EJBI 2014 ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 10 (2014), Issue 1

Editor

Jana Zvárová

An Official Journal of the European Federation for Medical Informatics

www.ejbi.eu

Aims and Scope

The **European Journal for Biomedical Informatics** reacts on the great European need to share the information in the multilingual and multicultural European area. The journal publishes peer-reviewed papers in English and other European languages simultaneously. This opens new possibilities for faster transfer of scientific-research pieces of knowledge to large international community of biomedical researchers, physicians, other health personnel and citizens. From time to time, the journal publishes articles on particular focus themes as part of a journal's issue.

The generally accepted translations of the English version of the abstract and keywords or full paper are to the European languages, which can be found at `http://www.ejbi.org/en/about/`.

Editors and Management

Editorial Board: National Members

Editorial Board: Representatives of Cooperating Journals

Publisher

EuroMISE s.r.o.
Paprsková 330/15
CZ-14000 Praha 4
Czech Republic
EU VAT ID: CZ25666011

Office

EuroMISE s.r.o.
Paprsková 330/15
CZ-14000 Praha 4
Czech Republic

Contact

Karel Zvára
zvara@euromise.com

Instructions to Authors

Please access http://www.ejbi.org/en/instructions/

EJBI Online

The online version of the full volume is available at http://www.ejbi.org/en/ejbi/.

Instructions to Authors

General Remarks

This journal follows the guidelines of the International Committee of Medical Journal Editors (http://www.icmje.org/index.html) and the Committee on Publication Ethics (http://www.publicationethics.org).

Authors should especially be aware of the following relevant issues in these guidelines:

Authorship

All authors should have made

(1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data;

(2) drafting the article or revising it critically for important intellectual content; and

(3) final approval of the version to be published.

Conflicts of interest

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their actions.

Protection of human subjects and animals in research

Authors who submit a manuscript on research involving human subjects should indicate in the manuscript whether the procedures followed were in compliance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects (http://www.wma.net/en/30publications/10policies/b3/).

European Journal for Biomedical Informatics does not publish material that has already appeared elsewhere. Submitted manuscripts should not be submitted in parallel to any other journal.

Manuscript preparation

Authors are kindly requested to carefully follow all instructions on how to write a paper. In cases where the instructions are not followed, the paper will be returned immediately with a request for changes, and the editorial review process will only start when the paper has been resubmitted in the correct style.

Authors are responsible for obtaining permission to reproduce any copyrighted material and this permission should be acknowledged in the paper.

Authors should not use the names of patients. Patients should not be recognizable from photographs unless their written permission has first been obtained. This permission should be acknowledged in the paper.

In general the manuscript text (excluding summary, references, figures, and tables) should not exceed 5 000 words.

Kindly send the final and checked source and PDF files of your paper to `manuscripts@ejbi.org`. You should make sure that the LaTeX and the PDF files are identical and correct and that only one version of your paper is sent. Please note that we do not need the printed paper.

Where appropriate, the paper should be organised into the following sections: *Abstract, Introduction, Objectives, Methods, Results, Discussion, Conclusions, Acknowledgments and References*. Apart from the main headings, subheadings should be used and may be numbered.

Authors are strongly encouraged to use LaTeX 2_ε for the preparation of manuscript. The LaTeX template `ejbi_template.tex` can be downloaded from www.ejbi.org/en/instructions/.

When you are not able to use LaTeX, please use MS Word or OO Writer and send us the unformatted text. Kindly follow just instructions about preparing figures, tables and references. We are going to convert your text into LaTeX instead of you.

If you use LaTeX together with our template file, `ejbi_template.tex`, your text is typeset automatically. Please do *not* change the preset fonts. Do not use your own macros, or styles.

Please use the commands `\label` and `\ref` for cross-references and the commands `\bibitem` and `\cite` for references to the bibliography, to enable us to create hyperlinks at these places.

Title page

The first page of the article should contain: title of the paper (also the shorter version for running heads), initials and last name of each author, to be followed with their institutional affiliations, the name, address, e-mail address and telephone of the corresponding author.

Abstract and Keywords

The abstract should summarize the contents of the paper and should not exceed 250 words. Authors are requested to write a structured summary, adhering to the following headings: Background (optional), Objectives, Methods, Results, Conclusions.

At the end of the Abstract, the contents of the paper should be specified by, at most, five keywords. We recommend using MeSH keywords.

Headings

Headings should be capitalized (i.e. nouns, verbs, and all other words except articles, prepositions, and conjunctions should be set with an initial capital) and should be aligned to the left. Words joined by a hyphen are subject

to a special rule. If the first word can stand alone, the second word should be capitalized.

Figures and Tables

Attach figures and tables as separate files. Do not integrate them into the text. Do not save your table as an image file or insert a table into your manuscript text document as an image. Figures and tables should be referenced in the manuscript by their numbers.

Annotations belong in a (self-)explanatory legend, do not use headings in the figure, explain abbreviations in the legend. Label all axes. Use a uniform type size (we recommend Arial 10 point), and avoid borders around tables and figures.

Submit graphics as a sharp printout as well as a file. The printout and the file must be identical. Submit the image file with clear labelling (e.g. Fig_1 instead of joint_ap).

Image resolution is the number of dots per width of 1 inch, the "dots per inch" (dpi). Printing images require a resolution of 800 dpi for graphics and 300 dpi for photographics.

Vector graphics have no resolution problems. Some programs produce images not with a limited number of dots but as a vector graphic. Vectorisation eliminates the problem of resolution. However, if halftone images ("photos") are copied into such a program, these images retain their low resolution.

If screenshots are necessary, please make sure that you are happy with the print quality before you send the files.

In the printed volumes, illustrations are generally black and white (halftones), and only in exceptional cases, and if the author is prepared to cover the extra cost for colour reproduction, are coloured pictures accepted. Coloured pictures are welcome in the electronic version free of charge. If you send coloured figures that are to be printed in black and white, please make sure that they really are legible in black and white. Some colours as well as the contrast of converted colours show up very poorly when printed in black and white.

Formulas

Displayed equations or formulas are centred and set on a separate line (with an extra line or halfline space above and below). Displayed expressions should be numbered for reference. The numbers should be consecutive within each section or within the contribution, with numbers enclosed in parentheses and set on the right margin.

Footnotes

The superscript numeral used to refer to a footnote appears in the text either directly after the word to be discussed or – in relation to a phrase or a sentence – following the punctuation sign (comma, semicolon, or period). Footnotes should appear at the bottom of the normal text area, with a line of about 2 cm set immediately above them.[1]

Program Code

Program listings or program commands in the text are normally set in a typewriter font, e.g. CMTT10 or Courier.

Acknowledgements

Scientific advice, technical assistance, and credit for financial support and materials may be grouped in a section headed Acknowledgements that will appear at the end of the text (immediately after the Conclusions section). The heading should be treated as a subsubsection heading and should not be assigned a number.

In case that a financial support of the paper development (e.g. sponsors, projects) is acknowledged, the fee of 50 EUR will be charged by Publisher. The accepted peer-reviewed papers with an acknowledgement of a financial support, where the fee was not paid, will be published free of charge, but the financial acknowledgement will be withdrawn.

References

The list of references is headed "References" and is not assigned a number. The list should be set in small print and placed at the end of your contribution, in front of the appendix, if one exists. Please do not insert a pagebreak before the list of references if the page is not completely filled. For citations in the text please use square brackets. In the text number the references consecutively in the order in which they first appear. Use the style, which is based on the formats used by the US National Library of Medicine in MEDLINE (sometimes called the "Vancouver style"). For details see the guidelines from the International Committee of Medical Journal Editors (http://www.nlm.nih.gov/bsd/uniform_requirements.html).

Examples:

[1] Schwartze J, HaarbrandtB, Fortmeier D, Haux R, Seidel C. Authentification systems for securing clinical documentation workflows. Methods Inf Med 2014; 53 (1):3-13.

[2] Kalina J, Seidl L, Zvára K, Grünfeldová H, Slovák D, Zvárová J. Selecting relevant information for medical decision support with application in cardiology. European Journal for Biomedical Informatics 2013; 9 (1): 2-6. Available from: http://www.ejbi.org/img/ejbi/2013/1/Kalina_en.pdf

[3] Hasman A, Blobel B, Zvárová J, editors. Data and Knowledge for Medical Decision Support. Amsterdam: IOS Press; 2013.

[1] The footnote numeral is set flush left and the text follows with the usual word spacing.

Multilingual Issue

The authors are asked to translate English version of Abstract and Keywords to at least one European languege. The translated versions of Abstract and Keywords should be send to `manuscripts@ejbi.org`

Checking the PDF File

Kindly assure that the Contact Volume Editor is given the name and email address of the contact author for your paper. The contact author is asked to check through the final PDF files to make sure that no errors have crept in during the transfer or preparation of the files. Only errors introduced during the preparation of the files will be corrected.

If we do not receive a reply from a particular contact author, within the timeframe given, then it is presumed that the author has found no errors in the paper.

Copyright Transfer Agreement

The copyright form may be downloaded from http://www.ejbi.org/en/downloads/. Please send your signed copyright form to the Contact Volume Editor, either as a scanned pdf or by fax or by courier. One author may sign on behalf of all the other authors of a particular paper. Digital signatures are acceptable.

EuroMISE Copyright Permission Policy

Written permission is required to reproduce material from EuroMISE s.r.o. publications in other publications, electronic products, or other media. To obtain a copyright permission please contact Jana Zvárová: zvarova@ejbi.org, fax: +420 241471337. You may fax or e-mail your request along with the full citation of the journal in which the paper appears in with volume number and page number(s) as well as what you are requesting to use the material for. Use of copyrighted material always requires proper citation.

Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND): for non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as credit the author(s) and provided they do not alter or modify the article.

Contents

Holistic View on Electronic Healthcare

Jana Zvárová[1]

[1] Editor-in-Chief, European Journal for Biomedical Informatics, Prague, The Czech Republic

Correspondence to:

Prof. Jana Zvárová, Ph.D., DSc., Editor-in- Chief,
European Journal for Biomedical Informatics,
Address: Paprsková 15, 140 00 Prague, The Czech Republic
E–mail: zvarova@ejbi.org

EJBI 2014; 10(1):1
published: February 28, 2014

Healthcare is an information intensive sector, generating huge volumes of data from hospitals, primary care, clinics, surgeries and laboratories. Information can be extracted from large databases, using nowadays information technologies, and to bring a new knowledge. However, in healthcare sector is working highly educated and knowledgeable personnel, providing the knowledge to processes of medical treatment and healthcare management. The inability to share information across systems and between care organizations is just one of the major obstacles towards quality, efficiency, security and cost-effectiveness of healthcare.

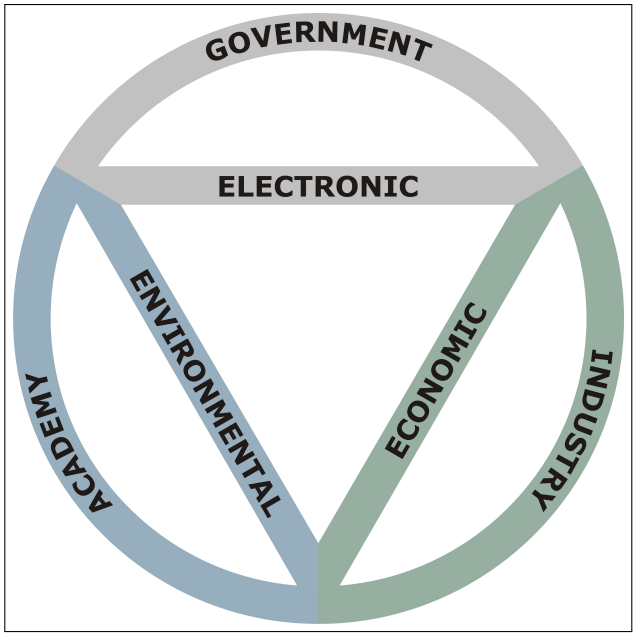

Figure 1: Scheme of holistic electronic healthcare.

There are many reasons for this state, including underinvestment in information technologies, lack of political will, fragmented markets with inadequate development of new systems, lack of standards, complexity of medical data, data entry problems, security and confidentiality. The gap between the demand for healthcare from an increasingly well-informed citizens and the ability of the government and healthcare organizations to meet this demand is widening all the time.

All these problems in electronic healthcare can be successfully overcome by a holistic approach. Good cooperation among Government, Industry and Academia is a basic prerequisite for building modern healthcare supported by three basic components electronic, economic and environmental [1]. The scheme of holistic electronic healthcare is displayed on Figure 1.

In the year 2014 European Journal for Biomedical Informatics (EJBI) welcomes original articles dealing with topics influencing electronic healthcare. Authors are not paying an article processing fee for the immediate release of peer-reviewed articles, but a small financial support is required in case that the support of projects or sponsors is acknowledged (see Instruction to authors). Due to the focus of the journal to semantic interoperability issues we ask authors for translation of structured abstract of their articles to at least one European language. EJBI provides immediate open access to peer-reviewed papers, which will be published in the running first issue of EJBI during each calendar year. The other issues of EJBI are special issues related to different biomedical informatics topics. Topics for special issues can be proposed to editor-in- chief of EJBI using the form Proposal of EJBI special issue for further processing. Topic for special issue is specified by an open call or by a special event. We invite you to propose special topics that would help to accelerate needed changes in electronic healthcare by easy transfer of a new information and knowledge for health care delivery.

References

[1] Zvárová J., Zvára K: e3Health: Three Main Features of Modern Healthcare. In: Moumtzoglou A, Kastania A (eds). E-Health Systems Quality and Reliability: Models and Standards, IGI Global Press, New York, 2010; 18–27

Optimization of Malaria's Treatment: An Approach of Medical Decision Analysis

Eustache Ayumba Muteba[1]

[1] Independent Researcher and Correspondent of IMIA in DR Congo

Abstract

Objective: Nowadays, it is recognized in most modern hospital and public health systems an increasing concern to measure the quality of care. The quality of care can be focused on the characteristics of hospital production and the indicators of performance. The indicators of performance can permit, generally, to decrease complication rate, morbidity, mortality and costs of care. Therefore, one of the ways to optimize the quality of care is to use medical decision support system.
Methods: The optimization of malaria's treatment is based on an automatic extraction of a geographic information system database that can store and provide relevant information on malaria's patient case of different regions. The method proposed is consisted of height main steps namely: specification of the case, indications or problems, actions or treatments strategies, estimative outcomes (benefit and risk), performance measure, decision, result and optimization.
Results: One of the most important outcomes of this work is an understanding of the requirements on a medical decision analysis formalism and system. The case study presented for the simulation constitutes a theoretical component that consolidates the validation of the formalism before the implementation.
Conclusion: The work embodied in this paper formed the second part of a research project called 'OMaT'. OMaT is an online system that aims to assist physician at medical consultation in order to optimize the quality of care of the patients with malaria disease.

Keywords

Optimization, Malaria's Treatment, Medical Decision Analysis

Correspondence to:

Eustache Muteba A.
Independent Researcher and Correspondent of IMIA in DR Congo,
Address: P. O. Box 14769, Kinshasa, DR Congo
E–mail: muteusta@maesoft1.net

EJBI 2014; 10(1):2–5
received: July 1, 2013
accepted: February 9, 2014
published: February 28, 2014

1 Introduction

Nowadays, it is recognized in most modern hospitals and public health systems, an increasing concern to measure the quality of care. The quality of care [1], can be focused on the characteristics of hospital production and the indicators of performance.

These indictors of performance will permit to decrease complication rate, morbidity, mortality and cost of care. One of the ways to optimize the quality of care is to use medical decision support systems [2, 3, 4, 5] based on eHealth and mHealth to serve the unserved [6].

Malaria was one of the most challenging infectious diseases caused by the parasite called plasmodium and localized mainly in areas of Asia, Africa, and Central and South America. The overall disease burden is devastating youth, women and health systems.

The technical capability to perform a correct and a timely diagnosis and an appropriate treatment of malaria infection in an ill patient is of critical importance since symptoms of complicated malaria may suddenly develop, possibly leading to death despite intensive care efforts. To decide what tests to order, what diagnoses to consider, and what treatments to administer, physicians draw on a large, rapidly growing body of knowledge. [7]

The concern of the present paper, is the development of a method of medical decision analysis specifically an optimization of malaria's treatment.

2 Methods

The essence of the adopted method is in one hand based on a differential diagnostic model since the signs and symptoms of malaria can be confused with others diseases. And in the other hand, it is based on an automatic information extraction of geographic information system of stored malaria's patient case for an optimization of malaria's treatment.

2.1 Medical Decision Analysis

In medicine, comes often situations of uncertainly on knowledge, facts and sometimes on the used language. Kenner et al [8] reveal that "for some diseases, definite and unique causes like certain infections may be found. For other diseases, multifactorial causes have to be assumed, mostly because of lack of knowledge." The rational approach to decision making for problems where uncertainty figures as a prominent element is a decision analysis.

Major information on the medical decision analysis model can be found in [9, 10, 11, 12]. The medical decision analysis method proposed is prescriptive, based on a multicriteria methodology and constructive induction method. It consists of height main steps: specification of the case, indications or problems, actions or treatments strategies, estimative outcomes (benefit and risk), performance measure, decision, result and optimization.

Step 1: Specification of the case. The specification of the case describes basics clinical information relative to a particular patient in consultation such as sex, age, weight, antecedents, allergies, ...

Step 2: Indications. The diagnostic indicates the problems found on the patient. Indication is a set of information related to problems concerning a particular patient.

Step 3: Actions. The actions are different possible treatments referring to the given indications.

Step 4: Estimative Outcomes. The estimative outcomes depend on the information related to similar patients' cases provided by clinicians and stored in the Geographic Information System. This information can be extracted at this step. Furthermore, it can be automatically updated at the optimization step and then increased the effective of population concerned by malaria. There are two kinds of estimative outcomes: the outcome with benefit and the outcome with risk.

Outcome with Benefit. The outcome with benefit expresses the degree to return to normal health. It is a value compute as frequency of reveal result at the optimization step.

Outcome with Risk. The outcome with risk expresses the complication or the death. It is a value compute as frequency of reveal result at the optimization step.

Step 5: Performance Measure. The performance measure is a benefit-risk ratio referring to the action chosen by the clinician.

Step 6: Decision. The benefit-risk ratio can permit the clinician to make a decision. Practically, if the ratio is > 1 then the action can give benefit otherwise, if the ratio < 1 then the action have a risk.

Step 7: Result. The result is the really consequence of the decision chosen at the light of the performance measure. The patient can be in the following situation: a. Benefit, b. Risk. The clinician can vote for one of the presented situation and the system automatically will be updated. This information may be considered sufficient and trusted.

Step 8: Optimization. The optimization can allow the clinician to analyze the results and if needed to readjust the actions.

2.2 Information Extraction

The pre-processing of optimization can be doing as following:

1. Each patient case at time t is represented by an attribute-value vector:

   P = [userId: V1, patientId: V2, sex: V3, averageAge: V4, averageWeight: V5, country: V6, status: V7, conditions/Diseases: V8, pastMedications: V9, allergies: V10, symptoms: V11, testResult: V12, treatments: V13, ...]

2. A patient case is a n-dimensional vector where each dimension corresponds to a distinct attribute and n is the total number of possible attributes.

3. Identification of different patient communities in a population of patient cases.

4. For that, two issues are suggested:

   - Determine meaningful subsets (communities/patients with similar case).
   - Determine meaningful concepts for each subset (stereotypes).

5. The communities' stereotypes are built up by trying to identify patterns.

6. Incrementally generates clusters (patient with common characteristics) representing patient communities as following:

   - Creating a new cluster.
   - Placing a new patient case into an existing cluster.

- Combining two clusters into a new one.
- Dividing an existing cluster in two or several new cluster.
- Extracting representative information.

2.3 Method Validation

A given specification case S and an indication $\{I_n, n = 1, \ldots, N\}$ area root of solution; where n is an integer between 1 and N.

An action $\{A_m, m = 1, \ldots, M\}$ is a set of treatments strategies possible and admissible knew as applicable, obtain by means of selected multiple criteria reflecting the specification of the indication:

$$max\langle S.I_n = A_m\rangle \quad (1)$$

The action A_m implies estimative outcomes; Let the outcome with benefit (OB) be a time series and recursive function defined as following:

$$OB_t = \left(A_m, \left(\sum_{t=1}^{T} OB\right)\right) \quad (2)$$

where t is an integer between 1 and T. And, let the outcomes with risk (OR) be a time series and recursive function defined as following:

$$OR_t = \left(A_m, \left(\sum_{t=1}^{T} OR\right)\right) \quad (3)$$

where t is an integer between 1 and T.

The decision to choose an action A depends on the benefit-risk ratio call the performance measure (PM); If the ratio is > 1 then the action can give benefit otherwise, if the ratio < 1 then the action have a risk;

Let the performance measure (PM) be a function that associate an action to the ratio of the estimative outcomes

Table 1: Case study of malaria.

Specification of the problem		
Sex: Female Age: 40 Weight: 72		
Antecedents : - Allergies: Chloroquine		
Associated conditions or diseases: 2 months of pregnancy		
Localisation: Central Africa		
Symptoms and signs: Febrile paroxysms with body ache, nausea		
Indication		
Type of Infection: Plasmodia Falciparum (CIM-10, B-52)		
Severity of infection: Typical malaria		
Status: No recurrent		
Action		
Typical *P. falciparum* malaria is treated by Coartem, Quinine, Clindamycin and Chloroquine		
Estimative Outcomes		
Action 1 Coartem	Action 2 Quinine + Clindamycin	Action 3 Quinine
Estimative Benefit 1 405 effectives	Estimative Benefit 2 677 effectives	Estimative Benefit 3 318 effectives
Estimative Risk 1 212 effectives	Estimative Risk 2 312 effectives	Estimative Risk 3 241 effectives
Performance Measure		
Performance Measure 1 1.9	Performance Measure 2 2.2	Performance Measure 3 1.3
Decision based on estimative outcomes and its performance measure at time t		
Decision 1 (t) Choice: No	Decision 2 (t) Choice: Yes	Decision 3 (t) Choice: No
Real Result at time t+1		
-	Real Result 1 (t+1) Voted: Risk	-
Optimization at time t+1		
Choice: Yes		

OB and OR:

$$PM = (A_m, (OB_t/OR_t)) \quad (4)$$

It comes therefore: If, $A_1, \ldots, A_m$ implies effectively, by explicit verification, the estimative outcomes OB or OR; Then if, $A_{k+1}, \ldots, A_{k+m}$ considers as giving the outcomes OB, then, A_{k+m+1}, the real result will be OB, necessary by constructive induction demonstration; Or then if, $A_{k+1}, \ldots, A_{k+m}$ considers as giving the outcomes OR, then, A_{k+m+1}, the real result will be OR, necessary by constructive induction demonstration.

3 Results

3.1 Simulation: A Case Study

Malaria is one of the world's most deadly diseases. Even though it is highly preventable and treatable. More information can be found in [13, 14, 15].

The case study presented in Table 1 allows the simulation of our system, the data set used is fictive but it is approximately a reality.

4 Conclusion

On the first phase of the development of the project OMaT [16, 17], we are only proposed generic decisions without optimizations. We have been using the HTML, the JavaScript and the CSS for programming the client-side or the interface. OMaT is an online system that aims to assist physician at medical consultation in order to optimize the quality of care of the patients with malaria disease. We envision to contribute also to the realization of malaria vaccine by providing relevant information for vaccine malaria research such as virulence, antigenicity, evolution, and gene and protein interactions.

The work embodied in this paper formed the second part of our research project and provides a theoretical approach of the optimization of malaria's treatment.

References

[1] Feld S. What is the Definition of Quality Medical Care? [Internet] 2007 [cited 2013 Sept 10]. Available from: http://stanleyfeldmdmace.typepad.com/repairing_the_healthcare_/2007/03/what_is_the_def.html

[2] Zvarova J., Heroutova H., Grünfeldova H., Zvara K., Buchtela D. Empowering Clinicians by eHealth Technologies in Decision-Making Tasks, In: Adlassnig K.P., Mantas J., Masic I., eds. Medical Informatics in a United and Health Europe. IOS Press; 2009. P.683-687 (Studies in Health Technology and Informatics; vol 150).

[3] Tribus M. Rational Descriptions, Decisions and Designs. New York: Pergamon Press, 1969.

[4] White D.J. Fundamentals of Decision Theory. New York: North-Holland, 1976.

[5] Karni E. Decision Making under Uncertainty, The Case of State-Dependent Preferences. Cambridge: Harvard University Press, 1985.

[6] Akter S., Ray P. mHealth – an Ultimate Platform to Serve the Unserved. Geissbuhler A., Kulikowski C., eds. IMIA Yearbook of Medical Informatics. Stuttgart: Schattauer; 2010. P. 94-100.

[7] Kahn C.E. Artificial Intelligence in Radiology: Decision Support Systems. In: Hendee WR, Trueblood JH, eds. Digital Imaging. Madison, WI: Medical Physics Publishing, 1993: 281-300.

[8] Kenner T., Kenner L. Risk Factors, Protective Factors and Medical Decisions. Scripta Medica. 2001; 74(1):5–10.

[9] Payne J.W., Bettman J.R., Coupey E., Johnson E.J. A Constructive Process View of Decision Making: Multiple Strategies in Judgement and Choice. Acta Psychologica 1992 Aug; 80(1–3):107–141.

[10] Roy B. MultiCriteria Methodology for Decision Aiding. Dordrecht: Kluwer Academic, 1996.

[11] Steuer R. E. Multiple Criteria, Theory, Computation and Application. New York: John Wiley, 1986.

[12] Sundaran R.K. A first Course in Optimization Theory USA: Cambridge University Press, 1996.

[13] Mbonye A.K., Bygbjerg I.C., Magnussen P. Intermittent preventive treatment of malaria in pregnancy: a new delivery system and its effect on maternal health and pregnancy outcomes in Uganda. Bulletin of the World Health Organization. [Internet] 2008 [cited 2013 Sept 10]; 86(2):81-160. Available from: http://www.who.int/bulletin/volumes/86/2/07-041822/en/index.html

[14] Schantz-Dunn J., Nour N.M. Malaria and Pregnancy: A Global Health Perspective. Rev Obstet Gynecol. [Internet] 2009 Summer; [cited 2013 Sept 10]; 2(3): 186–192. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2760896/

[15] WHO, International Classification of Diseases (ICD). [Internet] 2013 [cited 2013 Sept 10]. Available from: http://www.who.int/classifications/icd/en/

[16] Muteba E. OMaT Project. [Internet] 2013 [cited 2013 Sept 10]. Available from: http://www.maesoft1.net/omat.htm

[17] Muteba E. Towards Networked eHealth: OMaT Project (Phase I). [Internet] 2013 [cited 2013 Sept 10]. http://jhia-online.org/index.php/jhia/index

Formalization of Clinical Practice Guidelines: Nonalcoholic Steatohepatitis Diagnosis Model-Related Personalized Medicine

Nassim Douali[1], **Meriem Abdennour**[1,2], **Jean Daniel Zucker**[1,3], **Marie-Christine Jaulent**[1]

[1] Inserm URMS 872 Team 7 and 20, Pierre and Marie Curie University, Paris, France

[2] Echosens, RD Departement, Paris, France

[3] IRD, UMI 209, UMMISCO, IRD France Nord, Bondy

Abstract

Objective: Non-alcoholic fatty liver disease (NAFLD) is a recently recognized entity related to modern lifestyle and with expanded clinical importance because of the rising incidence of obesity and diabetes.
Methods: We have developed a framework for interacting with patient's heterogeneous data (omics, clinical and biological information) and formalizing medical knowledge.
Results: In this paper we present new diagnosis model to predicate NASH. We extracted 18 clinical concepts and these concepts are annotated with SNOMED CT concepts. We tested our diagnostic model with database of 36 patients. We have a performance of 91%.
Conclusion: This work represents a preliminary step in developing a CDSS and we'll use a clinical database to test this system and to compare it with others statistic reasoning methods.

Keywords

Non-alcoholic fatty liver disease, Liver Disease, clinical decision support system, knowledge representation, artificial intelligence, fuzzy logics

Correspondence to:

Nassim Douali
Inserm URMS 872 Team 7 and 20, Pierre and Marie Curie University,
Address: Paris, France
E–mail: nassim.douali@crc.jussieu.fr

EJBI 2014; 10(1):6–10
received: July 31, 2013
accepted: February 9, 2014
published: February 28, 2014

1 Introduction

Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of chronic liver disease in the world [1]. NAFLD is a clinical syndrome and is pathologically characterized by diffuse macrovesicular fatty change in the hepatocytes. NAFLD includes simple nonalcoholic fatty liver disease, nonalcoholic steatohepatitis (NASH) and hepatic cirrhosis [2]. NAFLD is a recently recognized entity related to modern lifestyle and with expanded clinical importance because of the rising incidence of obesity and diabetes. NAFLD is an increasingly recognized cause of liver-related morbidity and mortality and it is frequently associated with insulin resistance. While insulin resistance and hyperinsulinemia, are, in large part, metabolic consequences of obesity, the basis of diversity in severity and progression of inflammation and fibrosis is not known [3].

The presence of fat in the liver means the accumulation of triglycerides. This accumulation determines the evolution of the disease. Infiltration of 30% of hepatocytes is an incipient form, the moderate form means that 60% of hepatocytes were infiltrated and severe form induces over 60% of infiltration hepatocytes [4].

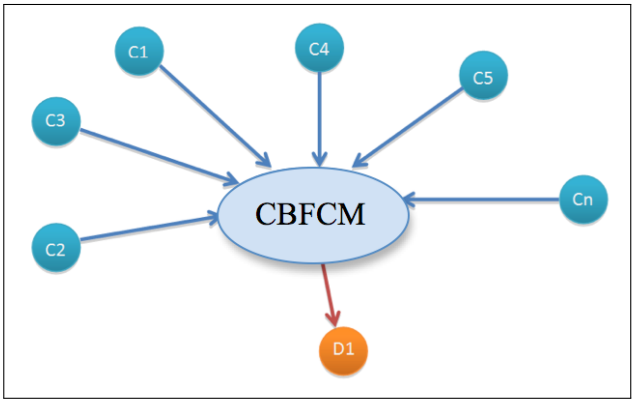

Figure 1: Diagram representing the CBFCM method.

NASH is a disease evolving under the influence of various stimuli still poorly understood. However in this disease, it is well know that insulin resistance is largely implied [5]. Risk factors for NASH/ Fibrosis are: Old over 45 years, Obesity, Diabetes/Insulin resistance, low platelets, low albumin, AST > ALT and imaging signs of hyper portal hypertension.

Several biological Nash prediction tests are developed [6, 7]. The evolution of NAFLD and NASH is variant for each patient and it is important to use all relevant information to diagnose the disease: clinical information, biological test, genomic information and imaging. In this paper we describe a new diagnosis support system based on validated knowledge from scientific literature and clinical practice guidelines (CPG) to diagnose NASH. We tested our diagnostic model with database of 36 patients generated randomly.

2 Methods

We have developed a framework for interacting with patient's heterogeneous data (omics, clinical and biological information) and formalizing medical knowledge.

2.1 CPG Formalization

In the present work we used not only CPG[1] but also scientific literature as sources of knowledge. We have developed a Fuzzy Semantic Web Reasoning approach using the RDF/N3 language to express the rules from the knowledge sources [8]. This approach was previously tested for Infectious diseases diagnosis decision support [9].

2.2 Case Based Fuzzy Cognitive Maps (CBFCM)

Case Based Fuzzy Cognitive Map (CBFCM) is a hybrid decision-making computing technique [10]. CBFCM is represented as nodes (concepts) that illustrate the different aspects of the system's behavior. Concepts may represent variables, states, events, inputs and outputs, which are essential to model a system [11]. The value of each node (concept) is represented as Fuzzy Set. Figure 1 is a graphical representation of CBFCM method. A patient N is described by a set of clinical parameters. These clinical parameters can be clinical signs, age, gender or biological results. CBFCM is validated in many medical areas [12, 13, 14, 15].

[1]World Gastroenterology Organisation Global Guidelines: Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis http://www.worldgastroenterology.org/assets/export/userfiles/2012_NASH%20and%20NAFLD_Final_long.pdf

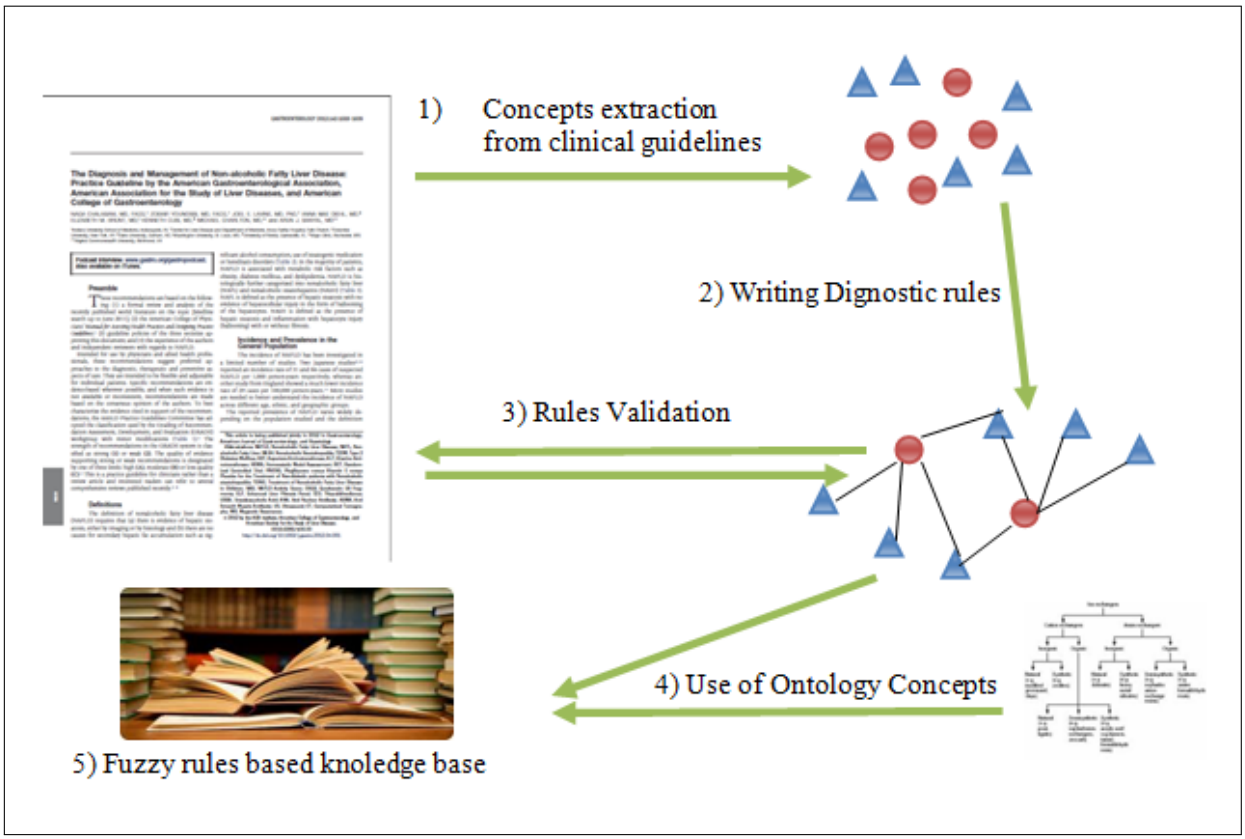

Figure 2: CPG fuzzy formalization method.

Table 1: Clinical concepts of NASH diagnosis model.

Clinical concepts	Label Description	SNOMED CT Concepts ID
C1	Hyperglycemia	80394007
C2	Hypertriglyceridemia	302870006
C3	HDL	9422000
C4	Hypertension	38341003
C5	BMI	60621009
C6	Waist circumference	276361009
C7	Fasting insulin	252251004
C8	Index HOMA-IR	237650006
C9	Alcohol consumption	160580001
C10	ALT	250637003
C11	AST	250641004
C12	Apolipoprotein	259599001
C13	GGT	60153001
C14	Haptoglobin	85294008
C15	α-fetoprotein	16236008
C16	Adiponutrine gene profil	413451007
C17	Old-Age	70753007
C18	Sexe	263495000
D1	NASH	197321007

2.3 Case Based Fuzzy Cognitive Map Reasoning Mechanism

The CBFCM reasoning process follows a number of steps till the system's equilibrium point. These steps can be found in [10, 12] and we briefly present them here. At first step, the initial state of the concepts is given either from experts or from the existing medical database. During reasoning the CBFCM iteratively calculates its state until convergence. The state is represented by a state vector $C^{(k)}$, which consists of real node values $C_i^{(k)} \in [0,1], i = 1, 2, \ldots N$ at an iteration k. The value of each node is calculated by the following equation:

$$C_i^{(k+1)} = f\left(C_i^{(k)} + \sum_{j=1, j\neq i}^{N} C_j^{(k)} \cdot W_{ji}\right) \quad (1)$$

Where f is a threshold (activation) function:

$$f(x) = \frac{1}{1 + e^{-m(x)}} \quad (2)$$

Where m is a constant parameter [13]. The parameter m determines how quickly the $f(x)$ approaches the limiting values of 0 and 1. The transformation function is used to reduce unbounded weighted sum to a certain range, which hinders quantitative analysis, but allows for qualitative comparisons between concepts.

In order to remove the spurious influence of inactive concepts (concepts with zero values) on other concepts, and to avoid the conflicts emerge in cases where the initial values of concepts are 0.5, as well as the missing data, a modified CBFCM reasoning formalism can be used. Based on this assumption, we reformulated eq. (1) as:

$$C_i^{(k+1)} = f\left((2C_i^{(k)} - 1) + \sum_{j=1, j\neq i}^{N} (2C_j^{(k)} - 1) \cdot W_{ji}\right) \quad (3)$$

For example, in Figure 1, the value of a concept "DiagnosisN" is obtained by multiplying the value of each of its input concepts, C_i, by their respective weights, W_{ij}, giving values in [0..1]. These values are then summed as in eq. (1) and eq. (3) and the nonlinear function f is used

```
48 (:PatientId0328 snomedct:266468003) fl:pi 0.999959699696969.
49 (:PatientId0329 snomedct:266468003) fl:pi 0.999912345521133.
50 (:PatientId0330 snomedct:266468003) fl:pi 0.998752019938721.
51 (:PatientId0331 snomedct:266468003) fl:pi 0.997674901812091.
52 (:PatientId0329 snomedct:266468003) fl:pi 0.988729171911901.
```

Figure 3: Example of decision rules.

to limit the range of possible output values. The inference follows an iteration process till the system convergence in a steady state (means a state where all the concepts do not change any more their values).

The simulation stops when a limit vector is reached, i.e., when $C_i^k - C_i^{k-1} \leq e$; where e is a residual, whose value depends on the application type (and in most applications is equal to 0.001) [13]. The conclusions based on CBFCM should be viewed together with existing scientific knowledge [14].

The construction of CBFCM [9] is consisting of three parts: (a) to determine concepts and (b) to determine the strength of cognitive relationships between concepts (c) to explicit fuzzy control rules (see Figure 2).

In the semantic web N3 formalism, the weights W_{ij} are in the range $[0, 1]$, each weight (concepts' influences) presents a degree of influence from 0 to 1 [8].

2.4 Information model

The Semantic Web framework based on CBFCM integrate heterogeneous data: clinical data (signs, symptoms...), Biological data (lab test...), Imaging and Omics data. To create a patient clinical profile we need to use all these types of data. These data are annotated with SNOMED CT concepts. We extracted a related part of SNOMED CT using UMLS.

3 Results

We extracted 18 clinical concepts and these concepts are annotated with SNOMED CT concepts (Table 1). We tested our diagnostic model with database of 36 patients. We have a performance of 91%.

Figure 3 is an example of result rules. This output is N3 triples:PatienIdxxx is the ID of our patients, snomedct:266468003 is a decision concept (NASH) and fl:pi is the confidence degree of the decision (the range value is (0-1))

4 Discussion and Conclusion

The CBFCM approach allowed us to integrate heterogeneous clinical data to perform a personalized patient profile. This method can identify causal relationships between clinical, biological, genetic concepts and decision concept (Diagnosis of NASH). The use of CBFCM enables to incorporate several sources of knowledge (several CPGs, knowledge from literature), which is of great advantage since all knowledge is rarely embedded in a unique CPG. Indeed, knowledge of a medical field is usually broad, complex and closely related to other areas so that several knowledge sources are needed to cover and modeled the medical domain in question.

We have implemented the knowledge bases, rules and databases in the same environment (RDF, N3, Euler...) without compatibility constraints; this is one of the advantages of using Semantic Web tools. The success rate of 91% shows the functionality of the model and its future usefulness in clinical practice.

The conducted study allowed us to test cognitive approaches reasoning to enable personalized medicine. The advantage of this approach is to enable the sharing and reuse of knowledge and simplify maintenance. This work represents a preliminary step in developing a CDSS and we'll use a clinical database to test this system and to compare it with others statistic reasoning methods.

References

[1] Farrell G.C. and Larter C.Z. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatology 2006; 43:S99–S112

[2] Sanyal AJ. AGA technical review on nonalcoholic fatty liver disease. Gastroenterology 2002;123:1705-25.

[3] Charlton M. Nonalcoholic fatty liver disease: a review of current understanding and future impact. Clin Gastroenterol Hepatol. 2004. 1048-58

[4] James OFW and Day CP. Non-alcoholic steatohepatitis (NASH): a disease of emerging identity and importance. J Hepatol 1998; 29:495 – 501.

[5] Medina J, Fernandez-Salazar LI, Garcia-Buey L, Moreno-Otero R. Approach to the pathogenesis and treatment of nonalco- holic steatohepatitis. Diabetes Care 2004 ; 27 : 2057-66.

[6] Poynard T, Ratziu V, Charlotte F et al. Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. BMC Gastroenterology 2006, 6:34

[7] Ratziu V, Massard J, Charlotte F, Messous D, Imbert-Bismut F etall. Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. BMC Gastroenterology 2006, 6:6

[8] Douali N, Papageorgiou E, De Roo J, Jaulent MC. Case Based Fuzzy Cognitive Maps: New method for medical reasoning - Fuzzy Systems (FUZZ), 2011 IEEE International Conference on, 844-850

[9] Douali N, De Roo J, Jaulent MC. Clinical diagnosis support system based on case based fuzzy cognitive maps and semantic web. Stud Health Technol Inform. 2012;180:295-9.

[10] Douali N, De Roo J, Jaulent MC. Decision support system based semantic web for personalized patient care. Stud Health Technol Inform. 2012;180:1203-5.

[11] Douali, N., Jaulent, M.Ch. "Genomic and personalized medicine decision support system." Complex Systems (ICCS), 2012 International Conference on. IEEE, 2012.

[12] Douali, N. et al., Diagnosis Support System based on clinical guidelines: comparison between Case-Based Fuzzy Cognitive Maps and Bayesian Networks, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.09.012

[13] Douali, N. et al. "New Semantic Web rules and new medical reasoning framework." Instrumentation and Measurement Technology Conference (I2MTC), 2013 IEEE International. IEEE, 2013.

[14] Douali, N. et al. "Noninvasive Diagnosis of Nonalcoholic Steatohepatitis Disease Based on Clinical Decision Support System." Studies in health technology and informatics 192 (2012): 1178-1178.

[15] Douali, N., Papageorgiou, E. I., De Roo, J., Sun, H., Colaert, D., Jaulent, M. C. (2013, January). Improve treatment of pneumonia and reduce adverse drug events. In Point-of-Care Healthcare Technologies (PHT), 2013 IEEE (pp. 89-92). IEEE.

Possibilities of Personal Health Status Monitoring

Jan Havlík[1], **Jan Dvořák**[1], **Jakub Parák**[1], **Matouš Pokorný**[1], **Lenka Lhotská**[1], **Petr Panýrek**[2]

[1] Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

[2] HIGH TECH PARK, Prague, Czech Republic

Abstract

Background: There are a lot of possibilities how to monitor the patient using wireless technologies and how to initiate the appropriate action in life-threatening situations. The paper focuses on those issues. The overview of basic principles and outline of significant advantages and disadvantages of the methods are presented.
Objectives: The main aim of these solutions is to ensure safety of the elderly and impaired people in their everyday life, and especially, to enable the possibility to stay in their natural environment (home, family, etc.) instead of institutionalizing them.
Methods: Two solutions for monitoring and providing better life quality conditions are presented in more detail. The specific solution for monitoring vital signs and classifying urgent states using telemedical system - the Intelligent Primer Nurse application - is presented in detail. The smart system combining the telemedical approach and home TV computer - Home Brain system - is described.
Results: Both solutions have been tested in real use and under different conditions. The Intelligent Primer Nurse system has been tested during long term vital signs monitoring including several activities - walking, running and idle standing. The Home Brain system has been evaluated during a pilot study with a group of elderly in their real life. The user experiences have been investigated by a moderated discussion.
Conclusions: Finally, we discuss the future development in this domain.

Keywords

Telemonitoring, telemedicine, vital signs monitoring, assistive technologies

Correspondence to:

Jan Havlík
Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague
Address: Technická 2, CZ–16627 Prague 6
E–mail: xhavlikj@fel.cvut.cz

EJBI 2014; 10(1):11–14
received: December 17, 2013
accepted: February 9, 2014
published: February 28, 2014

1 Introduction

Currently, there is an increasing demand for smart solutions helping and serving in daily life. Teenagers, students, young people, and also many people in productive age are the users of smart phones, tablets, netbooks, and other smart and portable devices. Although elderly people often mistrust these devices, it could be observed during the last years that there is an increasing interest in these solutions also in post-productive age. Elderly people become suddenly aware that the smart solutions could significantly help them managing their daily life.

Simultaneously with the decrease of mobile device prices, the number of mobile applications has been growing. Especially in the field of personal health state monitoring, the speed of development of applications is very rapid. Actually we can measure many physiological parameters from a human body in relatively easy way: electrocardiogram (ECG), heart rate (HR), breathing rate, body temperature, blood pressure, etc. [1]. Many clinical trials have been performed, e.g. [2, 3] assessing the usefulness and efficiency of telemonitoring systems.

Many of the existing solutions could be easily adopted and adapted to elderly and applied in real life. The main aim of almost all of these applications is to safeguard the elderly and impaired people in their daily life, and more specifically to enable them stay even longer in their own natural environment (home, family, etc.) rather than being institutionalized.

Based on the facts mentioned above, one of the most actual problem is to design and develop the smart solutions, embedded devices, and telemonitoring applications as user-friendly as possible [4, 5]. The paper presents two examples of solutions for monitoring health status and providing better life for elderly or impaired persons.

1.1 Methods

The first solution is a specific solution for monitoring vital signs and classifying urgent states using telemedical system - the so-called Intelligent Primer Nurse application. The second one is a smart system combining the telemedical approach and home TV computer - the so-called the Home Brain system.

1.2 Intelligent Primer Nurse

The Intelligent Primer Nurse is a device for a continuous monitoring of vital signs that acts as a personal and portable vital signs monitor (see the Fig. 1 for the laboratory realization). The device is able to activate alarms in case the vital signs are not any longer in the specified, pre-defined range. Chronically diseased, elderly, and other threatened persons including persons at risk for a stroke are the target group of users.

The device as such is based on the EvoPrimer [6] development kit. The basic kit is supplemented with the extension board, the input module, and communication modules. The device allows measuring electrocardiography (ECG), plethysmography (PPG) and activity (accelerometry) signals, displaying the values of the signals on the screen in real time, and processing these signals accordingly.

The alarm is activated whenever the heart rate is out of the pre-defined range or when the user is inactive for long time. The activity alarm could be inhibited by pushing the activity button. The device is able to send alarm signal wirelessly to the subordinate PC system or directly to the control center.

1.3 Home Brain

Home Brain is a unique system put into practice at the end of 2011 after more than five year of advanced research. The system is technically designed as a small device like a set-top-box connected to a standard TV set (see Fig. 2). The main aim of the device is to provide the gate functionality to the Internet, multimedia services, senior monitoring, health state monitoring, social networking, remote control of home devices, intelligent security system, etc. The Home Brain system allows networking without demand for complicated control in the easiest way possible. This approach significantly increases the ability of elderly to stay at their own home environment instead of moving them into centers of institutionalized care such as senior houses [7].

2 Results

Both systems presented above have been tested in real use and under different conditions.

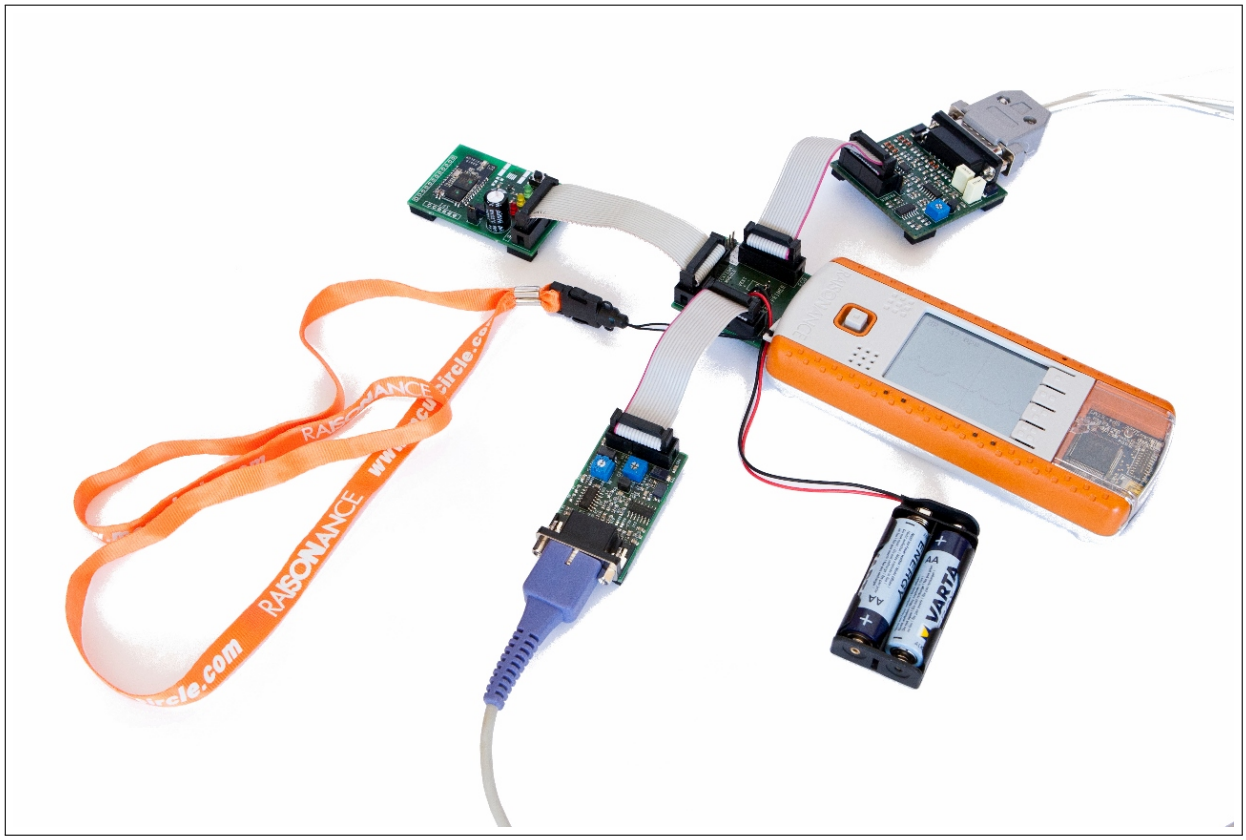

Figure 1: Intelligent Primer Nurse.

The Intelligent Primer Nurse application has been evaluated in simulated situations, the so-called test scenarios. The system has been tested during long term vital signs monitoring including several activities - walking, running and idle standing. The reaction to life-threatening situations like low heart rate, no activity (for example after the falling down) etc. has been also tested [8].

The Home Brain system has been evaluated during a pilot study with a group of elderly in their real life. During this study the Home Brain has been tested by 5 users, 4 women and 1 man, from 66 to 78 years old. The user experiences have been investigated by a moderated discussion. As a main advantage of the system the users high-lighted the simplicity and the intuitiveness of system control. The respondents had no manuals, they learned only by using the system. They agree that the most important issue is not to be afraid of using the Home Brain. As the most frequently used functions they activate watching TV, using a TV archive (instead video-recorder), listening to radio, Skype calling, instant messaging, using the photo archive, managing details about their home (payments, important decisions, dates of medical visits etc.) and evidence of health status. The respondents have been also asked for arguments for potential new users. As crucial arguments they mention the simplicity of the use, the comfortableness, and the enhancement of communication possibilities [9].

Recently, the Home Brain has been introduced to the market.

3 Conclusion

There are many possibilities to help elderly and impaired persons in their everyday life. Using the up-to-date techniques and applications such as smart phones, tablets and specialized portable devices, it is relatively cheap and easy to monitor the health status continuously, to send a message in case the health status is not all right, and to additionally network them with their family, friends, or caregivers. For further development we have to keep in mind the requirement on simple and intuitive control of these devices that can increase their acceptance in population. Another issue that is not yet satisfactorily solved is the interoperability of different solutions. The ultimate

Figure 2: Home Brain.

goal is to introduce the "plug-and-play" concept for easier extension of the applications.

The number of possibilities to monitor personal health status is increasing as fast as the prices of new technologies are getting down.

Acknowledgements

This work has been jointly supported by the grant No. G3 902/2013 of the University Development Foundation, by the grants No. SGS12/143/OHK3/2T/13 and SGS13/203/OHK3/3T/13 of the Czech Technical University in Prague, and by the research program No. MSM 6840770012 of the Czech Technical University in Prague (sponsored by the Ministry of Education, Youth and Sports of the Czech Republic).

References

[1] Xiao-Fei T., Yuan-Ting Z., Poon C., Bonato P.: Wearable medical systems for p-health. IEEE Reviews in Biomedical Engineering 2008; 1:62–74.

[2] Martin-Lesende I., Orruño E., Cairo C., Bilbao A., Asua J., Romo M., Vergara I., Bayn J., Abad R., Reviriego E., Larrañaga J.: Assessment of a primary carebased telemonitoring intervention for home care patients with heart failure and chronic lung disease. The telbil study. BMC Health Services Research 2011, 11(56).

[3] Kraai I., Luttik M., de Jong R., Jaarsma T., Hillege H.: Heart failure patients monitored with telemedicine: Patient satisfaction, a review of the literature. Journal of Cardiac Failure 2011, 17(8):684–690.

[4] Chan M., Esteve D., Escriba C., Campo E.: A review of smart homes - present state and future challenges. Computer methods and programs in biomedicine 2008, 91(1):55–81.

[5] Lhotská L., ŠtěpánkovJ. O., Pěchouček M., Šimák B., Chod J.: ICT and eHealth projects. Telecom World (ITU WT), Technical Symposium at ITU, Piscataway: IEEE; 2011; 57–62.

[6] EvoPrimer for STM32F103VE. STM [Internet]. 2012 [cited 2012 Feb 15]. Available from: http://www.stm32circle.com/resources/stm32Eprimer.php.

[7] HIGH TECH PARK: Project3B HomeBrain. High Tech Park [Internet]. 2011 [cited 2012 Feb 15]. Available from: http://www.htpark.eu/en/solutions/r1-project3b-homebrain/.

[8] Parák J., Dvořák J., Havlík J.: Device for long term measurement of heart rate. ACM Digital Library: Proceedings of 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies [CD-ROM], New York: ACM, 2011, 1–5.

[9] Příbová, M.: HOMEBRAIN Opinions and feelings of users about multifuncional device (in Czech). Unpublished research note, ILOL, 2010.

Design and Implementation of an Ontology for the Computable Representation of Clinical Prediction Rules

Derek Corrigan[1], **Adel Taweel**[2], **Tom Fahey**[1], **Theodoros Arvanitis**[3], **Brendan Delaney**[2]

[1] HRB Centre for Primary Care Research, Royal College of Surgeons in Ireland, Ireland

[2] Kings College London, United Kingdom

[3] University of Warwick, United Kingdom

Abstract

Objective: A lack of acceptance has hindered the widespread adoption and implementation of clinical prediction rules (CPRs). The use of clinical decision support systems (CDSSs) has been advocated as one way of facilitating a broader dissemination and validation of CPRs. This requires computable models of clinical evidence based on open standards rather than closed proprietary content.

Methods: The on-going TRANSFoRm project has developed ontological models of CPRs suitable for providing CPR based decision support.

Results: This paper describes the design and implementation of a generic ontology model for the representation of computable CPRs. The conceptual validity and implementation of the ontology is discussed using an illustrative example of a CPR in the form of the Alvarado Score for acute appendicitis.

Conclusions: We demonstrate how the model is used to query the structure of this particular rule, providing a generic computable representation suitable for the representation of CPRs in general.

Keywords

Clinical prediction rules, ontology, clinical decision support

Correspondence to:

Derek Corrigan
Department of General Practice,
HRB Centre for Primary Care Research, Beaux Lane House
Address: Mercer Street Lower, Dublin 2, Ireland
E–mail: derekcorrigan@rcsi.ie

EJBI 2014; 10(1):15–23
received: July 31, 2013
accepted: February 9, 2014
published: February 28, 2014

1 Introduction

Although many diverse examples of clinical prediction rules (CPRs) in primary care can be identified in research literature, their use has yet to gain widespread acceptance among clinicians [1, 2]. There are a number of valid concerns that influence why clinicians are reluctant to use them as part of their day-to-day clinical practice.

Despite the existence of an accepted development lifecycle for producing CPRs, many of them have traditionally focussed solely on the derivation phase of the CPR lifecycle [3]. Many derived CPRs are subject to poor or non-existent CPR validation and impact analysis. This lack of validation severely limits their perceived applicability to the same restricted patient populations defined in the original derivation research populations. Complications may arise when there are multiple rules derived by different researchers for any chosen clinical condition. As an example, a clinical condition such as Pulmonary Embolism has numerous variations of CPRs that may pertain to it [4]. This can lead to confusion and a lack of clarity about which CPR variations are the "correct" or "best" ones to use.

With some exceptions the format for dissemination of CPRs is largely literature based, putting an onus on clinicians to search literature for suitable CPRs [5]. This is compounded by the fact that literature based rules are by their nature static in content and do not provide for recording of versioned rule changes. This may have implications for the applicability of any particular CPR as changes take place over time in the demographics of the original rule derivation study population.

One suggested way of addressing these limitations is through development of clinical decision support systems (CDSSs) based on computable models of clinical evidence [6, 7, 8]. The ultimate vision is to provide for computable representations of CPRs that allow derivation, validation, dissemination, versioning and on-going revision from empirical sources of electronic primary care patient data. This can be complemented using extraction of patient cues

and demographics from electronic health records (EHRs) as a trigger for initiating appropriate rule execution.

The TRANSFoRm project has developed computable ontological models of CPRs to support their electronic derivation, implementation and validation [9]. We describe the models and conceptual validity through implementation of a well studied CPR, the Alvarado score [10, 11]. We demonstrate how clinical questions are expressed as ontological queries for use by a CPR based CDSS currently being developed by the TRANSFoRm project.

2 Definition and application of Clinical Prediction Rules

2.1 CPR Definition

It is necessary to clearly define at the outset what we mean when we talk about using a clinical prediction rule. A CPR "is a clinical tool that quantifies the individual contributions that various components of the history, physical examination, and basic laboratory results make toward the diagnosis, prognosis, or likely response to treatment in a patient" [12, 13]. The formal characteristics of a CPR can be clearly identified based on this definition. Typically a CPR is derived from a statistical model and will be constructed and structured based on the following distinct parts:

- A clinical outcome that relates to a defined diagnostic, prognostic or treatment outcome associated with a selected clinical condition.
- A set of diagnostic cues and associated criteria that is indicative of the clinical outcome being assessed by the rule.
- A statistically derived scoring scheme that quantifies the relative contribution of each cue where present to the clinical outcome.
- A threshold based scoring scheme that defines relative clinical interpretations of risk categories for all possible scores for the rule.
- An optional decision indicating a clinical action in response to each risk category to be recommended based on each of the defined threshold scores.

2.2 Application of CPRs as part of a defined Diagnostic Strategy

In order to understand how CPRs may be potentially applied as a diagnostic tool in clinical practice it is useful to place their use in a broader diagnostic context. A clinician needs to formulate and consider the evidence for all possible differential diagnoses when a patient first presents with a particular clinical complaint. This is done by considering each differential diagnosis and can involve "ruling out" differentials based on the underlying diagnostic cues as presented by the patient. CPRs can provide a useful tool to assist with these potential "rule outs" using the results of applying suitable CPRs obtained to any particular patient case [13]. Their appropriate use can be applied as a tool to reduce the possibility of diagnostic error at the outset through consideration of possible differentials [12, 13]. As an example a patient presenting with abdominal pain who scores less than 4 on the Alvarado score, could indicate a potential "rule out" for acute appendicitis for that patient.

3 Model Development Methodology

The development of the formal models of clinical prediction rules described here followed a number of distinct steps subsequently described in detail:

- Clinical use case development.
- Functional requirements definition of the CPR model.
- Model design based on functional requirements.
- Model construction and clinical evidence population.
- Clinical use case implementation and validation.

3.1 Clinical Use Case Development

The models presented here provide the backend knowledgebase to be used as part of a broader piece of work currently in progress to develop a functional diagnostic decision support system as part of the TRANSFoRm project. The CDSS will consume and ask clinical questions of the models described here that provide the underlying knowledgebase. This CDSS tool will be deployed and used by primary care practitioners to assist them in formulating and quantifying differential diagnoses to consider for patients presenting with three defined diagnostic conditions. The use of electronic CPRs will be deployed as part of the diagnostic strategy for ruling out of potential differential diagnoses. The three primary care patient safety use cases will be used to test and validate the fully functional CDSS being developed by TRANSFoRm.

The selected patient safety use cases focus on potential diagnoses relating to patients presenting with the general complaints of chest pain, abdominal pain or dyspnoea. These were chosen for the cognitive challenge they present in primary care with potential for diagnostic error [14, 15]. Reviews of evidence based sources identified CPRs supporting selected diagnoses for these patient safety use cases [10, 16]. In total 41 clinical prediction

rules were identified relating to 20 diagnostic conditions relating to the three patient safety use cases. In this paper we describe the model representation of a single CPR called the Alvarado Score relating to a diagnosis of appendicitis for a patient presenting with abdominal pain.

3.2 Functional Requirements Definition of the CPR Model

In considering the model design requirements it is useful to first consider the functional requirements of any application that will use those developed models. The models described here will be ultimately queried by the TRANSFoRm CDSS. The CDSS will want to query particular diagnostic conditions, retrieve associated CPRs for any condition and query all of the constituent rule structures for any selected CPR. We have therefore defined our model requirements based on the different CPR related questions it needs to be able to answer. The functional requirements can be stated as clinical questions we wish to be able to ask of our CPR model. We identified the following questions as general functional requirements that we want to able to answer using the finished CDSS tool:

- What are the differential diagnoses to consider for a selected patient reason for encounter (RFE)?
- What are the related CPRs associated with a selected diagnosis?
- What are the cues, criteria and associated scores of a selected CPR?
- What are the scoring interpretation schemes of a selected CPR?

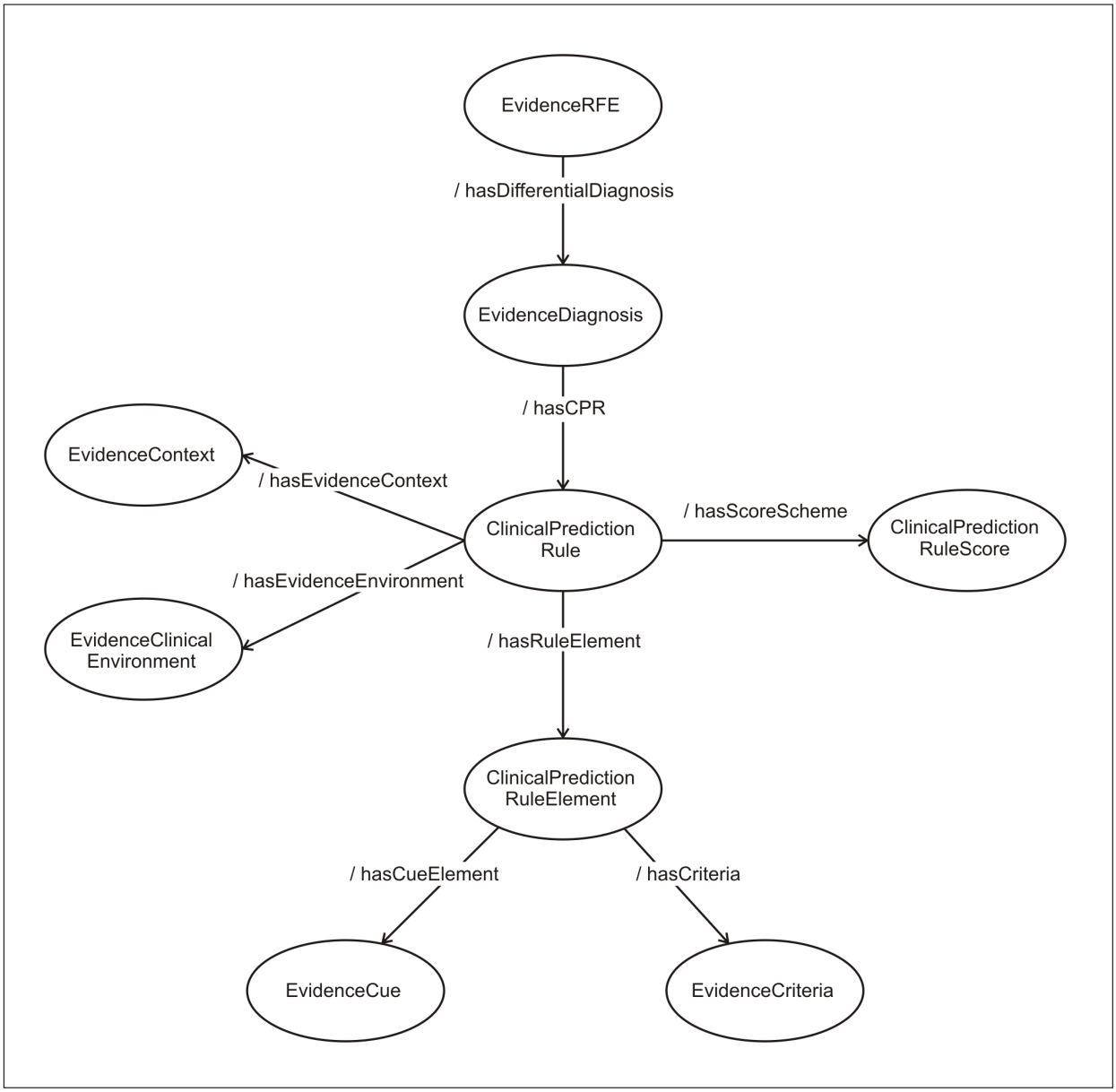

Figure 1: Relationship of CPR ontology concepts.

- What are the population characteristics associated for application of a selected CPR?
- What is the clinical setting associated for application of a selected CPR?
- What are the supporting literature sources for a selected CPR?
- What is the current version number of a selected CPR?

3.3 Model Design based on Functional Requirements Definition

An ontology representation was chosen as the basis for the CPR model to support dissemination of CPRs using open standards that support a simple underlying data structure. Many methodologies have been proposed for design and development of ontologies [17]. The approach we have selected uses an application focussed approach where ontology requirements are expressed as "competency questions" that can then be used as a set of functional requirements to validate ontology completeness [18]. In our example, the functional requirements we have already defined can be considered to also define suitable ontology competency questions. If our ontology is designed correctly we should be able to express all our competency questions as formal ontology queries that generate correct clinical results with respect to our selected clinical use cases when executed (in this case representation of the Alvarado Score for appendicitis). Competency questions were deconstructed to identify the required formal ontology concepts and defined relationships that exist between them. The ontology concepts and relationships identified are shown in Figure 1. Although named inverse relationships exist for all relationships within the constructed ontology, we have only shown relationships in one direction in the diagram for clarity. For example the relationship 'hasDifferentialDiagnosis' has a corresponding inverse relationship called 'isDifferentialDiagnosisOf' that is not explicitly shown.

These core CPR ontology concepts are described in Table 1 along with examples of clinical instances and associated attributes of those concepts.

A fundamental requirement of the TRANSFoRm project is the appropriate use of standard clinical vocabularies, terminologies and classifications to add semantic meaning to any ontology terms being used through binding of vocabulary terms.

Table 1: Core CPR Ontology Classes with Descriptions and Class Attribute Examples.

Class Name and Description	**Class Instance**	**Attribute Examples**
EvidenceRFE: The patient reported reason for encounter (RFE)	AbdominalPainRFE	hasUMLSCode C0000737 hasIPC2Code D06 hasICD10Code R10.0 hasReadCode XaA06 YaYkf
EvidenceDiagnosis: A differential diagnosis of a particular RFE	Appendicitis	hasUMLSCode C0003615 hasIPC2Code D88 hasICD10Code K35 hasReadCode J20..Y30Di
ClinicalPredictionRule: A versioned CPR associated with a particular diagnosis with links to supporting literature URLs	AlvaradoScore1_0	hasRuleVersion 1_0 hasSupportingLiteratureURL http://www.biomedcentral.com/content/pdf/1741-7015-9-139.pdf
ClinicalPredictionRuleElement: One individual element of the CPR that is associated with one cue and the criteria to apply to it for a particular CPR	AlvaradoScoreElement1	
EvidenceCue: An associated sign, symptom, risk or clinical test that may be associated and reused in more than one CPR	Nausea	hasUMLSCode C0375548 hasIPC2Code D09 hasICD10Code R11.0 hasReadCode X75qw.Y7Cjf
EvidenceCriteria: The criteria and weighted rule score associated with a ClinicalPredictionRuleElement. The presence or absence of the cue and score is indicated through the criteria attributes	AlvaradoElementCriteria1	isPresent = True hasScoreInterpretation 1
ClinicalPredictionRuleScore: A score range to be used for clinical interpretation of the rule along with the textual interpretation of that score level	AlvaradoScoreLevel3	hasStartScore 7 hasEndScore 10 hasScoreInterpretation "Surgery"
EvidenceContext: A group of classes that defines the evidence population demographics used to derive the rule	Adult, Male, Ireland	hasAgeGreaterThan 17 hasISOCode 1 hasISOCode "IE"
EvidenceClinicalEnvironment: The clinical setting or context in which the rule was derived and is suitable for application	PrimaryCare	

A TRANSFoRm vocabulary service has been developed to allow runtime access to a number of vocabularies through defined web service interface methods [19]. The Unified Medical Language System [UMLS] has been used as a pivot terminology from which mappings have been provided to others including the International Classification of Primary Care Version 2 (ICPC2), SNOMED Clinical Terms, the International Classification of Diseases Version 10 (ICD10) and Read Codes [20, 21, 22, 23].

The CPR ontology model provides attributes (as shown in Table 1) to allow association of selected terminological codes to instances of the EvidenceRFE, EvidenceDiagnosis and EvidenceCue concepts. Multiple code system terms can be associated to any instance. At present these terms are manually entered into the ontology.

In order to facilitate CPR execution based on coded RFEs or diagnostic cues extracted from individual patient EHRs, future development will focus on integrating the ontology models with the TRANSFoRm vocabulary service. This will allow querying at runtime using only UMLS associations to pivot to the appropriate terminology implemented by the EHR data. This can also provide for coded ontology content to be represented and populated dynamically into the ontology through application of data mining techniques to electronic sources of coded primary care data.

3.4 Model Construction and Clinical Evidence Population

This constructed ontology design has been expressed using the ontology language/resource description framework (OWL/RDF) representation and implemented using the Protégé 4.1 ontology designer [24, 25, 26]. It is hosted using a Sesame triple store for query formulation, testing and future dynamic programmatic update of ontology content [27, 28]. The clinical content for the ontology was manually populated as instances of the ontology concepts to reflect the structure of the Alvarado score as described in literature [10, 11].

3.5 Ontology Metrics

The CPR ontology model is part of a larger clinical evidence ontology model that also supports the general representation of diagnostic knowledge. The knowledgebase metrics for the full ontology are:

- Number of ontology classes = 43
- Number ontology relationships = 101
- Data of ontology attributes = 48
- Number of ontology class instances = 505

Table 2: Competency Questions 1-4 (from Table 1) Expressed as SPARQL Queries with Associated Results.

SPARQL (Protocol and RDF Query Language)	Query Result (Instance Relation Value)
SELECT ?DifferentialDiagnosis WHERE {?DifferentialDiagnosis isDifferentialDiagnosisOf AbdominalPainRFE .}	Appendicitis, BacterialEnteritis ChronsDisease, CorPulmonale EctopicPregnancy, Pyelonephritis UrinaryTractInfection
SELECT ?CPR WHERE {?CPR isCprOf Appendicitis.}	AlvaradoScore1_0
SELECT ?CueElement ?Property ?Value WHERE {?RuleElement isRuleElementOf AlvaradoScore1_0. ?CriteriaElement isCriteriaOf ?RuleElement. ?CueElement isCueElementOf ?RuleElement. ?CriteriaElement ?Property ?Value. ?Property rdf:type owl:DatatypeProperty. }	MigrationOfPain isPresent true MigrationOfPain hasScoreInterpretation 1 Anorexia isPresent true Anorexia hasScoreInterpretation 1 Nausea isPresent true Nausea hasScoreInterpretation 1 RightLowerQuadrantTenderness isPresent true RightLowerQuadrantTenderness hasScoreInterpretation 2 ReboundPain isPresent true ReboundPain hasScoreInterpretation 1 ElevatedTemperature isPresent true ElevatedTemperature hasScoreInterpretation 1 Leucocystosis isPresent true Leucocystosis hasScoreInterpretation 2 WhiteBloodCellShiftLeft isPresent true WhiteBloodCellShiftLeft hasScoreInterpretation 1
SELECT ?ScoreElement ?Property ?Value WHERE {?ScoreElement isScoreSchemeOf AlvaradoScore1_0 . ?ScoreElement ?Property ?Value. ?Property rdf:type owl:DatatypeProperty. } ORDER By ?ScoreElement	AlvaradoLevel1 hasScoreInterpretation "Discharge" AlvaradoLevel1 hasStartScore1 AlvaradoLevel1 hasEndScore 4 AlvaradoLevel2 hasScoreInterpretation "Observation/Admission" AlvaradoLevel2 hasStartScore 5 AlvaradoLevel2 hasEndScore 6 AlvaradoLevel3 hasScoreInterpretation "Surgery" AlvaradoLevel3 hasStartScore 7 AlvaradoLevel3 hasEndScore 10

4 Clinical Use Case Implementation and Validation

4.1 The Alvarado Score as a CPR example

A particular clinical example of a well studied CPR is the Alvarado Score which we will use as a clinical example to illustrate use of our models. This rule categorises the risk of patients having potential acute appendicitis based on the presence or absence of 8 diagnostic indicators. The risk of appendicitis is expressed as three score-based risk categories with associated recommended treatment options. This rule has been designed to be suitable for primary care and is based on the presence of diagnostic cues without the need for imaging [10]. Reviews have highlighted the importance of capturing the demographic context of the derivation study population. Clinical performance of the Alvarado score has been shown to vary in different populations depending on gender and age, performing best for adult males [11]. This demographic variability should be reflected in any model design.

Using the example of appendicitis and the Alvarado Score we identified the following questions as functional requirements that we want to able to answer using the finished CDSS tool:

- What are the differential diagnoses to consider for a reason for encounter (RFE) of abdominal pain?
- What are the CPRs associated with the differential diagnosis of appendicitis?
- What are the cues, criteria and associated scores of the Alvarado score?
- What are the scoring interpretation schemes of the Alvarado score?
- What are the population characteristics associated for application of the Alvarado score?
- What is the clinical setting associated for application of the Alvarado score?
- What are the supporting literature sources for the Alvarado score?
- What is the current version number of the Alvarado score?

4.2 Expression of CPR Model Queries

The competency questions previously defined as functional requirements were expressed as Protocol and RDF Query Language (SPARQL) ontology queries using the ontology concepts and relationships previously identified [28]. These queries were executed and results checked for consistency with respect to the clinical evidence sources used to populate the ontology. Queries and results are shown in Table 2 for four competency questions

4.3 Development of Clinical Evidence Service

The evidence defined in the ontology has been made available to the TRANSFoRm CDSS through a REST based web service [29]. This allows the CDSS to access ontology resources through defined URL constructs that are linked to programmatically implemented SPARQL queries. The Sesame infrastructure provides a programmable API that can be used to programmatically connect to and query the ontology using SPARQL queries. The rest interface was developed using Java implementing the Jersey REST implementation [30].

System interoperability is supported by allowing query results to be returned to any third party consumer tool in a number of supported data formats including XML, JSON and plain text responses. In addition, the Sesame infrastructure also provides its own REST based interface that can be used directly to execute SPARQL queries to return responses in native RDF data formats. The components of the evidence service are shown in Figure 2.

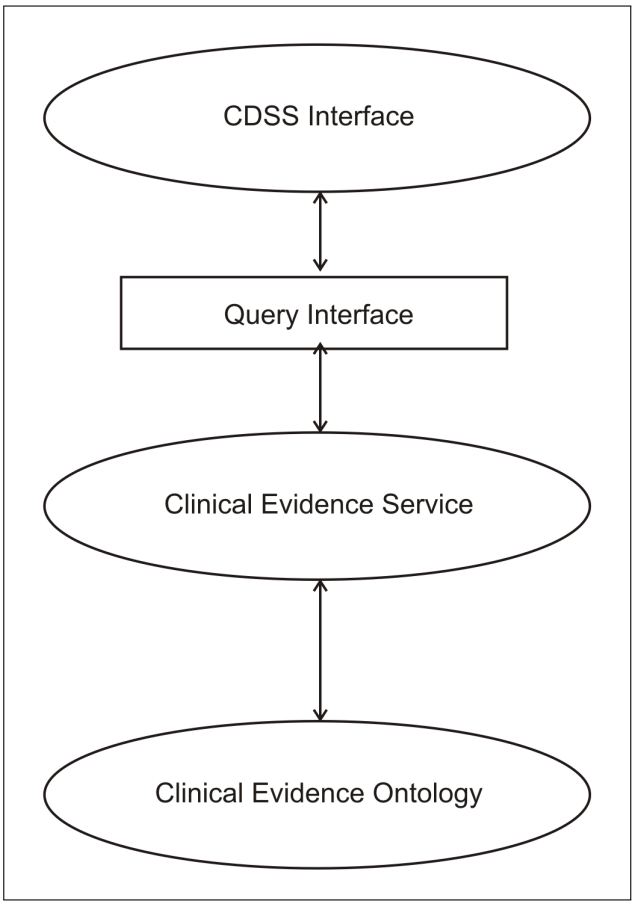

Figure 2: Evidence service components.

As an example, the structure of the Alvarado score can be accessed using the URI http://localhost:8080/ClinicalEvidenceRESTService /interfaces/query/cprs/AlvaradoScore1_0. A sample of the generated XML output is shown in Figure 3.

4.4 Implementation of a Diagnostic Strategy using Evidence Service Calls

We previously referred to the role of CPRs as part of a broader diagnostic strategy to "rule out" a potential diagnosis. The steps to be implemented would require:

- Obtaining the list of supported patient RFEs
- Obtaining a list of differential diagnoses to consider based on a presenting patient RFE
- Obtaining a list of the CPRs available for any differential diagnosis associated with the RFE
- Obtaining the cues, criteria and scores for any chosen CPR to apply
- Obtaining the scoring scheme and decisions for any chosen CPR
- Execution of the CPR based on a comparison to the patient cues provided to determine if a "rule out" may be appropriate

Using the example of a patient presenting with abdominal pain and an investigation of possible appendicitis, the clinical evidence service can be used to implement these steps using the following series of REST based calls to present results as XML, JSON, plain text or RDF data formats:
../ClinicalEvidenceRESTService/interfaces/query/rfes
../ClinicalEvidenceRESTService/interfaces/query /differentials/AbdominalPainLocalisedOtherRFE

Figure 3: XML generated from evidence service to provide criteria for the Alvarado Score rule.

../ClinicalEvidenceRESTService/interfaces/query
/cprs/Appendicitis
../ClinicalEvidenceRESTService/interfaces/query
/cprs/AlvaradoScore1_0
../ClinicalEvidenceRESTService/interfaces/query
/cprs/score/AlvaradoScore1_0

At present, the actual third party tool consumer would implement a web service client to provide the logic to compare appropriate EHR patient coded data RFEs and cues to the information returned by the service. Future work will parameterise the evidence service to allow submission of patient data as XML directly to the web service which will do the evidence comparison itself, returning a CPR result based on the patient data provided.

5 Discussion

A core requirement for the development of the CPR model was that it be a generalisable representation of the common structure of CPRs and not just suitable for the representation of specific examples of rules as found in literature. The efficacy of using CPRs as tools to be deployed in decision support systems has been shown to be effective but focussed on implementing specific instances of CPRs rather than supporting their more general usage through a service based knowledgebase [31]. We have used the model to represent 41 clinical prediction rules relating to 20 diagnoses including the Finnish Diabetes Risk Score [32], the Edwards Score [33] (tuberculosis) and the Little Symptom rule [34] (urinary tract infection).

In considering how this model relates to other initiatives to represent electronic clinical guidelines it is important to consider the original definition of a CPR previously provided. Each CPR is defined to be a discreet independent clinical tool to be used in its own right with respect to a particular patient. They do not attempt to define a complex clinical workflow or series of clinical steps to be implemented. From this point of view they are potentially useful tools to support decision-making in primary care where time pressures apply to consultations with each patient. As such, they could be considered to be either stand alone tools or are analogous to decision points found in more complex electronic guidelines that do define computable workflows, such as Guideline Interchange Format (GLIF) or the Guideline Elements Model (GEM) [35, 36].

The previous definition of a CPR also allows for an optional clinical decision or action to be taken based on the score outcome of the rule (sometimes then referred to as a Clinical Decision Rule). It was considered to be out of the scope of this work to represent these clinical decisions as computable entities in their own right and they have been treated as informational textual descriptions in the ontology e.g. "Surgery". It could be possible though to represent these decisions as separate concepts in their own right within the ontology e.g. CPRClinicalDecision. The workflow content of these clinical decisions could be modelled separately as GLIF or GEM based guidelines with an appropriate reference or link from our ontology concepts.

There are limitations to this work because the TRANSFoRm project as a whole is still a work in progress. The future development of the clinical decision support system that consumes our evidence service will be necessary to do a full clinical validation of the models that we propose here. What we have presented here is a conceptual validation of the ontology structure and the implementation in a way that supports system interoperability (through recognised data representation standards such as XML, JSON and RDF) along with semantic interoperability (through the use of the TRANSFoRm vocabulary service).

6 Conclusion

The research described in this paper can encourage the wider use and acceptance of clinical prediction rules by clinicians in three ways; by making CPRs more accessible and searchable than literature equivalents; through provision of a computable representation that allows for development of versioned rules from data mined sources of aggregated primary care data that are more sensitive to clinicians own patient populations; through provision of a web service allowing the deployment of CPRs as part of third party decision support tools linked to EHRs to facilitate easier use and execution.

Acknowledgements

TRANSFoRm is partially funded by the European Commission – DG INFSO (FP7 247787). This research was partly supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

This work was partially funded by the Health Research Board (HRB) of Ireland through the HRB Centre for Primary Care Research under Grant HRC/2007/1.

References

[1] Keogh C, Wallace E, O'Brien KK, Murphy PJ, Teljeur C, McGrath B, Smith SM, Doherty N, Dimitrov BD, Fahey T. Optimized retrieval of primary care clinical prediction rules from MEDLINE to establish a Web-based register. Journal of Clinical Epidemiology 2011;64(8):848-60.

[2] Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Annals of Internal Medicine 2006;144(3):201 .

[3] Wallace, E., Smith, S., Perera-Salazar, R., Vaucher, P., McCowan, C., Collins, G., Verbakel, J., Lakhanpaul, M. and Fahey, T. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). BMC Medical Informatics and Decision Making, 11, 1 2011), 62.

[4] Chunilal, S. D., Eikelboom, J. W., Attia, J. and et al. Does this patient have pulmonary embolism? JAMA, 290, 21 2003, 2849-2858.

[5] The Doctor's Toolbag [Internet]. British Medical Journal [cited 31 July 2013]. Available from http://group.bmj.com/products/mobile-apps/doctors-toolbag-iphone-app

[6] Lang ES, Wyer PC, Haynes RB. Knowledge Translation: Closing the Evidence-to-Practice Gap. Annals of Emergency Medicine 2007;49(3):355-363.

[7] Keogh C, Fahey T. Clinical prediction rules in primary care: what can be done to maximise their implementation? Clinical Evidence [Internet] 2010 [cited 31 July 2013]. Available from http://www.clinicalevidence.com/x/mce/file/05-10-10.pdf

[8] Roshanov P, Misra S, Gerstein H, Garg A, Sebaldt R, Mackay J, Weise-Kelly L, Navarro T, Wilczynski N, Haynes RB and others. Computerized clinical decision support systems for chronic disease management: A decision-maker-researcher partnership systematic review. Implementation Science 2011;6(1):92.

[9] TRANSFoRm Project Homepage [Internet]. TRANSFoRm project [cited 31 July 2013]. Available from http://www.transformproject.eu

[10] Alvarado A. A practical score for the early diagnosis of acute appendicitis. Annals of Emergency Medicine 1986;15(5):557-64.

[11] Ohle R, O'Reilly F, O'Brien KK, Fahey T, Dimitrov BD. The Alvarado score for predicting acute appendicitis: a systematic review. BMC Medicine 2011;9:139.

[12] McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature. JAMA: the journal of the American Medical Association 2000;284(1):79-84.

[13] Falk G, Fahey T. Clinical prediction rules. British Medical Journal 2009;339.

[14] Kostopoulou O. Diagnostic errors: Psychological theories and research Implications. In: Hurwitz B, Sheikh A, editors. Health Care Errors and Patient Safety: Wiley Online Library; 2009. p 95-111.

[15] Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care–a systematic review. Family Practice 2008;25(6):400-13.

[16] Wagner JM, McKinney WP, Carpenter JL. Does This Adult Patient Have Appendicitis? McGraw-Hill Medical, City, 2009. p 53-63.

[17] Fernandez-Lopez M, Gomez A. Overview and analysis of methodologies for building ontologies. The Knowledge Engineering Review 2002;17(02):129-156.

[18] Gruninger M, Fox MS. Methodology for the Design and Evaluation of Ontologies. In Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI (Vol. 95).

[19] TRANSFoRm Vocabulary Service [Internet]. University of Birmingham [cited 31 July 2013]. Available from http://www.eutransformvs.bham.ac.uk:8081/evs/

[20] Unified Medical Language System [Internet]. U.S. National Library of Medicine [cited 20 December 2013]. Available from http://www.nlm.nih.gov/research/umls

[21] International Classification of Primary Care Version 2 [Internet]. World Health Organisation [cited 20 December 2013]. Available from http://www.who.int/classifications/icd/adaptations/icpc2/en/

[22] SNOMED Clinical Terms [Internet]. U.S. International Health Terminology Standards Organisation [cited 20 December 2013]. Available from http://www.ihtsdo.org/snomed-ct/

[23] International Classification of Diseases Version 10 [Internet]. U.S. World Health Organisation [cited 20 December 2013]. Available from http://www.who.int/classifications/icd/en/

[24] Web Ontology Language [Internet]. World Wide Web Consortium [cited 31 July 2013]. http://www.w3.org/TR/owl2-overview/

[25] Resource Description Framework [Internet]. World Wide Web Consortium [cited 31 July 2013]. Available from http://www.w3.org/RDF/

[26] Protégé Homepage [Internet]. Stanford University [cited 31 July 2013]. Available from http://protege.stanford.edu/

[27] Sesame [Internet]. OpenRDF [cited 31 July 2013]. Available from http://www.openrdf.org/index.jsp

[28] SPARQL Standard [Internet]. World Wide Web Consortium [cited 31 July 2013]. Available from http://www.w3.org/TR/rdf-sparql-query/

[29] Fielding, R. T. and Taylor, R. N. Principled design of the modern Web architecture. ACM Transactions on Internet Technology (TOIT), 2, 2 2002), 115-150.

[30] Jersey Web Services in Java [Internet]. Oracle Corporation [cited 31 July 2013]. Available from https://jersey.java.net/

[31] McGinn, T.G., McCullagh, L., Kannry, J., Knaus, M., Sofianou, A., Wisnivesky, JP. Efficacy of an evidence-based clinical decision support in primary care practices: a randomized clinical trial. JAMA Internal Medicine 2013;173(17):1584-91

[32] Soriguer, F., Valdes, S., Tapia, M. J., Esteva, I., Ruiz de Adana, M. S., Almaraz, M. C., Morcillo, S., Garcia Fuentes, E., Rodriguez, F. and Rojo-Martinez, G. Validation of the FINDRISC (FINnish Diabetes RIsk SCore) for prediction of the risk of type 2 diabetes in a population of southern Spain. Pizarra Study]. Medina Clinica (Barc), 138, 9 (Apr 14 2012), 371-376.

[33] Narayan, S., Mahadevan, S. and Serane, V. T. Keith Edwards score for diagnosis of tuberculosis. Indian Journal of Pediatrics, 70, 6 (Jun 2003), 467-469.

[34] Bent, S., Nallamothu, B. K., Simel, D. L., Fihn, S. D. and Saint, S. Does this woman have an acute uncomplicated urinary tract infection? JAMA, 287, 20 (May 22-29 2002), 2701-2710.

[35] Wang, D., Peleg, M., Tu, S. W., Boxwala, A. A., Ogunyemi, O., Zeng, Q., ... Shortliffe, E. H. Design and implementation of the GLIF3 guideline execution engine. Journal of biomedical informatics 2004 37(5), 305-318.

[36] Guideline Elements Model [Internet]. Yale School of Medicine [cited 20 December 2013]. Available from http://gem.med.yale.edu/default.htm

Historical Mining Areas and Their Influence on Human Health

Stanislav Rapant[1,2], **Veronika Cvečková**[1], **Katarína Fajčíková**[1], **Milan Kohút**[1], **Darina Sedláková**[3]

[1] State Geological Institute of Dionyz Stur, Bratislava, Slovak Republic

[2] Goethe Uni Bratislava, Faculty of Tourism, Piešťany, Slovak Republic

[3] WHO Country Office in the Slovak Republic, Bratislava, Slovak Republic

Abstract

Aims: Impact of potentially toxic elements (PTE) on the health status of population of the Slovak Republic has been studied in two historical mining areas with ore extraction from Middle Ages (the Middle Slovak Neovolcanics, the Slovak Ore Mts.) and one historical mining area with more than hundred years brown coal mining (Upper Nitra region).

Methods: The contents of PTE were analysed in groundwater/ drinking water and soils. The health status of resident population was evaluated based on 43 health indicators classified according to the international classification of diseases (ICD, 10th revision), including mainly those indicators characterizing mortality on cardiovascular and oncological diseases. In these areas the health status of population living in municipalities with increased PTE contents (As, Pb, Zn, Cu, Cd, Hg and Sb) was compared with that in adjacent municipalities showing low PTE contents.

Results: A total of 138 contaminated and 155 non-contaminated municipalities of similar socioeconomic, natural and geochemical-geological character were compared. PTE contents in soils of polluted municipalities reported considerably increased levels – between 2 to 10 times higher in contrast to non-contaminated municipalities. On the other hand, PTE contents in groundwater were almost identical both in contaminated as well as non-contaminated areas and in majority of cases were below limit standard values for drinking water.

Conclusion: Based on the assessment of the health status of population (using 43 health indicators), no significant difference in the health status of population in contaminated and non-contaminated municipalities has been reported.

Keywords

historical mining areas, potentially toxic elements, health indicators, health status

Correspondence to:

Stanislav Rapant
State Geological Institute of Dionyz Stur,
Address: Mlynská dolina 1, 817 04 Bratislava, Slovak Republic
E–mail: stanislav.rapant@geology.sk

EJBI 2014; 10(1):24–34
received: May 13, 2014
accepted: September 30, 2014
published: October 31, 2014

1 Introduction

The connection between the geological environment and human health has been known since antiquity. Both excess and deficiency chemical elements in the environment can result in possible negative health effects. Most of the studies in medical geology and public health deal with the influence of potentially toxic elements (PTE) from geological materials (rocks, soil, water, stream sediments, ambient dust, etc.) in the environment on the health of local population in geogenically or anthropogenically contaminated areas. Research focused on the impact of increased arsenic contents, particularly in groundwater/drinking water, on human health is a typical example of these studies [1, 2]. Another important topic of concern is the historical mining sites throughout the world because of the potentially adverse impact on the local population of the mobilized geological materials containing PTEs [3, 4, 5, 6]. In these areas there is a higher probability of potential health risks with regard to considerably increased contents of various PTE particularly in soils and groundwater as a result of intense historical mining activities.

In Slovakia there are several historical mining sites as a heritage of long-term mining activities carried out in the past (exploitation and processing of Ag-Au-Sb ores, Pb-Zn-Cu ores, Hg ores, brown coal, etc.). The health status of population living in three historic mining areas of

the Slovak Republic, where higher PTE contents were observed in geological components of the environment, was compared with the health status of people living in adjacent areas of similar socioeconomic character and the same or similar geological structure with low or no PTE contamination. The main objective of this study was to assess how, and to what extent, the PTE contamination of the geological environment might influence the health status of residents living in the selected historical mining areas.

2 Material

The connection between the contaminated geological environment in the selected historical mining sites and the health status of the local population has been assessed based on the data from national databases of environmental and health indicators [7].

The environmental indicators are the contents of chemical elements, components or values of chemical parameters analysed and measured in the environment [7]. In this study we evaluate environmental indicators in groundwater and soils as these components of geological environment definitely show the most significant connection with human health [8]. In addition, it is probable that PTE contents in groundwater and soils can influence human health to a great extent.

In Slovakia, groundwater is used as the source of drinking water for more than 90% of Slovak inhabitants [9]. The soil is the base of food chain and represents that part of the environment, where human life directly takes place. Crops we eat are grown in soil and meat, eggs, and milk come from animals the life of which is integrally connected with the soil too. Moreover, children and also some adults are known to ingest soil as well. Thus, there are numerous ways that people can be exposed to the trace elements in soil. The soil and groundwater chemical contents are determined as "total contents" (way of digestion and sampling methods and chemical analysis [10, 11, 12].

The set of environmental indicators and their mean values for groundwater and soils in the Slovak Republic is summarized in Table 1 (according to [7]).

The total number of chemical analyses for groundwater was 20, 339 and for soils 10, 738. The data included analyses that have been collected since 1991, when the modern environmental-geochemical mapping of the Slovak Republic started under the IGCP 360 Geochemical Correlation Programme [13]. The density of groundwater samples was about one sample per 2.5 km^2 and of soil samples about one sample per 5 km^2.

The health indicators (the indicators health status and demographic development of population) are variables that can express the health status of people in society via direct measurement or observation [14]. We can say whether the assessed health is good or bad, only when a number of areas or time periods are subject to our evaluation. In addition, they must be compared to each other and to standard or published values for larger units within a sufficiently long period of time (i.e. decades).

There is no single comprehensive indicator which would capture all or majority of aspects of population health status. Therefore, a relatively large set of multiple indicators was used in the study.

With regard to sensitivity and, especially diversity of the data there is the need for a longer period of time in which health indicators are carefully monitored and evaluated. In our study we used a ten-year period (1994 - 2003) which still seems to be minimally sufficient particularly concerning small-sized or problematic municipalities (military districts). The data of health indicators were obtained from the database of the Statistical Office of the Slovak Republic [15]. We only used the data describing demography and mortality. The data evaluating the incidence of specific diseases are not available.

In order to assess the health status of population in contaminated or non-contaminated areas 43 health indicators were selected. It is expected that these indicators can be affected by the geological environment. The list of assessed health indicators with nationwide mean values is given in Table 5.

The selected health indicators describe relevant information on age, and particularly analyse mortality in many different ways. We have deliberately chosen only robust indicators that are stable, not rare, and do not alter suddenly. Out of the 43 health indicators subject to assessment four (1 to 4 in Table 5) are considered as positive, i.e. the most favourable values are the highest values. The remaining 39 health indicators are negative, i.e. their most favourable values should be as low as possible reaching even zero values.

3 Methods

3.1 Elaboration of Environmental Indicators

When elaborating and calculating the environmental indicators we adopted the method of geochemical data processing and such representation of environmental indicators so that they can be united with health indicators. Therefore, we had to transform environmental indicators into a form of health indicators, which represent one number for the assessed administrative unit of the Slovak Republic – a municipality or a district. Transformation of the geochemical data on chemical composition of soils and groundwater in the Slovak Republic was conducted in the same way. Thus, the environmental indicators were calculated for the basic territorial units of the Slovak Republic – municipalities (2,883 municipalities). Calculations of environmental indicators represented a determination of the mean value of an element or component for the evaluated territorial administration units (Slovak municipalities) based on the contents of all soil and water samples

Table 1: Characteristics of environmental indicators for the Slovak Republic (mean values).

GROUNDWATER (n=20 339)												
pH	**T.D.S.**	**COD_{Mn}**	**Ca+Mg**	**Li**	**Na**	**K**	**Ca**	**Mg**	**Sr**	**Fe**	**Mn**	**NH_4**
7.33	629.75	2.18	3.5	0.019	20.34	11.10	93.56	28.29	0.36	0.17	0.12	0.10
F	**Cl**	**SO_4**	**NO_2**	**NO_3**	**PO_4**	**HCO_3**	**SiO_2**	**Cr**	**Cu**	**Zn**	**As**	**Cd**
0.13	32.96	79.32	0.11	38.76	0.20	303.85	18.21	0.0013	0.0026	0.2673	0.0019	0.0010
Se	**Pb**	**Hg**	**Ba**	**Al**	**Sb**							
0.0010	0.0014	0.0001	0.0747	0.0297	0.0009		Note: Data except of pH in $mg.l^{-1}$, Ca+Mg in $mmol.l^{-1}$					
SOILS (n=10 738)												
Al	**As**	**B**	**Ba**	**Be**	**Bi**	**Ca**	**Cd**	**Ce**	**Co**	**Cr**	**Cu**	**F**
5.90	12.45	65.03	392.78	1.39	0.41	1.46	0.60	64.65	11.77	87.55	26.15	330.98
Fe	**Hg**	**K**	**Mg**	**Mn**	**Mo**	**Na**	**Ni**	**P**	**Pb**	**Sb**	**Se**	**Sn**
2.71	0.24	1.70	0.87	0.08	0.68	0.85	29.29	0.07	29.62	3.69	0.16	4.71
Sr	**V**	**W**	**Zn**	**pH_{H2O}**	**pH_{KCl}**	**carbonates**						
101.38	79.07	0.92	75.79	6.26	5.52	2.45	Note: macrocomponents in %, microcomponents in $mg.kg^{-1}$					

found in a related administration unit using the kriging method [7].

The data for municipalities located in the three areas with historical mining activities were selected from the nationwide geochemical data of environmental indicators for all municipalities in the Slovak Republic and then analysed (293 municipalities in total).

3.2 Elaboration of Health Indicators

All health indicators were calculated as a cumulative function for a period of years from 1994 to 2003, i.e. for a ten-year period, when all cases were summed and the numbers of inhabitants were taken as persons-per-years (number of inhabitants as of December 31 in a pertinent year) for each territorial unit (municipality) assessed.

Calculation methodology and standardization of health indicators was carried out according to recommendations of WHO [16, 17, 18, 19].

Selection of health indicators was based on the International Classification of Diseases by the WHO 10th revision [20]. Demographic indicators describing the age composition of municipalities express the average age of the population of the observed municipalities or areas. A percentage of elderly people over 60 years was calculated as 100 times the number of inhabitants aged 60 years and over per number of inhabitants. Indirect age-standardized mortality indicators were standardized to a Slovak standard (19 age groups). Relative mortality indicators are calculated as the number of deaths per 100,000 inhabitants (not involving the impact of the age of the inhabitants). Potential years of life lost are calculated as 100,000 times the sum of the years of people up to the age of nearly 65 years (deaths at age 1 to 64 years per number of inhabitants. Calculation methods of various health indicators or formulas used to calculate specific health indicators are given in Table 5. Subsequently, from the nationwide data of health indicators the data for municipalities located in the three selected areas with historical mining acitivities were selected and analysed.

3.3 Determination of Contaminated/Non-contaminated Areas

It is possible to determine the influence of PTEs from abandoned mining sites on human health only in three regions with historical mining activities, in the Slovak Republic – a relatively small country with the total land area of less than 50,000 km^2. The three selected regions are as follows – the Middle Slovak Neovolcanics, the Slovak Ore Mts. and the Upper Nitra region (Fig. 1). The first two regions represent the historical mining areas with ore extraction from the Middle Ages. Mining activities in these areas were completed at the end of the twentieth century. The third area is a territory characterized by exploitation of brown coal, which has been mined since 1909. Currently, brown coal extraction yields reach about 2 million tonnes per year. This coal is primarily used for domestic heating in the region and as a source of fuel in a local power plant. The coal is characterized an extremely high content of arsenic (about 800 ppm by weight) and sulphur (about 2% by weight).

Determination of contaminated and adjacent non-contaminated areas in individual regions mentioned above resulted from a basic criteria of a minimum number of 15 municipalities in each of them.

The definition of contaminated and non-contaminated areas was based on the limit values for the assessment of soil pollution and drinking water quality valid in the Slovak Republic (Table 2). Contaminated or non-contaminated municipalities were selected based on the PTE contents in soils, since the contents of health risk elements in soils are of higher variability than in groundwater. Due to relatively high pH of water reaching almost neutral pH values of (as a result of abundant presence of carbonates in the ore veins) PTE mobility is seen as relatively very low. Under these conditions PTEs are removed from the groundwater and bound in soils and sediments.

Table 2: Assessed elements and their limit values.

Soils („A“ reference values of MP SR resolution No. 531/1994-540)															
Element	**As**	**Ba**	**Be**	**Cd**	**Co**	**Cr**	**Cu**	**Hg**	**Mo**	**Ni**	**Pb**	**Se**	**Sn**	**V**	**Zn**
Limit [mg.kg^{-1}]	29	500	3	0.8	20	130	36	0.3	1	35	85	0.8	20	120	140
Groundwater (limit values of the Slovak government order No. 496/2010 of Collection of Laws) – drinking water															
Element	**TDS**	**NO_3**	**Cl**	**SO_4**	**F**	**NH_4**	**Na**	**Fe**	**Mn**	**Al**	**As**	**Cd**	**Cr**	**Cu**	**Hg**
Limit [mg.l^{-1}]	1000	50	100	250	1.5	0.5	200	0.2	0.05	0.2	0.01	0.003	0.05	1.0	0.001
Element	**Pb**	**Sb**	**Zn**												
Limit [mg.l^{-1}]	0.01	0.005	3.0												

Non-contaminated municipalities include those with PTE contents in soils not exceeding reference values for any of the assessed elements.

The determination of contaminated and non-contaminated areas in the three regions with historical mining activities on the basis of the above criteria is shown in Fig. 1. A total of 138 contaminated and 155 non-contaminated municipalities were selected and then compared within individual assessed regions by the concentration level of chemical elements/ compounds in groundwater and soils (environmental indicators) and health status of the population (health indicators).

4 Area Description

4.1 Geological Setting

The geologic structure of the evaluated areas is composed of various geological-tectonic units and thus is represented by rocks of varied petrographic and geochemical character [21]. While the geological environment of the Middle Slovak Neovolcanics consists predominantly of Neogene volcanics, geological structure in the other two areas the geology is more complicated and contains rocks of diverse geological character.

The eastern part of the Slovak Ore Mts. area consists mainly of Lower-Paleozoic (Cambrian to Carboniferous) weakly metamorphosed flysch metasediments (metasandstones, metagreywackes, phyllites) and metavolcanics – basaltoid, keratophyre and rhyolite character. The western part consists dominantly of Lower to Upper Paleozoic metamorphic rocks of crystalline basement with signs of migmatization, granitization, especially orthogneisses, paragneisses, migmatites, amphibolites, diorites and metacarbonates. Approximately 5 % of the area is represented by the Mesozoic (carbonatic) sedimentary cover, which consists of the Lower-Triassic quartzite, dolomite and limestone. Mainly metasomatic and ore mineralizations of Fe, Cu, Pb, Zn, Sb, Ag, Au and Hg have been mined in the Slovak Ore Mts. since the Middle Ages.

The Middle Slovak Neovolcanics are predominantly (over 95 %) consists of Neogene volcanics, particularly andesites, basalts (less by rhyolites and dacites), and their pyroclastics. Crystalline rocks (orthogneisses, gneisses, granites) together with the Mesozoic carbonatic rocks having manifestations of scarnization are locally found in the

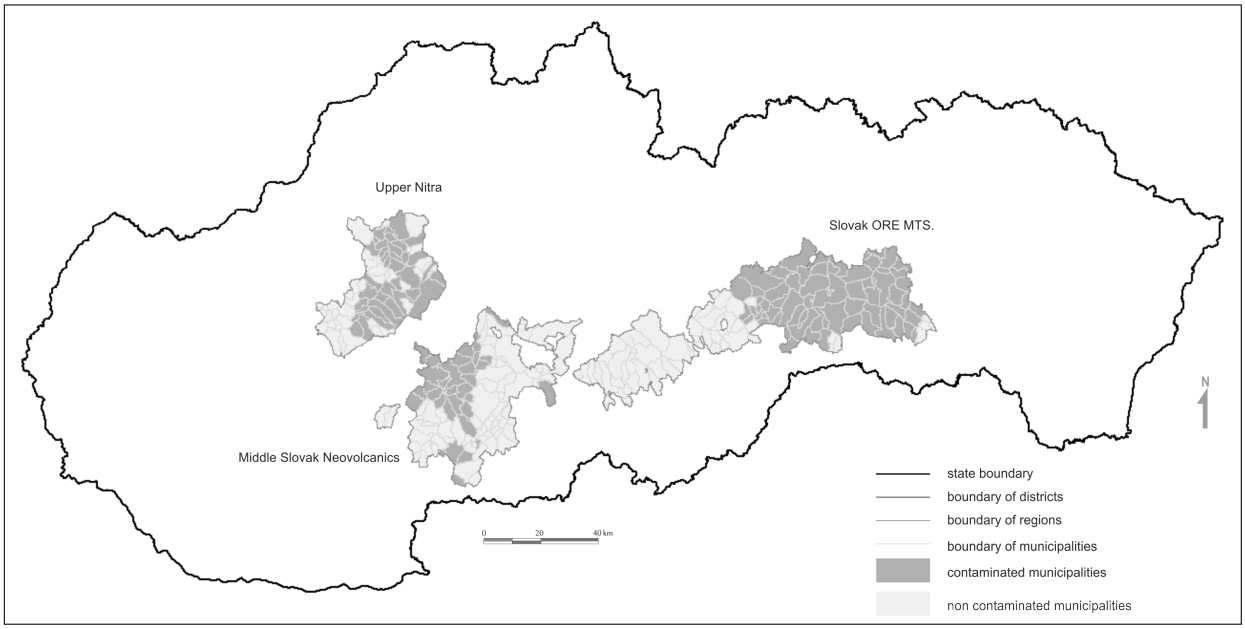

Figure 1: Contaminated and non-contaminated areas of the Slovak Republic.

Table 3: Unemployment rates in assessed areas in 2001 and 2011.

Region	unemployment rate in %			
	Contaminated area		Non contaminated area	
	2001	*2011*	*2001*	*2011*
Upper Nitra	19.09	14.81	19.14	15.20
Slovak Ore Mts.	27.32	25.65	32.20	25.78
Middle Slovak Neovolcanics	25.90	23.19	24.21	24.86
Slovak Republic	*2001*		*2011*	
	19.2		13.6	

Source: www. statistics.sk

form of xenolithes. In the past the area of the Middle Slovak Neovolcanics was well-known as an important metallogenetic region with exploitation of Au, Ag, Pb, Zn, Cu and Hg ores. Nowadays there is only a limited Au ore mining in the area.

The centre of the Upper Nitra region is represented by the Upper Nitra basin-shaped valley, typical intramontane the Tertiary depression of the Western Carpathians, which is surrounded by core and volcanic mountains. The basin area consists mainly of Paleogene nummulite sandy limestones and polymict and dolomitic breccias and conglomerates gradually passing into sandstones, siltstones and claystones. It is overlain by flysch sedimentation where mainly sandstones alternate with claystones and siltstones. The Neogene rocks are represented especially by the Eggenburgian sandstones and conglomerates, clays and Badenian volcaniclastics with coal seams being overlaying by basaltic andesites. These predominantly Tertiary sedimentary units constitute between 40 to 45 % of the studied area. The Mesozoic (mostly carbonates) complexes in surrounding core areas consist mainly of a number of limestones and dolomites with andstones, shales and quartzites covering about 20 % of the investigated area. The crystalline marginal core areas are primarily composed of acidic granitoid rocks with about 20% of the area covered by migmatites and gneisses. About 20% of the studied area, particularly in the Eastern border, consists of neovolcanic rocks – andesites, basalts and pyroclastics.

In the Upper Nitra region (Prievidza, Handlová, Nováky) brown coal and lignite exploitation has been carried out for more than 100 years.

4.2 Socioeconomic Characteristic

A number of studies analyzing the prevalence of health determinants (especially lifestyle risk factors, as well as poverty, education, employment, ethnicity and housing) in the selected districts has been carried out in the Slovak Republic in recent years [22, 23]. However, there are no consistent data evaluating health risk factors conditioned by non-optimal lifestyle of the local residents in the studied regions. The epidemiological studies indicate some differences in lifestyle within the districts of the Slovak Republic. Nevertheless, it is difficult to predict these differences when considering adjacent municipalities or those of a similar character in individual areas subject to assessment (rural population, mostly mountainous regions, about the same socio-economic level of the population, a similar lifestyle).

Thus, the lifestyle of people living in contaminated and non-contaminated areas is similar. It seems that comparison of unemployment rates may be the most accurate method of assessing the economic situation of people living in contaminated and non-contaminated areas (Table 3). According to Table 3, displaying unemployment rate figures in contaminated and non-contaminated areas for the three regions, it is evident that the unemployment rate is about the same comparing data for 2001 and 2011. Moreover, the unemployment rate in the Upper Nitra and the Slovak Ore Mts. regions is slightly higher in the non-contaminated municipalities. Based on this, it is clear that economic level in the assessed areas is practically the same, and apparently accounts for no significant impact on the health of the population in the studied areas.

5 Results and Discussion

Basic characteristic of the selected chemical elements in groundwater and soils (environmental indicators) of the studied areas is given in Table 4. The characteristics of population health status in the evaluated contaminated and non-contaminated areas are presented in Table 6.

Health status of the population, according to the WHO general declaration, is caused mainly by the four factors as follows: lifestyle (way of life and work) accounts for about 50 % of all factors; genetic factors and the level of health care is attributed to a 10-20 % share; and environment (particularly its geological component) represents about a 20 % share. If we assume that the impact of the first three factors in contaminated and non-contaminated areas of the Slovak Republic are about the same, the decisive influence should be put down to geological factors and different contamination levels of geological environment by PTE.

In all three regions PTE levels in soils are usually significantly higher in contaminated areas than in non-contaminated areas. The only exception is Cd content in the Upper Nitra region, which is slightly higher in the unpolluted area. This region, however, is especially con-

taminated by As as a result of brown coal combustion (by atmospheric deposition).With regard to the predominant character of polymetallic mineralization in the Middle Slovak Neovolcanics the most sizeable differences in Cd, Pb, Zn and Cu contents are evident in soils. In the Slovak Ore Mts. (with predominant polymetallic ores and Au-Sb ores) the most significant differences are shown in contents of Sb, Hg, Cu, Pb and As in soils.

In case of groundwater we have observed a considerable difference between contaminated and non-contaminated areas only in case of As in the Upper Nitra region. The contents of other PTE in groundwater in contaminated and non-contaminated areas of the assessed regions are, in general, very similar. There is a clear relation to low PTE mobility in groundwater in the given hypergenic conditions mentioned above. In addition, the fact that in geochemical mapping and water sampling we tried to avoid water sources such as discharge from the drainage tunnels, tailings ponds, etc., can also play an important role. We haven´t collected soil samples from waste tips, tailings ponds, remnants of ore treatment plants and other extremely contaminated soil sources, too.

Based on the results of health indicators (Table 6) it is evident that no significant differences between health indicators in contaminated and non-contaminated areas were observed in any of the three regions. According to the summary health indicator (sum_neg), the population health status in the Middle Slovak Neovolcanics and Upper Nitra regions are almost exactly the same (between 13,670 and 13,137; between 9,431 and 9,461). Moreover, in the Slovak Ore Mts. the health of the population in a contaminated area (11,679) shows even more favourable figures than in the non-contaminated one (13,012). A very similar situation is described by individual health indicators as well. With regard to the objective of our study, the age indicators (the first four indicators) seem to be of minor importance. They tend to be distorted by population migration, especially moving of young people to big cities because of better job opportunities. They are alike in all three regions for contaminated as well as non-contaminated areas.

The only noticeable difference in these demographic indicators is that the life expectancy in the Upper Nitra region is about 2-3 years longer than in the other two regions. However, it results from the fact that the geological environment of carboniferous strata prevails in this region and is more favourable to population health than the silicate rock environment (volcanic rocks, granites, metamorphic rocks). In terms of the influence of the geochemical background of the rock environment on the health of the Slovak population it has been proven that rock units of volcanics, granites and metamorphic rocks are much less favourable to human health than carboniferous rocks (limestones, dolomites, flysch sediments) Rapant et al. [7].

This stems particularly from the deficient content of Ca and Mg in groundwater/ drinking water in silicate geological units [7]. The average contents of Ca, Mg, water hardness and carboniferous composition of soils are significantly higher in the geological environment of the Upper Nitra region than in the other two regions. Moreover, a similar trend – less favourable values of health indicators in the areas of the Slovak Ore Mts. and Middle Slovak Neovolcanics compared to those in the Upper Nitra region – has also been reported in the rest of health indicators. It is reflected in a sizeable difference in the sum of negative health indicators (sum_neg), which is by about 2,000 to 3,000 more favourable in the Upper Nitra region than in the other two regions subject to our assessment.

Increased PTE contents are in the world literature associated mainly with cancer [24, 25, 26]. Neither basic indicators of oncological diseases (ReC, SMRC, PYLLC) nor any other specific cancer mortalities according to individual diagnoses (Table 6) show less favourable figures in the areas contaminated with PTE. Mortality due to cardiovascular disease has a similar trend. We have observed increased mortality in PTE contaminated areas only in the case of endocrine system diseases (ReE). Moreover, differences in the mortality become even more evident when taking into account the age of population (SMRE) reaching about 20 % figures in the silicate regions (the Slovak Ore Mts. and Middle Slovak Neovolcanics) but over 30 % in the Upper Nitra region. Adverse effects of PTE on mortality due to diseases of endocrine glands (especially diabetes, thyroid, diseases caused by malnutrition or excessive diet) have been described several times in the world literature [27, 28]. This issue will be addressed in the next stages of our research using higher statistics, especially neuron networks. The other group indicators (ReK - digestive system, ReJ – respiratory system, ReN - genitourinary system) show no noticeable differences between contaminated and non-contaminated areas.

The contents of macroelements demonstrate more significant influence on the health status of the Slovak population than those of PTE. It is evident that especially deficiency of Ca and Mg in silicate rock environment in groundwater/drinking water accounts for increased incidence of cardiovascular and cancer diseases [7]. In addition, the detrimental impact of Ca and Mg deficiency in groundwater/drinking water on cardiovascular diseases has been described many times so far [29, 30, 31]. In the world literature deficient contents of Ca and Mg in groundwater/drinking water have also been associated with an increased occurrence of cancer [32, 33, 34]. Since both diseases are the ultimate cause of deaths in Slovakia, they are most markedly manifested in all the health indicators as well. Therefore, the observed differences in the health status of population in contaminated and non-contaminated areas are mainly linked with different contents of Ca and Mg. For instance, Ca and Mg contents in the Slovak Ore Mts. are significantly higher in contaminated than non-contaminated area, and this is probably the reason for better health of the population living in PTE contaminated area.

Chemical elements in groundwater/drinking water occur mainly in dissolved form, which is the most avail-

able form to human beings. Therefore, it is the reason why groundwater/drinking water has probably the major influence on population health, considerably much more than the soils. Regarding the PTE contents in groundwater/drinking water in the evaluated areas, they are predominantly low and about the same even in contaminated and non-contaminated areas reaching levels below the limits of drinking water standards.

The results obtained from the comparison of health indicators between contaminated and non-contaminated areas are surprising and contrary to the current assumptions. In general, poorer health status is predicted in areas contaminated with PTE. Nevertheless, our results suggest that the health status of populations in both, contaminated and non-contaminated areas is at the same level or slightly better in contaminated areas. We explain this phenomenon as follows: the bio-available proportions of PTE in soils in these areas are very low, often well below 5 % [6, 35, 36]. Thus, only a small portion of the PTE enters the food chain. The PTE contents in groundwater/drinking water are relatively low as well due to neutral to alkaline environment in the area. Even if the local inhabitants use the local groundwater for drinking purposes, there is no increased intake of PTE doses, which would affect their health status.

The increased PTE levels in vegetables grown locally (carrots, potatoes, parsley) were documented in contaminated areas in all evaluated regions showing almost twice as high values as in non-contaminated areas. However, in terms of overall PTE ingestion from food, this proportion of contaminated vegetables is practically very small and almost insignificant [6, 35, 37, 38]. Moreover, biomonitoring results (including hair, nails, blood and urine) in the areas subject to our research also manifest slightly increased PTE levels in human materials in contaminated areas when compared to non-contaminated areas [35, 37, 38]. These levels, however, are in the vast majority under the set limit values for unpolluted environment and are rarely exceeded only in some municipalities [39].

Thus, it is evident that PTE enter the local food chain and their contents show increased values also in human biological materials. Nevertheless, these contents are probably not high enough to be reflected in the health status of the population. It seems that various adaptive mechanisms can apply in the population living in the contaminated areas. Even in these contaminated areas it is the overall geochemical background, mainly the contents of macroelements, which might crucially affect the health status of the population (probably much more than the PTE contents).

In terms of PTE contents and their impact on the population health status we consider low PTE levels in groundwater as the most relevant fact (Table 4). Groundwater, which is routinely used for drinking purposes in the observed areas, has relatively low PTE levels and is mostly below the drinking water limits.

Thus, we can conclude that the historical areas with recorded PTE contamination in soils or sediments, but not in groundwater or surface waters used for drinking purposes, represent much lower risk to the health of local people than has been thought about recently.

6 Conclusion

The main aim of this study was to objectively assess the potential impact of PTE on human health in historical mining areas. The health status of population in municipalities situated in contaminated and adjacent non-contaminated areas was compared across three studied regions of the Slovak Republic. Contamination of the studied areas has been documented mainly in soils, while the contents of PTE in groundwater/drinking water were approximately the same and below the limits of drinking water standards.

We found no significant impairment in the health of the population living in the areas with higher PTE contamination compared to non-contaminated areas. Surprisingly, no significant differences between the health status of population living in contaminated areas and that living in non-contaminated areas were observed. Finally, we can conclude that if groundwater/drinking waters used for drinking purposes show no PTE contamination, the local population inhabiting these historical mining areas might be at much lower risk than has been, in general, reported so far.

Acknowledgements

This research has been performed within the project LIFE10 ENV/SK/000086" The impact of geological environment on health status of residents of the Slovak Republic", which is financially supported by the EU's funding instrument for the environment: Life+ programme.

References

[1] Smedley PL, Kinniburgh, DG. A review of the source, behaviour and distribution of arsenic in natural waters. Applied Geochemistry 2002; 17: 517–568.

[2] Duker AA, Carranza EJM, Hale M. Arsenic geochemistry and health. Environment International 2005; 31: 631–641.

[3] Wcisło E, Ioven D, Kucharski R, Szdzuj J. Human health risk assessment case study: an abandoned metal smelter site in Poland. Chemosphere 2002; 47(5): 507–515.

[4] Peplow D, Edmonds, R. Health risks associated with contamination of groundwater by abandoned mines near Twisp in Okanogan County, Washington, USA. Environmental Geochemistry and Health 2004; 26(1): 69–79.

[5] Lim HS, Lee JS, Chon HT, Sager M. Heavy metal contamination and health risk assessment in the vicinity of the abandoned Songcheon Au–Ag mine in Korea. Journal of Geochemical Exploration 2008; 96(2-3): 223–230.

[6] Rapant S, Cvečková V, Dietzová Z, Letkovičová M, Khun M. Medical geochemistry research in SGR Mts. In: Environmental Geochemistry and Health 2009; 31(1): 11–25.

[7] Rapant S, Cvečková V, Dietzová Z, Fajčíková K, Hiller E, Finkelman RB, Škultétyová S. The impact of geological environment on health status of residents of the Slovak Republic. Environmental Geochemisty and Health 2014; 36(3): 543–561.

[8] Rapant S, Letkovičová M, Cvečková V, Fajčíková K, Galbavý J, Letkovič M. Environmental and health indicators of the Slovak Republic. Monograph, State Geological Institute of Dionyz Stur, Bratislava; 2010; 279. (in Slovak). Available from http://www.geology.sk/?pg=geois.ms_ezi_en

[9] Klinda J, Lieskovská Z. State of the Environment report of the Slovak Republic. Bratislava, Ministry of Environment of the Slovak Republic; 2010; 192.

[10] Rapant S, Vrana K, Bodiš D. Geochemical Atlas of Slovakia-part I. Groundwater. Monography, Ministry of the Environment of the Slovak Republic, Geological Survey of Slovak Republic, Bratislava; 1996; 127.

[11] Vrana K, Rapant S, Bodiš D, Marsina K, Lexa J, Pramuka S, Maňkovská B, Čurlík J, Šefčík P, Vojtaš J, Daniel J, Lučivian-sky L. Geochemical Atlas of Slovak Republic at a scale 1 : 1 000 000. Jurnal of Geochem. Exploration 1997; 60: 7–37.

[12] Čurlík J, Šefčík P. Geochemical Atlas of Slovakia-part V. Soils. Monography, Ministry of the Environment of the Slovak Republic, Geological Survey of Slovak Republic, Bratislava; 1999; 98.

[13] Darnley AG, Bjorklund A. et al. A Global Geochemical Database for Environmental and Resource Management. Earth Sciences. 19, UNESCO, Paris; 1995.

[14] Last JM. A Dictionary of epidemiology, Oxford University Press; 2001; 218.

[15] www.statistics.sk

[16] Beaglehole R, Bonita R, Kjellstrom T. Basic Epidemiology. Geneva: WHO; 1993; 682.

[17] Jeníček M. Epidemiology, The Logic of Modern Medicine. Epimed Montreal; 1995; 624.

[18] Bencko V, Hrach K, Malý M, Pikhart H, Reissigová J, Svačina Š, Tomečková M, Zvárová J. Biomedicínska statistika III., Statistické metody v epidemiologii. (1), Nakladatelství Karolinum, Praha; 2003; 236.

[19] Bencko V, Hrach K, Malý M, Pikhart H, Reissigová J, Svačina Š, Tomečková M, Zvárová J. Biomedicínska statistika III., Statistické metody v epidemiologii, (2), Nakladatelství Karolinum, Praha; 2003; 269.

[20] www.who.int/classifications/icd/en/

[21] Marsina K (ed.), Bodiš D, Havrila M, Janák M, Káčer Š, Kohút M, Lexa J, Rapant S, Vozárová A. Geochemický atlas Slovenskej republiky – Horniny. Monography, GSSR Bratislava; 1999; 134.

[22] Vilinová K. Zdravotný stav obyvateľstva Slovenska. UKF v Nitre, Edícia prírodovedec 2012; (495): 125.

[23] Michálek A, Podolák P. Selected deteminants of regional diferentation of life expectancy at birth in Slovakia. Geografický časopis 2007; 59(4): 305–322.

[24] Bako G, Smith ES, Hanson J. et al. The geographical distribution of high cadmium concentrations in the environment and prostate cancer in Alberta. Can J Public Health 1982; 73: 92–94.

[25] Fryzek JP, Mumma MT, McLaughlin JK et al. Cancer mortality in relation to environmental chromium exposure. J Occup Environ Med 2001; 43(7): 635–640.

[26] Cabrera HN, Gómez ML. Skin cancer induced by arsenic in the water. J. Cutan. Med. Surg. 2003: 106–111.

[27] Lai MS, Hsueh YM, Chen CJ, Shyu MP, Chen SY, Kuo, TL, Wu MM, Tai TY. Ingested inorganic arsenic and prevalence of diabetes mellitus. Am J Epidemiol 1994; 139: 484–492.

[28] Gupta SK, Khan TI, Gupta RC. Compensatory hyperparathyroidism following high fluorine ingestions-a clinico-biochemical correlation. Indian Pediatr 2001; 38: 139–146.

[29] Shaper AG, Packham RF, Pocock SJ. The British regional Heart Study: Cardiovascular Mortality and Water Quality. J. Environ. Pathol. Toxicol. 1980; 3: 89–111.

[30] Rylander R, Bonevik H, Rubenowitz E. Magnesium and Calcium in Drinking Water and Cardiovacular Mortality. Scand. J. Work Environ. Health 1991; 17: 91–94.

[31] Selinus O, Alloway BJ, Centeno JA, Finkelman RB, Fuge R, Lindh U, Smedley P. Essentials of Medical geology, Impacts of the natural environment on public health. Elsevier Academic; 2005; 793.

[32] Yang ChY. Pancreatic Cancer Mortality and Total Hardness Levels in Taiwan's Drinking Water. Journal of Toxicology and Environmental Health, Part A: Current Issues; 1999; 56(5): 361–369.

[33] Yang ChY, Chiu HF, Cheng BH, Hsu TY, Cheng MF, Wu TN. Calcium and Magnesium inDrinking Water and Risk of Death from Breast Cancer. Journal of Toxicology and Environmental Health, Part A: Current Issues; 2000; 60(4): 231–241.

[34] Chiu HF, Chang ChCh, Yang ChY. Magnesium and calcium in drinking water and risk of death from ovarian cancer. Magnesium Research 2004; 17(1): 28–34.

[35] Krčmová K, Rapant S. Environmental exposure to arsenic and associated health risk for residents in Horná Nitra region: A geochemical and medical research. Mineralia Slovaca 2007; 39: 75–80.

[36] Vaculík M, Jurkovič Ľ, Matejkovič P, Molnárová M, Lux A. Potential Risk of Arsenic and Antimony Accumulation by Medicinal Plants Naturally Growing on Old Mining Sites. Water, Air and Soil Pollution 2013; 224(5): 1546, DOI 10.10007/s11270-013-1546-9.

[37] Krčmová K, Rapant S. Trace elements in local food chain of residents in selected regions of Slovakia: Soil contamination and health implications. In: Mihály Szilágy, Klára Szentmihály (Eds.), 2009: Trace elements in the Food chain. Vol. 3 Deficiency of Excess of Trace Elements in the Environment as a Risk of Health. Working Committee on Trace Elements of the Complex Committee Hungarian Academy of Sciences (HAS), Institute of Materials and Environmental Chemistry of the HAS, Budapest, Hungary. 83–87.

[38] Rapant S, Letkovičová M, Cvečková V, Fajčíková K, Nikodémová D. Zhodnotenie potenciálneho vplyvu geochemického prostredia na zdravotný stav obyvateľstva banskoštiavnickej oblasti, regionálny geologický výskum. Záverečná správa, Geofond, Bratislava, Manuskript; 2010; 193. (in Slovak)

[39] Rapant S, Dietzová Z, Cicmanová S. Environmental and health risk assessment in abandoned mining area, Zlatá Idka, Slovakia. Environmental Geology 2006; 51: 387–397.

Table 4: Selected values of environmental indicators in contaminated and non-contaminated areas of the Slovak Republic (mean values for all municipalities).

	MIDDLE SLOVAK NEOVOLCANICS		UPPER NITRA		SLOVAK ORE MTS.	
	Contaminated area	Non contaminated area	Contaminated area	Non contaminated area	Contaminated area	Non contaminated area
Soils						
As	11.03	7.06	32.38	16.90	96.68	13.14
Cd	3.34	0.60	0.24	0.34	0.79	0.31
Cu	35.67	19.18	19.15	17.91	139.89	22.68
Hg	0.16	0.08	0.15	0.10	3.03	0.18
Pb	91.42	29.63	37.65	29.95	118.34	26.26
Sb	2.96	1.53	1.23	0.97	76.79	2.36
Zn	134.14	78.40	88.32	72.75	89.81	74.59
Ca	1.14	0.96	1.47	1.55	0.65	0.91
Mg	0.73	0.59	0.95	0.91	0.69	0.84
carbonates	0.86	1.21	1.74	2.14	0.62	0.22
Groundwater						
As	0.00194	0.00160	0.02096	0.00194	0.01217	0.00165
Cd	0.00139	0.00286	0.00444	0.00818	0.00054	0.00205
Cu	0.00263	0.00239	0.00129	0.00169	0.00413	0.00112
Hg	0.00014	0.00012	0.00015	0.00014	0.00016	0.00013
Pb	0.00198	0.00106	0.00107	0.00193	0.00163	0.00104
Sb	0.00024	0.00021	0.00019	0.00023	0.00941	0.00048
Zn	0.17592	0.25344	0.20046	0.15462	0.12486	0.12066
Ca	43.87	48.98	63.32	93.82	38.33	33.02
Mg	11.75	13.25	18.65	25.72	14.09	9.88
Ca+Mg	1.58	1.77	2.34	3.40	1.54	1.23

Note: contents of elements for groundwater in $mg.l^{-1}$, Ca+Mg v $mmol.l^{-1}$, for soils in $mg.kg^{-1}$, Ca, Mg in %

Table 5: Evaluated health indicators of the Slovak Republic.

No.	Indicator	Description of indicator	Method of calculation	Unit	Mean SR
Demographic indicators describing age structure of municipalities					
1	LEp	life expectancy at birth – population	cumulative calculation of all years of life during lifetime / No. of living persons at the beginning of the year	years	72.60
2	LEm	life expectancy at birth – men			67.44
3	LEw	life expectancy at birth – women			77.07
4	A60+	proportion of population at age 60 and more	100 x (number of people aged 60 and over / number of inhabitants)	%	15.38
Crude mortality, premature					
5	SMRp	population	indirect age-standardized mortality rate of inhabitants to the Slovak standard (19 age groups)	%	100
6	SMRm	men			100
7	SMRw	women			100
8	PYLL100	potential years of lost life	100, 000 x [the sum of the years of people up to the age of nearly 65 years (deaths at age between 1 to 64 years) / number of inhabitants]	years	4033.0
Relative mortality for selected cause of death					
9	ReC00-C97	malignant neoplasms	100 000 x [No. of deaths for selected cause / number of inhabitants]	No. of deaths per 100 000 inhabitants	212.79
10	ReC15-C26	malignant neoplasms of gastrointestinal system			76.14
11	ReC16	malignant neoplasms of stomach			15.20
12	ReC18-C20	malignant neoplasms of colon and rectum			24.24
13	ReC30-C39	malignant neoplasms of respiratory system			45.19
14	ReC50	malignant neoplasms of breast			24.80
15	ReC64-C68	malignant neoplasms of urinary system			11.25
16	ReC81-C96	malignant neoplasms of orgnas for haematopoiesis			13.28
17	ReC91-C95	all leukemia			6.20
18	ReC00-D48	all neoplasms			213.62
19	ReE00-E99	endocrine, nutritional and metabolic diseases			14.38
20	ReI00-I99	diseases of the circulatory system			531.05
21	ReI21-I25	ischaemic heart disease			269.82
22	ReI63-I64	cerebral infarction and strokes			63.57
23	ReJ00-J99	diseases of respiratory system			58.08
24	ReK00-K93	diseases of the digestive system			45.83
25	ReN00-N99	diseases of urinary and reproductive system			13.69
Standardized mortality for selected cause of death					
26	SMRC00-C97	malignant neoplasms	indirect age-standardized mortality rate of inhabitants to the Slovak standard (19 age groups)	%	100
27	SMRC15-C26	malignant neoplasms of gastrointestinal system			100
28	SMRC30-C39	malignant neoplasms of respiratory system			100
29	SMRC81-C96	malignant neoplasms of organs for haematopoiesis			100
30	SMRE00-E99	endocrine, nutritional and metabolic diseases			100
31	SMRI00-I99	diseases of the circulatory system			100
32	SMRI21-I25	ischaemic heart disease			100
33	SMRI63-I64	cerebral infarction and strokes			100
34	SMRJ00-J99	diseases of respiratory system			100
35	SMRK00-K93	diseases of the digestive system			100
36	SMRN00-N99	diseases of urinary and reproductive system			100
Potential years of lost life for selected cause of death					
37	PYLLC00-C97	malignant neoplasms	100, 000 x [the sum of the years of people up to the age of nearly 65 years (deaths at age between 1 to 64 years) / number of inhabitants]	years	1005.20
38	PYLLC15-C26	malignant neoplasms of gastrointestinal system			242.26
39	PYLLC30-C39	malignant neoplasms of respiratory system			186.2
40	PYLLI00-I99	diseases of the circulatory system			866.19
41	PYLLI21-I25	ischaemic heart disease			396.32
42	PYLLJ00-J99	diseases of respiratory system			172.69
43	PYLLK00-K93	diseases of the digestive system			334.80

Table 6: Characteristics of population health status in contaminated and non-contaminated areas (data recalculated according to number of inhabitants in respective municipalities).

	MIDDLE SLOVAK NEOVOLCANICS		SLOVAK ORE MTS.		UPPER NITRA	
	1*	2*	1*	2*	1*	2*
LEp	71.10	70.99	71.12	71.53	73.55	73.45
LEm	65.78	66.10	66.49	66.99	69.75	69.62
LEw	75.96	75.65	72.88	74.95	77.06	77.13
A60+	18.16	17.89	15.31	16.91	17.87	17.99
SMRp	112.40	112.15	112.25	110.32	94.98	94.38
SMRm	122.67	117.58	115.75	111.37	94.07	91.57
SMRw	105.94	107.21	110.60	109.88	94.74	96.03
PYLL100	5244.41	5049.83	4527.48	4985.29	3485.95	3504.16
ReC	252.60	240.31	211.78	229.03	223.96	238.11
ReC1526	85.26	96.23	70.73	72.94	77.28	94.21
ReC16	14.24	20.72	14.30	15.34	22.60	20.90
ReC1820	27.41	32.32	24.46	20.60	23.02	28.49
ReC3039	55.44	46.94	45.58	51.67	50.09	43.19
ReC50	21.46	29.31	24.51	33.53	23.87	24.96
ReC6468	16.07	8.46	12.02	13.40	9.60	10.98
ReC8196	14.07	13.98	12.75	15.26	11.66	12.74
ReC9195	6.05	8.11	6.13	6.74	4.78	5.21
ReC00D48	241.58	242.61	212.62	229.40	223.83	240.48
ReE	21.52	16.63	17.45	16.49	20.92	14.73
ReI	760.28	668.37	582.93	682.26	613.95	617.30
ReI2125	392.94	310.74	355.31	363.62	288.70	280.75
ReI6364	141.29	108.41	46.26	126.98	55.71	79.70
ReJ	82.12	101.82	73.29	79.76	52.40	49.87
ReK	87.79	74.58	42.05	48.67	52.40	42.30
ReN	17.62	15.99	11.57	17.83	12.21	10.66
SMRC	103.88	100.01	104.66	99.71	93.86	98.75
SMRC1526	98.08	112.33	97.17	88.23	90.88	107.90
SMRC3039	114.40	93.69	110.51	102.84	98.88	83.05
SMRC8196	91.45	92.29	97.83	110.76	79.44	87.30
SMRE	119.57	103.60	131.67	109.24	129.10	89.61
SMRI	119.99	108.75	114.98	116.24	98.25	98.33
SMRI2125	100.30	104.62	137.26	118.20	91.62	86.94
SMRI6364	168.81	140.90	74.89	174.76	75.50	102.84
SMRJ	113.31	146.33	132.39	129.35	77.81	72.34
SMRK	127.85	151.96	99.01	96.85	96.77	82.22
SMRN	114.16	101.85	89.87	118.10	78.32	63.89
PYLLC	1216.72	1101.53	1062.55	1126.65	925.65	975.03
PYLLC1526	306.04	277.58	220.23	272.70	201.33	280.05
PYllC3039	242.30	227.34	200.19	232.48	193.05	151.50
PYlll	1170.12	1182.35	1116.08	1365.40	778.44	839.03
PYLLI2125	578.20	555.64	596.50	728.38	360.04	350.23
PYLLJ	245.71	286.85	272.24	266.66	74.51	71.90
PYLLK	585.14	596.79	391.31	415.26	351.55	219.86
sum_neg	13670,19	13137,34	11679,16	13012,17	9431,74	9461,49

Note: 1* - Contaminated area, 2* Non contaminated area, sum_neg: SMRV - PYLLK

Recurrence Quantification Analysis: A Promising Method for Data Evaluation in Medicine

Jakub Schlenker[1], Tomáš Nedělka[1,2], Lucie Riedlbauchová[3], Vladimír Socha[1], Karel Hána[1], Patrik Kutílek[1]

[1] Czech Technical University in Prague, Faculty of Biomedical Engineering, Kladno, Czech Republic

[2] Department of Neurology, University Hospital Motol and 2nd Faculty of Medicine of the Charles University, Prague, Czech Republic

[3] Clinic of Cardiology, University Hospital Motol and 2nd Faculty of Medicine of the Charles University, Prague, Czech Republic

Abstract

Introduction: This paper describes principles of a promising method for medical data analysis called recurrence analysis, which is based on the chaos theory that better describes processes in the living organisms.
Methods: Phase space reconstruction and recurrence analysis are explained in brief. The main part of the work focuses on recurrence plot, which is a basic tool for recurrence analysis.
Conclusion: Possible clinical applications of recurrence analysis in the field of medicine are discussed and a short overview of pilot projects performed by our team on this topic is given.

Keywords

Nonlinear analysis, recurrence analysis, recurrence plot, phase space, heart rate variability, autonomic dysfunction.

Correspondence to:

Jakub Schlenker
Joint Department of Biomedical Engineering CTU
and Charles University in Prague, Czech Republic
Address: Studničkova 7/2028, 128 00 Prague 2, Czech Republic
E–mail: jakub.schlenker@gmail.com

EJBI 2014; 10(1):en35–en40
received: October 9, 2014
accepted: November 19, 2014
published: December 15, 2014

1 Introduction

Usage of nonlinear methods for data analysis is becoming increasingly popular in medicine due to the fact that they seem to be able to describe selected processes occurring in living organism more effectively than it is these days [1].

To increase the probability of full recovery or to minimize the health damages, it is important to detect diseases in their early or even in their subclinical stages. Since specific methods of nonlinear analysis seem to be sensitive enough to uncover these early phases of the disease development, their application in the data analysis may improve health care and help the physicians to understand better the physiological and pathophysiological processes occurring in the human body.

One such nonlinear method recently applied in medicine is the recurrence analysis. Method of so called recurrence analysis is derived from the chaos theory which describes the basic dynamics of a system with chaotic behaviour that can be found in every biological system [2]. Recurrence analysis has been successfully used in pilot projects in cardiology [4, 5] and neurology [7, 6, 9, 8, 10] where it was mainly used to describe dynamics of the heart rate and blood pressure regulation. These physiologic variables are under permanent control of the autonomic nervous system which may be viewed as an example of nonlinear deterministic system since the autonomic nervous system instantaneously changes its tone based on the actual demands and needs of the organism [11, 2]. Impaired function of the autonomic nervous system is therefore often associated with reduced variability of functions that this system controls, i.e. reduced heart rate variability. Due to this fact, the system (heart rate) tends to recur to a similar state and exhibit only limited changes in response to outer inputs when its control through the autonomic nervous system is harmed.

The effort of our team has recently focused on the evaluation of possible role of the recurrence analysis in the diagnostic of various diseases in their early phases (diseases origin of which is associated with the autonomic dysregulation) especially in the field of neurology [9, 10] and cardiology.

2 Methods

Recurrence analysis is one of the non-linear analysis of data derived from chaos theory. Recurrent graphs, ba-

sic tool recurrent analysis allow to visualize repetitive behaviour of dynamic systems. The method is also suitable for non-linear analysis of short-term and non-stationary data [12].

- The first step of analysis, as with most nonlinear techniques [12], is the reconstruction of the phase space.
- The second step is the formation of recurrent plot using the threshold distance.
- The last step in the analysis is to calculate the measures of recurrent plot.

2.1 Phase Space Reconstruction

The n state variables in time form vector (trajectory) in n-dimensional space (phase space). Trajectory in phase space represents all possible states of a system. Each state of the system corresponds to a specific point in the space phase. For N state variables N-dimensional phase space is created. However, very often it is not possible to observe more than one state variable of a system in the field of clinical medicine, because they are not known or it is difficult to measure them [12]. Nevertheless we can reconstruct the phase space trajectory from a single observation by using the Takens' theorem [12]. Takens' theorem is one of the most commonly used method for reconstruction of the phase space trajectory [12, 13]:

$$x_i = (y_i, y_{i+\tau}, \ldots, y_{i+(m-1)\tau}),$$

where m is the embedding dimension, τ is the time delay and y_i is a single observation.

An optimal set of embedding dimension and time delay is important for reconstruction of phase space that fully describes the system dynamics [14]. One of the frequently used approaches for choosing time delay is the first minimum of mutual information [15]. The mutual information measures mutual dependence of two variables A and B and can be defined using entropy as [15]:

$$I(A,B) = H(A) + H(B) - H(A,B),$$

where $H(A)$ and $H(B)$ are the entropies and $H(A,B)$ is the joint entropy of A and B.

Entropy is generally, expressed in terms of a discrete set of probabilities s_i as [15]:

$$H(S) = -\sum_{i=1}^{n} p_s(s_i) \log p_s(s_i),$$

where $H(S)$ is entropy of system S, which consists of a set of possible messages $(s_1, s_2, ..., s_n)$ and the associated probabilities $p_s(s_1), p_s(s_2), ..., p_s(s_n)$.

False nearest neighbour method (FNN) is often used for optimal embedding dimension setting. Embedding dimension is a parameter which tells us how many dimensional phase space will be reconstructed. FNN method is based on the fact that choosing of low embedding dimension causes crossing of phase space trajectory[14]. We use modified FNN method using ratio of Euclidian distances of two neighbour's states in m and $m+1$ dimensional space by Cao [16]:

$$\begin{aligned} a(i,m) &= \frac{\| y_i(m+1) - y_{n(i,m)}(m+1) \|}{\| y_i(m) - y_{n(i,m)}(m) \|}, \\ y &= 1, 2, \ldots, N - m \cdot \tau, \end{aligned}$$

where $\| \ . \ \|$ is Euclidian distance, $y_i(m)$ is i-th reconstructed vector with embedding dimension m, $y_{n(i,m)}(m)$ is the nearest neighbor of $y_i(m)$.

The mean value of all $a(i,m)$ is introduced by Cao [16]:

$$E(m) = \frac{1}{N - m\tau} \cdot \sum_{i=1}^{N-m\tau} a(i,m).$$

The average $E(m)$ is dependent only on dimension m and delay τ. Finally, variable $E_1(m)$ is used to investigate variations from dimension m to dimension $m+1$ [16]:

$$E_1(m) = \frac{E(m+1)}{E(m)}.$$

Value of $E_1(m)$ stops changing when dimension m is greater than the dimension of attractor m_0.

2.2 Recurrence Analysis

Recurrences in the nature were observed a long time ago [12], but the method of recurrence analysis was introduced just after the development of computer science in 1980s [19]. The basic tool of recurrence analysis is recurrence plot (RP) see Fig. 1 which can visualize recurrences in system dynamics using two-dimensional graph [19]. RP shows all moments in times when the phase space trajectory of the dynamical system visited roughly the same area in the phase space [17]. It is possible to visualise multidimensional phase space using two-dimensional graph (RP). We can demonstrate recurrences in a dynamical system, find interrelations between several systems or detect transitions between different states with RPs [6, 12, 17]. Mathematical expression of RP see equation 1.

$$R_{i,j} = \Theta(\epsilon - \|x_i - X_j\|), \text{ for } i,j = 1,2,...,N, \quad (1)$$

where N is the number of states, ϵ is a threshold distance, $\|.\|$ a norm and $\Theta(.)$ the Heaviside function.

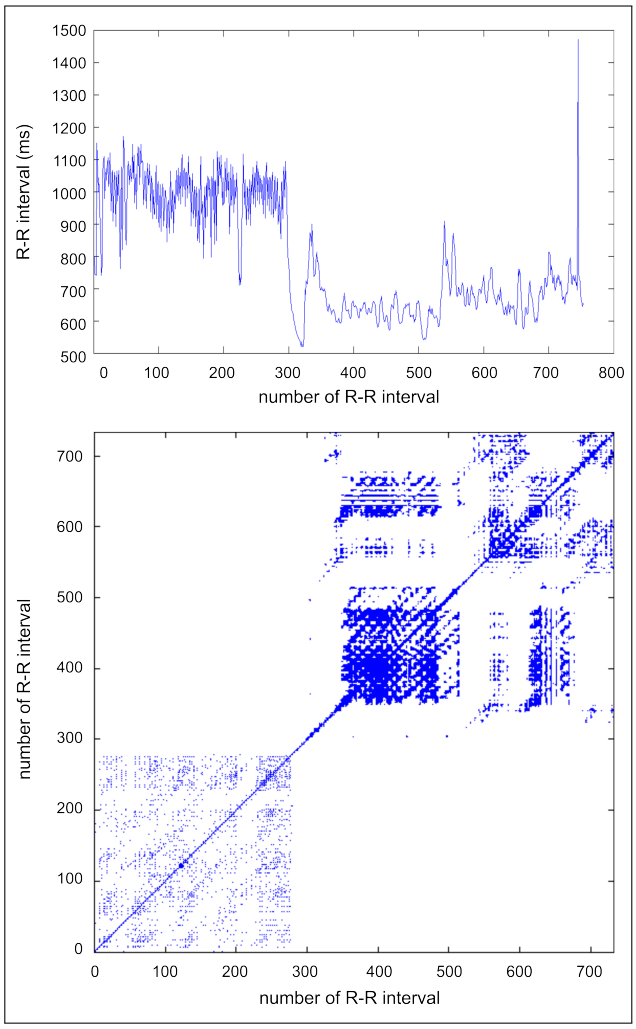

Figure 1: On the top figure, there is the input signal for recurrence analysis. The input signal is represented by the length of R-R intervals. On the bottom figure, there is the recurrence plot. On both figures we can distinguish two phases of the orthostatic test, resting in supine position and standup position, each for 5 minutes. In the second phase, after the verticalisation of the patient, we can recognise lower heart rate variability.

The most important parameter of recurrence analysis is the threshold distance. If the distance between two states on the phase space trajectory is smaller than a given threshold, the recurrence point in RP arises. Value of the point in recurrence matrix is one; otherwise it is zero [12]. This technique is called pair test and it is a pairwise test of all states (for N states we compute N^2 tests) [12].

With closer looking at the RP we can see single recurrence points, diagonal lines, vertical lines and horizontal lines. The structures formed by these elements are basis for recurrence quantification analysis (RQA). This analysis was introduced by Zbilut and Weber [19, 20, 21] to evaluate RPs quantitatively. RQA is a set of measures derived from diagonal and vertical structures of RP [17].

Diagonal line occurs when the trajectory visits the same area of the phase space at different times, it is when the system returns to the same or similar states at the different time points. Horizontal and vertical lines indicate how long the system stays in given state which does not change at all or just very slowly [17]. Single isolated recurrence point can represent a rare state [17].

The **Recurrence rate** is density of recurrence points in recurrence plot [17]:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}.$$

This value corresponds to probability that a specific state will recur.

Determinism is a measure which represents the percentage of recurrence points which form diagonal lines [17]:

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{i,j}^{N} R_{i,j}},$$

where $P(l)$ is the histogram of the lengths l of the diagonal lines.

Laminarity is the amount of recurrence points which form vertical lines. This variable is related with amount of laminar states [17]:

$$LAM = \frac{\sum_{v=v_{min}}^{N} v\ P(v)}{\sum_{v=1}^{N} v\ P(v)},$$

where $P(v)$ is the histogram of the lengths v of the vertical lines.

The average length of vertical lines is called **Trapping time** TT and it is related with laminarity time. This value contains information about frequency and length of laminar states [17]:

$$TT = \frac{\sum_{v=v_{min}}^{N} v\ P(v)}{\sum_{v=v_{m}in}^{N} P(v)}.$$

The low value of LAM and TT indicate high complexity of a system [6].

The **Longest diagonal line** L_{MAX} is length of the longest diagonal line [17]:

$$L_{MAX} = max(l_i; i = 1 \dots N_l).$$

Divergence DIV is the inverse of L_{MAX} and it is related with Kolmogorov-Sinai entropy [17].

$$DIV = \frac{1}{L_{MAX}}.$$

Other measures of RQA are average length of diagonal lines ($AVDL$), ratio ($RATIO$) between DET and RR, Shannon entropy ($ENTR$) and maximal length of vertical line (V_{MAX}) can be found in [17].

The advantages of RP include the ability to capture the chaotic properties of a system without a need of a long data series and the fact that it is relatively immune to noise and nonstationarity [17]. RQA is a sensitive tool for detecting any dynamic changes; however, it can be easily affected by inappropriate settings especially setting of the threshold distance where even a small change can dramatically affect the results of RQA [22, 23]. There are several methods for threshold distance setup [17, 22, 23]. One of the widely used method is setting of the threshold distance as a percentage of the maximum distance in phase space. Furthermore there is setting where values should not exceed 10% or 15% of the average or maximum distance in phase space [24, 18]. Quite frequently used method is called a fixed percentage of recurrent points. This means that we set a threshold distance value that guarantees the exact percentage of recurrent points [6]. Often is this value 1% [22, 24], but we can also find other values, such as 5% [6, 9, 25]. Selecting the optimal threshold distance appears to be a task for data mining [26].

3 Results

Our team currently evaluates possible applicability of the recurrence analysis in the field of medicine, specifically in neurology and cardiology. Our pilot study was focused on the early detection of the autonomic dysfunction in subclinical stages of diabetic cardiovascular neuropathy [9]. We found a significant increase in measures DET, LAM, $LMAX$ and TT in diabetic patients compared to control group. Based on the promising results of this pilot study, a bigger project on this topic has started, the aim of which is to confirm high sensitivity of the recurrence analysis in the early detection of diabetic autonomic disfunction and to identify typical values of RQA measures both in patients with diabetes and manifest autonomic disfunction and in the healthy control group. One of the goals is also the identification of optimal value of threshold distance for diagnosis of autonomic neuropathy. For now we reached promising results using a fixed percentage of recurrent points and our experimental setup using standard deviation:

$$\epsilon = \frac{0.02}{od},$$

where ϵ is threshold distance, od is a standard deviation of input signal.

This study is a joint project of Faculty of Biomedical Engineering CTU and University hospital Motol.

Next application of recurrence analysis in the field of neurology is the assesment of vasovagal syncopes. Results of a pilot project on this topic were recently presented by our team at the international conference in Slovakia [10]. We found reduced complexity of the heart rate control in young patients with vasovagal syncope using the recurrence analysis. RQA measurement showed significant differences in DET, $LMAX$, DIV and $ENTR$ see Fig. 2 However, future studies are needed in this particular field because of the limitations of our study - no gender match and small number of subject in groups.

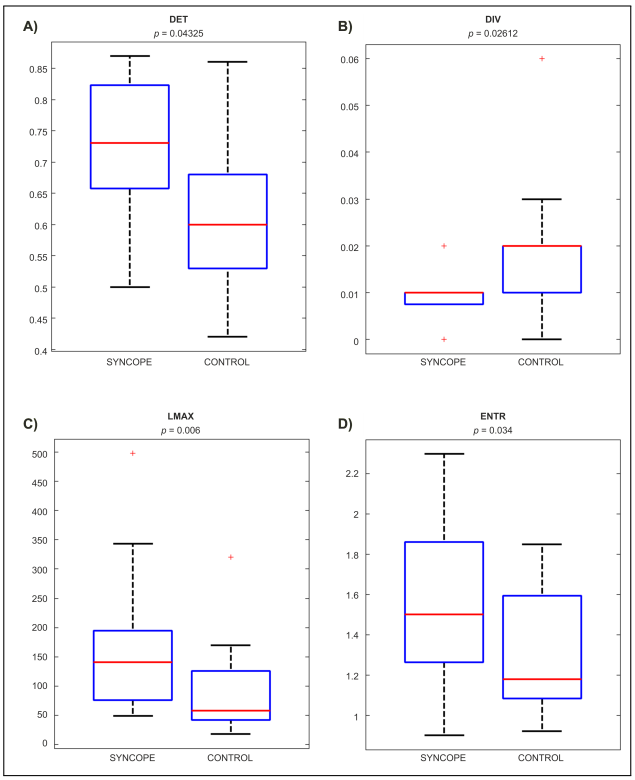

Figure 2: Boxplots[1] illustrating the comparison between patients with syncope and the control group. A – Determinism ($p = 0.04225$). B – Divergence ($p = 0.02612$). C – The length of the longest diagonal line ($p = 0.006$). D – The Shannon entropy ($p = 0.034$) [10].

Another cardiologic application of recurrence analysis that we are working on is the prediction of onset and termination of paroxysmal atrial fibrillation. It is a similar project to that of Mohhebi and Ghassemian [4] with the difference that we use real patient's Holter monitor data recorded at the Clinic of cardiology in the University hospital Motol. We hope, that our project will confirm capability of recurrence analysis in this field.

Currently a perspective applications of recurrence analysis focused on evaluation of psycho-physiological condition of pilots [27, 28]. When elevating pilots condition in stress situations, a change in physiological functions takes place, such as the change in heart rate, respiratory rate, blood pressure etc. The preliminary results show high potential of RQA for data evaluation of these type of measurement.

[1] Boxplots shows maximum, minimum, median, first and thirdquartile, cross points are outliers.

4 Conclusion

Detecting diseases in their early or even subclinical stages is important for increasing the probability of full recovery. Nonlinear methods seem to be useful for describing selected processes occurring in the living organism effectively. One of such methods is the recurrent analysis. Our pilot studies showed high potential of its use in medicine. However, there is a question about optimal setup of threshold distance and normative data are needed and results of pilot studies need to be confirmed in larger trials yet.

Acknowledgements

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant no. SGS14/169/OHK4/2T/17 and no. SGS14/170/OHK4/2T/17.

References

[1] Silipo R, Deco G, Vergassola R, Bartsch H. Dynamics Extraction in Multivariate Biomedical Time Series. Biol Cybern. 1998 Jul;79(1):15–27.

[2] Sharma V. Deterministic Chaos and Fractal Complexity in the Dynamics of Cardiovascular Behavior: Perspectives on a New Frontier. Open Cardiovasc Med J. 2009;3:110–123.

[3] Marwan N, Carmenromano M, Thiel M, Kurths J. Recurrence Plots for the Analysis of Complex Systems. Physics Reports. 2007;438(5-6):237–329. doi:10.1016/j.physrep.2006.11.001.

[4] Mohebbi M, Ghassemian H. Prediction of Paroxysmal Atrial Fibrillation Using Recurrence Plot-Based Features of the RR-interval Signal. Physiol. Meas. 2011;32:1147–1162.

[5] Naschitz JE, Rosner I, Shaviv N, Khorshidi I, Sundick S, Isseroff H et al. Assessment of Cardiovascular Reactivity by Fractal and Recurrence Quantification Analysis of Heart Rate and Pulse Transit Time. J Hum Hypertens. 2003 Feb;17(2):111–8.

[6] Javorka M, Trunkvalterova Z, Tonhajzerova I, Lazarova Z, Javorkova J, Javorka K. Recurrences in Heart Rate Dynamics are Changed in Patients with Diabetes Mellitus. Clinical Physiology and Functional Imaging. 2008;28(5):326-331. doi:10.1111/j.1475-097X.2008.00813.x.

[7] Acharya UR, Sree SV, Chattopadhyay S, Yu W, Ang PC. Application of Recurrence Quantification Analysis for the Automated Identification of Epileptic EEG Signals. Int J Neural Syst. 2011 Jun;21(3):199–211.

[8] Priano L, Saccomandi F, Mauro A, Guiot C. Non-linear Recurrence Analysis of NREM Human Sleep Microstructure Discloses Deterministic Oscillation Patterns Related to Sleep Stage Transitions and Sleep Maintenance. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:4934–7.

[9] Nedělka T, Schlenker J, Riedlbauchová L, Mazanec R. Recurrence Quantification Analysis of Heart Rate Variability in Early Diagnosis of Diabetic Autonomic Neuropathy (in Czech). Cesk Slov Neurol N. 2012;75/108(6):721–728.

[10] Schlenker J, Nedělka T. Recurrence Analysis in Patients with Vasovagal Syncopes In: Measurement 2013 - 9th International Conference on Measurement (Proceedings). Bratislava: Institute of Measurement Science, 2013;145–148. ISBN 978-80-969-672-5-4.

[11] Rilley R, Balasubramaniam R, Turvey MT. Recurrence Quantification Analysis of Postural Fluctuations. Gait & Posture. 1999;9(1):65–78.

[12] Traut MH. MATLAB® Recipes for Earth Sciences. Potsdam: Springer, 2007, 288 s. ISBN 978-3-540-72748-4.

[13] Marwan N, Wessel N, Meyerfeldt U, Schirdewan A, Kurths J. Recurrence-Plot-Based Measures of Complexity and Their Application to Heart-Rate-Variability Data. Physical Review E. 2002;66(2), doi:10.1103/PhysRevE.66.026702.

[14] Kodera J, Tran VQ. Vizuální Nelineární Rekurentní Analýza a Její Aplikace na Český Akciový Trh. Politická ekonomie. Praha: VŠE, 2009;57(3). ISSN 0032-3233.

[15] Fraser A, Swinney H. Independent Coordinates for Strange Attractors from Mutual Information. Physical Review A. 1986;33(2):1134-1140. doi:10.1103/PhysRevA.33.1134.

[16] Cao L. Practical Method for Determining the Minimum Embedding Dimension of a Scalar Time Series. Physica D: Nonlinear Phenomena. 1997;110(1-2):43–50. doi:10.1016/S0167-2789(97)00118-8.

[17] Marwan N. Encounters with Neighbours: Current Developments of Concepts Based on Recurrence Plots and Their Applications [PhD thesis]. Potsdam: University of Potsdam; 2003. Available from: http://www.recurrence-plot.tk/furtherreading.php.

[18] Sun R, Wang Y. Predicting Termination of Atrial Fibrillation Based on the Structure and Quantification of the Recurrence Plot. Medical Engineering. 2008;30(9):1105–1111. doi:10.1016/j.medengphy.2008.01.008

[19] Marwan N. A Historical Review of Recurrence Plots. The European Physical Journal Special Topics. 2008;64(1):3–12. doi:10.1140/epjst/e2008-00829-1

[20] Zbilut JP, Webber CL. Embeddings and Delays as Derived from Quantification of Recurrence Plots. Physics Letters A. 1992;171(3-4):199–203.

[21] Webbber CL, Zbilut JP. Dynamical Assessment of Physiological Systems and States Using Recurrence Plot Strategies. Journal of Applied Physiology. 1994;76(2):965–973.

[22] Ding H, Crozier S, Wilson S. Optimization of Euclidean Distance Threshold in the Application of Recurrence Quantification Analysis to Heart Rate Variability Studies. Chaos, Solitons & Fractals. 2008;38(5):1457–1467.

[23] Garcia-Gonzales MA, Fernandez-Chimeno M, Ramos-Castro J. Errors in the Estimation of Approximate Entropy and Other Recurrence-Plot-Derived Indices Due to the Finite Resolution of RR Time Series. IEEE Transactions on Biomedical Engineering. 2009;56(2):345–351. doi:10.1109/TBME.2008.2005951.

[24] Schinkel S, Dimigen O, Marwan N. Selection of Recurrence Threshold for Signal Detection. The European Physical Journal Special Topics: Special Topics. 2008;164(1):45–53. Available from: http://www.springerlink.com/index/10.1140/epjst/e2008-00833-5

[25] Javorka M, Turianikova Z, Tonhajzerova I, Javorka K, Baumert M. The Effect of Orthostasis on Recurrence Quantification Analysis of Heart Rate and Blood Pressure Dynamics. Physiological Measurement. 2009;30(1):29–41.

[26] Kalina J, D Tebbens J, Schlenker A. Robustness of High-Dimensional Data Mining. ITAT 2014, Part II, pp. 53—60.

[27] Kaľavský P., Socha L.. Výskum tréningových metód pilotov s využitím leteckých. In: FLYIN: aviation industry exhibition : FLYIN letecká konference VŠB-TUO 2013 : sborník příspěvků. 1st. ed. Ostrava: Vysoká škola báňská - Technická univerzita Ostrava, 2013. ISBN 978-80-248-3091-9.

[28] Socha V., Szabo S., Socha L., Kutílek P. 2014, "Evaluation of the variability of respiratory rate as a marker of stress changes", Proceedings of 18th International Conference. Transport Means, KAUNAS UNIVERSITY OF TECHNOLOGY, 23-24 October 2014 (in press)

