a toolset of eight server-side programs providing web-based access to the 38 databases of the Entrez system, including PubMed: the Entrez Programming Utilities (E-utilities) [19].

Using E-utilities the data acquisition module retrieves the structured result set of a given PubMed-query, parses the incoming string data and generates document data objects (DDO). The subsequent modules of the processing pipeline can directly access the PubMed document ID (PMID), publication date, title, author list, abstract text (if available), and the keywords list from a DDO. Documents marked as "indexed for MEDLINE" are uniformly assigned to keywords of the Medical Subject Headings by the National Library of Medicine (NLM).

The *preprocessing* module extracts selected metadata from the DDOs and generates data structures suitable for the subsequent analysis step. For the graph based discourse analysis reported here the preprocessing module extracts 1) PMID, 2) year of publication, and 3) MeSH keywords in normalized relational format. Each rows of the resulting table contains PMID, year, and a single keyword derived by decomposing the keyword list into multiple rows. Self-joining the keyword assignment table by PMIDs generates pairs of co-occurring keywords indicating research-related association. By counting duplicate pairs of keywords for the same publication year the module generates inter-keyword links, weighted by their multiplicity and associated with the respective year of publication. Based on this relation the module returns graph data structures for each year, where the keywords are represented as nodes and their weighted associations as edges of the graph, respectively.

Data/Network analysis uses network-theoretic measures calculated based on the keyword graphs: degree and betweenness centrality. Degree centrality (DC) is basically given by the number of edges attached to a node of the graph. In our context degree centrality measures the global frequency of a given keyword (disregarding cases, where an article is indexed by a single keyword only). Betweenness centrality (BC) of a given node counts the fraction of shortest paths connecting any two nodes of the graph passing

the node in question. For a graph g, its node set $V(g)$, nodes $v, i, j \in V(g)$ and $\sigma_{i,j}(v)$ being a shortest path connecting nodes *i, j* and contains *v,* the betweenness centrality of *v* is defined [20] as:

$$BC(v) := \sum_{i,j \in V(g), i \neq j \neq n} \frac{\sigma_{i,j}(v)}{\sigma_{i,j}} \quad (1)$$

Suitable *data visualization* by heatmaps enables visual inspection and subsequent interpretation. The union set of all keywords used in a given time interval labels the x-axis of the heatmap, the years included in the analysis label the y-axis. A cell of the heatmap is colored according to the value of the respective centrality measure (DC, BC). Hierarchic clustering of the keywords can start from the keywords' vectors of centrality values per year. Keywords with similar expression patters of DC and BC are clustered together.

Occurrence of keywords or hashtags in larger corpora and network centrality measures can be expected to follow a power law (see [21] for an investigation of power-law like distributions in Entrez Gene and PubMed). Thus, direct translation of values into a linear color-scale would result in visible contrasts between few high-performing nodes and all others, while existing differences between the majority of nodes/keywords in the long-tail of the power-law distribution would produce no discernible details in the diagram. Thus, a nonlinear color-scale is created in order to increase contrast between long-tail values.

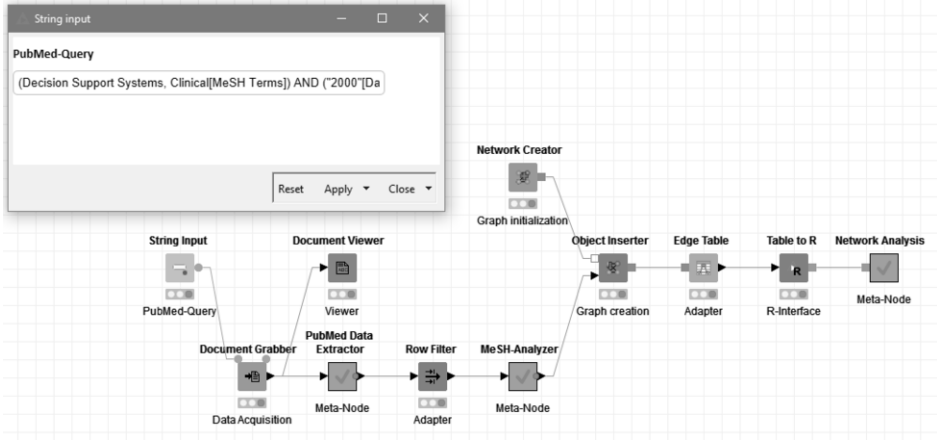


Figure 2. Implemented KNIME-workflow for network-based bibliometric discourse analysis - top level view with collapsed sub-processes (meta-nodes)

2.2. Target domain

Clinical decision support systems served as an exemplary domain for discourse analysis using our approach. Bibliographic data is acquired by the following PubMed query: "(Decision Support Systems, Clinical[MeSH Terms]) AND ("2000/01/01"[Date - Publication] : "2016/02/01"[Date - Publication])".

2.3. Software

The KNIME analytics platform (v3.1.0) served as a means for data management (data acquisition and preprocessing modules), while graph analysis and generation of the heatmaps was carried out using the R Project for Statistical Computing with packages igraph and ggplot2. Power law fitting and testing based on the bootstrapping approach proposed by Clauset et al. [22] was carried out using the powRlaw package [23].

3. Results

The processing pipeline described above was implemented using KNIME with integrated R-framework. Figure 2 shows an overview of the implemented KNIME workflow. Starting from a PubMed query to be entered by the user, the application acquires bibliographic data from PubMed, generates and analyses the keyword graphs and returns numeric output and heatmaps visualizing centrality results.

For the exemplary domain of decision support systems during the years from 2000 to present, the system retrieved and analyzed 5.094 DDOs resulting in a total of 174.663 weighted inter-keyword links to be included in the network analysis.

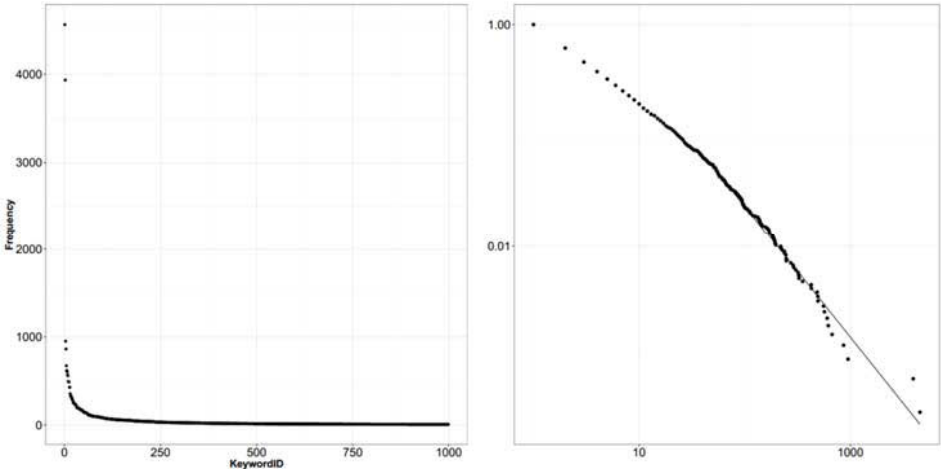


Figure 3. Frequency of keywords in the given domain (left) and corresponding Cumulative Density Function (CDF) plot including the line of best-fitting power-law distribution

Figure 3 shows the overall usage of keywords for indexing the given domain and the corresponding plot of the Cumulative Density Function (CDF) including the line, which corresponds to the (best-fitting) power-law distribution with parameters x_{min}= 48, and scaling parameter α=2.19. The goodness-of-fit test via bootstrapping recommended by Clauset et al. [22] yielded a gof-value of .037 and p=.459, thus indicating, that a power law is plausible (mind the comment on p-value-usage in [22], p.17).

Figure 4 shows the heatmap visualizing betweenness centrality for the top two percentiles of the keywords. The four keywords “Decision Support Systems, Clinical” (obviously induced by the initial query), “Humans”, “Male”, and “Female” represent the top permille and were excluded from the heatmap. The dendrogram on the left shows clusters of topics evolving similarly over time.

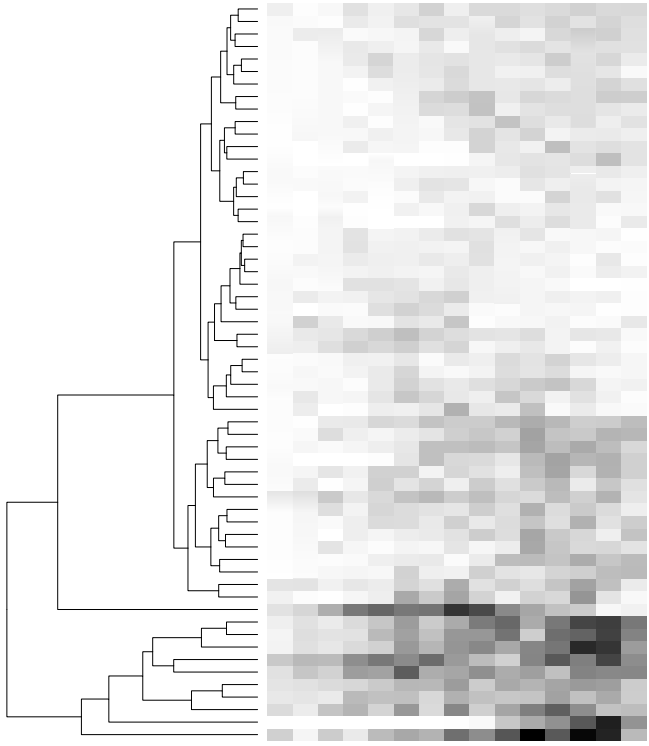


Figure 4. Heatmap visualizing betweenness centrality of the top two percentile nodes of the keyword graphs for publication years 2000-2014.

4. Discussion

While DC essentially measures the keyword frequency, BC detects topics serving as means of integration in the field. High BC indicates a topic connecting subfields of the discourse: If the respective concept was not present, the discourse network would be markedly less connected or even fell apart in separated subnets, i.e. areas of the scientific field having no conceptual intersection. Following this line of interpretation a topic as e.g. "Practice guidelines" induced concerted research effort in the field of clinical decision support for some years around 2005, with a slightly declining trend between 2008-2011 and reaching a plateau in the last years, whereas "Electronic Health Records" now strongly attracts interest from different areas of the field and serves as an integrational means. The latter observation is in good accordance with the implementation of the "meaningful use" incentive program [24], which has massively promoted the use of electronic health records for decision support since 2011 – and, therefore, supports the validity of the approach. Furthermore, the diagram reveals an increased focus on phases of life (assumingly of patients) with an emphasis on "Aged"/"Middle Aged" and the increasing importance of "Algorithms".

The dendrogram of clustered topics may provide additional insights, but should be interpreted with caution: The two clusters: A) "Decision Making, Computer-Assisted" –

"Hospital information Systems" and B) "Risk assessment" – "Diagnosis, Assisted" for instance exhibit complementary expression patterns over time. While cluster A includes focal points of previous (sobering) experiences with integrating expert system functionality into clinical information systems, Cluster B seems to represent a more recent trend to apply decision support and AI methodology in the context of evidence-based medicine, risk assessment and outcome orientation.

Our investigation found high plausibility for assuming a power law distribution of MeSH keyword assigned to publications in a given biomedical field, which has implications for an apt visualization of the heatmaps.

Compared to LDA used for topic analysis, our approach abstains from sophisticated text-mining. Instead we rely on the sound basis of MeSH indexing carried out by the NLM team and simply exploit this rich source a bit further. This may seem trivial from a text mining perspective, but one can argue, that precise and costly semantic aggregation as represented by MeSH keyword indexing should be exploited first, before adopting more experimental approaches.

The recent work of Ineichen and Christen [17] uses a very similar approach, but 1) starts from terms extracted from full-text and 2) uses layouted graphs and trendlines for visualization. From our own experience, multiple trendlines representing the developments e.g. of a centrality measure for many keywords soon become confusing and, therefore, do not allow exploration of more than a few selected or abstracted concepts, whereas heatmaps have proven to foster extensive data exploration e.g. in the field of gene expression analysis.

With only slight modifications the approach can process title words or abstract terms. Actually, a combination of part-of-speech tagging (POS tagging) and suitable stop-word filters has already been investigated in the course of this project. Nonetheless, without further processing based on medical classifications or ontologies the heterogeneity of terms used in titles and abstracts prevents a consistent overview supported by a considerable amount of related publications. Thus, extending the analysis to title and abstract terms is left for future effort.

Focusing instead on MeSH keywords clearly has the advantage of a relatively homogeneous assignment of publications to topics. The obvious drawback of using MeSH keywords is, that more detailed or newly introduced domain concepts are not included. As a consequence, the approach is not suitable for detecting fine-grained aspects of a domain of interest, but – in contrast – yields results showing a more general picture, which enables comparison across the border of a special domain.

4.1. Limitations

The KNIME framework enabled rapid prototyping and complete control on intermediate results. Nonetheless, the KNIME-to-R-integration turned out to be a performance bottleneck. Thus, the pipeline is being reimplemented now and will soon be replaced by a pure R-program using the RISmed package for data acquisition.

4.2. Conclusion

While tools for information/paper retrieval exist and are available for routine use, methods and tools for discourse analysis are still to be improved. Network-based bibliometric discourse analysis can be carried out efficiently using the processing

pipeline reported here and may add valuable insight on the thematic evolution of scientific domains.

References

[1] Merriam-Webster: Dictionary. [Internet] [Cited: 2016/02/08]. Available from:: http://www.merriam-webster.com/dictionary/discourse,

[2] Janssens ACJW, Gwinn M. Novel citation-based search method for scientific literature: application to meta-analyses. BMC Med Res Methodol. 2015;15:84.

[3] Ioannidis JPA. Measuring co-authorship and networking-adjusted scientific impact. PLoS ONE. 2008;3(7):e2778.

[4] Chen H, Martin B, Daimon CM, Maudsley S. Effective use of latent semantic indexing and computational linguistics in biological and biomedical applications. Front Physiol. 2013;4:8.

[5] Li DC, Thermeau T, Chute C, Liu H. Discovering associations among diagnosis groups using topic modeling. AMIA Jt Summits Transl Sci Proc. 2014;2014:43–9.

[6] Rosen-Zvi M, Griffiths T, Steyvers M, Smyth P. The author-topic model for authors and documents. Proceedings of the 20th conference on Uncertainty in artificial intelligence, 2004: 487-494.

[7] Li D-C, Rastegar-Mojarad M, Okamoto J, Liu H, Leichow S. A Bibliometric Analysis on Cancer Population Science with Topic Modeling. AMIA Jt Summits Transl Sci Proc. 2015;2015:102–6.

[8] Chien J-T. Hierarchical Theme and Topic Modeling. IEEE Trans Neural Netw Learn Syst. 2015, 99. DOI: 10.1109/TNNLS.2015.2414658.

[9] Li D, Okamoto J, Liu H, Leischow S. A bibliometric analysis on tobacco regulation investigators. BioData Min. 2015;8:11.

[10] Westgate MJ, Barton PS, Pierson JC, Lindenmayer DB. Text analysis tools for identification of emerging topics and research gaps in conservation science. Conserv Biol. 2015;29(6):1606–14.

[11] Neshati M, Hashemi SH, Beigy H. Expertise finding in bibliographic network: topic dominance learning approach. IEEE Trans Cybern. 2014;44(12):2646–57.

[12] Anitha P, Anbarasu A, Ramaiah S. Gene network analysis reveals the association of important functional partners involved in antibiotic resistance: A report on an important pathogenic bacterium Staphylococcus aureus. Gene. 10. Januar 2016;575(2 Pt 1):253–63.

[13] Luo Y-G, Wang D, Liu K, Weng J, Guan Y, Chan KCC, u. a. Brain Structure Network Analysis in Patients with Obstructive Sleep Apnea. PLoS ONE. 2015;10(9):e0139055

[14] Hsu Y-Y, Kao H-Y. CoIN: a network analysis for document triage. Database (Oxford). 2013;2013:bat076. DOI: 10.1093/database/bat076

[15] Bender ME, Edwards S, von Philipsborn P, Steinbeis F, Keil T, Tinnemann P. Using Co-authorship Networks to Map and Analyse Global Neglected Tropical Disease Research with an Affiliation to Germany. PLoS Negl Trop Dis. 2015;9(12):e0004182.

[16] Uddin S, Hossain L, Rasmussen K. Network effects on scientific collaborations. PLoS ONE. 2013;8(2):e57546.

[17] Ineichen C, Christen M. Analyzing 7000 texts on deep brain stimulation: what do they tell us? Front Integr Neurosci. 2015;9:52.

[18] Cui W, Liu S, Tan L, Shi C, Song Y, Gao ZJ, Tong X. Textflow: Towards better understanding of evolving topics in text. Visualization and Computer Graphics, IEEE Transactions on. 2011;17(12): 2412-2421.

[19] NCBI: Entrez Programming Utilities Help. Bethesda (MD), National Center for Biotechnology Information 2010-. [Internet] [Cited: 2016/02/07] Available from: http://www.ncbi.nlm.nih.gov /books/NBK25501/.

[20] Freeman LC. A set of measures of centrality based on betweenness. Sociometry. 1977;40:35–41.

[21] Su AI, Hogenesch JB. Power-law-like distributions in biomedical publications and research funding. Genome Biol. 2007;8(4):404.

[22] Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. SIAM Review. 2009;51(4):661-703.

[23] Gillespie CS. Fitting Heavy Tailed Distributions: The poweRlaw Package. Journal of Statistical Software, 2015;64(2):1-16. (URL: http://www.jstatsoft.org/v64/i02/).

[24] Centers for Medicare & Medicaid Services (CMS), HHS. Medicare and Medicaid programs: hospital outpatient prospective payment; ambulatory surgical center payment; hospital value-based purchasing program; physician self-referral; and patient notification requirements in provider agreements. Final rule with comment period. Fed Regist. 30. November 2011;76(230):74122–584.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-207

Evaluation of Game Engines for Cross-Platform Development of Mobile Serious Games for Health

Carina KLEINSCHMIDT[a,1] and Martin HAAG[a]
[a]Hochschule Heilbronn, GECKO Institut, Germany

Abstract. Studies have shown that serious games for health can improve patient compliance and help to increase the quality of medical education. Due to a growing availability of mobile devices, especially the development of cross-platform mobile apps is helpful for improving healthcare. As the development can be highly time-consuming and expensive, an alternative development process is needed. Game engines are expected to simplify this process. Therefore, this article examines the question whether using game engines for cross-platform serious games for health can simplify the development compared to the development of a plain HTML5 app. At first, a systematic review of the literature was conducted in different databases (MEDLINE, ACM and IEEE). Afterwards three different game engines were chosen, evaluated in different categories and compared to the development of a HTML5 app. This was realized by implementing a prototypical application in the different engines and conducting a utility analysis. The evaluation shows that the Marmalade engine is the best choice for development in this scenario. Furthermore, it is obvious that the game engines have great benefits against plain HTML5 development as they provide components for graphics, physics, sounds, etc. The authors recommend to use the Marmalade Engine for a cross-platform mobile Serious Game for Health.

Keywords. Mobile application, video games, education, medical.

1. Introduction

Studies have shown that Serious Games for Health can improve patient compliance and the chance of success of therapies [1,2]. Another application for Serious Games for Health is the field of medical education. As for example Boeker et al. (2009) [3] have shown with the game "Uro Island", it is possible to support the students' learning difficult medical subjects by wrapping the content in a Serious Game. Therefore, Serious Games for Health are a valuable contribution for improving healthcare. With the growing availability of mobile devices, especially Serious Games for Health as mobile apps are very useful.

The quality of the graphical UI plays an important role for the success of a Serious Game [4]. As many games for entertainment have highly realistic graphics, especially children and young adults have developed high demands in games graphics. Developing Serious Games with such good graphics can be very expensive and time-consuming, which is not acceptable especially in the healthcare sector. Game Engines are expected to solve this problem of development costs and time commitment.

[1] Corresponding Author: C. Kleinschmidt, Hochschule Heilbronn, Max-Planck-Str.39, 74081 Heilbronn, Germany. E-Mail: carina.kleinschmidt@hs-heilbronn.de.

Game Engines are data-driven architecture software pieces which contain certain codes needed for playing games and which can be reused for developing other games. Such codes consist for example of logic for collisions, physics and especially graphics [5]. For developing other games they supply frameworks, programming environments and tools. This article will evaluate if these Game Engines help to facilitate the development and to raise the quality of Serious Games for medical education and which Game Engine therefore provides the most benefit.

2. Methods

At first a systematic research for literature was conducted in different databases (MEDLINE, ACM, IEEE). Several publications about Serious Games for Health were found, but only a few mention the use of game engines [6] and only some of them what kind of game engine they used [7,8]. One publication aimed at a similar examination. Marks et al. (2007) [9] explored the use of game engines for simulated surgical training. However, this scenario is very specific and not applicable in general for Serious Games for Health.

Afterwards the scenario was determined for which the evaluation will be made. As Boeker et al. (2009) [3] found out that game-based e-learning in medical education is more effective than conventional methods, the scenario will contain the development of a Serious Game for medical education. To gain a larger target group of students who have many different mobile devices the game should be available for Android, iOS and different browsers. Therefore, the engines need to provide cross-platform development. To fully exploit the development process with a game engine a prototype will be implemented. For the graphics of this prototype, the images for the scene were designed with Inkscape [10]. It was intended to design an erythrocyte that swims in a blood vessel and has to avoid the collision with some “bad particles” that are represented by black squares. This idea derives from Barbosa and Silva (2011) [11]. They created a Serious Game called “OxyBlood” with the aim to teach young students the basic functioning of the circulatory system. A simplified version of this game without any medical knowledge covers the basic technical functions that should be evaluated with this prototype. It is desired that the erythrocyte can be controlled with the arrow keys on a PC and on a mobile device by touching the screen to point the direction, e.g. at the top of the screen for swimming up. This will be implemented as a sidescroller, which means the figure can only move left or right along the x-axis of a Cartesian coordinate system. The figure needs also to follow physical laws and a collision detection should be implemented for the “bad particles”.

The further evaluation beyond implementing will be examining various aspects ranging from costs, license and installation, to the aspects of development and deployment to different target platforms. Therefore, a utility analysis was conducted. This method belongs to decision theory and helps to quantify non-quantifiable aspects. [12]. Every aspect is furnished with an importance rating from 1-5, to consider major aspects more than others. Every engine receives points from 1-10 for every aspect. For deciding when to give the full points, a fulfillment norm is defined for every aspect. These will be multiplied with the importance rating and summed up, so that a clear result for every engine and the HTML5 development will be provided and compared. In **Table 1** the evaluation aspects and their importance rating can be seen. These aspects were partial derived from the process of cross-platform development and partial completed by

aspects from general software-tests of relevant magazines [13]. This is only an excerpt of the complete evaluation that can not be provided in this paper due to the lack of space. The complete evaluation including explanation of every aspect, the fulfilling norm and the definition of point's assignment can be found under this address [14].

Table 1: Excerpt of the evaluation: the aspects and their importance rating

Aspect	Cost	License	Expendable Functionality	Fulfills Scenario Needs	Problems with Installation
Rating	5	5	3	5	1
Aspect	Additional Tools for Inst. needed	Available OS	Clear Arrangement of the UI	First Orientation	Amount of Understanding needed
Rating	1	1	3	3	4
Aspect	Scripting Editor Provided	External graphic editor needed	Handling of the graphic editor	Support for Physic, collision detection and Audio available	Cross-Platform Testing
Rating	3	3	5	5	4
Aspect	Support for Installation of SDK's provided	Cross-Platform Deployment	Additional Software needed for deployment	Specialties of each Game Engine	
Rating	4	4	3	1	

As mentioned above, only a few publications reported whether or not they used a game engine and which one they used. For gaining a more complete overview about the available game engines, it was decided to start the selection process with a list of game engines from relevant resources from the Internet. For this article a list of game engines from the Wikipedia (2016) [15] was chosen, as this list contained 51 game engines at that time. This list was completed by two lists of relevant gaming magazines [16, 17]. Afterwards the list contained 79 game engines. As a next step, all engines that do not support cross-platform development for at least Android, iOS and common Browsers were excluded. Thus, the list was reduced to 14 game engines. Three engines were not available at this time (one will be released in spring 2016, the other two were stopped in development and are now no longer available). One game engine was removed from the list as the company does not provide maintenance anymore. With ten game engines left, a research in relevant gaming forums like for example the "unity-insider forum" [18] was conducted, to include experiences from users that are more skilled in this subject. Due to reasons of capacity it is not possible to evaluate all ten game engines left. As these ten game engines can be classified into three types of game engines, it was decided to examine one of each type to gain the best overview possible. One of these types are powerful 3D Game Engines that are mostly destined for developing 3D-First-Person-Shooter games to all kinds of target platforms. As these games require very realistic graphics, physics and sounds, the engines are very powerful. But it is also possible to develop a 2D mobile game with these engines. The second type that will be evaluated are game engines with focus on the development of 2D games. They do not provide the same amount of functionality like the first group of game engines. Instead they offer different functionalities to improve the comfort in developing. Additionally the third evaluated type are game engines without a graphical editor. These game engines provide a framework that contains functionality for supporting the developer in gameplay physics, collision detection, audio etc.. The usage of them is mostly provided via integration in other development technologies like for example in HTML5.

From these three types the game engines with the most recommendation from skilled users or magazines found are chosen for the evaluation. For the first type of powerful 3D game engines he engine of choice for this evaluation, will be "Unity" [18] as it is often recommended by experienced users to be the best choice for beginners and was mentioned in several publications [7,8] as the game engine in use. For the second type the "Marmalade" engine [19] is chosen, as it was especially recommended for developing mobile games. As an engine without a graphical editor the "Turbulenz HTML5 Game Engine" [20] was chosen. This game engine was recommended as it is fully open source and provides therefore high adaptability of the engine code.

3. Results

As mentioned in Chapter 2, the game engines Unity and Marmalade provide UI's for developing games. Whilst Unity comes with one UI that contains everything needed, like for example a graphic editor, a scripting editor and the possibility to simulate scenes, the Marmalade engine has several UI's for different purposes. The user starts with the so-called "hub". This management tool helps the user to dispense his projects, the dependencies, the different SDK's, the different simulators for testing, the iOS signing request and the IDE's. For developing in C++, Marmalade needs an IDE already installed, e.g. XCode for Mac. For developing in Lua, a powerful scripting language [21], Marmalade brings his own IDE. When starting a project, the hub opens the right IDE and uses the right SDK and simulators. For developing graphics Marmalade offers his own 2D graphics editor in a beta version. Scenes that were created with this editor can be imported to the Marmalade projects.

The Turbulenz HTML5 game engine is used in HTML5 projects by including some src-tags and then using the engine in the Javascript Code. Therefore, the engine provides no editor at all. The user decides which HTML5 editor he wants to use. The same goes with a plain HTML5 project.

When creating the scene for the prototype that was described in Chapter 2, first orientation was rather complicated in every engine. Unity and Marmalade provide several Documentations and Tutorials, as well as some test projects. Still it took time to understand different views and functions available, as they were not as self-explaining as expected. The Turbulenz engine provides a very detailed documentation on how to use the engine properly, but Tutorials were rare. It is expected, that the lack of Tutorials can be explained by the comparatively low distribution of the game engines without UI.

In Unity and Marmalade, the prepared images can simply be imported or dragged into the editor. They can as well be positioned by dragging them to the right position and alternatively by giving them numbers for the X and Y value. The objects are automatically declared as sprites. The creating of a scene in the Marmalade 2D Editor can be seen in Figure 1. This handling is very comfortable and saves a high amount of time versus the Turbulenz engine and the HTML5 project. When creating scenes, the images have to be positioned by tags. This can be very exhausting since after every change of the tags, the browser has to be reloaded in order to notice any changes.

As a next step, the scene had to be animated. Therefore, the erythrocyte is supposed to be able to move and the camera is supposed to focus on it when it moves. In Unity it is possible to use complete scripts that are provided by the engine. Marmalade provides functions for creating these animations. The usage of these both are quite easy to handle. The Turbulenz engine also provides functions and in HTML5 the users have to

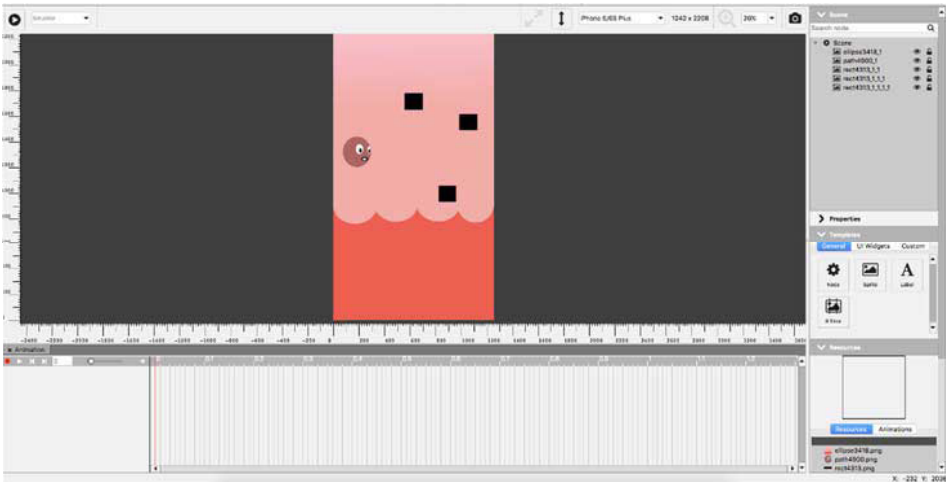


Figure 1: Creating a scene with the Marmalade 2D graphics editor

implement it all by themselves. Also in this point, the way of using them is not as comfortable as it is with the other two engines since for seeing any changes made in the code the browser has to be reloaded every time.
Apart from graphics, other important aspects of games that are not already implemented in the prototype, are physics and sounds. For realistic gameplay it is required that the objects follow physical laws like, for example, gravity, or that noise, e.g. rain, sounds realistic. All of these have to be implemented by the user, when creating a game in HTML5. Game engines normally provide these codes. In Unity this happens via characteristics of assets or via scripting code and is quite easy to handle. In Marmalade the objects can get the assignment of characteristics like gravity in the scripting code. The Turbulenz engine also provides the functionalities for physics and sounds. These have to be used in the scripting code.

When developing games cross-platform for mobile devices, a very important point is testing the game on different devices. This means not only the gameplay, but also the newly created UI, as every mobile device has a different format and the UI can appear different on each one of them. The Unity engine provides the simulation of the scene, but without considering different mobile devices. However, the user can adapt the camera view to a certain format e.g. the screen of an iPhone, but this will only adapt to one format and not to different formats of different devices. Testing on different devices can either be made with simulators or on real devices. Both of these require deployment to the target platform at first. This is also the way for testing games from the Turbulenz engine or HTML5. For testing in browsers, the Turbulenz engine comes with a testing environment that also contains a local server. For testing on mobile devices, a game has to be deployed to the target platform and then be tested either in a simulator or on a mobile device.

The Marmalade engine provides two kinds of simulations. In the 2D graphic editor, a simulator is provided that can simulate the appearance of the scene on common Android and iOS devices. In Figure 2, the current prototype scene is simulated on an iPhone 4/4S with the 2D editor. The other simulation is provided when scripting the gameplay. Here Marmalade offers the possibility not only to simulate the gameplay in the scene on the computer but also to use simulators for mobile devices. As mentioned

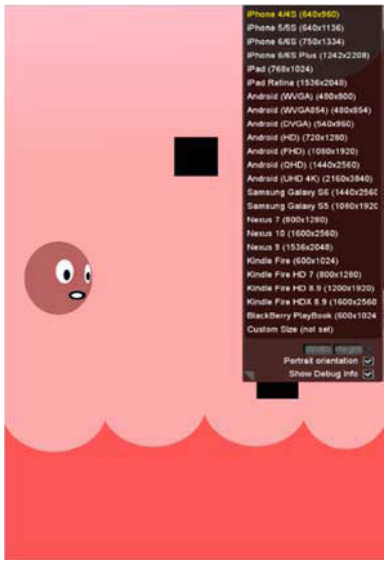


Figure 2: Simulating the scene with the Marmalade 2D graphics editor on iPhone 4/4S

before, in Marmalade the hub manages all projects. This includes also these simulators. If there are simulators installed on the PC the user can add them to the hub by pointing on their directory. Like this, the hub can use the simulator when running a project in the IDE. This is very practical and saves time as testing can be done before deploying and different formats are tested all by once.

In addition the deployment to different target platforms like iOS, Android and browsers can be very time-consuming. Marmalade and Unity support the user in this step, as the user can simply point the engines to the directory of the installed SDK's. The deployment afterwards is done by the engines. The Turbulenz engine does not provide any comfort in this step. Just as projects from plain HTML5, projects from the Turbulenz engine have to be deployed by the user with additional tools.

Finally, the results of the evaluation show a high benefit from game engines in general compared to the development in plain HTML5. Table 2 shows the overall points of each game engine and HTML5. The complete evaluation is provided under this address [14].

These results show that each game engine could reach more points than the development with HTML5. This is because HTML5 does not provide any support for the user for graphics, physics and sound components. Therefore, the game engines support the development process significantly. Additionally, the evaluation shows that game engines with graphical UI reached significantly more points than an engine without one. The graphic editor showed significant benefits while implementing, therefore the Marmalade and the Unity engine gained more points. When looking at the aspect of provides support for a great number of different mobile target devices and helps to manage the projects for them. It supports in deploying one project on different devices

Table 2: The results of the evaluation.

Game Engine	Marmalade	Unity	Turbulenz	HTML5
Total Points	585	550	402	329

cross-platform development, the engine of choice is the Marmalade engine. This engine and provides simulation of scene and gameplay on them. Therefore, as the aim of this evaluation was support in cross-platform development, the Marmalade engine got the most points in the evaluation and is the best choice for developing a cross-platform mobile Serious Game for Health.

4. Discussion

The evaluation shows that the development of games highly benefits from using game engines especially in the fields of graphics, physics and sounds. The game engines partial reached three times (Unity, Marmalade) the points of the development with HTML5. This can clearly be interpreted as major benefits from the game engines.

Also the cross-platform development can be highly supported by game engines. The evaluation aspects "cross-platform testing" and "cross-platform deployment" considered this mainly. The engine that provides the most support for this development is the Marmalade Engine that reached the most points in the complete evaluation as well (550). The engine supports simulation on different devices, management of projects, simulators, SDK's etc. This functionality, provided by the hub, can save a great amount of time, as testing and deploying do not have to be made for each device on its own. In addition, the simulation of the scene in the 2D-editor is very helpful to test the graphics on different screen sizes immediately.

The Unity engine also supports cross-platform development and reached the second most points (516). The game can be deployed to different target platforms, but as the simulation of the scene on a mobile device can only be done after deployment, the Marmalade engine is the better choice here. However, since the Unity engine is very powerful in 3D graphics, it may be a good choice when using 3D graphics.

The Turbulenz engine is not as comfortable to handle as the other engines and reached only the third most points (397). The main reason is the absence of a graphic editor, but as well to the lack of testing. The latter can be achieved in a provided testing environment with a local server, which is not as comfortable and effective as testing directly in the devices formats. The game has to be deployed as well first for being tested in the right simulator.

Therefore, the recommendation out of this evaluation is to use the Marmalade Engine for cross-platform development of Serious Games for health. As the market of game engines is highly dynamic it is definitely needed to repeat this evaluation in the future. If the capacity is provided it is also possible to evaluate other game engines that come into consideration.

The evaluation can also be expanded to more engines when the scenario of target platforms is changed. As there are many game engines available that can be used for developing for iOS and Android but not for browsers, the list of engines would be larger if browser would be taken off the scenario. However, this is not practical for a Serious Game for medical education, as it cannot be assumed for reasons of equalization that every student has access to mobile devices.

References

[1] Kato, P., Cole, S., Bradlyn, A., Pollock, B., A Video Game Improves Behavioral Outcomes in Adolescents and Young Adults with Cancer: A Randomized Trial, PEDIATRICS 122(2) (2008), 305-317.

[2] Bartolomé, N., Zorilla, A., Zapirain, B., Can Game-Based Therapies be trusted? Is Game-Based Education effective? A systematic review of the Serious Games for Health and Education, 16th International Conference on Computer Games (CGames) (2011), 275-282.

[3] Boeker, M., Andel, P., Seidl, M., Streicher, A., Schneevoigt, T., Dem, P., Frankenschmidt, A., Uro Islands I – Game-based E-learning in der Urologie, GMS Med Inform Biom und Epidemiol 5(1) (2009), Doc03.

[4] Ushaw, G., Davison, R., Eyre, J., Morgan, G., Adopting Best Pracitces from the Game Industry in Development of Serious Games for Health, Proceedings oft he 5th International Conference on Digital Health (2015), 1-8.

[5] Gregory, J, Game Engine Architecture, CRC Press, Boca Raton, 2014.

[6] Hassan, M., Hossain, M., Alamri, A., Hossain, M., A., Al-Qurishi, M., Aldukhayyil, Y., Ahmed, D., A cloud-based serious games framework for obesity, Proceedings oft he 1st ACM multimedia international workshop on Cloud-based multimedia applications and services for e-health(2012), 15-20

[7] Porcino, T., Strauss, E., Clua, E., Hugo against dengue: a serious game to educate people about dengue fever prevention, 2014 IEEE 3rd International Conference on Serious Games and Applications for Health (SeGAH) (2014), 1-5.

[8] Assigana, E., Chang, E., Cho, S., Kotecha, V., Liu B., Turner, H., Zhang, Y., Christel, M., Stevens, S., TF-CBT triangle of life: a game to help with cognitive behavioral therapy, Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play (2014), 9-16.

[9] Marks, S., Windsor, J., Wünsche, B., Evaluation of game engines for simulated surgical training, 5th International Conference on Computer Graphics and Interactive Technique (GRAPHITE) (2007), 273-282.

[10] Inkscape Community, Inkscape, https://inkscape.org/de/, last access: 19.02.2016.

[11] Barbosa, A.F.S., Silva, F.G.M. (2011): Serious Games – Design and Development of OxyBlood, Proceedings oft he 8th International Conference on Advances in Computer Enterteinment, Technologie (2011), 1-8.

[12] Keeney, R.L., Raiffa, H. (1976): Decisions with Multiple Objectives, Preferences and Value Tradeoffs, Camebridge, Camebridge University Press.

[13] Computer-Bild Spiele, http://www.computerbild.de/cbs/, last access:07.03.2016.

[14] Kleinschmidt, C., Haag, M. , Evaluation von Game Engines für die Cross-Plattform Entwicklung von Serious Games for Health: Detailübersicht, www.virtuellepatienten.de/Evaluation_Game_Engines.html, last access: 13.03.2016.

[15] Wikipedia, Liste von Spiele-Engines, https://de.wikipedia.org/wiki/Liste_von_Spiel-Engines, last access: 14.01.2016.

[16] Develop, The top 16 game engines for 2014, http://www.develop-online.net/tools-and-tech/the-top-16-game-engines-for-2014/0192302, last access: 11.01.2016.

[17] Appindex, The Big List of Mobile Game Development Tools, Engines, Guides and Resources, http://appindex.com/blog/big-list-mobile-game-development-tools-engines-guides-resources/, last access: 11.01.2016.

[18] Unity insider forum, http://forum.unity-community.de/, last access: 11.01.2016.

[19] Unity Technologies, Unity Engine, https://unity3d.com/, last access: 11.01.2016.

[20] Marmalade Technoligies Ltd., The Marmalade Engine, https://www.madewithmarmalade.com/, last access: 11.01.2016.

[21] Turbulenz Limited, The Turbulenz HTML5 Game Engine, http://biz.turbulenz.com/, last access: 11.01.2016.

[22] PUC RIO, About Lua, http://www.lua.org/about.html, last access: 18.01.2016.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-215

Integrated Patient Education on U.S. Hospital Web Sites

Edgar HUANG[a,1], Kerong WU[a] and Kelsey EDWARDS[a]
[a]*School of Informatics and Computing, Indiana University, USA*

Abstract. Based on a census of the 2015 Most Wired Hospitals, this content analysis aimed to find out how patient education has been integrated on these best IT hospitals' Web sites to serve the purposes of marketing and meeting online visitors' needs. This study will help hospitals to understand where the weaknesses are in their interactive patient education implementation and come up with a smart integration strategy. The study found that 70% of these hospitals had adopted interactive patient education contents, 76.6% of such contents were from a third-party developer, and only 20% of the hospitals linked their patient education contents to one or more of the hospital's resources while 26% cross-references such contents. The authors concluded that more hospitals should take advantage of modern information communication technology to cross-reference their patient education contents and to integrate such contents into their overall online marketing strategy to benefit patients and themselves.

Keywords. Patient education, hospitals, Web sites, integration.

1. Introduction

Strategically investing in information technology (IT) serves as a hospital's commitment to all parties in healthcare, including patients [1]. IT can impact patients' healthcare-related decisions and trust in caregiver [2]. By 2011, 47% of the U.S. hospitals had provided interactive patient education contents on their Web sites [3]. By examining the 2015 Most Wired Hospitals, this study aimed to find out how patient education has been integrated on these best IT hospitals' Web sites to serve the purposes of marketing and meeting online visitors' needs.

Patient education, provided by healthcare professionals to patients and their loved ones in the forms of text, still images, videos, games, and so on, has been understood as an important part of IT development in a hospital because it can improve patient care, reduce hospital readmission rates, gain more patients, meet the regulatory compliance, cut cost for the hospital, insurers and the patients' employers, decrease administrative tasks, and increase the overall efficiency within the healthcare market [4]. Many hospitals have made their online patient education interactive by using different kinds of online technology, such as relational databases, PHP, JavaScript, HTML forms, links, and so on so that an online visitor can easily interact with content, choose content, and make a decision by clicking a provided links [5]. Patient education can greatly influence patients during their decision-making process; the lack of such education can put a patient at risk and a hospital's revenue at risk [6]. The benefits of providing interactive patient education have been tested and documented by many studies [2,7]. Today,

[1] Corresponding Author: Edgar Huang, Indiana University, 535 W. Michigan Street, IT481, Indianapolis, IN 46074, USA, E-Mail: ehuang@iupui.edu

patients are encouraged by providers to be full partners and seen as an integral part of the healthcare team due to the equal accessibility to healthcare information for patients and for physicians [4,8]. In his study regarding the challenges in asthma patient education, Cabana [9] pointed out that "many of the recommended components of asthma care might not be effective without adequate patient education" (p. 1225). A 2008 survey among the Most Wired Hospitals found that patients favored hospitals with advanced IT [1].

Nevertheless, many hospital Web sites, for different reasons, still have not seen the benefits of incorporating patient education [10]. It was found that more than 40% of e-health sites operated by hospitals and healthcare systems offer little or no consumer-focused health information [11]. As a result, users may pass over most hospital Web sites due to lack of relevant or useful health information [12] Researchers are still attempting to find out to what extent hospital Web sites are helping hospitals to be the cornerstones of patient education [7]. This study will help hospitals come up with a smart strategy to integrate patient education with their overall marketing strategy so that they can best serve their patients and serve themselves.

2. Literature review

A hospital Web site has been traditionally regarded as a marketing tool [12,13]. Hastings and Saren [14] pointed out that hospitals should work with their patients to reach a mutually beneficial way forward instead of simply seeking hospitals' own interests. Huang [15] concluded, "Patient education provided on hospital Web sites is, in fact, implicitly altruistic marketing. Such information could show to the healthcare information seekers that the information provider cares about them" (p. 57).

Many hospitals today are allowing their online visitors to find specific patient education through interactive mechanisms and multimedia. Huang and Chang [3] found that 21 interactive e-health tools used by U.S. hospitals in 2011 and that 47% of the hospitals provided interactive patient education contents. They also found that overall, larger hospitals were more likely than smaller hospitals to provide interactive tools on their Web sites (Ibid.). Lustria [16] found that interactivity could significantly boost comprehension as well as attitudes toward a hospital's Web site. Studies over the years have repeatedly and positively correlate interactivity to customer satisfaction and conversion rates [17,18]. Interactive patient education provides the opportunity for increased user engagement with health information and best meets users' needs and expectations [19]. "In so doing, hospitals are increasingly seeking to take on the role of trusted adviser, a role that is closely aligned with the accountable care organization (ACO) model in which health care providers work to empower patients to improve population health" [20].

Scholars have promoted the integration approach when developing a hospital Web site [12,21]. Integrated Web design, Stoop, Riet & Berg [22] believe, can offer surplus value to the available education means. Kransnoff and Loubeau [23] argue, "The most effective Web sites are interactive and provide a wealth of assistance to patients while enhancing the institution's marketing effort" (p. 56); they maintain that "[t]he ultimate hospital Web site should be a single, comprehensive source of information that balances the consumer's need for quality information and interactivity with the hospital's desire to attract customers, increase market share and build the bottom line" (p. 58). From the organizational perspective, Campbell, Sherry and Sternberg [11] suggest that all

departments in a hospital, not just the marketing department, should demonstrate ownership to its Web site development; they maintain, "Integration means the e-health Web site is an integral part of the operating practices of the organization and is part of the daily activities of a large number of staff and physicians" (p. 40).

Some large hospitals, such as Mayo Clinic and Cleveland Clinic, have relied on themselves to develop numerous copyrighted patient education materials for their Web sites. Nevertheless, creating such materials is resource-intensive and time-consuming, and many hospitals do not have the manpower to do such a job; therefore, many healthcare marketers have spent $15,000 to $20,000 a year to license interactive patient education contents from third-party content developers or use free health information from MEDLINEplus, sponsored by the National Institutes of Health [21,24]. Huang and Chang [3] found that 10% of U.S. hospitals licensed such third-party patient education contents in 2011 and that large and medium-sized hospitals were more likely to use third-party contents than smaller hospitals did.

Based on the literature review, the general research question of this study was how the 2015 Most Wired Hospitals have integrated patient education on their Web sites.

Here are five specific research questions based on the general research questions:

1. How consistent are hospitals in naming patient education contents so that patients can easily identify and find such contents?
2. How many hospitals have incorporated interactive videos and tools in their patient education contents, and calendars for local patient education classes/events?
3. How many hospitals have integrated patient education contents, either internally or externally developed, into their marketing efforts?
4. How have hospital size and university affiliation status affected hospitals' patient education development and the integration of such contents in their marketing efforts?
5. What are the best practices in interactive patient education?

3. Methods

In this study, interactive patient education was defined as healthcare information that informs an online visitor regarding conditions, diagnosis, procedures, drugs, wellness, etc. and such non-hospital-specific information is presented under a menu name and in the form of 1) articles, 2) videos, or 3) patient education tools, including calculators (i.e. BMI calculator, ovulation date calculator), quizzes, health risk assessors, or animated navigator tools (e.g. anatomy navigator, conditions navigator).

Since the purpose of this study was to examine patient education contents on hospital Web sites, content analysis naturally became the research approach. Although there are close to 6000 hospitals in the United States, [25] this study aimed to find out how the hospitals that had best taken advantage of information technology in the United States had integrated interactive patient education on their Web sites; therefore, the study was based on a census of the 326 hospitals titled the 2015 Most Wired Hospitals. [26] Based on the number of beds, the hospitals were sorted into four categories: Small hospitals (1–200 beds), medium-sized hospitals (201–500 beds), large hospitals (501–2000 beds), and mega hospitals (more than 2000 beds).

As Hanif et al. [27] mentioned some differences regarding the quality of patient education contents developed by government/universities and commercial developers, consequently, in this study, all hospitals were coded as university-affiliated and non-

university-affiliated based on the relevant information presented on the hospitals' Web sites in an attempt to detect whether university-affiliated hospitals tended to develop their own contents.

Hospital size and university affiliation were the two independent variables. The dependent variables observed included what the menu names were for major patient education content, such as healthcare library, whether the hospital used its own patient education content or used the third-party content, whether patient education contents were associated with the hospital's departments or doctors, and whether patient education contents included videos, class announcements/calendars, or interactive tools, such as various kinds of health calculators, quizzes, health risk assessments, and symptom navigators.

The coding was conducted in the fall of 2015 by two coders. The coders went through multiple rounds of coding training and pilot studies both conducted by the correspondence author over five weeks to reach agreements. After the coding was completed, the coders compared the coding work and adjusted most of the coding that differed. Eventually, the Scott's Pi was on average 0.97.

Since the data were based on a census, both descriptive statistics and inferential statistics were employed in data analysis, which was conducted in SPSS.

4. Findings

Here are the demographic data regarding the hospital sizes (small: 27.3%, medium-sized: 31.3%, large: 34%, mega: 7.4%) and university affiliation type (university-affiliated: 30%, non-university-affiliated: 70%).

4.1. How consistent are hospitals in naming patient education contents so that patients can easily identify and find such contents?

Out of these 326 Most Wired Hospitals in 2015, 70% carried interactive patient education contents clearly listed under a menu name or immediately featured on the home page; 2% carried no text-based patient education contents listed under a menu name but showed patient education videos or interactive tools only. More hospitals (83.3%) posted local patient education class or event information on their Web sites. The following data are all based on the 228 hospitals (70%) that did carry interactive patient education contents.

In total, 77 menu names were used to present such contents. The most popular name was Health Library (20.9%), the second most popular was Health Information (7.6%); Health and Wellness (4.3%) and Health Resources (4.3%) were a tie as the third most popular.

4.2. How many hospitals have incorporated patient education videos and tools in their patient education contents, and calendars for local patient education classes/events?

Fifty-two percent of the hospitals used patient education videos (third-party: 82.3%, internally developed: 16.1%, both: 1.6%) and 64.5% provided patient education tools (third-party: 89.1%, internally developed: 9.5%, both: 1.4%).

4.3. How many hospitals have integrated patient education contents, either internally or externally developed, into their marketing efforts?

In terms of patient education articles, 23.4% hospitals produced their own while 76.6% used the articles written by a third-party company or, rarely, a government agency.

Twenty percent of the hospitals linked their patient education contents to one or more of the hospital's resources, including departments and services (13.3%), doctors (8.4%), local patient education classes and events (2.7%), and hospital locations (2.2%). More hospitals (26%) cross-referenced their patient education contents to patient education tools (21.6%), news and articles (20.7%), diseases and conditions (20.3%), drug and supplement information (18.1%), and tests and procedures (17.2%).

4.4. How have hospital size and university affiliation status affected hospitals' patient education development and the integration of such contents in their marketing efforts?

Multiple Chi-Square tests show that hospital size

- has no correlation with hospitals' rate of adopting interactive patient education,
- does not affect hospitals' ways of associating patient education contents to their hospital resources, and
- shows no significant difference between developing their own contents and adopting third-party contents.

On the other hand, significantly more non-university-affiliated hospitals used third-party patient education contents and significantly more university-affiliated hospitals developed their own contents (Chi-Square=3.6, df=1, $p<0.05$).

4.5. What are the best practices in interactive patient education?

The following hospitals can provide a glimpse of the best practices on using interactive patient education. UC San Diego Health (health.ucsd.edu) had 586 beds. Its interactive patient education is under the menu item Health Info, which housed classes and events, Health Encyclopedia, a video library, Interactive Tools, and so on. The Web site used 99 videos in its video library and more than 100 patient education tools. Although all patient education contents, except for classes and events, were licensed from a third-party company, the design of the Health Encyclopedia was contextual. On the right-hand side of most of the articles in the Encyclopedia, the hospital's physicians, specialists, departments, services related to a term were listed. In addition, other patient education elements related to that term, such as diseases and conditions, tests and procedures, articles, news, interactive tools, drug references, were also listed below the physician photos. Medical University of South Carolina (MUSC) (muschealth.org), with 710 beds, used almost an identical interactive patient education module as UC San Diego Health's, and both were university-affiliated hospitals; however, a significantly difference is that MUSC built its Health Library completely on the third-party site though the top menu could bring a user back to the hospital site easily.

Genesis Healthcare System (genesishcs.org) had 465 beds. As a non-university-affiliated hospital, it had also highly integrated third-party patient education contents on its own Web site. Numerous article pages in its Health Library were integrated for marketing purpose. For instance, if a user is surfing a topic on "Men's Health," the pictures and names of and links to the physicians from this hospital who were related to

men's health were listed on the right-hand side so that a user could pick one to continue to the next step. In addition, below and to the lower-right of the article were a wealth of more related information, including Health Topics, Medical Tests, Medications, Make a Wise Decision, Interactive Tools, and Symptom Checker. In short, a user's surfing experience could be highly navigated. Mary Greeley Medical Center (mgmc.org), another non-university-affiliated hospital with 190 beds, used almost an identical model to Genesis Healthcare System's except that below the physicians' photos, names, and links, there was a phone number for First Nurse.

5. Discussion and Conclusions

The 70% adoption rate for adopting interactive patient education contents among the 2015 Most Wired Hospitals is higher than the 47% adoption rate among all the U.S. hospitals in 2011, as revealed in Huang and Chang's 2012 study. However, this new adoption rate shows that interactive patient education on hospital Web sites is far from being ubiquitous among these advanced IT adopters; it is reasonable to deduce that there is still a long way to go for U.S. hospitals to compete with Google, WebMD or the like through interactive patient education for traffic. Hospitals need to more aggressively take the initiative to provide such basic information to cater to their online visitors' needs.

What menu name to use for interactive patient education contents may sound trivial, but online visitors sometimes shop around, and helping them easily find what they want probably is something hospitals can easily entertain. An Education & Events menu item that covers only local events can be misleading, but it is implemented so on a Most Wired Hospital's Web site. Additional Resources or Fast Health sounds vague and confusing. Sometimes, Education in a menu on a university-affiliated hospital Web site means degree programs and has nothing to do with patient education. Therefore, coming up with an industry standard for naming for interactive patient education contents could help.

Huang [28] found that, by the end of 2008, 42% of the hospitals that presented online videos contained patient education videos. After seven years of development, even the 2015 Most Wired Hospitals did not seem to get much of an edge with a 52% adoption rate. Of all the digital and social media tools available to online users, hospital marketers say online video is one of the most effective tools because, after patients have watched a video from a credible source, 60% went on to make a direct contact with the physician or hospital featured in the video [29]. Therefore, integrating videos in patient education contents can do both patients and hospitals a big favor, and the adoption of patient education videos on hospital Web sites should be expedited.

So far, no data from earlier studies can serve as benchmark for the adoption of patient education tools on hospitals' Web sites. The 64.5% adoption rate has certainly left much room for growth. Since several third-party developers have extensively developed such very useful tools for users and widely adopted by these Most Wired Hospitals (89.1%), any hospital can just include them in their licensing.

It is understandable that most of the hospitals do not have the manpower to develop their own patient education contents such as health encyclopedia, and it is probably unnecessary for so many hospitals each to revamp wheels. Therefore, licensing a health library is the right decision for most organizations. Nevertheless, both the low rates of cross-referencing patient education contents (26%) and integrating such contents into a hospital's marketing effort (20%) are crying for big improvement. Integration will both help a hospital find the patients it needs and help patients find the doctors and other

service information they need. Integration brings a win-win situation. O'neill [47] maintains: "While attaching a health content library is a necessary step to providing a truly integrated experience, settling only for an attached library is a compromise that provides minimal utility to site visitors and a poor return on the investment you've made licensing quality health content." O'neill [47] suggests that such third-party content be customized to best fit a hospital's needs and brand. Today's content management systems (CMS) can use taxonomy to automatically promote the hospital's physicians and departments most related to the concept mentioned in a health library article (Ibid.), and no manual work is needed. Campbell, Sherry and Sternberg [20] said, "Each component of the marketing mix must be considered—from public promotion of the site to creating awareness among hospital staff to the site's appearance and navigation" (p. 42). Therefore, more hospitals should take advantage of modern information communication technology to cross-reference their patient education contents and to integrate such contents into their overall online marketing strategy while maintaining the integrity of a hospital's branding by keeping all third-party contents on the hospital's own server.

The best practice examples from this study show that, whether a hospital is large or small or whether it is affiliated to a university or not, it can come up with a user-friendly interactive patient education package to attract and serve its online visitors. While the hospitals provide abundant local patient education classes and events, they should consider enhancing their online patient education presence to reach farther than their geographic restriction.

This study has its limitations. The data are based on the elite hospitals in using IT and cannot be extrapolated to all the hospitals in the United States. Most crucially, no study has investigated how online visitors have used interactive patient education contents on hospital Web sites, but such a study is very much needed in the near future.

Acknowledgement

The authors would like to thank Stephanie Huang, a student at the University of Houston Law Center, for her proofreading and valuable insight into the improvement of this manuscript.

References

[1] A. Solovy, S. Hoppszallern, S. B. Brown, Are patients Happier at Most Wired Hospitals?, *Hospitals & health networks/AHA* ***82***(7) (2008), 30--45.

[2] S. Fox, S. Jones, The social life of health information: Americans' pursuit of health. Pew Internet, http://www.pewinternet.org/Reports/2009/Generations-Online-in-2009.aspx, last access: 10.10.2015.

[3] E. Huang, C. A. Chang, Patient-oriented interactive e-health tools on US hospital Web sites, *Health marketing quarterly*, ***29*** (4) (2012), 329.

[4] A. Mukherjee, J. McGinnis, E-healthcare: An analysis of key themes in research, *Journal of Pharmaceutical and Healthcare Marketing* **1** (4) (2007), 349-363.

[5] T. Hampton, Media lab uses videos, comics, and more to help people understand health issues, *JAMA* **307** (16) (2012), 1679-1680.

[6] H. A. Liszka, T. E. Steyer, W. J. Hueston, Virtual medical care: How are our patients using online health information?, *Journal of community health*, **31** (5) (2006), 368-378.

[7] L. M. Gallant, Promoting participatory medicine with social media: New media applications on hospital Websites that enhance health education and e-patients' voices, http://www.jopm.org/evidence/research/2011/10/31/promoting-participatory-medicine-with-social-

media-new-media-applications-on-hospital-websites-that-enhance-health-education-and-e-patients-voices, last access: 10.10.2015.
[8] W. B. Lober, J. L. Flowers, Consumer empowerment in health care amid the Internet and social media, *Seminars in oncology nursing* **27** (3) (2011), 169-182.
[9] M. D. Cabana, T. T. Le, Challenges in asthma patient education, *Journal of allergy and clinical immunology* **115** (6) (2005), 1225-1227.
[10] R. Andis, Online and on target: Hospitals add sophistication to their websites, improving value to patients, *Modern Healthcare* **41** (7) (2011), 32-33.
[11] G. S. Campbell, D. Sherry, D. J. Sternberg, A hospital Web site that works, *Marketing Health Services* **22** (2) (2002), 40-42.
[12] E. Huang, D. Bolchini, J. F. Jones, Users' consumption of healthcare videos on hospital web sites, *International Journal of Pharmaceutical and Healthcare Marketing* **5** (1) (2011), 43-53.
[13] L. Bishop, B. J. Holmes, W. McEnroe, Hospital web marketing takes off, http://www.forrester.com/Research/Document/Excerpt/0,7211,37998,00.html, last access: 10.10.2015.
[14] G. Hastings, M. Saren, The critical contribution of social marketing: Theory and application, *Market Theory* **3**(3) (2003), 305-322.
[15] E. Huang, Six cases of e-health videos on hospital web sites, *E-Service Journal 6* (3) (2009b), 56-71.
[16] M. L. A. Lustria, Can interactivity make a difference? Effects of interactivity on the comprehension of and attitudes toward online health content, *Journal of the American Society for Information Science and Technology* **58** (6) (2007), 766-776.
[17] H. Lin, The impact of Website quality dimensions on customer satisfaction in the B2C e-commerce context, *Total Quality Management & Business Excellence* **18**(4) (2007), 363-378.
[18] K. Singh, Top interactive Web design features, http://www.site-reference.com/articles/Website-Development/Top-Interactive-Web-Design-Features.html, last access: 10.10.2015.
[19] L. Gallant, C. Irizarry, G. L. Kreps, User-centric hospital Web sites: A case for trust and personalization, *e-Service Journal* **5** (2) (2006), 5-26.
[20] T. R. Huerta, J. L. Hefner, E. W. Ford, A. S. McAlearney, N. Menachemi, Hospital Website rankings in the United States: Expanding benchmarks and standards for effective consumer engagement, *Journal of Medical Internet Research* **16** (2), e64 (2014), http://doi.org/10.2196/jmir.3054, last access: 10.10.2015.
[21] R. Cosentino, S. Haimowitz, Surfers & seekers: What to put on your hospital Web site, *Radiology Today* ***5*** (19) (2004), 22.
[22] A. P. Stoop, A. V. Riet, M. Berg, Using information technology for patient education: Realizing surplus value?, *Patient Education and Counseling* **54**(2004), 187-195.
[23] A. R. Kransnoff, P. R. Loubeau, Analysis of New York area hospital web sites reveals good basic features but ample room for improvement, *Healthcare Informatics* **18** (10) (2011), 55-59.
[24] M. O'neill, The eight steps of a successful health content library integration, http://geovoices.geonetric.com/2014/07/health-content-library-integration, last access: 10.10.2015.
[25] American Hospital Association, Fast facts on US hospitals, http://www.aha.org/research/rc/stat-studies/fast-facts.shtml, last access: 10.10.2015.
[26] American Hospital Association, HealthCare's Most Wired Awards, http://www.hhnmostwired.com/winners/index.dhtml, last access: 10.10.2015.
[27] F. Hanif, J. C. Read, J. A. Goodacre, A. Chaudhry, P. Gibbs, The role of quality tools in assessing reliability of the Internet for health information, *Informatics for Health and Social Care* **34** (4) (2009), 231–243.
[28] E. Huang, Diffusion of videos on US hospital web sites, *International Journal of Pharmaceutical and Healthcare Marketing* **3**(4) (2009a), 347-362.
[29] P. L. Dolan, Online videos may offer most bang for medical marketing buck [American Medical News], http://www.amednews.com/article/20110822/business/308229968/5/, last access: 10.10.2015.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-223

User Interface Design in Medical Distributed Web Applications

Alexandru SERBAN[a,b,1], Mihaela CRISAN-VIDA [b],
Leonard MADA[c] and Lacramioara STOICU-TIVADAR [b]
[a]*Department of Medical Informatics,*
University of Medicine and Pharmacy Victor Babeş Timişoara
[b]*Department of Automation and Applied Informatics,*
University Politehnica Timisoara, Romania
[c]*Syonic SRL, Timisoara, Romania*

Abstract. User interfaces are important to facilitate easy learning and operating with an IT application especially in the medical world. An easy to use interface has to be simple and to customize the user needs and mode of operation. The technology in the background is an important tool to accomplish this. The present work aims to creating a web interface using specific technology (HTML table design combined with CSS3) to provide an optimized responsive interface for a complex web application. In the first phase, the current icMED web medical application layout is analyzed, and its structure is designed using specific tools, on source files. In the second phase, a new graphic adaptable interface to different mobile terminals is proposed, (using HTML table design (TD) and CSS3 method) that uses no source files, just lines of code for layout design, improving the interaction in terms of speed and simplicity. For a complex medical software application a new prototype layout was designed and developed using HTML tables. The method uses a CSS code with only CSS classes applied to one or multiple HTML table elements, instead of CSS styles that can be applied to just one DIV tag at once. The technique has the advantage of a simplified CSS code, and a better adaptability to different media resolutions compared to DIV-CSS style method. The presented work is a proof that adaptive web interfaces can be developed just using and combining different types of design methods and technologies, using HTML table design, resulting in a simpler to learn and use interface, suitable for healthcare services.

Keywords. User interface design, web interface design, application analysis, HTML

1. Introduction

Modern Web applications have DIV technology based layout design, combined with CSS (Cascade style sheet) programming language, which gives properties to HTML elements in the document. Many developers prefer this technology because it uses a simplified HTML code. On the other hand, CSS files containing classes, styles and properties can become very complex in coding lines, especially when the layout is developed to be adaptable to different resolutions, like concept of “cross browser”, "responsive design", "fluid design" [1]. HTML technology TD (table data) and TR (table rows) is considered by many obsolete, due to the many HTML lines of code needed for building a web layout interface. Besides is considered outdated, it is still used to build and maintain a fixed resolution graphical interface that is NOT adjustable on mobile

[1] Corresponding Author: Alexandru SERBAN, University Politehnica Timisoara, Piata Victoriei,nr 2, Timisoara, Romania, E-Mail: alex.serban81@yahoo.com

tablet or smart phone [2]. One of the advantages of TD versus HTML DIV, is that it can be very well optimized for mobile applications without using multiple properties CSS files or CSS classes for each HTML element. DIV tag technology requires creating a separate class for each style or CSS property of a column or line within the HTML document. This is not necessary in TD HTML based design.

The existence of table attributes makes for a rather flat learning curve because the developer doesn't have to use a separate style sheet. With DIVs, the developer must use the style attribute or an external style sheet, because the div tag doesn't have any attributes attached to it, compared to TD that has his own attribute properties [3]. Also, tables don't break when the content is too wide. Columns are not squeezed under other columns as they are in a div-based structure. This adds to the feeling that using tables is safe [4].

Using a DIV for structure can make a page more fragile when content is pushing the div to its limit. It can also force columns to fall under each other. But this is usually only true for older browsers (particularly IE6); newer browsers make content flow to the next column.

Every extra DIV the developer adds, makes layout code harder to read by browser and difficult to edit by another designer. More lines of code lead to longer download times and so on. On the other side, HTML table based design, despite that is more complex in terms of coding, is more developer friendly, because each table has his own properties like "width and height" and the developer is not forced to edit the CSS files to change the attributes and styles of a table. Besides that, table based design code can easily edited using web editors like Adobe Dreamweaver, that is a very professional web editor used by most of the web developers in present days [5].

Using HTML table design can provide a more efficient way of flexibility with different media devices, as long as the design is separated from the content and inserted in the CSS style sheet [6,7]. Also a table design layout is easy to maintain because offers the possibility to move or insert new rows and columns without touching the CSS code. It also provides a reliable way to manipulate the objects and image placeholders in the document, using only the TD properties. This things can't be applied in DIV based designs, since DIV is dependable on CSS styles [4].

DIV is more less code required to achieve the same result. DIVs also give you more control of how things should look. But it does take more time to decorate it and require more testing with different browser, such as Firefox, IE and Chrome. Different tests need to be done with different version of browsers, as well as Operating systems. A page could look different on a Mac with Firefox or on Windows with Firefox using DIV based design. However because DIV based layouts are dependable on external CSS language, getting the CSS to act the way that you want, can be a real challenge [8].

The TABLE based design, is a still a good way to design and build web interface layouts, because is more user friendly, each table has his own property and is not dependable on CSS styles [9].

icMED is a web-based medical information patient management system, which works using interconnection between offices, clinics, medical centers, hospitals, laboratories, pharmacies and emergency centers. icMED provides performance indicators, which are budget based information input by physicians, then centralized on a web server. Other advantages of application icMED are the automatic completion of medical records, electronic patient records, reports to insurance offices etc. icMED is an innovative concept and an effective tool for managing patient treatments by healthcare professionals, clinics or laboratories. This application, by its nature and complexity,

offers users full support for deployment in optimal conditions and fair activities, using a common database of unique electronic files. icMED is a product that provides professional services for healthcare providers, making it an efficient modern tool, rather than the classic methods of pen writing specific documents, tracking their epidemiological indicators, activity or budget.

icMED is a widely system used in Romania by medical centers, hospitals or private practice that supports exercise healthcare system in public and in private centers. The online application offers technical support via a call center at no additional cost. It is currently the only Romanian project which has been designated as "good practice" in the European program PEOPLE. PEOPLE program rewards the best initiatives and projects in the field of improving workflow, increased service quality and improving the economic situation for all levels of stakeholders.

The aim of the current study is to prove that using HTML table design to build adaptive media web interfaces is a reliable alternative to the more complicated DIV tag layout structures. From the designer perspective, using HTML table design to develop web layouts is easier, in relation to the structure of the code and the ease to edit, in comparison to DIV structures.

2. Methods

2.1. Analysis of the current interface

A screenshot of current index layout page of the icMED medical application is shown in Figure 1. Currently, the icMED web interface application is developed using DIV tags and javascript codes and scripts. The current interface has been designed 8 years ago and has been constantly improved. Given that 8 years ago, adaptability to different media resolution was at the beginning, the online application was designed to be viewed especially from desktop computers and laptops. Although being an online application, which can be accessed also from media devices, it was developed as a standard HTML fixed web page, with fixed resolution. The graphical interface is designed to be viewed from a standard resolution of 1024x768 pixels. Being a very complex application, with many subpages, it is addressed especially for users from the medical system, as doctors, specialists etc.

Many dynamic elements of the current application are incompatible with the modern media technologies, as the FLASH elements incompatible with mobile operating systems. If the application is accessed from a mobile device, flash elements are not displayed correctly. Moreover, mobile browsers do not even load the page containing embeded objects that cannot be interpreted by the browser.

The application is designed on a framework based on the source graphic files, which are stored on a web server. These items are loaded into the interface, expanding access time. At the same time, using graphic images for the interface design, the application code becomes very rich, and increases the risk of encountering display errors in different browsers. Features like colors, frames, text or icons are present in the interface as separate graphics. A real disadvantage for users is that these items are not properly interpreted by the entire browser, especially on mobile browsers, making the application harder to navigate, especially on mobile devices. This is a real disadvantage for the mobile based users.

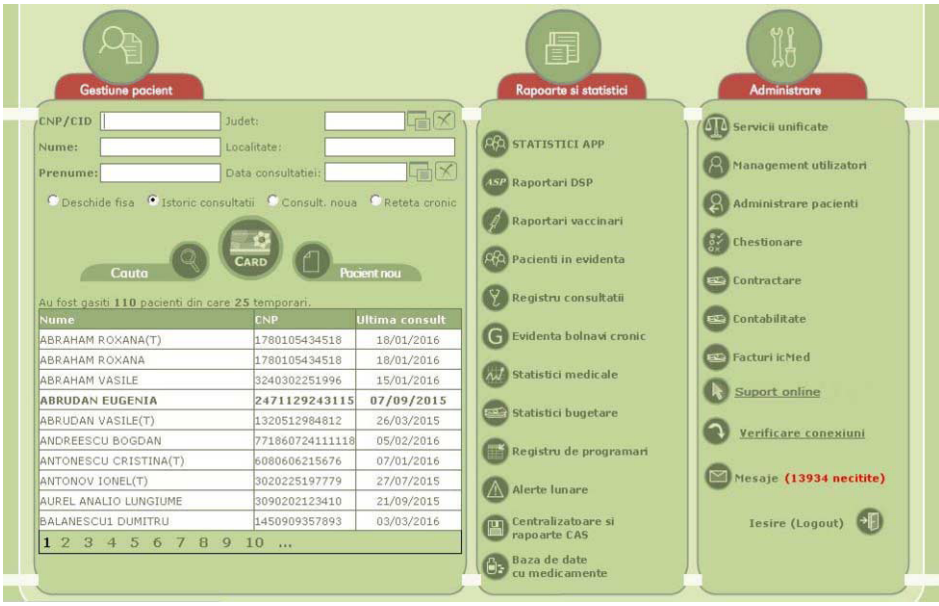


Figure 1. Current Index layout page of the icMED medical application

8.000 users are currently using the icMED application on a daily basis, and many consider the application to be outdated, in terms of layout design and adaptability to media screens (see Figure 1). The most frequent usability errors encountered by the users are: broken links, missing images sources and HTML overlay text in some subpages, long page loading time and interpretation errors on mobile browsers. Therefore in collaboration with the icMED research department, a new prototype layout was proposed using HTML TD technique in order to develop a more efficient layout that fits on every media screen (fluid design, responsive design, cross-browser design) and is developed using only HTML code for design, instead of using graphic image sources.

2.2. New Layout proposal

After a short technical analysis of the current interface, in which improving engineering opportunities were highlighted in terms of design and development, a new layout proposal has been made for a modern interface with "fluid design" technology, using HTML table coding in combination with CSS3, to be more dynamic and efficient. Using HTML table and CSS3 technology resulted in a graphical interface developed exclusively from code, without additional graphics or images in the source, improving the interface dynamics in terms of accessing time and display across different types of browsers, especially on mobile devices.

2.3. Coding and scripting

The prototyped interface is developed using HTML and javascript. HTML language offers a large scale of possibilities in relation to code and optimization, making it very easy for developers. Being a prototype interface based on HTML table technology, has a very rich source of code lines, reaching up to 300 lines of code / page. The DIV

technology design results in a reduced number of code lines, but needs the rich CSS classes' files and styles for each type of DIV, complicating the development process. The interface contains only one CSS file, called *frames.css*, with few classes and styles. Every style or class can apply to one or more tables, reducing the time of the interface development process. Another improvement in the prototype layout is the tables property. All the tables had the *width* and *height* properties set in percent (%). This ensures that the entire layout is adaptable on different resolutions, making it more *cross browser responsive,* compared with DIV technology, where all the cells properties are written in CSS styles. In this case it was not necessary to write CSS styles especially for rows or columns properties. Also the colors of the cells and even the graphical shadows are set directly from code, without using any external graphical files. For graphical shadows, to be well interpreted by all the browsers, a predefined CSS code was used*: webkit-box-shadow* for Google Chrome, *-moz-box-shadow:* for Mozilla Firefox and *box-shadow*: for Internet Explorer, or Microsoft Edge browser.

In order to be displayed properly on most mobile devices, the prototype contains a *@viewport{}* script in the HTML code, with maximum and minimum resolution (**Figure 2**.). The minimal resolution is 640 pixels and the maximum 800 pixels. Most modern mobile devices have a minimum resolution of 640 pixels. In this way the interface is displayed correctly on the devices and the side scrolling will not appear in navigation. To increase readability, the *@viewport{}* script uses lines of code to optimize the text displayed on screen devices, increased by an amount of 1.5 – 2.5 of its standard size.

Figure 2. Simulated prototyped icMED layout adaptation on smart phone

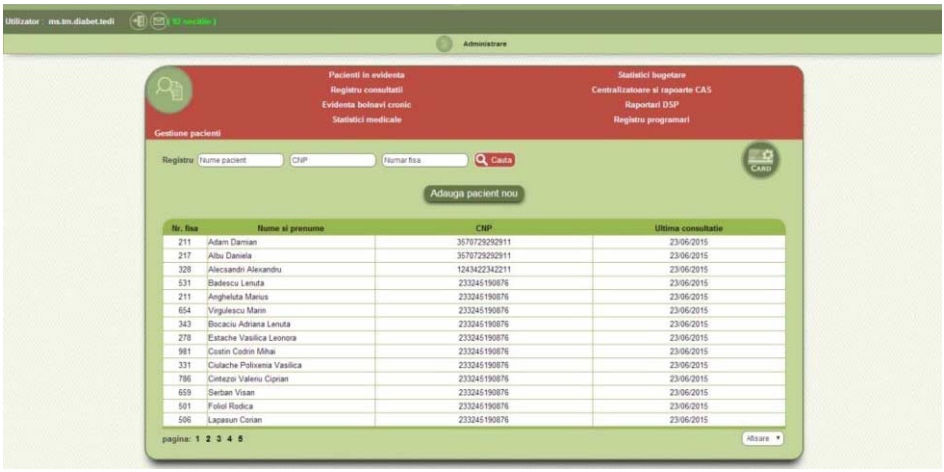


Figure 3. The new responsive layout proposal for the icMED web application on a resolution of 1600x900 pixels

3. Results

A prototype web interface was developed, for an online medical application (icMED) dedicated to medical professionals, patients and laboratories, using HTML table technology and CSS3 language, as an alternative to HTML DIV technologies. Several advantages and methods where highlighted when using HTML table design for developing dynamic and adaptive user interfaces for various media devices.

3.1. The new prototype interface

In the design process, the aim was to achieve a graphical interface based on intuitive user experience (UX), using different techniques and methods of the field. The absence of image sources design resulted in an improved source code quality used while simplifying the number of scripts required to achieve the graphical interface. The only common elements kept in the current interface, shown in **Figure 3**, are the same color range and icons, as identifiers of application's sections.

The links to subpages have been preserved entirely, but repositioned on the main page, and divided into primary and secondary links. The primary links have been relocated in the center of the page with a red background for a better identification. The secondary links were "hidden" in a section called "Management", because they are actually the links to the personalization module of the application. This improvement has provided a larger work area for the main interface, compared to the current interface where all the links are displayed in two rows. Another improvement, in terms of UX, is the use of text fields with text identification inside (**Figure 4)**. This was made possible by coding value *"value=" identification text"* in the html code assigned field. This method is more efficient than the classic one, in which the identification text is separated from the text field itself. Using this technique, more design space was saved for other necessary interface elements. This improvement is especially effective when the application is used on mobile devices, where low resolution does not allow expansion of the interface.

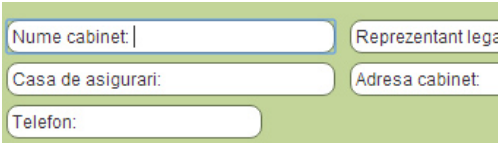


Figure 4. Identification text inside dynamic text field

Currently having a consistent group of medical staff using the application (aprox. 8,000) the new interface is in testing and collecting of feedback comments. During the testing process the icMED's users will provide information for the following parameters: ease of navigation, display or interpretation errors, intuitive navigation improvements and adaptability errors to media screens.

4. Conclusion

The paper presents new methods and techniques for design and implementation of user interfaces using HTML table technology in complex medical applications. This method represents a viable alternative to DIV technology, which is widely used today by most developers for building web interfaces. The methods and technique presented is a proof that dynamic web interfaces and adaptive can be developed combining different types of design methods and technologies, using HTML table design. The work presented in the paper is the first step, the technological improvement that has to be confirmed also from the users' perspective in a future step.

References

[1] Palmer, J.W., Web site usability, design, and performance metrics. Information systems research, 2002. 13(2): p. 151-167.
[2] Hermes, D., UI Design Using Layouts, in Xamarin Mobile Application Development. 2015, Springer. p. 45-104.
[3] Belchin, M. and P. Juberias, Implementing Design Patterns and Creating Web Components, in Web Programming with Dart. 2015, Springer. p. 389-406.
[4] Marriott, K., P. Moulder, and N. Hurst, HTML automatic table layout. ACM Transactions on the Web (TWEB), 2013. 7(1): p. 4.
[5] Meyerovich, L.A. and R. Bodik. Fast and parallel webpage layout. in Proceedings of the 19th international conference on World wide web. 2010. ACM.
[6] Conner, M.H., M.D. McClain, and L. Xu, Method for providing a visual representation of dynamic HTML table attributes. 2005, Google Patents.
[7] Chen, Y., W.-Y. Ma, and H.-J. Zhang. Detecting web page structure for adaptive viewing on small form factor devices. in Proceedings of the 12th international conference on World Wide Web. 2003. ACM.
[8] Fugate, J.K. and R.J. Vokurka, Current trends in web page design. International journal of innovation and learning, 2004. 1(2): p. 158-165.
[9] Wang, Y. and J. Hu. A machine learning based approach for table detection on the web. in Proceedings of the 11th international conference on World Wide Web. 2002. ACM.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-230

Medical Informatics and Information Technology Supporting Oral Medicine

Janat ZVÁROVÁ[a,1], Karel CHLEBORÁD[b] , Karel ZVÁRA[a] and Taťjána DOSTÁLOVÁ[b]
[a]*Institute of Hygiene and Epidemiology, 1st Faculty of Medicine, Charles University Prague, Czech Republic*
[b]*Department of Stomatology, 2nd Faculty of Medicine, Charles University Prague and University Hospital in Motol, Czech Republic*

Abstract. Electronic healthcare documentation is the key element of electronic healthcare (eHealth). Electronic oral health record (EOHR) supporting oral medicine is discussed. To provide dentists with a methodology and instrument to create oral health documentation in more efficient way, support information exchange and integration in dental domain and to ease dental decision-making and forensic dentistry identification tasks.The proposed methodology is used to model lifelong EOHR based on a small specific ontology where the use of other classification systems and nomenclatures, e.g. SNODENT, is possible. EOHRwith Lifelong DentCross user interface was developed and it has been supporting dental care at the University Hospital in Prague-Motol. The user interface is working in four languages and controlled by voice or keyboard. Lifelong DentCross user interface is reflecting the way of the work in dentistry and the EOHR can provide both structured and free text information to oral medicine.

Keywords. Oral Medicine, Electronic Health Record, User Computer Interface, Medical Informatics

1. Introduction

A broad overview of eHealth concepts is presented alongside an in-depth analysis of some of the most important related issues is in the book *Managing eHealth* [1]. The concept of Information-based Holistic Electronic HealthCare (IHE-HC) was introduced in [2]. It focuses on information in the form of data or knowledge that can be stored, processed or transmitted in an electronic form (Fig. 1). IHE-HC is connecting the e3 Health concept [3] with the application in a given healthcare system via co-operation among three main stakeholders (State, Academy and Industry). The inability to share information across these main stakeholders is just one of the major obstacles towards quality, efficiency, security and cost-effectiveness of healthcare. The gap between the demand for IHE-HC from an increasingly well-informed citizens and the ability of the government and healthcare organizations to meet this demand is widening all the time.

The electronic health records that are based on a common information architecture with highly standardized data definitions will play the key role in the electronic healthcare. Main goal of EHRs is the support of continuing, efficient and high quality integrated healthcare by sharing patient health information among authorized users. However, EHRs allow collection of data for other reasons than for direct patient care,

[1] Corresponding Author: Prof. Dr. Jana Zvárová, Charles University, 1st Faculty of Medicine, Katerrinska 32, 12108 Prague, Czech Republic, E-Mail: jana.zvarova@lf1.cuni.cz

such as quality improvement, outcome reporting, resource management and public health communicable disease surveillance, see e.g. [4-12].

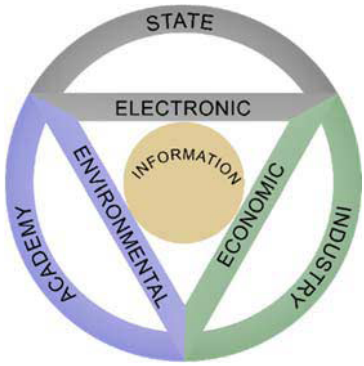


Figure 1: Information-based Holistic Electronic Healthcare

The part of EHR focused on oral health is called electronic oral health record (EOHR), see [13]. EOHR can fulfill the need of dentistry for more detailed description of problems and procedures using new granular and sophisticated coding sets. In advanced healthcare environment a dental clinics can connect oral health information to a hospital EHR. It means that past medical history, laboratory tests and other information can be retrieved from the hospital EHR to EOHR directly.

New strategy to advance the consistency of data in dentistry is emerging. Systematized Nomenclature of Dentistry (SNODENT), an official subset of SNOMED CT, is a vocabulary designed for the use in the electronic environment. The SNODENT can be considered as rising ontology for dentistry that should in future to standardize coding and present lists of problem and procedure terminology in a detail and granular way with the goal to cover all possibilities [14]. In spite of the general support for ontologies within the biomedical community, there are relatively few ontologies available to be used by the dental community at the present time. What ontology engineering efforts have been undertaken have largely been directed to the provision of small, special purpose and application of specific ontologies; large-scale dental ontologies with broad coverage of the dental domain are currently absent. Apart from specific ontologies like Tooth Positional Ontology [17], two large-scale ontologies are under development the Ontology for Dental Research (ODR) [18] and the Oral Health and Disease Ontology (OHD) [19].

We developed the small specific ontology for permanent dentition based on about 60 different actions, treatment procedures or tooth parameters. The ontology was developed by Czech dentists from University Hospital in Prague-Motol for promoting consistency of collected data in dental care. The ontology made possible to develop the first version EOHR with the DentCross interactive user interface in 2005. The user interface enabled recording fully structured dental information in a consistent and user-friendly way. The main requirement for EOHR was the structured way of data storage combined with free text with possibility of dynamic extension and modification of the set of collected attributes without any change. We used the DentCross user interface for collecting data for computer-supported treatment of patients with temporomandibular joint disorder parafunction [15] and in forensic dentistry [16].

The research objective of our approach is to provide dentists with a methodology and instrument to make lifelong oral health documentation by more effective way, support information exchange and integration in the dental domain and use collected information in dental decision-making and identification problems in forensic dentistry.

2. Methods

Based on our extensive experience with the interactive DentCross user interface of EOHR for permanent dentition, we have created a new object-oriented EOHR model compatible with HL7 RIM and implementable using EHRcom hierarchical structure that includes the possibility of entering the data not only for permanent, but also for mixed and deciduous teeth. We call this model Electronic Oral Health Record in Dentistry.

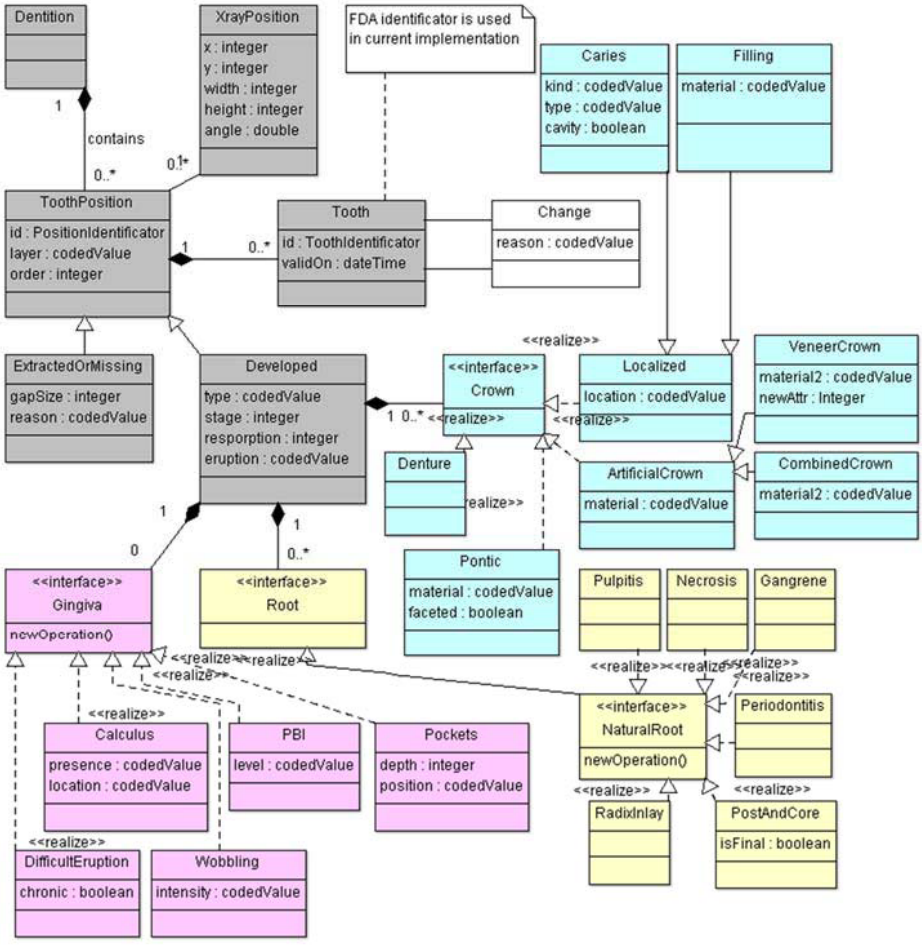


Figure 2: Model of Electronic Oral Health Record

The EOHR model (Fig. 2) is meant to represent the ontology of basic human dental structures. It is based on the premise that the model should be able to describe all valid situations and not to lose any substantial information. Model instance (in the form of linked objects) itself represents a static view of patient's dentition at the given time. Coded values are shown in Table 1. The model includes the possibility of entering the data not only for permanent, but also for mixed and deciduous teeth. Each existing tooth is described using the basic anatomical structures – a crown, root, suspension system of the tooth.

The model has its root at *Dentition* class that holds tooth *Positions*. Two-dimensional X-ray images may be mapped to individual positions. Any position may be "occupied" by a tooth. The tooth may exist on that position or may be missing. Any existing tooth is described using its basic properties and three parts: the gingiva, the root and the crown. The gingival properties include eruption problems, dental calculus, wobbling, papilla bleeding index (PBI) and pocket. The root interface is extended by specialized interface *NaturalRoot* that represents natural root conditions as pulpitis, necrosis, gangrene, periodontitis, root canal treatment and root inlay. The R*oot* interface is directly implemented by the *Implant* class that represents an implant. The crown interface supports localized properties such as Caries and Filling that are bound together in order to not lose the information on what caries have been fixed by the filling. The *Crown* interface is implemented by *Denture* class, *Pontic* class and is extended by the *Artificial Crown* interface. The *Localized* class *location* property is divided up to 9 square parts to further specify location. The *Artificial Crown* interface is implemented by *Veneer Crown* and Combined Crown interfaces to store the information about materials used. Tooth and Crown interfaces may be linked with their other "instances", specifically with their previous and next states (if available).

Table 1: List of attributes with codedValue in EOHR model

Class	Attribute	codedValue	Class	Attribute	codedValue
ToothPosition	level	deciduous, permanent, denture	Caries	kind	primary, secondary, relapsing
ExtractedOrMissing	reason	diagnosed, agenesis, unknown	Caries	type	acute, chronic, stopped, initial carious lesion
Developed	type	incisor, canine, premolar, molar	Filling	material	amalgam, composite, compomer, glassionmer, provisional
Developed	eruption	none, partial, full	Pockets	position	mesial, distal, vestibular, oral, interradicular
Calculus	presence	present, absent, unknown	Localized	location	mesial, distal, vestibular, oral, interradicular
Calculus	location	supragingival, subgingival	VeneerCrown	material2	resin, composite, base metal, gold, ceramic
Wobbling	intensity	none, level I, level II, level III	CombinedCrown	material2	resin, composite, base metal, gold, ceramic
PBI	level	0, 1, 2, 3, 4, 5	ArtificialCrown	material	resin, composite, base metal, gold, ceramic

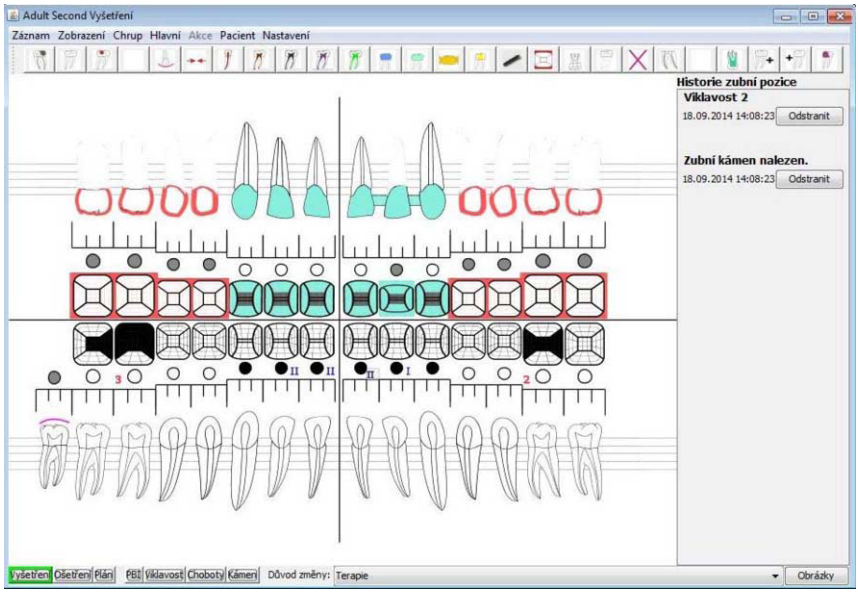


Figure 3. Lifelong DentCross user interface of EOHR

3. Results

The EOHR model was implemented using the new graphical design of user interface Lifelong DentCross (Fig. 3).

Lifelong DentCross user interface is dominated by a graphical dental cross. In addition to the basic information needed to identify the patient (name, surname and ID number) one can also find user interface controls. These are chosen so that the user can easily enter the history of treatment, treatment plan and the dentist can also use other forms of periodontal detailed examination, such as recording the presence of tartar, record the depth of periodontal pockets, and looseness of the teeth and gums condition called PBI (papilla bleeding index). We can demonstrate a record number of complex examinations including periodontal compared with intra-oral radiographs. In the right part of the screen we can find two options – history and legend. History brings us to the chronological treatment of the individual. Legend contains mainly a variety of different materials in the color scale to facilitate orientation in the user record. Individual teeth are called a double-digit number by quadrants, for example the Tooth 27 is the second upper molar upper left.

The crowns of the teeth are divided into seven fields to locate carious lesions or restorations, individual fields are marked according to the anatomic custom M – mesial, D – distal, O – occlusal, I – incisal, R – oral, V – vestibular, C – cervical. Size of the lesion seeing a number 1 to 4, which is to Mount classification of carious defects. By the selection of the filling different colors denote the type of material used. The blue color indicates photo composite, black amalgam and the green color indicates glassionomer dental restorations. So the embedded information from the examination and X-ray documentation becomes much more valuable than photographs of dental state or mere X-ray image. Dentists can also add dental reports in free text to all this structured data.

Nowadays, the Lifelong DentCross interactive user interface for EOHR can be controlled by voice or keyboard and it is running in four languages. These languages are Czech, English, German, and Spanish. Each language has its set of definitions, which contains about 300 dental terms.

4. Discussion

The most important source of information for biomedicine and healthcare is the data. We can see a huge amount of data collected in an unstructured form, e.g. in narrative medical reports. The unstructured data is very difficult to analyze and most information in the unstructured data is used rarely or never. With penetration of ICT in healthcare comprehensive lifelong EOHR combined with appropriate representing, accessing and visualizing health data have been developed. We tried to minimize unstructured data in EOHR to make information reusable for other tasks.

EOHR with the interactive Lifelong DentCross user interface has been used in dental care at the University Hospital in Prague-Motol. Lifelong DentCross user computer interface has been supporting data recording in dentistry very well. .Unfortunately, the retrieving data from hospital information system to EOHR is not supported by hospital information system. In 2016 the Czech Republic became the member of International Health Terminology and Standards Organization (IHTSO) that opens the possibility to use SNODENT classification in EOHR. The usability of EOHR with the DentCross user interface was partially studied in [20]. The comparison of three methods for time-consuming data entry in dentistry was performed on 126 patients: a) dental registration in the WHO card, b). EOHR controlled by keyboard and c) EOHR controlled by voice. In the study on 126 patients dentists added textual comments very rarely. Due to sufficient information in structured data there was no need to enrich the EOHR model with part extracting information from narrative dental reports [21]. In clinical practice, it is required to find ways to avoid the manual operation using a keyboard, mouse or touch screen. Therefore, the added automatic voice recognition allows the dentist to use the software without having to touch. This eliminates the need for a second person making the entry to the computer or redundant hygienic procedures (washing hands, changing gloves, etc.). Dental expertise occupies an important position in the identification process and is necessary especially when the priority is a recognition, the results of DNA analysis and fingerprint comparison. Properly conducted dental documentation is often the key to successful identification of unknown persons or skeletal remains. Not only the teeth, but also the analysis of the materials used to rehabilitate our teeth with the help of modern technology can lead to a clear identification of the individual. The EOHR can support a rapid identification of unknown individuals as well as decision making in dentistry, see e.g. [22], [23].

Acknowledgement.

The present study was supported by the grants PRVOUK P28/LF1/6, P29/LF2 and SVV 260267 of Charles University in Prague.

References

[1] Rosenmoller M, Whitehouse D, Wilson P. (Eds.) Managing eHealth - from Vision to Reality, Palgrave M8cmillan 2014

[2] Zvarova J. Holistic electronic healthcare. Int J on Biomedicine and Healthcare 3(2).; 2014: 2–8

[3] Zvarova J, Zvara K: eHealth: Three main features of modern healthcare, In: A. Moumtzoglou, A. Kastania (Eds.) E-Health Systems Quality and Reliability: Models and Standards, IGI Global; 2011:18–27

[4] Häyrinenn K. Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. Int J Med Inf 2008, 5: 291–304

[5] Hammond K, Helbig S, Benson C, Brathwaite-Sketoe B. ,Are Electronic medical records trustworthy? Observations on copying, pasting and duplication. AMIA Annual Symposium Proceedings, 2003; 269-273

[6] Boonstra A, Versluis A, Vos J.: Implementing electronic health records in hospitals: a systematic literature review. BMC Health Services Research 2014, 14:370

[7] Blobel B: EHR/PHR systems today and in the future. Int J on Biomedicine and Healthcare 4 (1) 2016::

[8] Ceusters W, Elkin P, Smith B. Negative findings in electronic health records and biomedical ontologies: a realist approach. Int J Med Inf 2007; 76:326–333.

[9] Ludwick DA, Doucette J. Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. Int J Med Inf 2009; 78:22–31.

[10] Van Ginneken AM, Stqam H, van Mulligen EM, de Wilde M, van Mastrigt R, van Bemmel JH. ORCA: the versatile CPR. Methods Inf Med 1999; 38:332–38.

[11] Van Ginneken AM. Considerations for the representation of metadata for the support of structured data entry. Methods Inf Med 2003, 42,226-235

[12] Blobel B, Hasman A, Zvarova J (eds): Data and Knowledge for Medical Decision Support, IOS Press, Amsterdam, 2013

[13] Heid W.D., Chasteen J., Forrey A.W.: The electronic oral health record, J Contemp Dent Pract 2002, 3:43–54.

[14] Torres-Urquidy MH, Schlezer T. Evaluation of the Systemaic Nomenclature of Dentistry (SNODENT]) using Case Reports: Preliminary Results. AMIA Symposium Proceedings, 2006, 1124

[15] Dostalova T., Seydlová M., Zvarova J., Hanzlicek P., Nagy M.: Computer-supported treatment of patients with the TMJ parafunction. In: Blobel B., Pharow P. Zvarova J., Polez D. (Eds) eHealth: Combining health telematics, telemedicine, biomedical engineering and bioinformatics to the edge. CEHR Conference Proceedings 2007, Berlin, 171-8

[16] Zvarova J., Dostalova T., Hanzlicek P., Teuberova Z., Nagy M., Pies M.: Electronic health record for forensic dentistry. Methods Inf Med. 2008, 47: 8–13

[17] Park SG, Kim HG, Kim HK, Tooth Positional Ontology Represented in OWL, In: Kuhn K, Warren J, Leong T (Eds) MEDINFO 2007, IOS Press 2007, 2288-2289

[18] Smith B, Goldberg IJ, Ruttenberg A, Glick M. Ontology and the future of dental research informatics. JADA 2010, 141(10): 1173-11

[19] Schleyer TK, Ruttenberg A, Duncan W, Haendei M, Tornial C, Acharya A, Song M, Thyvalikakath TP, Liu K, Hernandez P. An ontologybased methods for secondary use of electronic dental record data. AMIA J Summit Transl Sci Proc. 2013,234-238

[20] Chleborad K, Zvara K Jr, Dostalova T, Zvara K, Hippmann R, Ivancakova R, Trmal Z, Psutka J. Evaluation of voice-based data entry to an electronic health record system for dentistry. Biocybernetics and biomedical engineering 2013, 33:204-210.

[21] Zvara K, Svatek V: Amenability of Czech Medical Reports to Information Extraction. Eur J Biomed Inform 2012; 5:43–7.

[22] Hippmann R, Dostalova T, Zvarova J, Nagy M, Seydlova M,, Hanzlicek P, et al. Voice supported electronic health record for temporomandibular joint disorders. Methods Inf Med 2010; 49:168–72.

[23] Hippmann R, Nagy M, Dostalova T, Zvarova J, Seydlova M. Electronic system for data record and automatic diagnosis assessment in the temporomandibular joint disorder. Eur J Biomed Inform 2011;7:11–6

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-237

Continual Screening of Patients Using mHealth: The Rolling Score Concept Applied to Sleep Medicine

Claudio ZLUGA[a,1], Robert MODRE-OSPRIAN[a], Peter KASTNER[a] and Günter SCHREIER[a]
[a]*AIT Austrian Institute of Technology, Austria*

Abstract. Continual monitoring of patients utilizing mHealth-based telemonitoring applications are more and more used for individual management of patients. A new approach in risk assessment called *Rolling Score Concept* uses standardized questionnaires for continual scoring of individuals' health state through electronic patient reported outcome (ePRO). Using self-rated questionnaires and adding a specific *Time Schedule* to each question result in a movement of the questionnaires' scores over time, the *Rolling Score*. A text-processing pipeline was implemented with KNIME analytics platform to extract a *Score Mapping Rule Set* for three standardized screening questionnaires in the field of sleep medicine. A feasibility study was performed in 10 healthy volunteers equipped with a mHealth application on a smartphone and a sleep tracker. Results show that the proposed *Rolling Score Concept* is feasible and deviations of scores are in a reasonable range (< 7%), sustaining the new approach. However, further studies are required for verification. In addition, parameter quantification could avoid incorrect subjective evaluation by substitution of questions with sensor data.

Keywords. Risk assessment, screening, questionnaires, ePRO, telehealth

1. Introduction

Continual monitoring of patients using mHealth applications are more and more used for personalized management of patients with chronic diseases like congestive heart failure (CHF) [1], [2]. Sleep-related breathing disorders (SRBD) are identified as one contributing factor to the excess morbidity and mortality in CHF population [3]. The prevalence rate of obstructive sleep apnea (OSA), a common type of SRBD, in American adults between the ages of 30 and 70 years is estimated at 26 percent [4].

The gold standard for diagnosing SRBD is polysomnography, a time-consuming and expensive method with limited capacity for testing considering the large number of affected patients [5]. Screening questionnaires provide a reliable, concise and cost-effective method for individuals' health state and rapid risk assessment. Screening tools like the Berlin Questionnaire (BQ) [6] and the STOP-BANG Questionnaire [7] are validated and easy-to-use methods for identification of persons in a community who are at high risk for OSA. In addition, the quality of sleep can be assessed by the Pittsburgh Sleep Quality Index (PSQI) self-rated questionnaire [8].

Using self-rated questionnaires and adding a specific *Time Schedule* to each question leads to movement of the questionnaires' score over time. Such a "Rolling Score" (RS)

[1] Corresponding Author: Claudio Zluga, AIT Austrian Institute of Technology, Reininghausstraße 13/1 8020 Graz, E-Mail: Claudio.Zluga.fl@ait.ac.at

can be used to continually assess the individual health state of patients easily through electronic patient reported outcome (ePRO) and mHealth-based telemonitoring. The present paper deals with a new approach in risk assessment called *Rolling Score Concept* using standardized questionnaires for continual scoring of individuals' health state. In addition to the concept presentation, a technical feasibility study on the risk assessment approach has been performed in the field of sleep medicine.

2. Methods

2.1. The Rolling Score Concept

Standardized questionnaires have an inherent structure divided into three levels, where each level is defined by a corresponding scoring. Figure 1 illustrates these three levels from Questionnaire Level I, to Category Level II to Item Level III. Since some standardized questionnaires show no separation between Level I and Level II, each item is defined as a category C_i. An item in this context is termed as a single question. Each item I_i is assigned to a defined score interval. The transition between different score levels is determined by prescribed rule-sets 1 and 2.

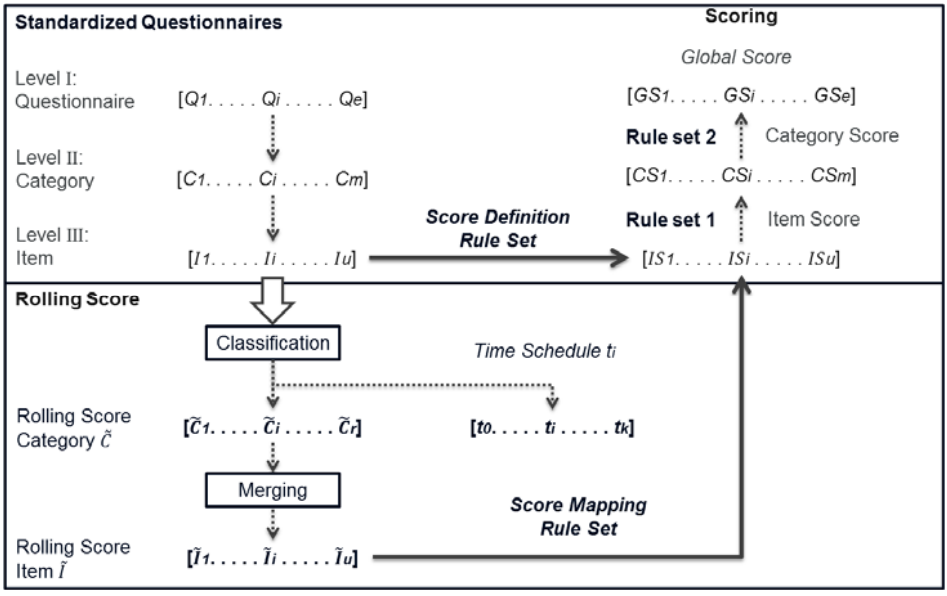


Figure 1: The *Rolling Score Concept*: Scoring, classification, merging and mapping of questionnaires

The concept of the RS used standardized questionnaires Q_i to create a time-dependent global score GS_i. All I_i were classified with respect to temporal and thematic context for assignment to a unique RS category $\tilde{C}_i$ and a corresponding *Time Schedule* t_i. Items within the same category $\tilde{C}_i$ were merged to avoid semantically-identical items. Each category $\tilde{C}_i$ consisted of a particular number of RS items $\check{I}_i$. The respective item score IS_i was recalculated using a *Score Mapping Rule Set*.

As an initial reference point, all items of a standardized questionnaire had to be answered. Subsequently, all categories were queried according to the *Time Schedule*. After answering, the *Score Mapping Rule Set* assigned the item $\check{I}_i$ to the original IS_i, followed by a calculation of the category score CS_i and the global score GS_i using rule set 1 and 2, respectively. This resulted in a continual adaptation of the GS_i over time, thus we name it the *Rolling Score Concept*. At the end of the study, we asked the standardized questionnaire again to get the *End Score* in order to compare the RS and the score of the standardized version. Considering that the responses of items show an intrinsic variability and to take into account that the risk assessment of the RS should provide similar results compared to the standardized questionnaire, we expect to obtain differences ≤ 10% as defined in [9]. These differences are defined by the measure γ (Formula 1) with n as the total number of observations. Subjects with a compliance of 75% or greater were included in calculations.

$$\gamma = \frac{\frac{1}{n}\sum_{k=1}^{n} |GS_k - End\ Score_k|}{Maximal\ Score\ Range} * 100\ [\%] \quad (1)$$

2.2. Classification and Merging

Generic classification of questionnaires was done using the text processing tools of the open source analytics platform KNIME. Classification was performed by the Score Mapping Rule Set consisting of the Content-related Context Rule and the Temporal Context Rule. A pipeline was set up for extraction of the Content-related Context Rule, consisting of the following steps: preprocessing, keyword-extraction, annotation and automatic classification of items (Figure 2).

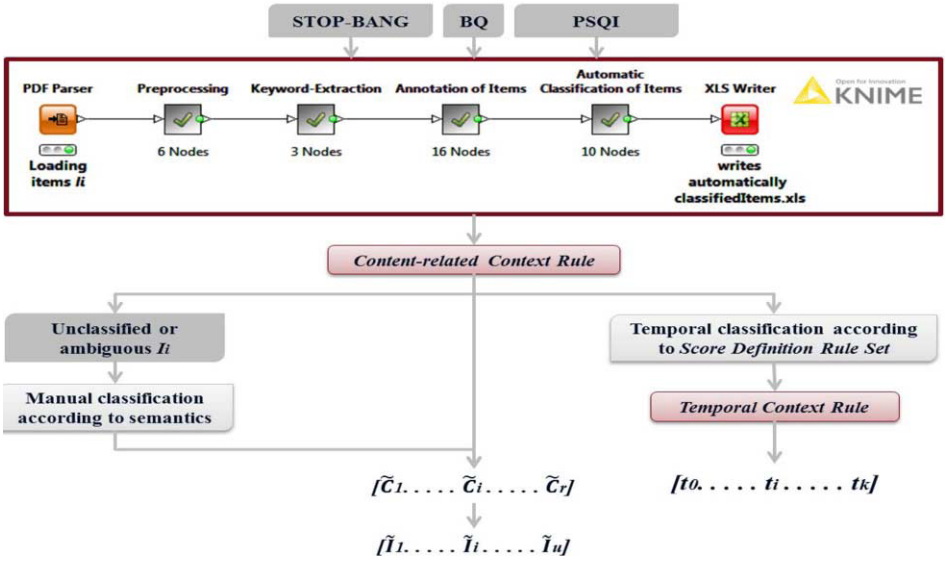


Figure 2: Extraction of *Score Mapping Rule Set*

Annotation of items was performed using extracted keywords of predefined categories, to ensure an independent classification outcome. Items assigned to more than one category $\tilde{C}_i$ or unclassified items were assigned manually to a unique category $\tilde{C}_i$ according to semantics. Using the *Content-related Context Rule*, the categories $\tilde{C}_i$ and respective $\tilde{I}_i$ were implemented and assigned to a unique and coherent category $\tilde{C}_i$. Applying a temporal classification using the *Score Definition Rule Set* extracted the *Temporal Context Rule* to set the *Time Schedule* t_i.

Since items were already grouped with respect to time and content, a merging of them within the same category $\tilde{C}_i$ was possible. There were two merging possibilities: One implied a complete or partial replacement of one item by two or more analogous items regarding the content. A second method used a disjunction to extent one item with the content of the replaced one. If the response format between the merged items did not match, the greatest score range was set as response format. As an item $\tilde{I}_i$ could contain information about two or more original items, the *Score Mapping Rule Set* mapped the IS_i to the corresponding items.

2.3. Feasibility Study: Rolling Score applied to Sleep Medicine

The present concept has been applied to the questionnaires BQ, STOP-BANG and PSQI in the field of sleep medicine. The *Rolling Score Concept* was applied to each single questionnaire as well as to a joined version of all three questionnaires.

A feasibility study over a one month time interval has been designed and conducted with a group of ten healthy volunteers. An existing mHealth-based telemonitoring platform solution (KIOLA eHealth platform, AIT Austrian Institute of Technology) was used for collecting data of the subjects at their homes. The system consisted of a mHealth application on a smartphone, a sleep tracker (Withings Pulse O_x) and a web-based backend solution. Subjects were instructed to complete all questions and to use the sleep tracker during their sleep for one month. As data was collected, subjects gave written informed consent. All data has been transmitted to a backend for analysis and compliance check. The compliance in this context was defined as the ratio between all sent categories to the number of received categories.

3. Results

3.1. Rolling Score

By applying the classification and merging procedures, categories $\tilde{C}_i$, items $\tilde{I}_i$ and the *Time Schedule* t_i were created (Table 1). A total of 13 categories were reduced to 8 categories $\tilde{C}_i$. The categories *Snoring*, *Sleep Disturbances*, *Daytime Somnolence*, *Sleep Medication* and *Weekly Quantifiable* had a weekly *Time Schedule*, distributed over the week from Monday to Friday to reduce the daily effort. *Blood Pressure*, *Body Mass Index (BMI)*, *Neck Size* and *Sleep Quality* were asked every 28th day after the beginning based on the *Score Definition Rule Set*. Category *Age/Gender* was asked only once at the beginning.

We asked one questionnaire as a combination of all three (joined RS) and additionally, each single questionnaire as RS separately. All questionnaires were asked according to the same *Time Schedule* t_i.

Table 1: All categories C_i, $\tilde{C}_i$ and corresponding *Time Schedule* t_i

Category C	Number of Items I	Category $\tilde{C}$ (Number of Items I)	Number of Items $\tilde{I}$	*Time Schedule* t_i
Snoring	1/5*	Snoring (8)	4	Mon.
Observed	1	Snoring (8)	4	Mon.
Sleep Disturbances	9	Sleep Disturbances (8)	8	Tue.
Sleep Medication	1	Sleep Medication (1)	1	Wed.
Daytime Somnolence	1/3/2*	Daytime Somnolence (6)	4	Thu.
Sleep Latency	2	Weekly Quantifiable (5)	5	Fri.
Sleep Duration	1	Weekly Quantifiable (5)	5	Fri.
Sleep Efficiency	2**	Weekly Quantifiable (5)	5	Fri.
Blood Pressure	2	Monthly Quantifiable (4)	3	Every 28th day after start
BMI	1	Monthly Quantifiable (4)	3	Every 28th day after start
Neck size	1	Monthly Quantifiable (4)	3	Every 28th day after start
Sleep Quality	1	Sleep Quality (1)	1	Every 28th day after start
Age/Gender	2	Age/Gender (2)	2	At start

* Order: STOP-BANG/BQ/PSQI ** One item of category *Sleep Duration* for GS_i calculation necessary

The *Initial Score* served as a reference point for the score updating and corresponds to the standardized questionnaire. All three RS of the joined questionnaires from subject 03 are illustrated in Figure 3, showing GS_{BQ}, $GS_{STOP\text{-}BANG}$ and GS_{PSQI} over a period of 29 days. The maximal range of GS_{BQ}, $GS_{STOP\text{-}BANG}$ and GS_{PSQI} is from 0 to 3, 0 to 2 and 0 to 21, respectively. In order to get a number of the $GS_{STOP\text{-}BANG}$, the following risk scale was defined: low risk ≙ 0, intermediate risk ≙ 1 and high risk ≙ 2. The questionnaires did not distinguish between one month and four weeks, whereas the time period was set at four weeks [8]. At the end of the study period (29th day) we asked each participant to complete the standardized questionnaire again to generate the *End Score* in order to find differences between this *End Score* and the joined and single questionnaires' GS_i. Figure 3 indicates that all GS_i of subject 03 were at health level, showing a temporal variance as all GS_i slightly decreased or remained at zero. In general, all GS_i of the subjects were at low risk with a minimal temporal variance. Only subject 06 and 09 showed a temporally increase of the $GS_{STOP\text{-}BANG}$ at high risk level.

Table 2 shows the mean value, standard deviation and γ over all subjects of each GS_i and the *End Score*. All subjects had the required compliance of 75%.

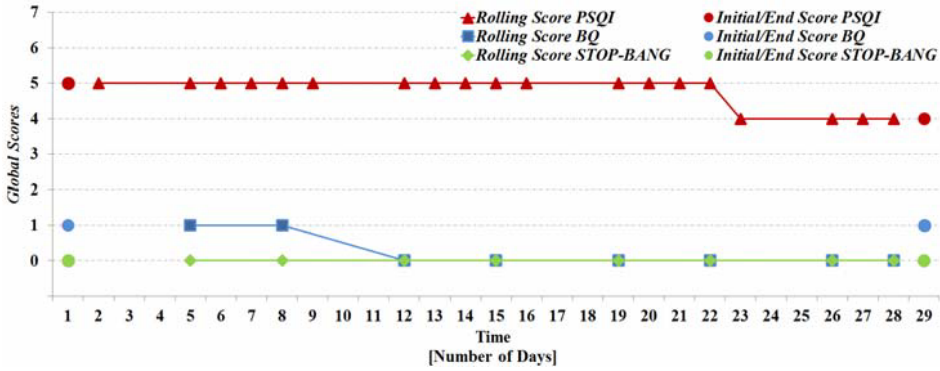


Figure 3: The *Rolling Score* of joined questionnaires from subject 03

Table 2: Mean, standard deviation and γ of joined and single GS_i and *End Score*, * joined questionnaires

	*GS**	*GS*	*End Score*	γ*	γ
BQ	0.4 ±0.5	0.4 ±0.5	0.5 ±0.5	3.3%	3.3%
STOP BANG	0.4 ±0.8	0.4 ±0.8	0.4 ±0.8	0.0%	0.0%
PSQI	4.7 ±2.2	3.9 ±2.3	4.5 ±2.1	6.7%	6.7%

3.2. Classification and Merging

We applied the classification pipeline (Figure 2) to a total of 35 items from the three considered questionnaires. Text of prescribed response possibilities was not considered in classification. Due to automatic classification, 11 items were classified in two categories and 7 items remained unclassified. These items have been assigned manually to a unique category $\tilde{C}_i$ showing equal semantics. Due to the merging possibilities, a total of 12 categories and 35 items have been summarized to 8 $\tilde{C}_i$ and 28 $\tilde{I}_i$, respectively. Figure 4 shows three items from different questionnaires with respective IS_i before classification and merging was applied. The outcome was $\tilde{I}_3$. To recalculate each IS_i the *Score Mapping Rule Set* was used (Figure 5). The *Score Definition Rule Set* determines the *Time Schedule* and, as for $\tilde{I}_3$, the cut-off value in Figure 5 labels exemplarily this point.

BQ I_7	BQ IS_7	STOP-BANG I_2	STOP-BANG IS_2	PSQI I_9	PSQI IS_9
During your waking time, do you feel tired, fatigued or not up to par?		Tired? Do you feel Tired, Fatigued or Sleepy during the daytime[…]?		During the past month, how much of a problem has it been for you to keep up enough enthusiasm to get things done?	
o Nearly every day	1	o Yes	1	o No problem at all	0
o 3-4 times a week	1	o No	0	o Only a very slight problem	1
o 1-2 times a week	0			o Somewhat of a problem	2
o 1-2 times a month	0			o A very big problem	3
o Never or nearly never	0				

Figure 4: Three items before classification and merging

Category $\tilde{2}$ *Daytime Somnolence* $\tilde{I}_3$	*Score Mapping Rule-Set*		
During your waking time, do you feel tired, fatigued or not up to par?	1:1 Mapping	Threshold	Arithmetic Mean
o Nearly every day	1	1	3
o 3-4 times a week	Cut-off → 1	1	3
o 1-2 times a week	0	0	2
-			
o Never or nearly never	0	0	0

Figure 5: Merged $\tilde{I}_3$ and corresponding *Score Mapping Rule Set*

The PSQI questionnaire assesses and scores sleep quality and disturbances during a one-month interval. Due to the *Time Schedule* the categories were asked weekly, the information of three weeks before was not considered. Therefore, the *Score Mapping Rule Set* was applied in order to regain information of the last four weeks. The mapping algorithm corresponds to an arithmetic mean of the IS_i at the present day t_i and the item scores of the past three weeks.

3.3. Feasibility

A feasibility study was conducted between December 2015 and January 2016. Ten healthy volunteers participated and were equipped with a smartphone and an app for questionnaire completion and a sleep tracker. A study coordinator registered them with a unique subject number to avoid identification among them.

During the feasibility study, 523 categories were sent to the frontend without errors. The backend received 486 categories, whereas 37 were missing due to a lack of compliance. The overall compliance was 91.8%. Two technical errors occurred and were solved without disturbing the workflow. The categories *Weekly*- and *Monthly Quantifiable* included the following parameters: blood pressure, BMI, neck size, sleep latency, sleep duration and sleep efficiency. In order to compare subjective and objective values, the mean value of all measurements within one week was computed. For one subject no data was available. The average difference and standard deviation with min-max values [min value, max value] for the sleep parameters were: Sleep efficiency of 9.1 $\pm$5.4 % [2.9, 20.8], sleep duration of 0.8 $\pm$0.8 hours [0.0, 3.5] and sleep latency of 11.6 $\pm$20 minutes [0.8, 96.8].

4. Discussion

We developed a generic approach using standardized questionnaires in order to continually assess individuals' risk based on the movement of the questionnaires' scores over time. Modifying standardized questionnaires is a critical procedure because in general, this leads to a loss of validation. We defined a *Score Mapping Rule Set* to determine the way of mapping processed items $\tilde{I}_i$ back to the original items IS_i. Since order and wording of questions are relevant aspects for providing respondents with the same stimuli [10], the *Content-related Context Rule* ensures a grouping of items $\tilde{I}_i$ dealing with equal stimulus. The *Temporal Context Rule* takes care of modifications concerning different time intervals between items and their corresponding item score.

Segmentation of a questionnaire according to a specific *Time Schedule* leads to a loss of time-related validation. The *Score Mapping Rule Set* compensates this loss to a certain degree as can be seen by the results in Table 2. The categories *Sleep Disturbances* and *Daytime Somnolence* show a variability, which causes the deviations of the scores. The small temporal variance of the GS_i is also caused by this variability. The temporal variability of these categories was actually expected, since tiredness and sleep disturbances (example given: *going to toilet during sleep* or *frequent awakenings*) can strongly alternate from week to week and even depend on daily state. At GS_{PSQI} we found the highest γ at 6.7% due to these categories and the large scoring range (0 to 21). In general, γ was below the expected value of 10%. We conclude, applying a specific *Time Schedule* to questions has only a small impact on the validity of the global scores. Therefore, the *Rolling Score Concept* is a new promising method for continual screening of patients using standardized questionnaires. Nevertheless, further studies are required in order to verify the concept.

Merging of items could be difficult with respect to time aspects. Some items show time references in its responses and others in the question only. This different time aspect has to be considered in the *Score Mapping Rule Set* with a specific algorithm as, for example was done for the *Daytime Somnolence* category (Figure 5).

Since subjects were instructed to use sensors for items' quantification, sensor data may had an influence on the answers of the corresponding items. This may have introduced bias into the scoring of questionnaires. Nevertheless, we found differences between the subjective estimation and objective measurement of sleep efficiency, sleep duration and sleep latency.

5. Conclusion and Outlook

Standardized questionnaires provide an easy-to-use, concise and cost-effective screening tool for quick assessment of individuals' health state. Using self-rated questionnaires with mHealth-based telemonitoring enables continual risk assessment at home. The results show that the application of the proposed *Rolling Score Concept* in the field of sleep medicine is feasible and deviations of scores are in a reasonable range smaller than 7%, although a general validation has still to be approved. As we found quantifiable parameters, sensor data could be used to substitute corresponding items. However, mapping these measurements to the scoring may require a specific algorithm. In future work the *Rolling Score Concept* will be embedded in a collaborative network for heart failure patients in order to get real world experience with this new approach to continual risk assessment.

References

[1] S. C. Inglis, R. A. Clark, R. Dierckx, D. Pietro-Merino, and J. G. F. Cleland, Strucutred telephone support or non-invasive telemonitoring for patients with heart failure, *Cochrane Database of Systematic Reviews* **10** (2015), p. CD007228.

[2] A. Von der Heidt, E. Ammenwerth, K. Bauer, B. Fetz, T. Fluckinger, A. Gassner, W. Grander, W. Gritsch, I. Haffner, G. Henle-Talirz, S. Hoschek, S. Huter, P. Kastner, S. Krestan, P. Kufner, R. Modre-Osprian, J. Noebl, M. Radi, C. Raffeiner, S. Welte, A. Wiseman, and G. Poelzl, HerzMobil Tirol network: rationale for and design of a collaborative heart failure disease management program in Austria, *Wien klinische Wochenschrift* **126**(21-22) (2014), 734-741.

[3] M. R. Costanzo, R. Khayat, P. Ponikowski, R. Augostini, C. Stellbrink, M. Mianulli, and W. T. Abraham, Mechanisms and Clinical Consequences of Untreated Central Sleep Apnea in Heart Failure, *Journal of the American Colege of Cardiology* **65**(1) (2015), 72–84.

[4] National Healthy Sleep Awareness Project, American Academy of Sleep Medicine, http://www.aasmnet.org/, last access: 16.01.2016.

[5] N. Boehning, A. Blau, B. Kujumdshieva, A. Staubitz, and W. Boehning, Preliminary results from a telemedicine referral network for early diagnosis of sleep apnoea in sleep laboratories, *Journal of Telemedicine and Telecare* **15**(4) (2009), 203–207.

[6] M. J. Thurtell, B. B. Bruce, D. B. Rye, N. J. Newman, and V. Biousse, The Berlin questionnaire screens for obstructive sleep apnea in idiopathic intracranial hypertension, J*ournal of Neuroophthalmology* **31**(4) (2011), 316–319.

[7] F. Chung, B. Yegneswaran, P. Liao, S. A. Chung, S. Vairavanathan, S. Islam, A. Khajehdehi, and C. M. Shapiro, STOP Questionnaire A Tool to Screen Patients for Obstructive Sleep Apnea, *Anesthesiology* **108**(5) (2008), 812–821.

[8] D. Buysse, C. Reynolds, and T. Monk, The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research, *Psychiatry Research* **28** (1989), 193–213.

[9] K. Bölenius, C. Brulin, K. Grankvist, M. Lindkvist, and J. Söderberg, A content validated questionnaire for assessment of self reported venous blood sampling practices, *BMC Research Notes*, **5**(1) (2012), 39.

[10] I. M. Crawford, *Questionnaire Design*, in: Marketing Research and Information Systems. Food and Agriculture Organization of the United Nations, Rome, 1997. pp. 40-50.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-245

Secondary Use of Claims Data from the Austrian Health Insurance System with i2b2: A Pilot Study

Florian ENDEL [a,1] , Georg DUFTSCHMID [b]
[a]*Vienna University of Technology*
[b]*Medical University of Vienna, Center for Medical Statistics, Informatics, and Intelligent Systems*

Abstract. Background: In conformity with increasing international efforts to reuse routine health data for scientific purposes, the Main Association of Austrian Social Security Organisations provides pseudonymized claims data of the Austrian health care system for clinical research. Objectives: We aimed to examine, whether an integration of the corresponding database into i2b2 would be possible and provide benefits. Methods: We applied docker-based software containers and data transformations to set up the system. To assess the benefits of i2b2 we plan to re-enact the task of cohort formation of an earlier research project. Results: The claims database was successfully integrated into i2b2. The docker-based installation approach will be published as git repository. The assessment of i2b2's benefits is currently work in progress and will be presented at the conference. Conclusions: Docker enables a flexible, reproducible, and resource-efficient installation of i2b2 within the restricted environment implied by our highly secured target system. First preliminary tests indicated several potential benefits of i2b2 compared to the methods applied during the earlier research project.

Keywords. Health Information Systems, Medical Records Systems

1. Introduction

Reuse of routinely collected administrative claims data for clinical research gains momentum all over the world (e.g. [1-3]). This is also the case in Austria, where the Main Association of Austrian Social Security Organisations [2] (HVB) provides a collection of claims data called GAP-DRG[3] for research purposes. Based on this database numerous research projects have been conducted (e.g. [4-7]), which typically relied on manually developed database queries and individually applied statistical procedures.

The goal of the present work was to examine whether an integration of the GAP-DRG database into the i2b2 clinical research framework [8-10] would be possible. This should provide a consistent and more efficient working environment for the before-mentioned kinds of projects. Due to the size (amongst other data, there are nearly $5 * 10^8$ ambulatory services and about $1{,}9 * 10^8$ medication prescriptions and dispensings of about $11 * 10^6$ patients documented for two years) and the complex structure (59

[1] Corresponding Author: Florian Endel, Vienna University of Technology, Institute for Analysis and Scientific Computing, Wiedner Hauptstraße 8-10, 1040 Wien. E-Mail: florian.endel@tuwien.ac.at
[2] in German "Hauptverband der österreichischen Sozialversicherungsträger"
[3] GAP-DRG is an abbreviation of one of the first larger research projects based on the data collection called "General Approach for Patient-oriented Ambulant DRGs"

normalized tables with cleanly defined relational constraints for two years' worth of data) of the GAP-DRG database, we expected the process to be challenging. Also, GAP-DRG is embedded in a highly secured hardware infrastructure with several restrictions that would likely complicate our work.

In the present work we report on the experiences we have made so far with integrating the GAP-DRG database into i2b2.

1.1. The GAP-DRG database

The GAP-DRG database comprises pseudonymized claims data from the Austrian healthcare system covering about 97% of the Austrian population[4]. Pseudonymized data about medication prescriptions and dispensings, sick leaves, ambulatory visits and inpatient episodes are linked and accompanied by demographics as well as various metadata (e.g. spatial data and diagnostic schemas). Additionally, derivated information resulting from previous projects, like for example diagnoses determined from prescriptions data [11], are integrated tightly.

The database was set up with data from all 19 Austrian social insurance institutions covering the years 2006 and 2007. Much knowledge about the data generating process, i.e. the Austrian healthcare system and it's reporting procedures was gathered during this first era of GAP-DRG. Among other challenges, methods for data quality assessment [12] and record linkage [13,14] had to be explored and developed. Additionally, most elements of the infrastructure supporting the database, including secured severs with remote desktop environments for researchers, a virtual private network with two-factor authentication and a wiki system for documentation, were created. In the course of a second delivery, data from Lower Austria's insurance carrier from the years 2008 to 2011 were added. Data quality is controlled routinely and project-wise [12,15].

Overall, the research database GAP-DRG covers the majority of all healthcare services of nearly the entire Austrian population for the years 2006 and 2007. However, some shortcomings (e.g., lacking diagnoses from the outpatient sector) and blind spots (e.g., lacking data from ambulatory care in hospitals) are known and have to be minded in research projects.

Since GAP-DRG constitutes a system that has evolved over time, incorporating various data sources, the resulting data model became rather complex, hard to handle and update. Furthermore, admission control on a user level to improve data governance is hardly possible in the current system.

1.2. The i2b2 framework

Informatics for Integrating Biology and the Bedside (i2b2) [8] is the name of the NIH[5] funded National Center for Biomedical Computing (NCBC) in Boston, Massachusetts and one of their products, a clinical data informatics framework[6] [16]. The source code of the i2b2 platform was released in 2007 under their own open software license.

i2b2 was originally developed for reusing electronic health records (EHR) and other patient centered data for research by providing integration and secure presentation of

[4] About 3% of the population are covered by other insurance carriers (e.g., municipalities, religious orders, unemployment service) and are not included in GAP-DRG.

[5] National Institute of Health: nih.gov

[6] In the context of this paper, the term i2b2 refers to the software framework only.

these data. It is designed to be easily extendable by splitting it up into several modules called cells, which communicate via a web service based on SOAP and REST calls [9]. The conglomeration of these loosely coupled cells constitutes the i2b2 Hive [17]. Data and configuration information are stored in a database system. Additionally, web and desktop clients as well as sample data and VMware images with pre-installed software are available.

Even for basic functionality, several cells must be deployed. Most important for this paper, the CRC ("Clinical Research Chart") cell is the central data repository and the ONT ("Ontology") cell manages the specific hierarchical metadata structure. Additionally, there are cells for e.g. project management, user authentication and system configuration. The central part of the CRC cell holding the core data is modelled as an entity, attribute, value (EAV) [18] schema. Several dimension tables containing information about patients, providers, visits, concepts and modifiers are related with the central fact table. The ONT cell requires a specific hierarchical metadata structure describing the fact and dimension tables as well as providing the foundation for the graphical representation in client software. Besides these basic cells, several specialized cells, e.g. for handling genomic data and natural language processing (NLP), are available.

2. Methods

Our first step was the implementation of the i2b2 platform under the specific circumstances and requirements of the GAP-DRG infrastructure. The second step was to transform the GAP-DRG data into the format prescribed by i2b2 and to load it into the platform. In the third step, which we are currently working on, we aim to examine the benefits of i2b2 in comparison to applying manually developed database queries for typical tasks of clinical research projects. For this purpose, we plan to re-enact an earlier project [19] that was conducted without i2b2.

2.1. Installation of i2b2

The installation process of the i2b2 hive is known to be complex [16] and for our desired combination of operating system (Linux) and database (PostgreSQL) the official installation guide is affected by several errors. The installation is further complicated by the fact that software components of rather specific and partly long outdated releases are required to work together seamlessly.

Furthermore, the infrastructure surrounding the GAP-DRG database, where we aim to deploy the i2b2 platform, consists of highly secured and therefore conservatively managed Linux (CentOS 5) servers without internet access and administration rights for the authors. According to system administration policy we were not allowed to install i2b2 in a traditionally virtualized environment or even use the readily available image.

In recent years operating system level virtualization, also known as software containers, like docker[7] [20,21] evolved rapidly. A docker container is an instance of a docker image running (ideally) exactly one single application. Images can be defined and built using Dockerfiles, which consist mostly of Linux shell scripts installing and configuring the requested application. Docker images require the docker environment

[7] docker.com

only and are largely independent of the underlying Linux system. Therefore, the deployment of containers consists primarily of exporting the prepared images and importing them on the targeted server. Additionally, some minor adjustments of the docker-compose configuration (a tool for container orchestration) have to be implemented, considering aspects of the hardware resources and network configuration.

Main advantages of a containerized solution are the reproducible and automatic building process, the low resource overhead of a running container and integrated options for serialization and transportation of images. After considering all these aspects, we arrived at the decision that software containers based on docker would be a favourable solution for bringing i2b2 to the GAP-DRG infrastructure.

2.2. Extraction, transformation, and loading GAP-DRG data into i2b2

Data from the GAP-DRG database was extracted, transformed and loaded (ETL process) for development and testing. As the development environment did not fulfil the security requirements for handling the original GAP-DRG data, data exported for testing were randomly distorted (gender, year of birth, dates of service utilization, association between individuals and claimed services). This allowed us to test with insensitive data that still resembled the original structure and size. Except for the random distortion, the developed ETL procedures were designed and tested as reproducible processes. This subsequently allowed us to re-enact the ETL process with the original data on the target system.

2.3. Examination of i2b2's benefits in a real-world GAP-DRG research project

In an earlier GAP-DRG based research project we analyzed by means of manually developed database queries and statistical methods whether young parenthood may be seen as a risk factor for myocardial infarctions [19]. We now plan to re-enact the task cohort formation and exploration for this project within the new i2b2 framework. This should help us to assess the benefits of i2b2 in this context by means of a "before-after comparison".

At the moment, manually developed SQL scripts are the most commonly used way to form and explore cohorts in GAP-DRG. We aim to examine how users experience working with i2b2, to what extent the currently applied methods that require skilled and experienced programmers can be facilitated or even replaced by the new graphical user interface provided by i2b2, and whether the group of GAP-DRG users may be broadened by means of i2b2. As the current user group is limited, the assessment of i2b2's benefits will be conducted with a small sample of persons, namely (i) two medical doctors being experts on the Austrian healthcare system without programming experience but knowledge of GAP-DRG, (ii) two experienced database developers with knowledge of GAP-DRG, and (iii) two persons with computer skills but without prior knowledge of GAP-DRG will take part in a black-box testing approach. After a short introduction to the system, they will try to execute basic cohort formation and exploration tasks with i2b2.

Several aspects concerning accessibility, ease of use, required knowledge, total amount of time taken to acquire a result, ratio of productive work and undirected searching, and number of deadlocks where help is required will be extracted from a screen recording and a short interview after the tasks have been accomplished. Additionally, the database developers will be asked to implement the same tasks with

current tools to gather hard facts and personal feedback on actual differences between both approaches.

Summarizing, we plan to gather information on the differences between current workflows and i2b2 by involving three types of possible users. Advantages and shortcomings of the new system should allow us to find a suitable niche for the new tool, give information on an integrated workflow and might even help to improve both established and novel approaches.

3. Results

3.1. Installation of i2b2

i2b2 was successfully installed and tested data using software containers based on docker. As typical docker containers only run a single application, we separately prepared three main components of the i2b2 framework. We accompanied the Dockerfiles with several supporting shell scripts and developed and tested them to handle all i2b2 1.7.04 to 1.7.06 releases. These scripts allow various combinations of software components and versions to be chosen during the building process, preparing for prompt and flexible reactions to new software releases.

First, we deployed i2b2 to the PostgreSQL database of GAP-DRG. We created appropriate PostgreSQL instances for development and testing. To cope with the expected amount of data, a columnar storage [22] extension (cstore_fdw from citusdata) that was implemented as foreign data wrapper, was integrated with the official docker images of PostgreSQL 9.3 and PostgreSQL 9.4 as well as in the database of the GAP-DRG platform. Routines to load different versions of i2b2's test data and procedures were added as well.

Second, i2b2's web client and administration interface was installed on an official docker image of Apache httpd and PHP 5.6. Several extensions (i.e. ExportXLS, GIRI, CARE Concept Demographic Histograms, ARE Concept Observation Tally Histograms, WISE Annotator, WISE Searcher, IDRT Web Client Plugin [23]) were integrated and configured during the build process.[8]

Third, the i2b2 hive was prepared using a vanilla CentOS 7 docker image. In addition, Ubuntu 14.10 and baseimage-docker from Phusion were equipped and tested to run the readily installed hive as well. An i2b2 image including the very specific but largely outdated software products from the installation manual was implemented as intended by the developers. Additionally, docker images providing an i2b2 installation based on all later main releases of Java, JBoss Application Server known as Wildfly, Apache Axis and Apache Ant in various combinations were prepared and tested successfully. All i2b2 core cells, the patients count plugin for the CRC cell as well as the GIRI (Generic Integration of R into i2b2) [24] and IDRT [23] extension were integrated and configured during the build process.

Especially the integration of the R extension GIRI required special attention. GIRI was deployed as an i2b2 cell inside the hive. It exchanges data and plots with the webclient extension using the local file system by writing directly to the webserver's folders. As the hive and the webserver were split into different docker images in our installation, we had to employ sharing volumes between both containers.

[8] The IDRT plugin could not be tested successfully even after receiving help from the original authors.

In addition to these three core images, several supporting images were created to ease the database setup, configuration of several runtime parameters, create and manage new projects as well as create backups. Also, monitoring and logging software as well as database clients were prepared as docker images. docker-compose was applied to calibrate and coordinate all these independent components, configure shared volumes, connect network interfaces and allocate hardware resources.

We made some inconvenient experiences with existing public i2b2 resources. Besides unclear and erroneous sections of the installation guide and hard to interpret error messages, we had to tackle and patch actual software bugs. By providing means to flexibly combine various software and database versions as well as test patches in a reproducible way, our docker-based building routines helped to find and correct the flaws.

A task that we are still working on is the installation of i2b2's desktop client. Whereas the existing Windows 7 version works fine, no official build exists for Ubuntu Linux that we run our test system on. A self-prepared build proved unstable. On the GAP-DRG CentOS 5 Linux system we have not yet been able to build an executable desktop client due to lacking administrator privileges. This is planned for the next system update cycle by the GAP-DRG administrator. As a workaround we used plain database queries for the ETL process and the i2b2 webclient as main user interface.

3.2. Extraction, transformation, and loading of GAP-DRG data into the i2b2 platform

All data that were used in [19] were exported from GAP-DRG. The extract was then transformed to the CRC cell's entity-attribute-value (EAV) data model [18]. As required by i2b2, contacts (in our case claimed healthcare services such as hospital episodes) were modelled as entities, whereas their properties (e.g. time interval, diagnoses, etc.) were represented as attributes and corresponding type-specific values. Also modifiers, which allow to include additional information about an attribute (e.g., whether a diagnosis is regarded as main or additional diagnosis) were applied. Further, metadata from the GAP-DRG database (e.g., diagnosis and medication-related terminologies, demographics, and details of the financing system) were transformed into the defined schema.

Next, the ontology and the hive's configuration were transferred to the targeted database inside GAP-DRG. After completing the ETL test phase with the distorted data in our test environment, all project related data (stored in tables of the CRC cell) were extracted, transformed, and loaded without any distortion inside the secure target environment. Finally, a working i2b2 instance including several extensions and a project based on data from the GAP-DRG research database was prepared for deployment. After successful tests, the deployment was carried out on a separate development server without problems.

3.3. Examination of i2b2's benefits in a real-world GAP-DRG research project

Currently we are still working on re-enacting the tasks of cohort formation and data export for statistical analysis for the project described by Endel et al. [19] within the new i2b2 framework. First preliminary tests of the task cohort formation indicated several useful features of i2b2 such as iterative and interactive discovery of research cohorts, collaboration between distant individuals by enforcing role based access limitation, security policies and automatic documentation at the same time. We aim to present a detailed comparison of the handling of the two tasks with either conventional manually developed database queries or the i2b2 framework at the conference.

4. Discussion

The goal of this work was to examine, whether a large database containing claims data from the Austrian healthcare system could be integrated into the i2b2 framework and to clarify resulting benefits compared to previous methods of data handling. Despite extensive available i2b2 background resources, the installation process proved to be complex. Besides having to deal with erroneous documentation and outdated official i2b2 system components, we had to deviate from standard installation procedures due to restrictions implied by our highly secured target hardware infrastructure. Under the given circumstances, we decided to organize the different i2b2 components in separate docker-based software containers. Hereby, a flexible, reproducible, and resource-efficient installation of i2b2 was achieved. To ease the installation process for other novel i2b2 adopters, we plan to make the complete source code of our docker-based approach available as git repository on github.com/FlorianEndel.

The transformation of the original data model into i2b2's EAV model was complicated due to unclear documentation but overall rather straight forward. The size of the GAP-DRG database and the resulting vast amount of data accumulated in the fact table required the application of novel technical approaches utilizing PostgreSQL's foreign data wrappers and a column oriented storage engine. The creation of the metadata tree for the ontology cell and the integration of other annotating information proved to be cumbersome. More documentation on the semantics and interaction between tables and cells would have been of great help. After gathering first experiences, extending and optimizing the integrated data was a much swifter process.

The comparison of data handling with i2b2 compared to previous methods is currently work in progress. We aim to report corresponding strengths and limitations of i2b2 at the conference.

References

[1] H. H. Pham, D. Schrag, A. S. O'Malley, B. Wu, and P. B. Bach, "Care patterns in Medicare and their implications for pay for performance," *New England Journal of Medicine*, vol. 356, no. 11, pp. 1130–1139, 2007.

[2] S.-H. Cheng, Y.-F. Hou, and C.-C. Chen, "Does continuity of care matter in a health care system that lacks referral arrangements?," *Health Policy and Planning*, vol. 26, pp. 157–162, Mar. 2011.

[3] F. Chini, P. Pezzotti, L. Orzella, P. Borgia, and G. Guasticchi, "Can we use the pharmacy data to estimate the prevalence of chronic conditions? a comparison of multiple data sources," *BMC Public Health*, vol. 11, p. 688, Sept. 2011.

[4] G. Duftschmid, W. Dorda, G. Endel, K. Fröschl, W. Gall, W. Grossmann, and M. Hronsky, "Fragmentation of diabetes treatment in Austria - an indicator for the need for shared electronic health record systems," *Studies in Health Technology and Informatics*, vol. 180, pp. 667–671, 2012.

[5] S. Thurner, P. Klimek, M. Szell, G. Duftschmid, G. Endel, A. Kautzky-Willer, and D. C. Kasper, "Quantification of excess risk for diabetes for those born in times of hunger, in an entire population of a nation, across a century," *Proceedings of the National Academy of Sciences*, Mar. 2013.

[6] S. K. Sauter, L. M. Neuhofer, G. Endel, P. Klimek, and G. Duftschmid, "Analyzing healthcare provider centric networks through secondary use of health claims data," in *Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on*, pp. 522–525, IEEE, 2014.

[7] C. Rinner, S. K. Sauter, L. M. Neuhofer, D. Edlinger, W. Grossmann, M. Wolzt, G. Endel, and W. Gall, "Estimation of severe drug-drug interaction warnings by medical specialist groups for Austrian nationwide eMedication:," *Applied Clinical Informatics*, vol. 5, no. 3, pp. 603–611, 2014.

[8] S. N. Murphy, M. E. Mendis, D. A. Berkowitz, I. Kohane, and H. C. Chueh, "Integration of clinical and genetic data in the i2b2 architecture," in *AMIA Annual Symposium Proceedings*, vol. 2006, p. 1040, American Medical Informatics Association, 2006.

[9] S. N. Murphy, M. Mendis, K. Hackett, R. Kuttan, W. Pan, L. C. Phillips, V. Gainer, D. Berkowicz, J. P. Glaser, I. Kohane, and others, "Architecture of the open-source clinical research chart from Informatics for Integrating Biology and the Bedside," in *AMIA annual symposium proceedings*, vol. 2007, p. 548, American Medical Informatics Association, 2007.

[10] V. G. Deshmukh, S. M. Meystre, and J. A. Mitchell, "Evaluating the informatics for integrating biology and the bedside system for clinical research," *BMC Medical Research Methodology*, vol. 9, no. 1, p. 70, 2009.

[11] P. Filzmoser, A. Eisl, and F. Endel, ATC –> ICD: Determination of the reliability for predicting the ICD code from the ATC code. 2009.

[12] F. Endel, "Understanding data quality in linked administrative data," in *International Health Data Conference 2014*, (Vancouver), 2014.

[13] F. Endel, G. Endel, and N. Pfeffer, "PRM34 Routine Data in HTA: Record Linkage in Austrias GAPDRG Database," *Value in Health*, vol. 15, p. A466, Nov. 2012.

[14] F. Endel, G. Endel, B. Weibold, and H. Katschnig, "Health service record linkage in a situation of multiple social health insurance institutions: the case of Austria," in *Proceedings of The methodological challenges of record linkage*, (University of St. Andrews, Scotland), Sept. 2011.

[15] F. Endel and H. Piringer, "Data Wrangling: Making Data Useful Again," in *MATHMOD 2015 Vienna -Abstract Volume*, vol. 8, (Vienna University of Technology, Vienna, Austria), pp. 111–112, ARGESIM and ASIM, German Simulation Society, Div. of GI – German Society for Informatics / Informatics andLife Sciences, 2015.

[16] E. K. Johnson, S. Broder-Fingert, P. Tanpowpong, J. Bickel, J. R. Lightdale, and C. P. Nelson, "Use of the i2b2 research query tool to conduct a matched case–control clinical research study: advantages, disadvantages and methodological considerations," *BMC medical research methodology*, vol. 14, no. 1, p. 16, 2014.

[17] S. N. Murphy, G. Weber, M. Mendis, V. Gainer, H. C. Chueh, S. Churchill, and I. Kohane, "Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2)," *Journal of the American Medical Informatics Association*, vol. 17, pp. 124–130, Mar. 2010.

[18] P. M. Nadkarni, "QAV: querying entity-attribute-value metadata in a biomedical database," *Computer Methods and Programs in Biomedicine*, vol. 53, pp. 93–103, June 1997.

[19] F. Endel, S. Sauter, L. Koller, A. Niessner, and G. Duftschmid, "Reusing claims data to assess parenthood as risk factor for myocardial infarction," *Studies in Health Technology and Informatics*, vol. 210, pp. 979–979, 2015.

[20] J. Fink, D. S. Librarian, and M. University, "Docker: a Software as a Service, Operating System-Level Virtualization Framework," *The Code4Lib Journal*, July 2014.

[21] D. Merkel, "Docker: Lightweight Linux Containers for Consistent Development and Deployment," *Linux J.*, vol. 2014, Mar. 2014.

[22] D. J. Abadi, P. A. Boncz, and S. Harizopoulos, "Column-oriented database systems," *Proceedings of the VLDB Endowment*, vol. 2, no. 2, pp. 1664–1665, 2009.

[23] C. R. K. D. Bauer, T. Ganslandt, B. Baum, J. Christoph, I. Engel, M. Löbe, S. Mate, S. Stäubert, J. Drepper, H.-U. Prokosch, A. Winter, and U. Sax, "Integrated Data Repository Toolkit (IDRT): A Suite of Programs to Facilitate Health Analytics on Heterogeneous Medical Data," *Methods of Information in Medicine*, vol. 54, Nov. 2015.

[24] Weinlich, B., Mate, S., Prokosch, H.U., Ganslandt, T., and Toddenroth, D., "„R-Scriptlets" für i2b2-Endanwender," 2014

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-253

How Many Patients Could Benefit From Pre-emptive Pharmacogenomic Testing and Decision Support? A Retrospective Study Based on Nationwide Austrian Claims Data

Wolfgang KUCH[a], Christoph RINNER[a], Walter GALL[a] and Matthias SAMWALD[a,1]
[a] *Center for Medical Statistics, Informatics, and Intelligent Systems. Medical University of Vienna*

Abstract. Pre-emptive pharmacogenetic (PGx) testing combined with clinical decision support is a promising new strategy for making pharmacotherapy safer and more effective. To estimate the number of patients whose therapies could be guided by this approach, we analysed claims data for patients in Austria in the years 2006 and 2007. We calculated the number of patients receiving one or several drugs for with pharmacogenomic guidelines are available (PGx drugs). The cohort consisted of 6,761,034 patients and was split into four age groups. Patients in the age group >= 65 were prescribed the most PGx drugs, with 72% of the patients receiving at least one PGx drug. 39.1% of all people over 65 received at least one drug metabolized by the three most frequent cytochrome P450 enzymes. Our data indicate that a sizable fraction of elderly patients could profit from the implementation of pre-emptive PGx testing and decision support.

Keywords. pharmacogenomics, clinical decision support systems, clinical guidelines, drug safety, claims data, re-use of patient data

1. Introduction

Pharmacogenomic (PGx) testing of patients offers the potential of making drug therapy safer and more effective by adapting drug dosing to individual genetic profiles of patients.

Previous research shows that actionable genotypes are not uncommon in patient populations. For example, a study at the Vanderbilt University Medical Center conducted PGx tests in 9,589 patients receiving pharmacotherapy [1]. They found that the fraction of patients with an actionable genetic profile was 28.5% for clopidogrel, 25.7% for simvastatin, 69% for warfarin, 9.1% for thiopurines and 23% for tacrolimus [2].

Several organizations made it their priority to publish PGx guidelines for clinical use. The most established are the Clinical Pharmacogenetics Implementation Consortium (CPIC) in the United States and the Dutch Pharmacogenetics Working Group (DPWG) in Europe [3,4].

While PGx testing is now becoming available at very low cost, its widespread adoption is hindered by a lack of information on its cost-effectiveness and uncertainty about how to integrate pharmacogenomic testing most efficiently into existing clinical

[1] Corresponding Author. Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, BT88, Floor 3, Spitalgasse 23, 1090 Vienna, Austria. E-Mail: matthias.samwald@meduniwien.ac.at

workflows. A promising approach for implementation is *pre-emptive* PGx testing. In pre-emptive testing, a genetic test yielding results for all important pharmacogenes is done once. With the assistance of clinical decision support software, the results can then be used to guide drug therapy with a wide variety of common medications in later patient care episodes. Several solutions for making PGx data and decision support tools are available, such as systems integrated into electronic health records or solutions based on mobile technologies [5].

The guidelines for specific combinations of genetic variants and pharmaceuticals are classified based on the clinical significance of potential adverse events. In guidelines from the DPWG, clinical significance is scored on a seven-point scale, with mild effects without clinical relevance classified as *AA* (lowest impact), and potentially lethal side-effect classified as *F* (highest impact) [4].

The aim of this study was to estimate the potential of pre-emptive PGx tests in Austria using drug claims data. Statistics on the usage of drugs for which PGx guidelines are available were analyzed for various age groups. These data can be used to estimate the potential reach of different PGx implementation scenarios in clinical settings.

2. Methods

2.1. Patient data source

We queried health claims data from the *General Approach for Patient-oriented Outpatient-based Diagnosis Related Groups* (GAP-DRG) database operated by the Main Association of Austrian Social Security Institutions. The database contains pseudonymized health claims data of patients who were enrolled in any public Austrian social insurance. Data are available for claims made in the years 2006 and 2007. The database was accessed through SQL queries. To access this database, a VPN connection secured with a password and single-use token was used to ensure data security.

2.2. Guidelines and data preparation

Our analysis was based on the guidelines from the DPWG [6]. Depending on the clinical relevance of outcomes for each drug-gene pair, two lists were generated.

The first list, named ‘Highly significant’, included all drug-gene pairs with a clinical significance rated C-F, while the other list, ‘DPWG’, included all drug-gene pairs irrespective of clinical significance of potential adverse outcomes. The study cohort was split up into four age groups (i.e. 0-13, 14-39, 40-64 and >= 65 years).

Table 1 lists all genes and the related drugs mentioned in the DPWG Guidelines and with a clinical relevance C-F (i.e., the ‘Highly significant’ list).

Since most drugs are metabolized by one of three cytochrome P450 enzymes (CYPs: CYP2D6, CYP2C9 and CYP2C19) and targeted genotyping for only these three genes might be easier to implement than for the entirety of PGx genes, we also calculated separate statistics for drugs metabolized by these specific genes only.

To identify claims referring to PGx medications within the GAP-DRG database, a document with all ATC codes for the drugs named in the guidelines was generated with the official ATC/DDD-Index 2007 [7]. 167 ATC codes were mapped to the related drugs. Table 2 shows an example with tramadol, metoprolol and clopidogrel.

Table 1. All drug-gene pairs in the DPWG guidelines which contain treatment recommendations of high clinical significance based on PGx test results

Gene	Substances in DPWG guideline ('highly-significant list')
CYP2D6	Amitriptyline, aripiprazole, clomipramine, codeine, doxepin, haloperidol, imipramine, metoprolol, nortriptyline, paroxetine, propafenone, risperidone, tamoxifen, tramadol, venlafaxine
CYP2C9	Acenocoumarol, glimepiride, phenprocoumon, phenytoin
CYP2C19	Clopidogrel, sertraline
UGT1A1	Irinotecan
TPMT	Azathioprine, mercaptopurine, thioguanine
HLA-B44	Ribavirine
HLA-B*5701	Abacavir
CYP3A5	Tacrolimus
VKORC1	Phenprocoumon
Factor V Leiden	Estrogen-containing OC
DPYD	Fluorouracil, capecitabine

Table 2. Examples of mappings between drugs and related ATC-Codes from the ATC-Index 2007

Drug	ATC-Codes
Tramadol	N02AX02, N02AX52
Metoprolol	C07AB02, C07AB52, C07FB02, C07CB02, C07BB02, C07BB52
Clopidogrel	B01AC04

Prodrugs were included, topical preparations of pharmaceuticals were excluded.

2.3. Statistics

The number of distinct drugs mentioned in PGx guidelines prescribed within the two-year period was calculated for each patient. Two hypothetical scenarios were considered:

First, a pre-emptive scenario, in which genetic testing is performed at the very beginning of the 2 years (of the available claims data) for all patients.

Second, a mixed 'reactive pre-emptive' scenario, in which a genetic test of all PGx genes is performed as soon as and exclusively for those patients who were first being prescribed a PGx drug within the timespan covered by the dataset. Each patient receiving a PGx test would also receive at least one PGx drug.

For both scenarios the number of patients receiving a multitude of different PGx drugs within the analyzed time period was calculated.

The data for this study has been extracted from GAP-DRG Database using SQL-Statements and was then transferred to a MS Excel Sheet for analyzing. For each age-group and PGx-druglist a separate Table was used.

3. Results

The database contained data on 6,831,128 patients in total. After exclusion of patients without a known date of birth, 6,761,034 patients (52% female) were included in our study cohort.

Table 3. Number of patients (absolute and percent of age group) with at least one PGx drug from the DPWG or "Highly significant" lists, grouped by age.

	0-13 y	14-39 y	40-64 y	>=65 y
DPWG-list	39,557 (4.6%)	547,441 (25.7%)	1,190,242 (51.1%)	1,037,127 (72.0%)
"Highly-significant"- list	33,483 (3.9%)	178,809 (8.4%)	599,811 (25.7%)	680,052 (47.2%)

3.1. DPWG list versus 'Highly significant' list

22.07% of the total included patients received at least one drug of the 'Highly significant' list and 41.63% received at least one drug of the 'DPWG' list (Table 3).

The most frequently prescribed drugs from the "DPWG"-List were proton pump inhibitors (i.e., pantoprazole, lansoprazole, omeprazole), while the most frequently prescribed drugs from the "Highly-significant"-list were tramadol, estrogen-containing drugs and metoprolol. Table 4 shows more details.

Table 4. Three most frequent prescribed drugs from DPWG and "Highly significant" lists and % of the included patients receiving each drug within the study period.

	DPWG	Highly-significant
Rank 1	Pantoprazole (15.1%)	Tramadol (6.1%)
Rank 2	Lansoprazole (8.6%)	Estrogen-containing drugs (5.9%)
Rank 3	Omeprazole (6.8%)	Metoprolol (4.1%)

In the pre-emptive simulation, 22.07% of all included patients received at least one PGx drug. In the cohort aged >= 65, even 47.24% received a PGx drug. 6.38% of the total population received at least two PGx drugs.

In the 'reactive pre-emptive' simulation, all of the 1,060,860 patients who received at least one drug are tested. With this selective inclusion to PGx testing, 22% receive two or more drugs, compared to 6% in the pre-emptive simulation.

3.2. Testing for CYP genes only

13.48% of the total included patients received at least one drug which is metabolized by the enzyme CYP2D6. Considering all three CYPs (CYP2D6, CYP2C9, CYP2C19) the number rose to 17.58%. In the >= 65 age group, 41.53% received at least one drug of these three CYPs. Table 5 illustrates these numbers with all age groups and only with the patients over 65.

Table 5. Percentage of total population in highly significant list in all age groups and only age group >= 65y

CYP-Enzymes	% of total population in highly-significant list all age groups	% of total population in highly-significant list with age group >= 65y
CYP2D6	13.5%	30.4%
CYP2C9	3.7%	12.8%
CYP2C19	3.5%	8.5%
All three CYPs	17.6%	41.5%

Table 6. Pre-emptive and 'reactive pre-emptive' simulations with the DPWG list and 'Highly significant' list with the percentage of the total study population who received PGx drugs. R-P: result for 'reactive pre-emptive' setting, P: result for pre-emptive setting.

	Pre-emptive simulation	**"Reactive pre-emptive" simulation**
At least one PGx drug	R-P: 41.2%, P: 22%	R-P: 100%, P: 100%
At least two PGx drugs	R-P: 19.5%, P: 6.4%	R-P: 48.4%, P: 28.9%
At least three PGx drugs	R-P: 9.9%, P: 1.7%	R-P: 23.9%, P: 7.9%

Pharmacogenetic testing for these 3 CYPs would already allow to optimize the prescriptions of about one third of the total population aged over 65.

Based on the DPWG list, a PGx test for the population over the age 65 for all three CYPs would include approximately 14.8% of the total study population and 69.59% of the study population over age 65.

3.3. Pre-emptive versus 'reactive pre-emptive' setting

In the pre-emptive setting, around 41% of the total included patients receives at least one drug from the DPWG guidelines and about 20% receives at least two drugs. In comparison to the reactive pre-emptive simulation, about 48% of all patients, who received at least one drug, are prescribed at least two drugs. Table 6 compares the two settings for both the DPWG and "Highly significant" lists.

4. Discussion

Patients who received two or more PGx drugs ranged between 6.38% in the 'Highly significant' list and 19.95% in the DPWG-list.

A pharmacogenetic test for patients aged 65 and over, including the three most frequent CYP-Enzymes (CYP2D6, CYP2C9, CYP2C19), covers about 39.1% of all patients in this cohort using the DPWG list.

If we take the amount of clinically actionable genotypes for clopidogrel CYP2C19 (28.5%) from the PREDICT study and put them into perspective with our results for all drugs in our test, about 20 % of all included patients aged 65 or older have at least one actionable result, for which the guidelines of DPWG recommends a different prescribing procedure [2]. The most frequently prescribed drugs are proton pump inhibitors (PPI), followed by the pain reliever tramadol. Using only the "Highly significant" list, the most frequently prescribed drugs are the analgesic tramadol followed by the cardiologic pharmaceutical metoprolol.

The main focus of this analysis was on the use of PGx drugs in the 'Highly-significant' list. About a quarter of the total study population and about the half of the patients over age 65 received drugs received drugs with already existing guidelines for avoiding adverse drug events by accounting for pharmacogenetic diversity.

This study has several limitation. The impact of pharmacogenomic testing varies depending on the demographic population and this study used Austrian data only. Also, the costs of genetic testing and the frequency of adverse drug events varies between regions. Finally, the scope of the dataset is limited to the two years covered by the GAP-DRG database. Future work could employ additional databases such as GAP-DRG2, which contains Austrian claims data from the years 2008-2011.

Acknowledgement

We want to thank the Main Association of Austrian Social Security Institutions and especially Gottfried Endel for granting us access to their database.

References

[1] Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, et al. Operational implementation of prospective genotyping for personalized medicine: The design of the Vanderbilt PREDICT project. Clin Pharmacol Ther. 2012 Jul;92[1]:87–95.

[2] Van Driest SL, Shi Y, Bowton EA, Schildcrout JS, Peterson JF, Pulley J, et al. Clinically actionable genotypes among 10,000 patients with preemptive pharmacogenomic testing. Clin Pharmacol Ther. 2014 Apr;95[4]:423–31.

[3] Caudle KE, Klein TE, Hoffman JM, Müller DJ, Whirl-Carrillo M, Gong L, et al. Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. Curr Drug Metab. 2014 Feb;15[2]:209–17.

[4] Swen JJ, Wilting I, de Goede AL, Grandia L, Mulder H, Touw DJ, et al. Pharmacogenetics: from bench to byte. Clin Pharmacol Ther. 2008 May;83[5]:781–7.

[5] Samwald M, Adlassnig K-P. Pharmacogenomics in the pocket of every patient? A prototype based on quick response codes. J Am Med Inform Assoc JAMIA. 2013 May 1;20[3]:409–12.

[6] Swen JJ, Nijenhuis M, de Boer A, Grandia L, Maitland-van der Zee AH, Mulder H, et al. Pharmacogenetics: From Bench to Byte— An Update of Guidelines. Clin Pharmacol Ther. 2011 May;89[5]:662–73.

[7] Wissenschaftliches Institut der AOK (WIdO): Amtlicher ATC-Code [Internet]. [cited 2016 Jan 18]. Available from: http://www.wido.de/amtl_atc-code.html

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-259

A Smartwatch-Based Assistance System for the Elderly Performing Fall Detection, Unusual Inactivity Recognition and Medication Reminding

Markus DEUTSCH[a] and Harald BURGSTEINER[a,1]
[a]*Institute for eHealth, Graz University of Applied Sciences, Graz, Austria*

Abstract. The growing number of elderly people in our society makes it increasingly important to help them live an independent and self-determined life up until a high age. A smartwatch-based assistance system should be implemented that is capable of automatically detecting emergencies and helping elderly people to adhere to their medical therapy. Using the acceleration data of a widely available smartwatch, we implemented fall detection and inactivity recognition based on a smartphone connected via Bluetooth. The resulting system is capable of performing fall detection, inactivity recognition, issuing medication reminders and alerting relatives upon manual activation. Though some challenges, like the dependence on a smartphone remain, the resulting system is a promising approach to help elderly people as well as their relatives to live independently and with a feeling of safety.

Keywords. Falling, Machine Learning, Neural Networks, Mobile Applications, Reminder Systems.

1. Introduction

Elderly people face a number of challenges in their daily lives which makes it necessary for many of them to be looked after by their relatives or health care professionals. Such challenges include forgetfulness and the tendency to fall as well as the fact that elderly people tend to live alone after the death of a spouse. The number of people suffering from dementia in e.g. Austria is expected to double from 120,600 in 2010 to 262,200 in 2050 [1]. The numbers may scale to other industrial countries as well. Forgetfulness is especially risky when those affected are prescribed to a medical therapy, because forgetfulness is one of the main factors that contribute to low medical adherence [2], which is associated with decreased therapeutic success and higher mortality [3]. Their tendency to fall is another risk for elderly people. 30% of people aged 65 years or older fall at least once every year, and 20% of those require medical treatment afterwards [4]. The physical injuries sustained from such incidents are not the only consequence: 60% of elderly people with a history of falling develop a fear of falling [5], which is associated with a loss of physical capabilities which ultimately leads to a decrease in mobility level, therefore making the person more likely fall again [5][6]. In case of falling or any kind of emergency it is important that elderly people have easy access to help, like a phone or a spouse that is able to assist them. With growing age the number of people living alone

[1] Corresponding Author: Harald Burgsteiner, Graz University of Applied Sciences, eMail: harald.burgsteiner@fh-joanneum.at

is growing [7], which makes falling an even more serious danger, because half of elderly people cannot get up by themselves after a fall, even if they are not injured [8].

The advance of smart technologies in recent years offers new ways of enabling elderly people to live a safe and self-determined life with minimal dependence on other people even at a high age. Smartphones and smartwatches make it easy to get in touch with relatives or emergency services if required, and allow them to send notifications to the user's wrist. Furthermore, smartwatches can track the behaviour of their users even if the smartphone is out of reach, detect unusual behaviour and react accordingly.

The purpose of this paper is to investigate the possibilities and challenges when creating an assistance system for elderly people based on consumer devices. It proposes a smartwatch-based assistance system which is able to recognize emergency situations by performing fall detection and inactivity recognition, besides offering the possibility to set medication reminders and get help from relatives on the push of a single button.

2. Methods

To implement the smartwatch-based assistance system ("Carrie") a suitable smartwatch was selected and fall detection, inactivity recognition as well as medication reminders were developed based on current research in the respective areas.

2.1. Smartwatch

An investigation of the smartwatch market showed that a number of watches for iPhones and Android devices exist. For the use in the scenarios described above, a smartwatch with three dimensional acceleration sensors is needed. Additionally, the battery life should supports many days of activity before recharging is required. For the implementation of the smartwatch-based assistance system the Pebble smartwatch was chosen because it is capable of serving both Android- and iOS-powered devices, has a battery runtime of up to 7 days according to the manufacturer, is available for less than 100 Euros and can be customized by creating own apps with an actively maintained SDK. Its ability to vibrate and the built-in acceleration sensor make it possible to alert the user and track his or her activity while wearing the watch. Other smartwatches are usually more expensive, only available for one platform or have a poor battery capacity.

2.2. Fall detection

The objective was to detect falls based on the acceleration data captured by the smartwatch. A lot of research exists in the field of accelerometer-based fall detection and [9] distinguishes between analytical and machine learning-based approaches to fall detection. While the former is mostly threshold-based, the latter uses mathematical models for the automatic classification of movements after a training period. For Carrie an approach based on machine learning utilizing a multilayer perceptron was chosen. To train the multilayer perceptron it was necessary to collect data of activities of the daily life and falls. Seven healthy subjects (aged 14, 23, 25, 68, 83 and two women of 76) were asked to walk, run, get up and sit down while the watch was recording their movements. The three youngest participants were also asked to perform falls to the left, right, front and back. Additionally data was recorded while the watch was vibrating and while the people were idle. Data was collected at 50Hz by the smartwatch's 3D acceleration sensor

and transmitted to the smartphone for further processing, because the smartwatch lacks the computing power necessary to perform fall detection [10]. Since the watch is not able to continuously transmit the x-, y- and z-readings of the sensor at a rate of 50 times per second, the Euclidean norm of the values was calculated and sent to the phone where the features needed for training the classifier were extracted:

$$Euclidean\ norm := \sqrt{x^2 + y^2 + z^2} \tag{1}$$

By using the Euclidian norm the information about the direction of the movement and acceleration gets lost. Tests showed that this information is not needed to only detect whether a fall occurred or not. A default input signal (DIS) was extracted from the data stream coming from the smartwatch by identifying the largest deviation from the resting potential (the Euclidean norm at rest is 1000) in the data recorded during the previous 7 seconds. The DIS consists of three seconds of data starting 1.5 seconds before the identified peak and is used to perform further feature extraction.

Three algorithms for fall detection based on a multilayer perceptron with one hidden layer and 10 output neurons (one for each type of movement: idle, walking, running, getting up, sitting down, vibration, falling forward, backward, left and right) were compared. Similar to [11], approach 1 used the DIS directly as network input, thus using 150 input neurons (three seconds of data with 50 samples per second). Approach 2 extracted features from the DIS and used the maximum (A in Figure 1) in the DIS which marks the impact, the minimum value that occurred before A which is the suspected end of the fall (B), the difference in acceleration and time between A and B (intensity of fall and duration of impact), and the duration of the fall which is defined as the time between B and the maximum value before B which is the suspected start of the fall (C). Therefore approach 2 uses 5 input neurons. Approach 3 uses additional features to further improve fall detection: the mean and average values of the DIS were added to incorporate a global

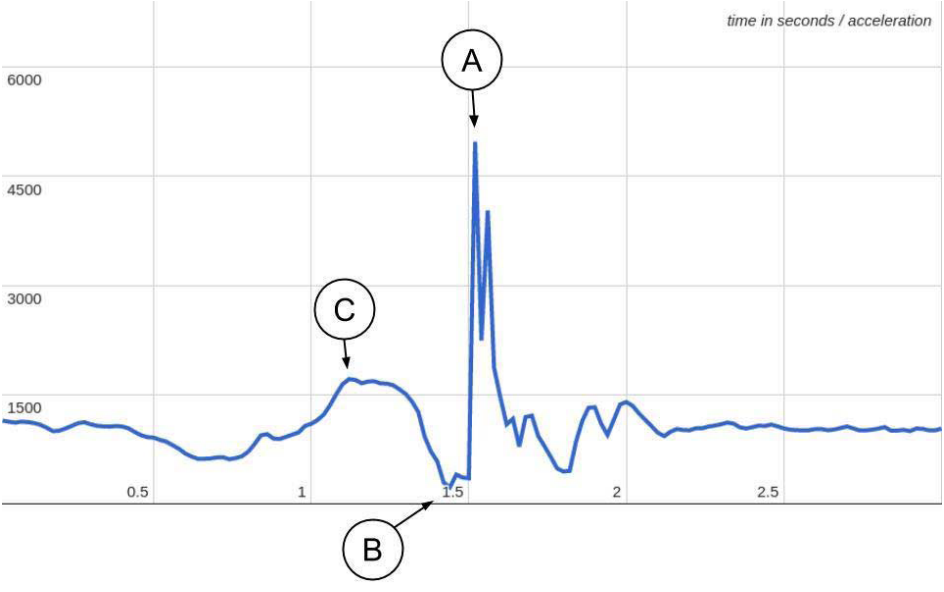


Figure 1: From the DIS features like the suspected impact (A), the end of the fall (B) and the begin of the fall (C) are extracted.

perspective on the sample as well as the difference in acceleration between start and end of the fall (B and C). Eight input neurons were therefore used for classifier 3.

All three algorithms were implemented and an individual neural network was trained and evaluated for each algorithm respectively. Sensitivity and specificity were determined for every algorithm. We also studied the effect of different numbers of neurons on the hidden layer. Algorithm 3 showed slightly better accuracy and significantly better sensitivity compared to the other approaches. It was outperformed slightly by algorithm 1 in terms of specificity, but performed better than all other algorithms when using 10 neurons in the hidden layer. The network was therefore implemented using approach 3 with 8 input neurons, 10 neurons on the single hidden layer and 10 neurons in the output layer.

WEKA (see [12]) was used to create the neural network. Since WEKA does not work on Android out of the box it was stripped of all classes and functionalities that caused compile-time errors when building it for Android. The resulting WEKA package can be used in an Android project to detect falls in real time. After providing WEKA's multilayer perceptron implementation with a network input it returns output values for each possible output class. The class with the highest output value determined the detected movement or fall ("winner-takes-all"). Falls were only reported if the highest resulting output value was associated with a fall and was at least 89%.

2.3. Inactivity recognition

In order to detect suspicious inactivity it is necessary to know the time that has passed since the last activity and usual behaviour patterns of the user. The approach used for inactivity recognition is very similar to the one described by Cuddihy et al. (see [13]), who record the duration of ongoing inactivity twice every hour and match it against a threshold which is calculated based on past inactivity durations that have been recorded on the same time of the day. It was modified to record the duration of ongoing inactivity four times every hour and respect weekly recurring behavioural patterns.

Cuddihy et al. calculate the threshold using the configurable parameters *MP*, *UBP* and *VBP* which are described with the values found reasonable by Cuddihy et al. in Table 1. The interval weights W_r have been adapted to include four more weights in order to respect the hour before and after the given time. The minimum threshold has been set to 30 minutes so that inactivity is never flagged as suspicious if it is 30 minutes or less.

Table 1: Configurable parameters used for inactivity recognition. All parameters and values (except parameter MT and the values for the weights) have been suggested by Cuddihy et al.

Parameter	Description	Value
Maximum Percentile (MP)	The percentile of data considered when determining the threshold. It is used to eliminate outliers.	0.97
Uniform Buffer Percentage (UBP)	The percentage by which the maximum inactivity should always be increased when calculating the threshold.	0.30
Variable Buffer Percentage (VBP)	Determines how much the surrounding intervals affect the threshold. A low VBP increases sensitivity at usually active times.	0.40
Interval Weights (W_r)	Controls the influence of each of the surrounding intervals, where r is relative to the current interval: -4 <= r <= 4.	1,2,3,3,4, 3,3,2,1
Minimum Threshold (MT)	The minimum threshold in minutes that needs to be exceeded in order to create an alert.	30 minutes

The threshold for a time interval i (where i = 0 for 00:00 o'clock, i = 1 for 00:15 o'clock) is calculated using the formula

$$threshold_i = MAX(M_i+UB_i+VB_i, MT) \quad (2)$$

where M_i is the longest inactivity duration that has been recorded one hour before and after the current interval and is calculated as follows:

$$M_i = MAX(m_{i-4}., m_{i-3}, m_{i-2}, m_{i-1}, m_i, m_{i+1}, m_{i+2}, m_{i+3}, m_{i+4}) \quad (3)$$

m_i is the 97th percentile of the previously recorded inactivity durations at interval i and is used instead of the real maximum in order to eliminate extremes.

Cuddihy et al. describe UB as a buffer that is proportional to the longest inactivity in the current timeframe and makes sure that inactivities that are slightly longer than the recorded maximum are allowed:

$$UB = UBP * m_i \quad (4)$$

Since UB should actually be different at every interval a slightly modified formula was used to reflect the values at each time interval i:

$$UB_i = UBP * m_i \quad (5)$$

VB_i is a variable buffer that is calculated by computing the weighted sum of variable buffers of all relevant intervals. Each timeframe is assigned a weight W_i that depends on the position of the timeframe relative to the current interval.

$$VB_i = \frac{1}{\sum W_r}[\sum_{r=-4}^{4}(VBP * m_{i+r} * W_r)] \quad (6)$$

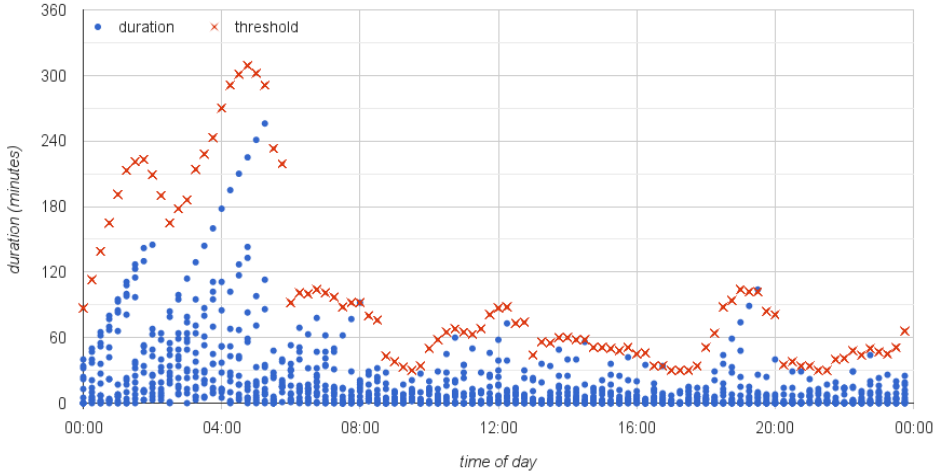


Figure 2: Durations of ongoing inactivity are represented as blue dots. The calculated thresholds are shown as red crosses.

Inactivity durations are recorded together with the number of the interval at which they occurred. The resulting data can be visualized as in Figure 2 where the calculated threshold for each time of day is visible as well. To calculate the threshold, data recorded on the 45 previous days is used. In order to make up for different behaviour of the user on different days of the week (like sleeping longer on weekends) same days of the week are given a higher chance of being used in the calculation: of 45 inactivity recordings up to 27 were recorded on the same weekday. The remaining inactivity recordings are taken from the most recent days. Thresholds are calculated once every day at midnight for the following day.

2.4. Energy efficiency

The constant stream of acceleration data from the Pebble smartwatch to the smartphone via Bluetooth resulted in the smartwatch's battery being drained after approximately 17-19 hours. This is much less than the anticipated seven days of normal operation without fall detection. An analysis showed, that the highest share of energy consumption comes from the wireless transmission of data. Thus, the application on the watch was reconfigured to not transmit a continuous stream of data but to decide about possible falls on the watch itself and restrain data transmission to these suspicious values only. For this pre-detection we retain the acceleration data of the last 6 seconds and only start transmitting the data if a certain threshold has been exceeded. The threshold was set to the Euclidean norm of 1,700 which can be calculated from the raw acceleration data. If a value higher than that is observed the smartwatch starts transmitting data for 30 seconds, sending both data that was recorded before, during and after the possible fall. This had the effect that battery runtime could be increased by at least 180%, from 17-19 to 50-60 hours, depending on the user's activity. The time of the last reception of acceleration data is also used on the smartphone to determine the currently ongoing inactivity for inactivity recognition.

3. Results

The resulting system is capable of detecting falls and suspicious inactivity, as well as providing a way of manually calling for help and issuing medication reminders using a smartwatch and a smartphone.

The assistance system consists of a Pebble smartwatch that communicates with an Android smartphone via Bluetooth. In case of emergency situations alerts are sent to configurable emergency contacts via text messages. It is essential that the smartwatch and smartphone are connected at all times as the smartwatch cannot perform fall detection on its own or send notifications because it does not have any network interfaces except Bluetooth.

Emergency detection is performed using fall detection, inactivity recognition and manual activation. A user can also manually issue an alert by pressing a dedicated button on the smartwatch. To reduce the risk of accidentally issuing a manual alert the smartwatch starts to vibrate and display a notification for 30 seconds during which the user has the chance to cancel the alert by pressing a specific button on the watch. Upon cancellation the currently ongoing inactivity duration is reset and fall alerts are inhibited for a period of one minute in case the cancelled alert originated from a misdetected fall.

If the user does not cancel the alert a predefined list of emergency contacts is notified of the situation. The alert contains the type of emergency and when an accurate GPS signal is available it also contains the user's address and a link to open Google Maps at the user's current geolocation so that the system is also of use if the user is outdoors, for example going for a walk. In case none of the contacts respond via call or text message within 5 minutes they are notified again. Emergency contacts can be configured on the smartphone by picking contacts from the contacts list or manually entering numbers. Contacts can be configured to only be notified if other contacts did not respond within 5 minutes. Hence, it is also possible to notify official help (e.g. an ambulance) if none of the primary helpers react.

To further enhance the usefulness of our application, the Carrie app on the smartphone can also be configured to issue medication reminders on different times of day. Alerts are then pushed at certain times to the smartwatch which starts to vibrate and shows a medication symbol. We decided that no information about which medication to take is included because the display can be hard to read for elderly people and it is assumed that either the user knows which pills to take or the medication has been prepared for them.

4. Discussion

The combination of fall detection, inactivity recognition and manual activation creates a smartwatch-based emergency detection system that promises high reliability because the detection mechanisms complement one another. For example if falling detection does not work the user can still manually trigger an alert or, if unable to do so, suspicious inactivity will be reported after some time has passed. A major weakness in the system's architecture is the heavy dependence on the availability of a smartphone which should therefore always be within a few meters of the user (signal range of the Bluetooth connection).

Research by Bagalà et al [14] has shown that machine learning algorithms that had been trained using staged falls tend to be less accurate when applied to real falls. Therefore, in the next phase of the project, real movement data of active elderly people will be collected in controlled environments like nursing homes to be able to better train the neural network. Additionally, the fall detection algorithm might yield better results when it is customized for the user. Therefore it might prove beneficial to walk the user through a setup process when first using the system, where he or she is instructed to perform activities of daily live which, together with pre-recorded falls, are used to train the neural network online. Furthermore, if a fall is mistakenly detected and the user cancels the alert, the signal that caused the alert could be used to further train the network so the same movement won't be classified as fall in the future. Further improvements to the fall detection could include monitoring a period of time after a suspected fall. If there are no normal activities an alert can be issued, while it can be cancelled if it is detected that the user is for example walking again. Future research should focus on evaluating and improving the emergency detection capabilities in field tests and clinical studies, as well as testing the system's acceptance and usability on elderly people.

References

[1] Höfler, Sabine; Bengough, Theresa; Winkler, Petra; Griebler, Robert (Hg.) (2015): Österreichischer Demenzbericht 2014. Bundesministerium für Gesundheit und Sozialministerium, Wien.
[2] Vervloet, M., Linn, A.J., van Weert, J.C., De Bakker, D.H., Bouvy, M.L., & Van Dijk, L. (2012). The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. Journal of the American Med. Inf. Assoc., 19(5), 696-704.
[3] Rasmussen, J.N., Chong, A., & Alter, D.A. (2007). Relationship between adherence to evidence-based pharmacotherapy and long-term mortality after acute myocardial infarction. Jama, 297(2), 177-186.
[4] Kannus, P., Sievänen, H., Palvanen, M., Järvinen, T., & Parkkari, J. (2005). Prevention of falls and consequent injuries in elderly people. The Lancet, 366(9500), 1885-1893.
[5] Legters, K. (2002). Fear of falling. Physical therapy, 82(3), 264-272.
[6] Vellas, B.J., Wayne, S.J., Romero, L.J., Baumgartner, R.N., & Garry, P.J. (1997). Fear of falling and restriction of mobility in elderly fallers. Age and ageing, 26(3), 189-193.
[7] Statistik Austria, Lebensformen, http://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/haushalte_familien_lebensformen/lebensformen/index.html, last access: 2015-08-02.
[8] Alexander, N. B., Ulbrich, J., Raheja, A., & Channer, D. (1997). Rising from the floor in older adults. Journal of the American Geriatrics Society, 45(5), 564-569.
[9] Noury, N., Fleury, A., Rumeau, P., Bourke, A.K., Laighin, G.O., Rialle, V., & Lundy, J.E. (2007). Fall detection-principles and methods. In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE (pp. 1663-1666). IEEE.
[10] Maglogiannis, I., Ioannou, C., Spyroglou, G., & Tsanakas, P. (2014). Fall Detection Using Commodity Smart Watch and Smart Phone. In Artificial Intelligence Appl. and Innovations (pp. 70-78). Springer.
[11] Sengto, A., & Leauhatong, T. (2012). Human falling detection algorithm using back propagation neural network. In Biomedical Engineering International Conference (BMEiCON), 2012 (pp. 1-5). IEEE.
[12] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I.H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.
[13] Cuddihy, P., Weisenberg, J., Graichen, C., & Ganesh, M. (2007). Algorithm to automatically detect abnormally long periods of inactivity in a home. In Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments (pp. 89-94). ACM.
[14] Bagalà, F., Becker, C., Cappello, A., Chiari, L., Aminian, K., Hausdorff, J. M., Zijlstra, W. & Klenk, J. (2012). Evaluation of accelerometer-based fall detection algorithms on real-world falls. PloS one, 7(5), e37062.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-267

ICD-11 (JLMMS) and SCT Inter-Operation

Marzouk MAMOU[a,1], Alan RECTOR[d], Stefan SCHULZ[e], James CAMPBELL[c], Harold SOLBRIG[f], Jean-Marie RODRIGUES[a, b]
[a]*INSERM U1142, LIMICS, Paris, France*
[b]*Department of Public Health and Medical Informatics, Univ.Jean Monnet of Saint Etienne, France*
[c] *University of Manchester, United Kingdom*
[d] *IMI, Medical University of Graz, Austria*
[e]*Department of Internal Medicine University of Nebraska Medical Center Omaha USA*
[f]*Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA*

Abstract. The goal of this work is to contribute to a smooth and semantically sound inter-operability between the ICD-11 (International Classification of Diseases-11th revision Joint Linearization for Mortality, Morbidity and Statistics) and SNOMED CT (SCT). To guarantee such inter-operation between a classification, characterized by a single hierarchy of mutually exclusive and exhaustive classes, as is the JLMMS successor of ICD-10 on the one hand, and the multi-hierarchical, ontology-based clinical terminology SCT on the other hand, we use ontology axioms that logically express generalizable truths. This is expressed by the compositional grammar of SCT, together with queries on axiomsof SCT. We test the feasibility of the method on the circulatory chapter of ICD-11 JLMMS and present limitations and results.

Keywords. Alignment; ICD-11; Interoperation; Ontology; SNOMED CT; Terminology

1. Introduction

This work is a contribution to the interoperability between two terminology standards for healthcare.

- ICD-11 JLMMS [1], the final output ofthe eleventh ICD (International Classification of Diseases) revision, intended to replace the present ICD-10, and
- SNOMED CT (SCT), an international clinical terminology standard, developed and maintained by the IHTSDO (International Health Terminology Standards Development Organization), which aims to cover the whole field of healthcare by codes, terms and logical formalisms, in order to represent the details of the health care process [2,3,4] .

ICD-10 and in the future ICD-11 [1] is available in several official languages of the WHO (World Health Organization) including French, while SCT [2] is fully available in four languages (US/UK English, Spanish, Danish and Swedish),with localisation

[1] Corresponding author: Marzouk MAMOU, INSERM U1142, LIMICS, Paris, France, 20 rue Arthur GROUSSIER, 75010 Paris, France. E-mail: marzouk.mamou@outlook.com

projects underway for French (Canada, Belgium), Dutch (Belgium), and with Lithuanian, and others ones planned [3, 4].

The present, ICD-10 (Classification of Diseases, 10th revision) developed by WHO, the leading standard for mortality and morbiditystatistics, is also used in other health contexts like healthcare documentation and billing within national modifications and extensions.

The current efforts of alignment between ICD-11 and SCT occur at a time when documentation specialists, epidemiologists, healthcare administrators, and health services researchers are identifying more and more use cases where SCT is used in parallel to ICD-10. One reason for this is the need to increase the granularity of clinical content, taking into account the expansion of the resulting medical knowledge, including genomics and its related research.

This parallelism is addressed by the institutional agreement between WHO and IHTSDO, signed in 2010, aiming at harmonizing, between the multi-components architecture of ICD-11 [5] and SCT [6, 7, 8]. In an ongoing process to lay the grounds for semantic interoperability between ICD-11 and SCT, a Joint Advisory Group (JAG), put in place by both organisations, has designed a semantic alignment method, based on the 1998 foundation works [9, 10,] and several ontology design methodologies developed since then [11, 12, 13, 14].

We report in this paper on the application, limitations and results of this method based on previous works on the ICD-11 Foundation Component [15], in order to establish interoperability between the SNOMED CT and the ICD-11 Joint Linearization for Mortality and Morbidity Statistics (JLMMS).

2. Materials and Methods

2.1. Materials

In 2007, WHO launched an ambitious revision processfor ICD [5].After the establishment of the JAG in 2010, there was consensus within the JAG to base the harmonization around a common ontology, according to the widely acknowledged principles [9-14]. The ICD-11 revision was designed as a multi-component architecture [6-8], of which a component, named “Foundation Component” (FC), contains the entirety of knowledge assembled in the revision process, arranging the ICD classes in a poly-hierarchical structure. This repository is intended as the conceptual basis for the generation of so-called linearizations, i.e. mono-hierarchical classifications of mutually exclusive and exhaustive classes, as known from earlier ICD versions. Priority had been given to the “Joint Linearization for Mortality and Morbidity Statistics” (JLMMS), intended to replace the current ICD-10, and is comparable in scope and granularity. The interoperability between ICD-11 – JLMMS and SCT is the primary target of the harmonization process steered by the JAG.

SCT is distributed in relational tables, from which description logics (DL) based version using the OWL-EL profile [14], compatible with the “Short Normal Form” (SNF) of SCT's legacy compositional grammar [16]. The Common ontology was named ICD-11–SCT-CO [6-8],

Fig.1 illustrates the current harmonization architecture.

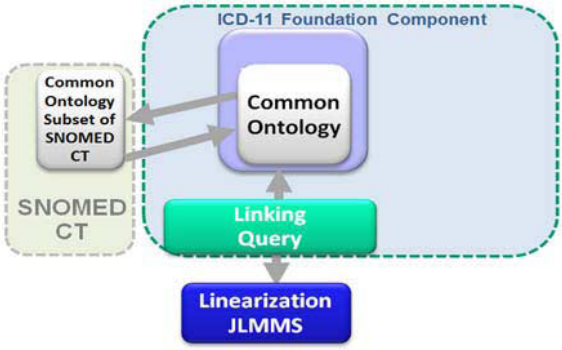


Figure 1. The global view of the new ICD-11 Architecture and its relations with SCT.

JLMMS was extracted from the foundation component to meet the criteria of exhaustiveness, mutually exclusive classes and mono-parenthood relationship. Moreover, to ensure the exhaustiveness, it was necessary to introduce some categories called "Other" and "unspecified". Finally, to be aligned with the organization of ICD-11 in chapters, each chapter has its rules of inclusions and exclusions. For example, the chapter of « circulatory system» excludes infections, neoplasms, endocrine and congenital diseases called "developmental", which have their own chapters.

Our work focus on the JLMMS and its logical representation by SNOMED CT compositional grammar, not on harmonization between the foundation component and SNOMED CT exposed elsewhere [6-8].

In the circulatory chapter of ICD-11 JLMMS, we excluded from the study residual classes ("Other" and "unspecified") because they cannot be represented for they are undefined and the Arrhythmia sub-chapter (Cardiac arrhythmia disorders), because the logical definitions of SCT corresponding concepts are almost all called "Primitives ", meaning they do not provide a complete ontological representation of the SCT concepts descriptions named Fully Specified Names (FSN).

2.2. Methods for the semantic alignment CIM-11 JLMMS –SCT

1. For the defined subset of ICD-11 linearization on the circulatory system chapter, identify the correspondence between the ICD-11 classes and the concepts of SCT hierarchy "Clinical findings", "Situations", "Events" or "Social context":considering the fully specified name (FSN) of SCT, the short definition of ICD, the logical definition of SCT expressed by compositional grammar [16].
2. For ICD-11 classes which have a full match with SCT content, verify that the logical definition of SCT provides a complete representation of the ICD-11 class: category M Table 1.
3. For ICD-11 classes which have not a full match with SCT concept, develop when possible a pre-coordinated concept (a logical expression formed by two or more concepts of SCT, that are defined by the compositional grammar) that will make correspondence between ICD class and SCT compositional grammar. Verify that the logical definition of SCT compositional grammar provides a complete representation of the ICD- 11 class: category O/A Table 1.
4. If it is not possible to create correspondence through pre-coordination of SCT concepts, try to find a part of the ICD-11 class representation in SCT through post-

coordination (by modifying one or more existents SCT's concepts, with respecting the compositional grammar of SCT). Verify that these logic definitions which respect the compositional grammar of SCT provide a complete representation of the ICD-11 class: category O/E Table 1.

5. If it is not possible to create correspondence even with a part of ICD-11 class, develop a logical expression which respects the compositional grammar of SCT. Verify that these logic definitions which respect the compositional grammar of SCT provide a complete representation of the ICD-11 class: category O/R Table 1.
6. Once all the ICD-11 JLMMS classes have logical representations with the compositional grammar of SCT, write queries on the SNOMED CT expression constraints language reflecting from one hand the difference between the logical definitions of ICD-11 JLMMS classes and the SCT concepts, and from the other hand the ICD-11 JLMMS exclusions and inclusions rules [17].
7. The residual classes (Other, unspecified) are not included in this work, because of their meanings that are not delimited by edges. So, it is impossible actually to find logical representations for them.

As a summary, we used the method developed to align the ICD-11 Foundation Component to SCT [6-8] but to align ICD-11 JLMMS and SCT logical representation.

3. Results

We present the results of the first five steps of our work.

The types of correspondence between the ICD-11 JLMMS classes and the common ontology are shown in Table 2. If we compare this ICD-11 JLMMS results with those presented in [8] for the Foundation component (FC) of ICD-11, we notice that there is 80,3 % of full match for JLMMS against 49,8 %,for FC as well as 96,4 % of global matching /direct or indirect for JLMMS against 85 % for FC.

Table 1. The types of correspondence between the ICD-11 JLMMS classes and logical definitions of SCT's concepts.

Match type and meaning	Action	Ontology	Queries
Full match (M).	Take the SNF representation of SCT.	The Short Normal Form (SNF) which exists.	None: if no ICD-11 exclusion.
No full match, but pre-coordination possible (O/A).	Add a pre-coordinate representation through the SNF representations of SCT.	The new pre-coordinated SNF expression.	Ensuring the pre-coordination and exclusions.
No full match, no pre-coordination possible (O/E).	Create the logical post-coordinate expression from the SNF of SCT and respecting its compositional grammar.	The new post-coordinated SNF expression.	Ensuring the post- coordination and exclusions.
No match but SCT compositional grammar (O/R).	Create the logical expression respecting the compositional grammar of SCT.	The new logical expression which respects the compositional grammar of SCT.	Exclusions.
The residual classes (Other, unspecified) excluded.	None.	None.	

Table 2.The types of correspondence between the ICD-11 JLMMS classes and the common ontology.

Types ofcorrespondence	Numbers %	Common Ontology
Full match (M).	244 (80.3 %)	*Short Normal Form* (SNF) of SCT.
No full match, but pre-coordination possible (O/A).	34 (11.2 %)	The new SCT representation of SNF expression through pre-coordination.
No full match, no pre-coordination possible (O/E).	8 (2.6 %)	The new SCT's representation of SNF expression through post-coordination.
No full match, no pre-coordination or post-coordination (O/R).	7 (2.3 %)	The new logical expression which respects the compositional grammar of SCT.
No match or require revision.	11 (3.6 %)	No common ontology: concepts need more clarification.

Table 3 shows some examples of correspondence between ICD-11 JLMMS and SCT.

4. Conclusion

This work meets some limitations:

On the one hand, a number of SNOMED CT's logical definitions for concepts are not complete. These concepts are called "primitives ". On the other hand, specifications of queries that will take into consideration the exclusions are still a work in progress using the SNOMED CT Expression Constraint Language [17]. The vague definition of categories named "other…" or "…unspecified" have not been taken into account in this study.

Nevertheless, the study covers most of the cases. Most classes in the JLMMS can best be represented with the compositional grammar of SNOMED CT much more effectively than with the Foundation (FC) [8]. It, therefore, seems that the methods initiated by this work in this paper can contribute to improve the interoperability between these two health terminologies, despite their different uses cases structures and details.

Acknowledgments

This work has benefited from the provision of JLMMS by WHO, and the method of ontology developed by the JAG (WHO/IHTSDO).

Table 3. Examples of matching between ICD-11 and the logical definitions of SCT

Types of correspondence	ICD rubric	Common Ontology
M	*Coronary artery ostial stenosis*	64572001 |Disease (disorder)| :{ 116676008 |Associated morphology (attribute)| = 415582006 |Stenosis (morphologic abnormality)|, 363698007 |Finding site (attribute)| = 55537005 |Structure of ostium of coronary artery (body structure)| }
O/A	*Acute myocardial infarction, STEMI, anterior wall*	401303003 | Acute ST segment elevation myocardial infarction | +54329005 | Acute anterior myocardial infarction
O/E	*Aortic aneurysm secondary to congenital heart disease*	67362008 | Aortic aneurysm |: {42752001 | Due to | = 13213009 | Congenital heart disease |}

References

[1] The World Health Organization. The International Classification of Diseases 11th revision JLMMS. http://apps.who.int/classifications/icd11/browse/l-m/en#/ (last accessed 11 Jan 2016).
[2] IHTSDO browser : http://browser.ihtsdotools.org/ (last accessed 11 Jan 2016).
[3] International Health Terminology Standards Development Organisation. SNOMED CT Starter Guide; Version: 2014-12-02, 12.
[4] International Health Terminology Standards Development Organisation. SNOMED CT Technical Implementation Guide; Jan 2015: 52-465, 560.
[5] Jean-Marie Rodrigues, Anand Kumar, Cédric Bousquet, Béatrice Trombert. Using the CEN/ISO standard for categorial structure to harmonise the development of WHO international terminologies, Studies in Health Technology and Informatics 2009; 150:255-259.
[6] Jean-Marie Rodrigues, Stefan Schulz, Alan Rector, Kent Spackman, Tevfik Bedirhan Üstün, Christopher G. Chute, Vincenzo Della Mea, Jane Millar, Kristina Brand Persson. Sharing Ontology between ICD-11 and SNOMED CT will enable seamless re-use and semantic interoperability. Studies in HealthTechnology and Informatics 2013; 192: 343-346.
[7] Jean-Marie Rodrigues, Stefan Schulz, Alan Rector, Kent Spackman, Jane Millar, James Campbell, Bedirhan Üstün, Christopher G. Chute , Harold Solbrig, Vincenzo Della Mea, Kristina Brand Persson. ICD-11 and SNOMED CT Common Ontology: Circulatory System. Studies in health technology and informatics, 2014;205: 1043-1046.
[8] Jean-Marie Rodrigues, David Robinson, Vincenzo Della Mead, James Campbell, Alan Rector, Stefan Schulz, Hazel Brear,Bedirhan Üstün, Kent Spackman, Christopher G. Chute, Jane Millar, Harold Solbrig, Kristina Brand Persson. Semantic Alignment between ICD-11 and SNOMED CT, Proceedings Med info 2015, 790-794.
[9] Barry Smith. Applied Ontology: A New Discipline is Born, PhilosophyToday, 1998; 29(12) : 5-6.
[10] Alan Rector, Angelo Rossi, Mori F. Consorti, P.Zanstra. Practical development of re-usable terminologies: GALEN-IN-USE and the GALEN Organisation, International Journal of Medical Informatics 1998; 48: 71-84.
[11] Michael Ashburner et al. Gene Ontology: tool for the unification of biology. Nature Genetics 2000; 25: 29, 1-5.
[12] Stefan Schulz, Ronald Cornet, Kent Spackman. Consolidating SNOMED CT's ontological commitment, Applied Ontology 2011; 6: 1-11.
[13] OBO foundry principles: http://obofoundry.org/ontology/ma.html (last accessed 18 Jan2016).
[14] W3C OWL Working Group: OWL 2 Web Ontology Language Document Overview 2012, http://www.w3.org/TR/owl2-overview/ (last accessed 17Jan 2016).
[15] The World Health Organization . The International Classification of Diseases 11th revision Foundation Component http://apps.who.int/classifications/icd11/browse/f/en#/ (last accessed 17Jan 2016).
[16] IHTSDO/SNOMEDCT languages. http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_CompositionalGrammarSpecificationAndGuide_Current-en-US_INT_20150522.pdf?ok
[17] IHTSDO/SNOMEDCT languages. http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_ExpressionConstraintLanguageSpecificationAndGuide_Current-en-US_INT_20150820.pdf?ok

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-273

A Diabetes Self-Management Prototype in an AAL-Environment to Detect Remarkable Health States

Denise SCHINDELBOECK[a,b,1], Friedrich PRAUS[b] and Walter GALL[a]
[a]*Medical University of Vienna, Center for Medical Statistics, Informatics and Intelligent Systems*
[b]*University of Applied Science Technikum Wien*

Abstract: Every year life span is increasing and simultaneously the proportion of people with one or more chronic diseases. This paper presents an implementation of a prototype with a decision tree to detect dangerous health conditions for Diabetes Type 1 and Diabetes Type 2. With the information we collect from Personal Health Devices and data from the Active-Assisted-Living environment, we are in the position to customize thresholds and to get individual results. With the help of a modified Glucose-Insulin Model (based on the minimal model of Stolwijk & Hardy) we predicted the future glucose concentration of the patient. We validated our model with an intention-to-treat pilot study including 8 subjects and obtained a significantly better ($p < 2.2^{-16}$) result than the original model.

Keywords. Decision Tree, Diabetes Mellitus, Glucose-Insulin Model, Stolwijk & Hardy model, AAL

1. Introduction

In the coming years the aging population in Europe will become a major challenge. In Austria the proportion of the older section (over 64 years) of the population grows every year by 0.25% (2012: 17%, 2020: 20%, 2060: 29%) [1]. In Europe 86% of deaths are caused by chronic diseases [2]. In center of focus is not only the individual health of the population, but are also many economic motives. If the potential of self-management (in a context of mobile Health) is fully exhausted by 2017, there is the possibility to save about €99 billion in healthcare costs in the EU (with an investment of €6.2 billion). Following the trend of self-managing health conditions, the World Health Organization (WHO) sees patient self-management as an essential component for chronic disease care [4]. Reasons are: early diagnosis and better treatment; patients care for their own health and have a healthier lifestyle; increased prevention; more effective and sustainable healthcare; health care professionals will be in the position to save 30% of their time spent on accessing and analyzing information [3].

The aim of this study was to implement and validate a routine to detect remarkable health states of patient with diabetes mellitus (DM). In the following chapters we will present the decision tree and the message types it use. In addition we describe the

[1] Corresponding author: Denise SCHINDELBOECK, aMedical University of Vienna, Center for Medical Statistics, Informatics and Intelligent System, Vienna, Austria. Email: n0925067@students.meduniwien.ac.at

methods we used to modify the existing original glucose-insulin model and the results of the validation study.

There are two projects with a similar targets that is to empower patients with Active-Assisted-Living (AAL)-technology: EMPOWER[5] and MODULAAR[6].

2. Methods

In our overall study we implemented two additional checking routines (Hypertension and Heart Failure). In this paper we focus only on the glucose concentration routine.

2.1. Checking Routine

The routine generates four different message types *1) Alert; 2) Warning; 3) Information; 4) Activity* news and gets data from different health devices in the AAL environment, like in our case a glucose meter and temperature meter.

The routine uses a knowledge-based method, decision trees. We chose this method, because it is commonly used in medical sector and easy to understand for care staff and patients. The checking routine follows the example of Vukovic [7].

Glucose concentration Routine: We implemented the check-up of the glucose value, delivered by a glucose meter. Every time a physical health device (PHD) sends a new value, the implemented modified glucose-insulin model delivers a prediction of the future glucose values for the rest of the current day. The routine checks if any value in the prediction is above or below the thresholds, if so a message is generated.

2.2. Glucose-Insulin Model selection & modification

To create an alarm system for DM we *needed* to predict the future blood glucose concentration values. There are some well documented and commonly used models, but none of them use data from an AAL-Environment to improve their results. To our knowledge, no other glucose-insulin model has been altered with more than one modification. Based on the information from Knudson[8] we chose the minimal model of Stolwijk & Hardy[9] (Equation 1 and 2) as the basis for our six modifications*, which are:* 1) Oral one time glucose intake [10]; 2) one time i.v. bolus insulin [14]; 3) alcohol intake [11-13]); 4) fever/infections [15-16]; 5) exercise [17-22]; 6) mental stress [23] We *chose* these six, because of their well-documented influence on blood glucose concentration..

$$\frac{dG(t)}{dt} = \frac{Q_G - \lambda * G(t) - \nu * G(t) * I(t) - \mu * [G(t) - A]^+}{VD} \quad (1)$$

Q_G [mg/min] is the constant glucose inflow from the liver, the constant insulin independent glucose absorption λ [ml/min], ν [ml^2 / (mU * min)] is the constant insulin dependent glucose absorption and μ [mg/min] is the constant glucose clearance.

$$\frac{dI(t)}{dt} = \frac{-\alpha * I(t) + \beta * [G(t) - B]^+}{VD} \quad (2)$$

Table 1. Start parameter values

Parameter	Healthy Person	DM Type 1	DM Type 2
Q_G [mg/min]	140	140	140
λ [ml/min]	41.17	41.17	41.17
ν [ml² / (mU*min)]	2316.67	2316.67	463.33
μ [mg / min]	120	120	120
A [mg / ml]	2.5	2.5	2.5
B [mg / ml]	0.51	0.51	0.51
α [(ml*mU) / (mg*min)]	126.67	126.67	126.67
β [(ml*mU) / (mg*min)]	23.83	4.77	23.83

α [ml/min] and β [(mU * ml) / (mh * min)] in Equation 2 are constants for insulin absorption/production. The corresponding thresholds are *A* (glucose clearance) and *B* (insulin production).

In Table 1 we show the start parameter values according to Kumar [24] we used in the implementation of the model.

2.3. Model validation:

To validate our model we set up a study with 8 subjects (7 DM Type 1 and 1 DM Type 2 patients). The inclusion criteria *were*: 1) men and women between 25 and 99 years; 2) DM Type 1 or Type 2; 3) physically and mentally capable of doing this study. The exclusion criteria are: 1) other chronic diseases then DM Type 1 or Type 2; 2) pregnancy. We ask to record every 1) intake of food; 2) intake of alcohol; 3) intake of insulin; 4) change of basal insulin input rate by pump; 5) timeframes of moderate or dramatic stress; 6) change in health conditions; 7) exercise intensity for three days. The days *need not to be* batched. Additionally the subjects *had* to measure their blood glucose concentration at least 6 times a day and *completed* a questionnaire about demographic information and *we checked* the inclusion criteria. The questionnaire included the following questions: 1) Name; 2) PatientID (assigned by the study responsible); 3) birthday; 4) weight; 5) height; 6) sex; 7) chronic diseases; 8) drinking habits; 9) acute diseases; 10) insulin needed?; 11) using of insulin pump and 12) medication. *As we did*n*'t know* the patients *glucose and insulin concentration of* on *the first day at midnight, we set the start value to G(0) = 120 and I(0) = 1 (DM Type 1) / 10 (DM Type 2).*

Our null hypothesis is that the difference of the measured value and the *predicted* value of the modified model is greater or equal than the difference of the measured value and the *predicted* values of the original model of Stolwijk & Hardy. The alternative hypothesis is that *predicted* glucose concentration values have smaller variance to the measured values than the original model of Stolwijk & Hardy. We decide to take the absolute value of the difference between the measured values and the *predicted* values for the statistical analysis.

The study was following the rules of an intention-to-treat study [25]. We choose the one-side, depending t-test for paired samples with $\alpha = 0.05$ to test our hypothesis. The study protocol was approved by the Ethics Committee of the Medical University of Vienna (EK 1684/2016).

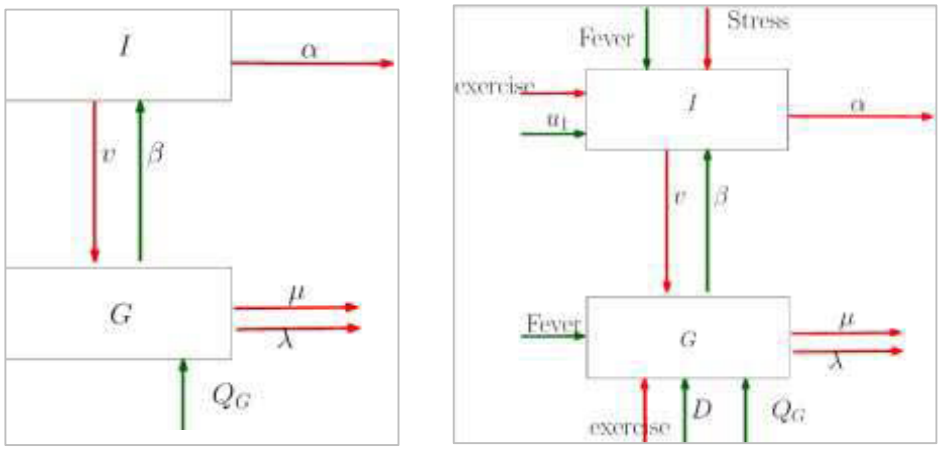


Figure 1. Original Model Stolwijk & Hardy (left) and modified model (right)

3. Results

3.1. Modified Glucose-Insulin Model

Figure 1 (left) shows the original minimal model by Stolwijk & Hardy [9]. In Figure 1 (right) you can see our modifications. The green arrows increase the concentrations in the compartment the red one will decrease them. Equation 3 and 4 show the new differential equations.

$$\frac{dG(t)}{dt} = G_e + \frac{Q_G*\left(1-C_g(t)\right)-\left(\lambda*G(t)-\nu*G(t)*I(t)-\mu*[G(t)-A]^+\right)*H(t)*S(t)+D(t)}{VD} \quad (3)$$

$$\frac{dI(t)}{dt} = -I_e(t) + C_i(t) + \frac{-\alpha*I(t)*H(t)+u_1(t)+\beta*[G(t)-B]^+}{VD} \quad (4)$$

D(t) describes the oral glucose intake of a meal.
$C_i(t)$ and $C_g(t)$ are the effects of alcohol on the glucose and insulin value.
H(t) is the reduced disappearance rate in case of an infection. In the implementation of the prototype we take 0.43.
S(t) is the factor to influence the glucose metabolism in case of mental stress. We decide to have two grades of stress level. Moderate stress S(t) = 1.17 and dramatic stress S(t) = 1.34.
$G_e(t)$ and $I_e(t)$ integrate the effect of exercise; $u_1(t)$ is the intake of insulin.

3.2. Parameter Prediction:

After a new glucose measurement is stored, we used a non-linear regression model to *predict* two parameters in our model: β and ν. We *chose* these two because they are the only two parameters which differ between *a healthy person and one that suffers of DM Type 1 or 2*. In our study we *predicted* β values between 0 and 28.847 and ν values between 0 and 2316.7 (lower bound was 0).

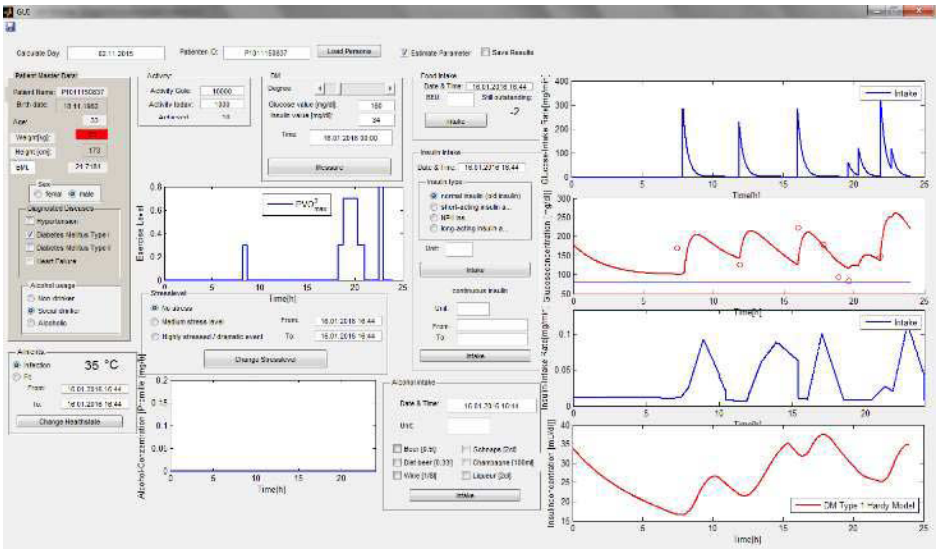


Figure 2. Matlab GUI Prototype

3.3. Prototype Implementation:

The main focus of the implemented prototype in Matlab R2012b was to give a clear overview of the input and output parameters of the modified model. This prototype *helped* to test and verify the model during the implementation process and study evaluation. In Figure 2 the GUI is depicted, *in which all* input parameters can be inserted.

3.4. Pilotstudy: Validation of the Glucose-Insulin Model:

As previously mentioned, we started a study to validate if the *predicted* glucose values of the modified model are closer to the measured values than the *predicted* values of the original model of Stolwijk & Hardy [9]. Our study participants did 221 glucose measurements with a mean value of 149.39 mg/dl ± 55.78 mg/dl. We decided to analyze the absolute distance between the measured values and the *predicted* ones, which lead to a mean range of 59.23 mg/dl ± 52.6 mg/dl for the modified model and 105.38 mg/dl ± 59.79 mg/dl.

As illustrated in Figure 3 the density of the predicted values of the modified model seems to match the density of the measured value much closer than the density of the original model. The variation of the absolute differences of both models is normally distributed, so we use the t-test to check our hypothesis: $p < 2.2\text{-}16$ $[-\infty, -38.0916]$. The modified model has a significantly smaller variance between the measured value and the predicted one.

4. Discussion

We have implemented a checking routine to show the user if he/she is in a dangerous health condition and we also indicated that our modified model is significantly more

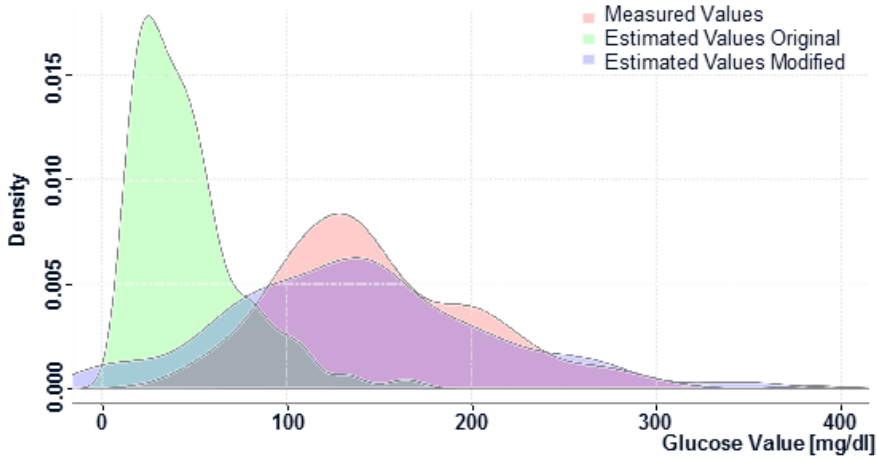


Figure 3. Density of the measured values and the predictions of both models in mg/dl

exact than he original model. However, there are some limitations to be taken into consideration:

4.1. The chosen simple Glucose-Insulin Model

Not all hormones which influence the glucose level are represented in the model.[9] Furthermore we use step functions in the model, though in nature no biological processes are like. A sigmoid function would be closer to the biologic reality. Another critical point is that insulin production is not an ongoing process, it is produced in squalls. [26]

4.2. Modifications of the Glucose-Insulin Model

In this model we don't distinguish between different types of sugar (e.q. fructose, sucrose), but according to Viljoen et al [27] there are *differences* in the absorption rate and elimination rate of alcohol. Additionally we choose two parameters to *predict*, but it would be better to *predict* all of them and use additional factors (e.g. glucose intake efficiency, insulin intake efficiency). Due to the fact that we don't have enough data to do that, the trial to *predict* all known parameters *ended* in some cases with unrealistic results.

There are a few ways to solve differential equations numerically. We choose one of the simplest, because the variation between the simple solution and the costly one (needs 175% more time) was only 61.5 mg/dl during one day (= 0.0427 mg/dl each minute). There is a permitted variation of 20% of any glucose meter [28], so we know that the collected data is impured by some variation as well. We therefore decided to accept the error produced by the numeric calculation of the differential equation.

4.3. Validation of the Glucose-Insulin Model

All values were measured from the patients themselves. Thus we rely on this data to operate the PHDs correctly and to note all relevant figures accurately. At the same time

the number of subjects is too small to make a comprehensive prediction about the significance of the model, but can give a precedent for new studies in this research area. Additionally, we had only one volunteer with DM Type 2. Unfortunately, we don't have the starting glucose and insulin values for the first day at midnight, so we had to set a fixed starting value.

It has to be mentioned that the influence of the i.v. bolus insulin on the glucose concentration is unnatural high in the original model. Most of our subjects use insulin or medication to lower their glucose concentration. This might be the answer, why the mean predicted glucose concentration is that low.

Nevertheless the parameter for the modifications are easy to monitor and the risk state detection is a helpful addition in an working AAL-Environment. Of course we try to eradicate some of this limitations during our project to improve the actual results even more.

Acknowledgment

The authors would like to thank the municipal department MA23 of the City of Vienna for funding the research project ViTAL under Grant Number 15-03, and the project members of the team ViTAL for their effort and emotional support. In addition, we would like to thank Ao.Univ.-Prof. Mag. Dr. Karch and Dr. med Andrea Vincenzo Braga, MBA for the valuable technical input.

References

[1] Statistisches Jahrbuch, 2015. Statistik Austria
[2] WHO. Preventing chronic diseases: a vital investment, 2005
[3] PricewaterhouseCoopers. Socio-economic impact of mHealth, June 2013
[4] European Commission. mHealth, what is it? – Infographic. https://ec.europa.eu/digital-agenda/en/news/mhealth-what-it-infographic, Accessed: 12.03.2016
[5] EMPOWER. Website, http://www.empower-fp7.eu/ Accessed: April 2015
[6] Modulaar project, June 2015. http://www.modulaar.at/ Accessed: 15.06.2015
[7] Vukovic M. Decision support for therapy optimization of heart failure patient based on retrospective home telemonitoring data and expert knowledge. PhD thesis, Technical University of Vienna, February 2014
[8] Knudsen I. Comparative analysis of mathematical model describing glucose-insulin regulation in vivo. Master thesis, University of Stavanger, 2013
[9] Stolwijk J.E. Hardy J.D.. Regulation and Control in Physiology. Medical Physiology 13, 1974
[10] Fisher M.E. A semi-closed loop algorithm for the control of blood-glucose level in diabetics. IEEE Transaction on Information Technology in Biomedicine, 38:57-61, 1991
[11] Wort & Bild Verlag Konradshöhe GmbH & Co. KG. Alkohol und Diabetes: Achtung, Unterzucker. Homepage, http://www.diabetes-ratgeber.net/Ernaehrung/Alkohol-und-diabetes -Achtung-Unterzucker-58356.html Accessed 12.03.2016
[12] Buetzer P. Alkohol. http://www.buetzer.info/unterlagenskripte.html Accessed: 26.01.2016
[13] Batt R.D. Watson PE. And Watson R. Total body water volumes for adult males and females estimated from simple anthropometric measurements. The American Journal of Clinical Nutrition, 33:27-39, January 1980
[14] Pflegewiki. http://www.pflegewiki.de/wiki Accessed: 12.03.2016
[15] Keusch G.T. Brothers M.J. Nechemias C. Smith H. Rayfield E.J. and Ault M.J. Infection and Diabetes: The case for glucose control. American Journal of Medicine, 72(3):439-450, March 1982
[16] Wannemacher R. Dinterman R. Rayfield E. Beisel W. Effect of glucose infusion on the concentration of individual serum free amino acids during sandfly fever in man. American Journal of Clinical Nutrition, 30:573-578, April 1977

[17] Wasserman D.H. Cherrington A.D. Hepatic fuel metabolism during muscular work: role and regulation. American Journal of Physiology, 260(6):811-24, June 1991
[18] Ahlborg G. Wahren J. Felig R. Jorfeldt L. Glucose metabolism during leg exercise in man. Journal of Clinical Investigation, 50(12):2715-25, December 1971
[19] Anirban R. and Parker R.S. Dynamic modeling of exercise effects on plasma glucose and insulin level. Journal of Diabetes Science and Technology, 1(3), May 2007
[20] Astran I. Aerobic work capacity in men and women with special reference to age. Acta Physiol Scand Suppl., 49(169):1-92, 1960
[21] Wahren J. Ahlborg G. Felig P. Hagenfeldt L. Hendler R. Substrate turnover during prolonged exercise in man. The Journal of Clinical Investigation, 53(4):1080-90, April 1974
[22] Wahren J. Felig P. Fuel homeostasis in excise. New England Journal of Medicine, 293(21):1078-84, November 1975
[23] Seematter G. Guenat E. Schneiter P. Cayeux C. Jequier E. Tappy L. Effects of mental stress on insulin-mediated glucose metabolism and energy expenditure in lean and obese women. American Journal of Physiologic Endocrinol Metab, 279:799-805, 2000
[24] Kumar D. Kumar S. Mathematical model for glucose-insulin regulatory systems of diabetes mellitus. Advances in Applied Mathematical Bioscience, 2(1):39-46, 2011
[25] Schulgen G. Schumacher M. Methodik klinischer Studien. Springer Berlin Heidelberg. ISBN: 978-3-540-43306-4. 2002
[26] Cobelli C. Federspil G. Pacini G. Salvan A. Scandellari C. An integrated mathematical model of the dynamics of blood glucose and its hormonal control. Mathematical Bio, 58:27-60, November 1980
[27] Viljoen D.L. Berman P.A.M. Baumgarten I. Effect of oral fructose on ethanol elimination from the bloodstream. South Africa Journal of Science, (99):47-50, 2003
[28] Bundesärztekammer. Richtlinien der Bundesärztekammer zur Qualitätssicherung laboratiumsmedizinischer Untersuchungen. August 2014

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

doi:10.3233/978-1-61499-645-3-281

Electronic Patient Reported Outcomes in Paediatric Oncology – Applying Mobile and Near Field Communication Technology

Katharina DUREGGER[a,b], Dieter HAYN[a, 1], Michael NITZLNADER[a], Martin KROPF[a], Markus FALGENHAUER[a], Ruth LADENSTEIN[c] and Günter SCHREIER[a]
[a]*AIT Austrian Institute of Technology, Austria*
[b]*Graz University of Technology, Austria*
[c]*CCRI, Children's Cancer Research Institute, Austria*

Abstract. Background: Electronic Patient Reported Outcomes (ePRO) gathered using telemonitoring solutions might be a valuable source of information in rare cancer research. Objectives: The objective of this paper was to develop a concept and implement a prototype for introducing ePRO into the existing neuroblastoma research network by applying Near Field Communication and mobile technology. Methods: For physicians, an application was developed for registering patients within the research network and providing patients with an ID card and a PIN for authentication when transmitting telemonitoring data to the Electronic Data Capture system OpenClinica. For patients, a previously developed telemonitoring system was extended by a Simple Object Access Protocol (SOAP) interface for transmitting nine different health parameters and toxicities. Results: The concept was fully implemented on the front-end side. The developed application for physicians was prototypically implemented and the mobile application of the telemonitoring system was successfully connected to OpenClinica. Future work will focus on the implementation of the back-end features.

Keywords. Mobile health, ePRO, paediatric oncology, telemonitoring, eHealth, mHealth, pseudonymisation

1. Introduction

Cancer is a rare disease in children and young adolescents. Less than 1% of all cancer patients in Austria are younger than 14 years [1]. In Europe, the neuroblastoma represents the most often diagnosed embryonic cancer [2]. Neuroblastoma treatment often involves long hospitalizations caused by all or a selection of surgery, chemotherapy, radiotherapy and immunotherapy. As in other rare diseases, neuroblastoma patients are registered in medical research networks and are frequently treated within clinical studies. Aim of the European Network for Cancer Research in Children and Adolescents (ENCCA) project is to connect different research networks in paediatric oncology in order to standardize and increase the quality of treatment of cancer patients [3]. Therefore, patient data from different contexts should be connected to find answers to arising research questions. Horizontal integration of patient data means that data from different studies are connected. In case of vertical integration, study data is connected with e.g. biosample data of biobanks [4].

[1] Corresponding Author: Dieter Hayn, AIT Austrian Institute of Technology, Reininghausstr. 13, 8020 Graz, Austria. Email: dieter.hayn@ait.ac.at.

When connecting patient data from different sources, personal data must at least be pseudonymized [5]. Therefore, within ENCCA a dedicated concept is applied as described in [6]. A unique and inaccessible ENCCA Unified Patient Identifier (EUPID) is assigned to each patient newly registered within the research network. For every context, in which patient data needs to be collected, a context-specific pseudonym (PSN) is created for the patient, which is further assigned to the patient's EUPID. This approach allows merging of patient data but preserving the possibility of re-identification in case new research findings would lead to an improved treatment. For creation of PSNs, the patients' identity data (IDAT), including first name, last name or date of birth, are used.

Patient Reported Outcomes (PRO) can be defined as "*any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else*" [7]. PRO might include subjective health parameters or parameters only assessable in the patients' domestic environment. By considering the patient's perception, the severity of adverse events is staged with higher accuracy and a more comprehensive dataset can be reached [8, 9]. Furthermore, gathering PRO for medical research purposes increases patient empowerment [10].

By applying modern communication methods and mobile technology for capturing electronic Patient Reported Outcome (ePRO) barriers caused by paper-based capturing of PRO might be overcome [8]. By the use of ePRO in combination with telemonitoring solutions during uncritical treatment phases, patients could potentially be discharged and treated at home at least for several days or weeks, which could significantly improve patients' and parents' quality of life.

Several research groups worldwide are currently investigating telemedical ePRO solutions for paediatric oncology [11-13]. Unfortunately, these approaches are not yet linked to research infrastructures, and therefore patient registration, data inspection at hospitals etc. show limited levels of data security and/or usability and they do not support secondary research.

The objective of the present paper was to develop a concept and to implement a prototype for integrating ePRO into an existing neuroblastoma research network. Therefor, Near Field Communication (NFC) technology in combination with mobile applications should be applied. The following two major tasks evolved:

- Providing patients and parents with a suitable telemonitoring system for capturing ePRO and transmission to the Electronic Data Capture (EDC) system currently used within the SIOPEN-R-NET, i.e. the European neuroblastoma research network.
- Development and implementation of a process for registering patients within neuroblastoma studies (i.e. creating PSNs) and providing patients with user credentials for authentication when transmitting telemonitoring data.

2. Methods

Fig. 1 shows the general approach for introducing ePRO into the existing neuroblastoma research network. Physicians were provided with a dedicated mobile application for registering patients within certain contexts by creating PSNs. A pre-manufactured patient ID card could be linked to a newly generated PSN and was provided to the patient together with a PIN. Patients used the previously developed telemonitoring system described in chapter 2.4 for capturing vital parameters and transmission to the back-end EDC system OpenClinica (OpenClinica, LLC, OpenClinica Community, version 3.5, Waltham, Massachusetts) described in chapter 2.5.

Figure 1. Approach for introducing ePRO into the existing neuroblastoma research network.

2.1. Near Field Communication (NFC) Technology

Near Field Communication (NFC) is a state-of-the-art communication technology for contactless transmission of small amounts of data between mobile devices or reading the content of Radio Frequency Identification (RFID) tags. NFC operates at 13,56 MHz and at a distance of approximately 4 cm. [14]

2.2. Quick Response (QR) Codes

In contrast to barcodes, Quick Response (QR) Codes can hold larger amounts of information, depending on the type. Among other fields of application, today QR Codes are commonly used for merchandising, in production management or providing links to websites. Due to the built-in error correction pattern, QR Codes are more robust than barcodes. [15]

2.3. Cross-platform Application Development

A mobile application for registering patients was developed for different mobile operating systems. Therefore, the framework Apache Cordova (Cordova) [16] for cross-platform application development was used. Cordova applications are programmed using Hypertext Markup Language (HTML5), Cascading Style Sheets (CSS) and JavaScript. By avoiding native programming languages and using Cordova, the development effort is decreased and further operating systems can easily be added.

Cordova is supported by Microsoft Visual Studio (Visual Studio). Visual Studio also provides tools for remotely debugging projects on iOS devices attached to an Apple MacBook within the same network.

2.4. Telemonitoring System for Neuroblastoma Patients

In a previous work [17] a telemonitoring system for paediatric oncology was developed. The system consisted of the following components:

- Mobile phone with an Android application called MoKi for capturing measurements from NFC enabled point-of-care devices
- Patient ID card for starting MoKi
- NFC-enabled blood pressure meter and body weight scales
- Dedicated smart poster for capturing subjective health parameters

Figure 2. Telemonitoring system for neuroblastoma patients [18].

The system could be used for capturing nine different vital parameters and toxicities often occurring during cancer treatment: wellbeing, blood pressure and heart rate, body weight, body temperature, C-reactive protein, white blood cell count, pain level, nausea level and skin alterations. Fig. 2 shows the telemonitoring system developed with respect to the requirements of neuroblastoma patients.

2.5. OpenClinica

OpenClinica is currently used for data capture in some neuroblastoma studies. The Community edition of the EDC system is free to use and offers web services for scheduling events and data entry via Simple Object Access Protocol (SOAP) requests. Patient data are stored in a PostgreSQL database.

For developmental purposes an OpenClinica instance as well as the OpenClinica web service package were set up on an Apache Tomcat server.

3. Results

Fig. 3 depicts the developed overall concept. Within this work two tasks were completed:

- A prototype of the mobile application EUPID Mobile for patient registration and providing user credentials was developed for Android and iOS devices.
- The patient ID card of the telemonitoring system was modified in order to hold a unique serial number. The monitoring application MoKi was connected to OpenClinica by establishing a SOAP interface.

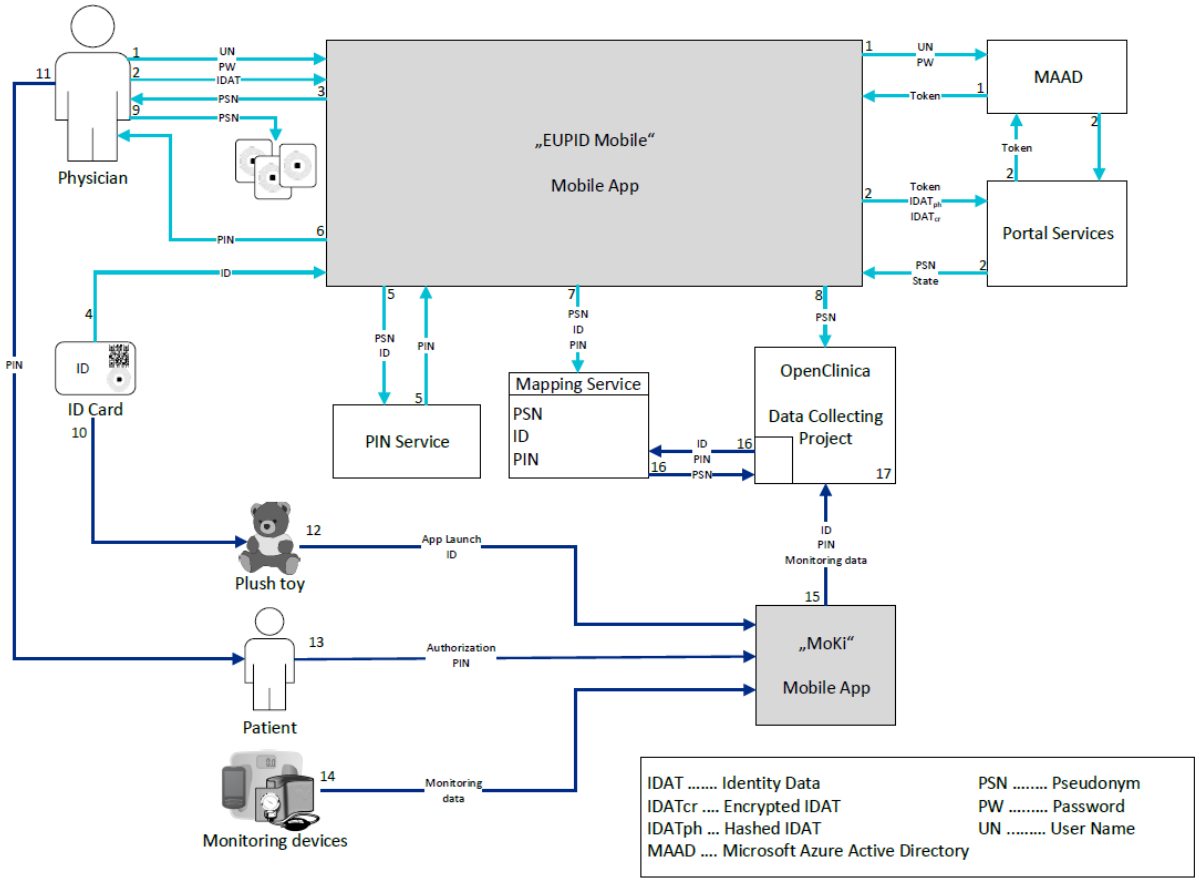


Figure 3. Concept for introducing electronic Patient Reported Outcomes (ePRO) into the existing neuroblastoma research network with the two mobile applications EUPID Mobile and MoKi as central elements.

3.1. Implementation Concept

As shown in Fig. 3, when starting EUPID Mobile users with the role 'physician' logged on to the EUPID system using their user name and password. After selecting the context, in which a patient should be registered, and entering the patient's IDAT, a PSN was returned by the EUPID service. Whether a patient had already been registered was determined by calculating phonetic hashes of the IDAT as described in [6]. In case the patient had not yet been registered within the selected context, a new PSN was created. If a similar patient had already been registered, physicians could force the registration of the patient and a PSN was returned. For already registered patients the existing PSN was returned by the EUPID service.

Next, a patient ID card could be linked to the PSN. Therefore, users read the patient card ID via NFC or QR Code (when using iOS devices) and obtained a corresponding PIN via a EUPID PIN service. Further, patients were enrolled in OpenClinica with their PSN. PSN, patient card ID and PIN were also entered into a mapping table. On Android devices physicians were also able to write the PSN to any number of NFC tags for labelling patient-related documents or biosamples.

After obtaining patient ID card and PIN for authentication, users with the role 'patient' were able to transmit telemonitoring data to the back-end OpenClinica.

3.2. User Interface of EUPID Mobile

Fig. 4 shows the user interface of EUPID Mobile on Android devices. The main screen showed the options for registering a patient and directly linking a patient ID card to an already existing PSN. Since, in contrast to iOS devices, Android devices supported the use of NFC in custom applications, it was also possible to write PSNs to NFC tags. The patient ID card could in case of Android devices either be read via NFC or QR Code.

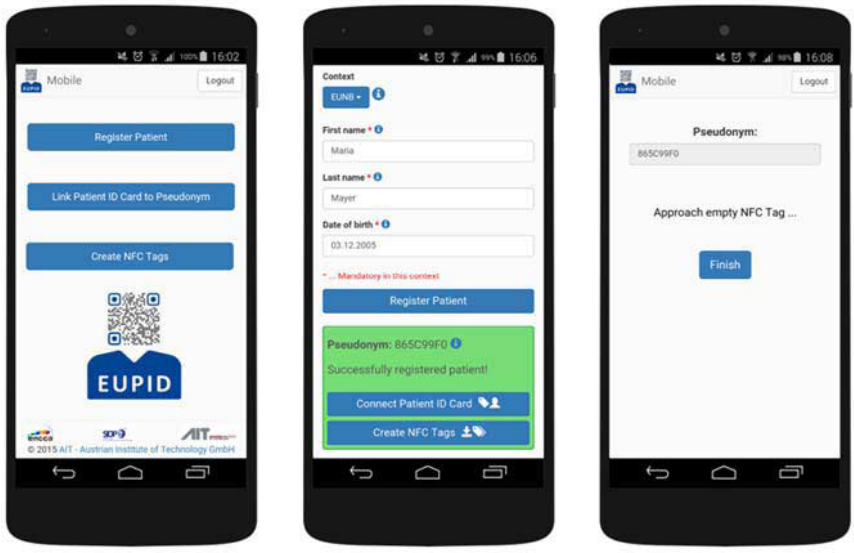


Figure 4. EUPID Mobile application for Android devices [18]. Left: Main screen, Middle: Successful patient registration, Right: Writing a pseudonym to NFC tags.

3.3. User Interface of MoKi

MoKi could be started by patients via NFC using the ID card generated by the EUPID Mobile application, which might be put into some kind of plush toy in order to create some game-like environment for children. When entering observations, patients were asked to touch their patient ID card and to enter the PIN. After performing a measurement, the observation was saved and transmitted to OpenClinica.

OpenClinica used the mapping table to establish a link between the patient or patient ID card, respectively, currently sending data and the corresponding PSN, which was previously used for enrolling this patient in a certain OpenClinica study.

4. Discussion

ePRO might be a valuable additional source of information in rare cancers. In such scenarios, usability and security are often contradictory requirements. With the EUPID Mobile application, physicians are able to easily register patients within the neuroblastoma research network by using their own iOS- or Android-powered smartphone, while providing a high level of security. Linking patient ID cards is possible simply by touching the ID card with the mobile phone or scanning the QR Code and without the need for manually typing in serial numbers.

In the present work the interfaces between the EUPID Mobile application and the services for obtaining the PIN, enrolling the patient in OpenClinica and making an entry into the mapping table were implemented as mockups. Therefore, future work will have to focus on the design of the PIN service, on establishing workflows for lost ID cards or forgotten PINs and on the implementation of the back-end services.

The existing telemonitoring system for paediatric cancer patients was extended by a SOAP interface and allowed patient authentication at the back-end via ID card and PIN. The ID card might also play an important role in (long-term) patient follow-up. One aim of the ENCCA project is to provide patients with a ‘survivorship passport’ containing the cancer treatment history, which might be important in future medical interventions.

Further, the NFC tags generated with EUPID Mobile might be used for labelling blood or tumour samples or similar items when sending them to laboratories, other research facilities or biobanks.

Introducing ePRO as an additional source of information in cancer research not only increases patient empowerment but could also improve the quality of research data and, therefore, help to improve the treatment of children suffering from neuroblastoma or other rare diseases. Our solution might help to integrate ePRO in existing research infrastructures supporting patients, parents, physicians and researchers in secondary use scenarios.

Acknowledgements

The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7 2007-2013) under the project ENCCA grant agreement no. 261743.

References

[1] STATISTIK AUSTRIA, Childhood and adolescent cancer, http://www.statistik.at, last access: 25.01.2016.
[2] G. Gatta et al., Embryonal Cancers in Europe, *European Journal of Cancer* **48**(10)(2012), 1425-1433.
[3] European Network for Cancer Research in Children and Adolescents, Mission, http://www.encca.eu, last access: 25.01.2016.
[4] N. Hochedlinger, Standardisierter Datenaustausch in medizinischen Forschungsnetzwerken mit Anwendung in der pädiatrischen Onkologie, Master's thesis, 2015, TU Graz.
[5] M. Nitzlnader, Konzept eines IHE-basierten Medizinischen Forschungsnetzwerkes in der Cloud, Master's thesis, 2014, UMIT – Private University of Health Sciences, Medical Informatics and Technology.
[6] M. Nitzlnader and G. Schreier, Patient Identity Management for Secondary Use of Biomedical Research Data in a Distributed Computing Environment in: *eHealth 2014 - Health Informatics Meets eHealth,* volume 198 of *Studies in Health Technology and Informatics*, 2014, 211-218.
[7] Food and Drug Administration, Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, http://www.fda.gov, 2009.
[8] A.V. Bennett et al., Electronic Patient-Reported Outcome Systems in Oncology Clinical Practice, *CA: A Cancer Journal for Clinicians*, **62**(5)(2012), 336-347.
[9] S.J. Coons et al., Recommendations on Evidence Needed to Support Measurement Equivalence between Electronic and Paper-Based Patient-Reported Outcome (PRO) Measures: ISPOR ePRO Good Research Practices Task Force Report, *Value in Health*, **12**(4)(2009), 419-429.
[10] P.R. Deshpande et al., Patient-reported outcomes: A new era in clinical research, *Perspectives in Clinical Research,* **2**(4)(2011), 137-144.
[11] N. Bradford et al., Home Telemedicine for Paediatric Palliative Care in: *Global Telehealth*, volume 161 of *Studies in Health Technology and Informatics*, 2010, 10-19.
[12] M.E. Bensink, Using videotelephony to support paediatric oncology-related palliative care in the home: from abandoned RCT to acceptability study, *Palliative Medicine*, **23**(3) (2009), 228-237.
[13] N. Bradford et al., A pilot study of the effectiveness of home teleconsultations in paediatric palliative care, *Journal of Telemedicine and Telecare*, **18**(8) (2012), 438-442.
[14] NearFieldCommunication.org, http://nearfieldcommunication.org, last access: 26.01.2016.
[15] DENSO WAVE INCORPORATED, http://www.qrcode.com, last access: 26.01.2016.
[16] Apache Software Foundation, https://cordova.apache.org, last access: 26.01.2016.
[17] K. Duregger, An mHealth System for Toxicity Monitoring of Paediatric Oncological Patients using Near Field Communication Technology, *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6848 – 6851, 2015, Milano, Italy.
[18] K. Duregger, Mobile and Near Field Communication (NFC) Technology for Electronic Patient-Reported Outcomes in Paediatric Oncology, Master's thesis, 2016, TU Graz.

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

Subject Index

Health Informatics Meets eHealth
G. Schreier et al. (Eds.)

Author Index