1

VOLUME 3 | ISSUE 1 | 2015

mefanet JOURNAL

ISSN (print) 1805-9163 | ISSN (on-line) 1805-9171

Indexed in ERIH PLUS,
Directory of Open Access Journals,
Index Copernicus,
Google Scholar,
Bibliografia medica Čechoslovaca,
Bibliografia medica Slovaca

http://mj.mefanet.cz

CONTENTS

MEFANET Journal | Periodicity twice a year | Registration code of Ministry of Culture of the Czech Republic MK ČR E 21223| ISSN (print) 1805-9163| ISSN (on-line) 1805-9171| Title abbreviation Mefanet J | Publisher Facta Medica, Ltd., Srbská 2186/19, 612 00 Brno, Czech Republic, Company identification number 28298110, GSM +420 737 985 593, +420 737 287 512; email fama@fa-ma.cz | Editor-in-charge Boris Skalka | Copy-editing Jakub Gregor | Graphic design and typesetting Radim Šustr (Institute of Biostatistics and Analyses, Faculty of Medicine at Masaryk University, Czech Republic) | Composed in Skolar, typeface designed by David Březina in 2011 | On-line version available at WWW <http://mj.mefanet.cz/>

PREFACE

With a great pleasure we present the first 2015 issue of the MEFANET Journal (MJ). MJ is dedicated to provide readers around the world with high quality peer-reviewed articles on a wide variety of topics related to applications of computer science and technology-enhanced learning in medical education. Its mission is to become the premier vehicle for disseminating information about MEdical FAculties NETwork (www.mefanet.cz), which covers all Czech and Slovak medical faculties as well as schools or faculties of health care sciences.

This first issue of the third volume comprises two original articles, two review articles and one detailed editorial report on the last MEFANET conference 2014. The presented articles can be categorized into the following groups:

- information science and evidence-based medicine,
- modeling and simulation.

The paper by Barrick et al. reviews the current state of big data in pharma. The topic is very interesting and sensitive when it comes to handling data concerning patients – and the manuscript brings into light valuable information. Volšička & Blaha exemplify five different methods for performing full-text search in medical records and they introduce Apache Solr platform as a search engine which is robust enough to handle various pitfalls hidden in the Czech language. Tučková et al. shares experience with PIRD method (P = population, problem, patient; I – index test; R – reference test; D – diagnosis of interest) and demonstrates an example of evidence synthesis in diabetes mellitus diagnostic tests for children. I believe that the article will be noteworthy especially for readers with the interest in courses of the information literacy for medical students. The original article by Tachecí and Ryška brings a new view on the development and usage of virtual patients or cases. The authors present their approach to extending the number of available virtual cases with the use of a pseudo-random generator. They point out that the cases solved by students are subsequently discussed also within tutorial lessons – enabling students to get a feedback from their teacher – enriching the learning process with the debriefing experience. The final editorial material by Gregor J. et al. recall the most important memories from the last 8th year of the MEFANET conference, which took place in Brno in the end of November 2014. The MEFANET conferences have always been providing a vibrant meeting place for delegates from medical and healthcare faculties, computer scientists as well as medical teachers and students from the Czech Republic, Slovakia and other countries as well. In 2015, the upcoming 9th year of the MEFANET conference (25th, 26th November 2015 – Brno, Czech Republic) will be focused on technology-enhanced learning and teaching in psychiatry, neurology and neuroscience.

I would like to extend my sincere appreciation to the editorial members and reviewers, without whom this issue would not have been possible. It is my hope that this fine collection of articles will be another valuable resource for medical education stakeholders and will stimulate further research into this exciting area of science. Readers are encouraged to submit both comments on these articles as well as their own relevant manuscripts.

August 2015

Daniel Schwarz
Editor-in-chief

REAL WORLD EVIDENCE: A FORM OF BIG DATA, TRANSFORMING HEALTHCARE DATA INTO ACTIONABLE REAL TIME INSIGHTS AND INFORMED BUSINESS DECISIONS

Uttam Kumar Barick[1]*, Daniel Schwarz[2], Behsad Zomorodi[3], Arun Gowda[1], Rituraj Mohanty[1], Martin Komenda[2]

[1] *Phamax Analytic Resources Pvt. Ltd, Karnataka, India*
[2] *Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic*
[3] *Phamax AG, Oberwil (BL), Switzerland*
* *Corresponding author: uttam.barick@fs-researchcenter.com*

ARTICLE HISTORY
Received 1 April 2015
Revised 27 May 2015
Accepted 30 May 2015
Available online 8 June 2015

KEYWORDS
RWE
BIG DIP
big data
data mining
real time insights
business decisions
visual analytics

ABSTRACT — *Data has always played an important role in assisting business decisions and overall improvement of a company's strategies. The introduction of what has come to be named "BIG data" has changed the industry paradigm altogether for a few domains like media, mobility, retail and social. Data from the real world is also considered as BIG data based on its magnitude, sources and the industry's capacity to handle the same. Although, the healthcare industry has been using real world data for decades, digitization of health records has demonstrated its value to all the stakeholders with a reaffirmation of interest in it. Over time, companies are looking to adopt new technologies in linking these fragmented data for meaningful and actionable insights to demonstrate their value over competition. It has also been noticed that the consequences of not demonstrating the value of data are sometimes leads regulators and payers to be severe. The real challenge though is not in identifying data sets but transforming these data sets into actionable real time insights and business decisions.*

Evidence and value development frameworks need to work side by side, harnessing meaningful insights in parallel to product development from early phase to lifecycle management. This should in-turn create evidence and value-based insights for multiple stakeholders across the industry; ultimately supporting the patient as the end user to take informed decisions that impact access to care.

This article attempts to review the current state of affairs in the area of BIG data in pharma OR BIG DIP as it is increasingly being referred to.

BACKGROUND

Data in all its various formats has historically played a significant role in the design of any business model and thus companies are mining data, hoping to increase efficiency and gaining a competitive advantage in the market to outperform their peers [1,2]. This is also true for pharmaceutical firms who for decades have been managing vast amount of data which are critical either in the demonstration of clinical and economic value, in product development or commercialization decisions [3]. However, considering the fact that most of this data involves people or patient level data, incremental costs are being added along with stringent regulations for the generation and handling of such data. In order to minimize the cost of these data while simultaneously maximizing the value, pharma companies are looking to adopt new technologies over time; alongside the struggle to integrate various types and sources of data globally [4]. In order to capitalize the usefulness "It is critical to collaborate with researchers and the technology ecosystem to develop innovative solutions to seemingly intractable problems emerging in healthcare and life sciences today" [5]. Unlike other industries where this might be the norm, the challenge in the pharma industry is the ability to look into the data generated differently and transforming it into scientifically acceptable evidence.

Solution to this may be managing BIG data. The term BIG data is a "blanket term" for any data set which is big in nature and too complex to process using the traditional management tools or data processing applications [6]. The impression of BIG depends on how an organization manages its data. For some companies,

facing hundreds of gigabytes for the first time might imply BIG data while for others it may require 100s of terabytes before considering information as BIG data [7].

BIG data is at the center of many ongoing discussions about data and its next avatar; so much that it attracts a lot of media attention and makes it difficult to differentiate reality from hype. An increasing trend though is the acceptance of BIG data's importance and integration into growth strategies across all the industries [8], including life sciences and pharmaceuticals.

Concurrently, healthcare stakeholders have started to see the value of evidence based data and its role in optimizing the patient journey or access to care. Over time, payers are becoming skeptic of the incremental benefit demonstrated in controlled settings by new drugs and also the value for money for patients and the economy. Payers are therefore increasingly demanding stronger evidence of value. Stakeholders at a local level have also started investing and developing their own evidence sources to assist decision making. Earlier, data was collected in silos, leading to a fragmented picture of the patient experience, limiting the ability of stakeholders to answer key questions. Now, stakeholders are no longer taking any risk associated with new products/interventions and have started demanding and mining more data to prove value. This marks the beginning of the "Prove It" era and the science of real world evidence has gained importance to bring together disconnected data sources in a meaningful way to answer key questions that various stakeholders pose. Such collaborations are supported by advances in technology; enabling secure and effective ways to link the data sources without compromising the integrity of data. The result – evolution of disconnected data sources into real world evidence leading to real time insights and actions. Real world evidence data sets are often called as BIG data due to their similarities in magnitude and sources [6].

This article attempts to review the current state of affairs in the area of BIG data in pharma or BIG DIP as it is increasingly being referred to.

DISCUSSION

How BIG is BIG DIP (data in pharma)?

The ability to deliver new lifesaving drugs to patients in a timely and cost effective way is dependent on the capacity to manage the huge amount of data generated throughout the phases of a product lifecycle [9]. However, the pharma industry has multiple functional areas and each contributes unique sets of insights to understand the mechanisms of progression of disease. This is making such data management increasingly difficult; pharmaceutical and healthcare firms are managing large data sets from the research records and patient information to utilization details, and supply chain monitoring. Around 70% of a typical pharmaceutical project now involves simply managing data [10] and this involves assimilation, transmission and cleaning before any actual analysis can begin. As per Thomson Reuters' survey, the biggest opportunity for BIG data in pharma industry is from early-stage discovery (i.e., drug discovery and development) ~41.20 percent, followed by understanding the market (i.e., real world) ~26.5 percent [11]. Data sources are always associated with the three forms of "V", i.e., volumes, varieties and velocities of data [12,13]. In the same time veracity and vocabulary are equally important, especially as social media is becoming a vital source for gathering of BIG data, primarily due to self-reported data. With the adaptation of semantic-based methodology [14], RWE studies can facilitate harnessing actionable business insights in real time manner.

What follows are the current analytical and technological trends with examples of transforming RWE to real time insights.

What's the BIG deal in BIG DIP? Data mining perspective

The pharmaceutical industry is well known for performing quantitative analyses for clinical and market research. In the marketing departments, data mining (DM) applications are used for sales force planning and direct marketing to doctors and consumers. Data mining techniques are used to a variety of critical business decisions in the pharmaceutical industry [15]. The role of DM methodologies in pharmacological domain is evolving. The field of DM covers varied powerful tools like "association, clustering, segmentation and classification", which provide better manipulation of the data and help the pharma sector compete on lower costs while improving the quality of drug discovery and delivery methods. "A deep understanding of the knowledge hidden in the pharma data is vital to a competitive position and organizational decision-making" [16]. In general, DM algorithms may significantly assist pharmacological scientists to discover potentially relevant drug-event associations [17]. The obtained results generated by the use of various DM techniques should be viewed as hypothesis proposals and should be evaluated in the context of other relevant data. Alternatively these signals can then be used as a bases for intervention as appropriate [18]. According to Poluzzi et al., the application of data mining techniques in clinical pharmacology can be broadly grouped into two following areas:

- "*Identification of new effects of drugs* (mostly adverse reactions, but sometimes also new therapeutic effects, and effects in special populations)";
- "*Appropriateness in drug use* (e.g., frequency of use in patients with contraindications, concomitant prescriptions of drugs known for the risk of clinically relevant interactions)" [19].

With respect to the increasing trend of volume (amount of data) and variety (range of data types) an application of advanced mining approaches is required. Ranjan [20] presented a set of relevant data sources, which are typical for the pharma industry: rules of general guidelines; clinical data (patient data, pharmaceutical data, medical treatments, length of stay); administrative data (staff skills, overtime, nursing care hours, staff sick leave); financial data (treatment costs, drug costs, staff salaries, accounting, cost-effectiveness studies); and organizational data (room occupation, facilities, equipment). Nevertheless, there is concern about the lack of systematic, objective validation of the methods in this particular area. Unfortunately, a gold standard to validate DM methods does not exist, although various imperfect reference standards may be used to obtain useful insights on the performance of any DM method [16]. The whole process is not only about data analysis; but taking informed decisions from the mined information and fulfillment of the need to communicate analytic outcomes simply and clearly. This is where the practical use of visual analytics approaches, which bring an innovative and effective way to deliver the knowledge from a particular domain to an individual user, is helpful. In general, the power of two robust scientific fields, data mining and visual analytics, can be successfully used to identify and display novel, valid and potentially useful patterns mined from huge databases and warehouses. Moreover data visualization helps in providing high level of understanding and trust to a user not initiated to such techniques [21]. In the most cases it brings transparent and understandable overview on achieved results for different groups of involved stakeholders.

Visual analytics is a new and progressive interdisciplinary field of study that calls for a more structured scientific approach to understanding the effects of interaction with complex graphical displays on human cognitive processes [22]. The higher-level visualization summaries can provide a framework for understanding the immense volumes of data and that reveal unexpected relationships have come to the forefront and it brings the most effective way to present and understand large sets of data [23]. Health care offers many potential applications for visual analytics including pharmaceutical development [24]. A survey presented by G. D. Sun et al. [25] reviews and classifies recent work in this particular domain into a set of application categories in many sectors of human interests and gives evidence about the importance and practical usefulness of visual analytics in real life.

Below the various case studies associated with extraction of unknown and potentially useful information from different areas of pharmaceutical domain are described. Usually DM techniques are used to support the clinicians at the point of care delivery, the controlling of clinical treatment pathways, the administrative and management tasks, and efficient management of organizational and financial data [20]. Parkinson et al. applied advanced informatics and data-mining tools to identify plausible preclinical gastrointestinal (GI) effects that may be associated with nausea and that could be of potential use in its prediction. The objective of this analysis was to assess whether preclinical GI observations and/or measurements can be used to evaluate the risk of drug-induced nausea in man. This work highlighted the usefulness of visual tools for displaying a large amount of data. For example, using a visual network to plot many preclinical effects that are associated with drug-induced nausea in man can help to quickly identify the most relevant findings [26]. Coulter et al. examined the relation between antipsychotic drugs and myocarditis and cardiomyopathy with the use of neural network architecture. They calculated the strength of dependency between a drug and adverse reaction using a logarithmic measure of disproportionality. The final analysis suggested that antipsychotic drugs other than clozapine may be associated with myocarditis and cardiomyopathy [27]. Harpaz et al. showed that a rich and diverse portfolio of data-mining approaches aligned to different strategies and objectives are now available for the analysis and detection of post approval adverse drug events. Published results offer unique prospects that collectively can advance the science of drug safety surveillance [28].

Decision makers in the pharma industry start to recognize the relevance of the definition of drugs and products in relation to management information. In the confusion between costs, care-results and patient satisfaction the right balance is needed and can be found in upcoming effective design and evaluation of graphical information systems that better support cognitive processes in areas as diverse as scientific research and emergency management [16, 29].

Answering the basic questions through novel journey of discovery

There is no doubt that BIG data and technology have enabled providing answers to questions that in the past were only answered by primary research. For example the introduction of EMRs (Electronic Medical Records) has helped in the transformation of clinical data into "liquid" usage, and further simplified mining the data for specific queries and needs [30]. This marks the introduction of information technology in the healthcare industry and eventually complements existing technology like business intelligence and data ware housing practices. The advances in the technical capability can help the industry to combine the unstructured clinical and non-clinical (e.g., claims or insurance) data without compromising the patient privacy [30]. EMR's have changed the way we understand and map the patient journey from diagnosis through adherence of treatment and to the desired or undesired endpoint. This data ranges from baseline assessment at a family physician or GP level, to specialist assessment

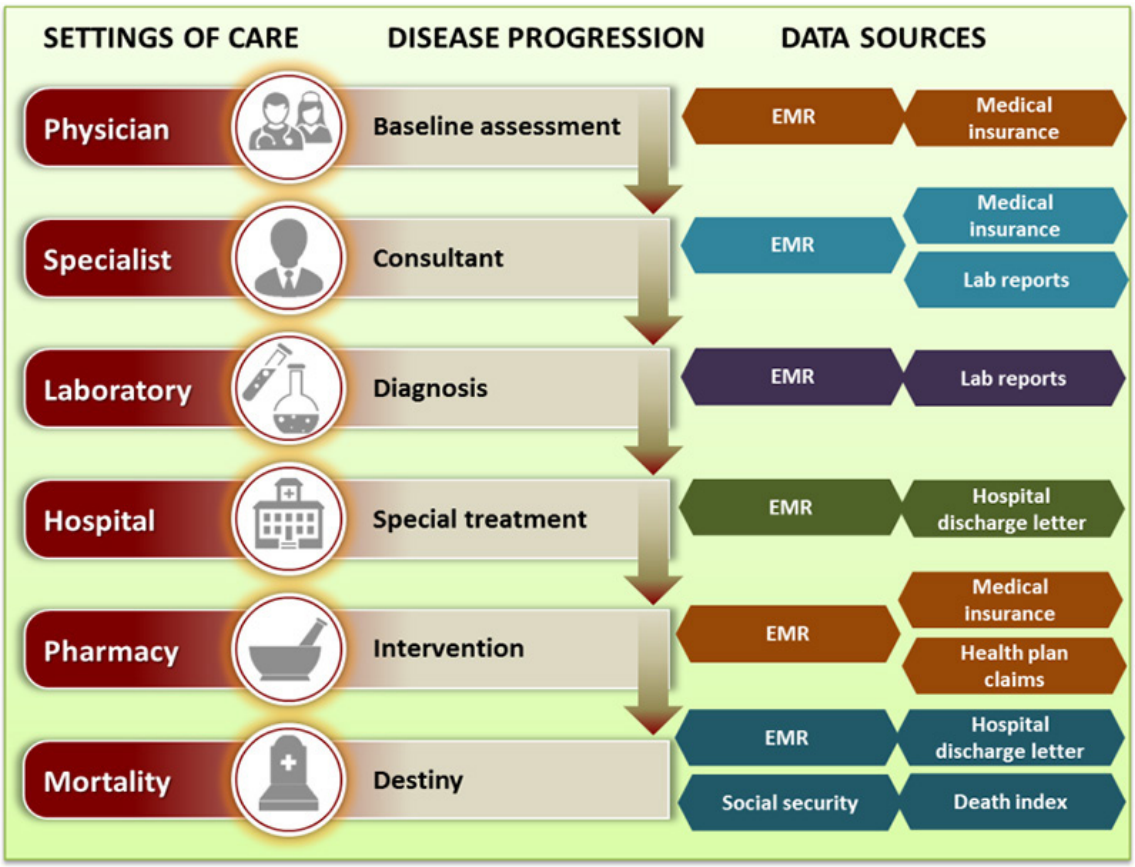

FIGURE 1 Multiple data sources in a patient journey

at a consultant level, to diagnostic data from the local and/or specialized lab reports, to treatment, and adherence data from claims.

Figure 1 is an attempt to map the patient journey to different sources of real world data that form individual robust databases from various settings of care throughout the progression of disease.

A combination of these datasets can generate evidence based patient journeys by disease, intervention or product that will answers key research questions such as:

a) Number of patients that have gone through this journey
b) Number of patients who never get drug treatment
c) Adherence ratio to the treatment regimen
d) Patients availing multiple treatments options, and most importantly
e) Cost burden and its implications on patient behavior [31]

Assessment of a combination of these data sets could potentially outline the cracks in the healthcare delivery system for the providers to fix. The potential of this is immense and the limit seems to be the ability of the human mind to arrive at novel ways of visualizing the hidden information in these datasets.

Transforming RWE into business insights and actions

Due to rise in the market competition, the pharma industry is under pressure to create and deliver the right treatments at the right time, for the right patients, at the right cost [10]. Therefore, big data in healthcare is venturing beyond the realms of improving profits and reducing overheads and is being used at advanced levels to predict epidemics, cure diseases, improve quality of life and tackle avoidable deaths [32]. Thus pharmaceutical companies need to measure and monitor metrics and trends of all types, faster than ever before. BIG data from real world provides the necessary foundation for greater insight in less time, shaping a more effective complementary process of "value" evidence generation' both in and outside the clinical drug development process [33]. A well designed study is always based on its foundation of methodology development and setting up of objective segmentation. Implementing a RWE is not much different than implementing any other clinical studies. However; a special attention is required in setting up of the objectives of the study to generate core data, identifying the patient outcomes or endpoints

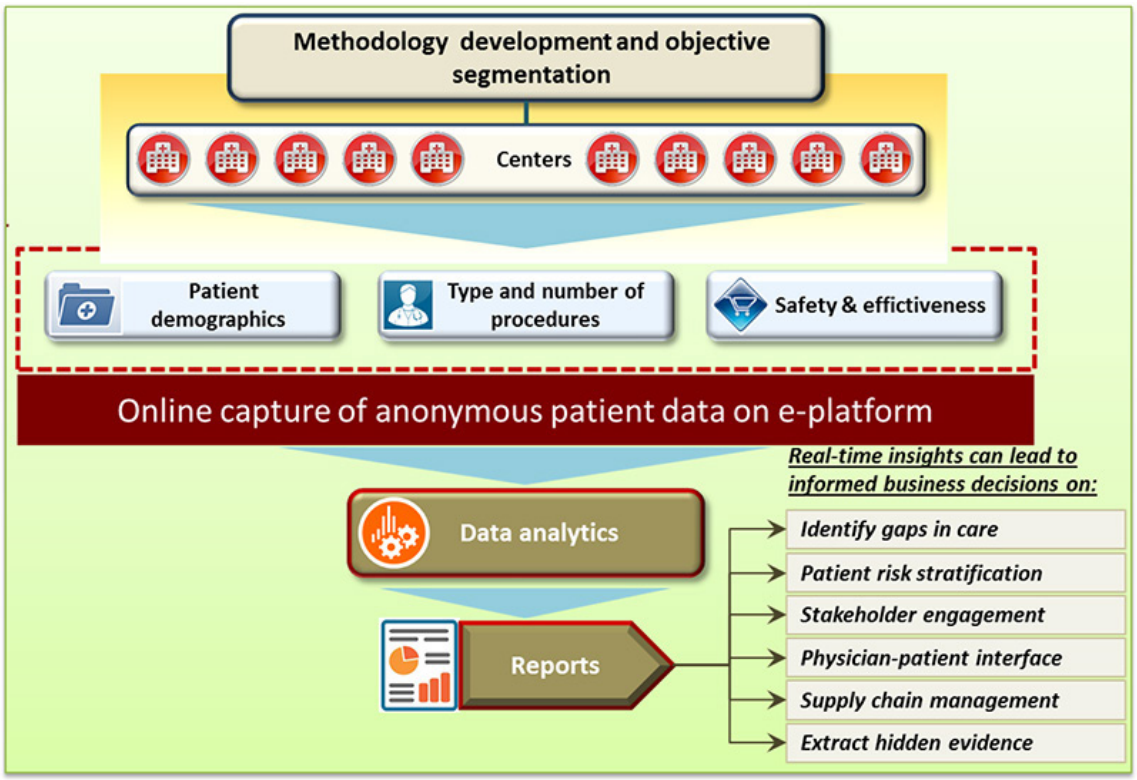

FIGURE 2 Transforming RWE into business insights and actions

from the define target population. These anonymous data points can be collected through online electronic portal in a real time basis to generate reports which eventually answers the informed business questions.

Figure 2 is the pictorial overview, illustrating-how analytics based decision support can convert real world evidence to real-time insights for informed business decisions in a hospital network setting.

FUTURE PERSPECTIVE

Pharma companies strive to innovate, create and deliver the right treatments at the right time to the right patients at the right cost. RWE is widely foreseen to be one of the instruments towards achieving this. At the same time, RWE could potentially fulfill the demand for evidence from health care decision makers who are continuously seeking value for money along with clinical outcomes involving doctors and their patients. By interlinking healthcare providers, payers, patients and other relevant stakeholders, RWE is likely to become central for establishing effective healthcare frameworks in future.

CONCLUSION

The buzz around BIG data and the hype generated around it is obvious and everyone is interested to acquire this new art of data understanding. However, evidence and value development frameworks need to work side by side, harnessing meaningful insights in parallel to product development from early phase to life-cycle management. This in turn creates value-based insights with robust foundational evidence for informed business decisions, post launch coverage and reimbursement evaluations.

According to Toby Leete and Dr. Ekta Sood, the ultimate utility of RWE is in improving the R&D cycle so that we can get safer drugs into the hands of patients quicker to allow them to achieve better outcomes [34]. Going forward there will be an increasing pressure on pharma companies in terms of reducing the cost of treatment with the demonstration of real value and in this situation "RWE will be vital throughout the life-cycle of every product" [35].

BIG DIP in the form of numerous datasets is omnipresent due to the nature of the healthcare industry in general. The real challenge is not the lack of data, but the awareness of where this data exists and establishing frameworks to mine these. The varied nature and size of these datasets makes it all the more important to be able to convert the data into visually appealing and digestible formats. These should in-turn facilitate the ability of pharma managers, R&D, payers, physicians and ultimately patients themselves to take informed decisions that impact patient's access to care. Caution however needs to be applied with each step as again the nature of these data mean there is

confidential information involved and the integrity of data has to be maintained. Whether BIG DIP derived from the real world evidence dispersed across the health delivery ecosystem will be able to revolutionize the time to discovery and access to a drug or intervention remains to be seen. But irrespective of how this evolves we are in the eye of the storm when it comes to transformation in the pharma industry from reliance on structured controlled data to opening up to unconventional yet valuable datasets.

EXECUTIVE SUMMARY

BIG data in pharmaceutical industry

The historical importance of data in the design of any business model is immense and thus companies are mining data to increase efficiency and gaining a competitive advantage in the market. Unlike other industries where this might be the norm, the challenge in the pharma industry is the ability to look into the data generated differently and transforming it into scientifically acceptable evidence.

It is believed that most of this data involves people or patient level data that requires special measurement to ensure the privacy of the patient. In addition to this, there are multiple functional areas within the industry that contributes unique sets of insights to understand the mechanisms of progression of disease. This is making data management increasingly difficult. Based on the volumes, varieties and velocities, pharma data sets are termed as BIG DIP.

Dealing BIG DIP in perspective of data mining

The role of data mining methodologies in pharma domain is evolving. The field covers varied powerful tools which provide better manipulation of the data and help the pharma sector compete on lowering the overall costs of the treatment while improving the quality of drug discovery and delivery methods. DM techniques are also used to support the clinicians at the point of care delivery, controlling of clinical treatment pathways and efficient management of organizational and financial data. Visual analytics is a new interdisciplinary field of study which provides a framework for understanding the large volumes of data and it brings the most effective way to present and understand large sets of data.

Answering the basic questions through novel journey of discovery

The introduction of electronic medical records has helped in the transformation of clinical data into "liquid" usage, and further simplified mining the data for specific queries and needs. EMR's have changed the way we understand and map the patient journey from diagnosis through adherence of treatment and to the desired or undesired endpoint. On integration of these data sets can generate evidence based patient journeys by disease, intervention or product that will answers key research questions.

Transforming RWE into business insights and actions

BIG data from real world provides the necessary foundation for greater insight in less time, shaping a more effective complementary process of "value" evidence generation' both in and outside the clinical drug development process.

BIG DIP in the form of numerous datasets is omnipresent due to the nature of the healthcare industry in general. The real challenge is not the lack of data, but the awareness of where this data exists and establishing frameworks to mine these and convert the data into visually appealing and digestible formats.

Uttam Kumar Barick

REFERENCES

[1] Big data harnessing a game-changing asset, a report from the Economist Intelligence Unit limited. [On-line] 2011. Available at WWW: http://www.sas.com/resources/asset/SAS_BigData_final.pdf
[2] McGuire T, Manyika J, Chui M. Why big data is the new competitive advantage. [On-line] 2012. Available at WWW: http://iveybusinessjournal.com/publication/why-big-data-is-the-new-competitive-advantage/
[3] McClearn C, Croisier T. Big pharma's market access mission. [On-line] Delotte University Press, 2013. Available at WWW: http://d2mtr37y39tpbu.cloudfront.net/wp-content/uploads/2013/08/DUP436_Big_Pharma2.pdf
[4] Brown B, Chui M, Manyika J. Are you ready for the era of 'Big Data'? [On-line] McKinsey Quarterly, McKinsey Global Institute, 2011. Available at WWW: http://www.mckinsey.com/insights/strategy/are_you_ready_for_the_era_of_big_data
[5] MedTech Media. Leveraging Big Data and Analytics in Healthcare and Life Sciences: Enabling Personalized Medicine for High-Quality Care, Better Outcomes. [On-line] Report on the Intel Healthcare Innovation Summit, 2012. Available at WWW: http://www.intel.in/content/dam/www/public/us/en/documents/white-papers/healthcare-leveraging-big-data-paper.pdf

[6] Berger M, Axelsen K, Subedi P. The Era Of Big Data And Its Implications For Big Pharma [On-line] Health Affairs Blog, 2014. Available at WWW: http://healthaffairs.org/blog/2014/07/10/the-era-of-big-data-and-its-implications-for-big-pharma/
[7] Guterman J. Release 2.0: Issue 11. Big Data. O'Reilly Media: Sebastopol CA 2009.
[8] Garg S. The New Frontier for the Pharmaceutical and Life Sciences Industry: Real Big Value from Big data. [On-line]. Available at WWW: http://www.tcs.com/SiteCollectionDocuments/White%20Papers/Pharmaceutical-Industry-Big-Data-1113-2.pdf
[9] Schultz T. Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle. Bull Am Soc Inform Sci Technol 2013; 39(5): 34–40.
[10] Informatica executive brief. Big data for the pharmaceutical Industry – leveraging data for greater innovation and improved outcomes. [On-line] 2013. Available at WWW: http://www.informatica.com/Images/02341_big-data-pharmaceutical-industry_eb_en-US.pdf
[11] Thomson Reuters. Big Data and the Needs of The Pharma Industry. [On-line]. Available at WWW: http://blog.thomsonreuters.com/index.php/big-data-and-the-needs-of-the-pharma-industry/
[12] Zikopoulos P, Eaton C. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. [On-line] 2011. Available at WWW: http://dl.acm.org/citation.cfm?id=2132803
[13] Beyer MA, Laney D. The Importance of 'Big Data': a Definition. Gartner, 2012.
[14] Larkin M. Social media for pharma – an expert's view: Even with risks and regulations, pharma needs to engage with patients and learn more about them. [On-line] Elsevier Connect, 2014. Available at WWW: http://www.elsevier.com/connect/social-media-for-pharma-an-experts-view
[15] Deshpande MS, Thakare DV. Data mining system and applications: A review. Int J Distrib Parallel Syst 2010; 1: 32–44.
[16] Ranjan J. Applications of data mining techniques in the pharmaceutical industry. Technol J Theor Appl Inf 2005; 3(4): 61–67.
[17] Hauben M, Madigan D, Gerrits CM, Walsh L, van Puijenbroek EP. The role of data mining in pharmacovigilance. Exp Opin Drug Saf 2005; 4(5): 929–948
[18] Almenoff J, Tonning JM, Gould AL, Szarfman A, Hauben M, Ouellet-Hellstrom R, et al. Perspectives on the use of data mining in pharmacovigilance. Drug Saf 2005; 28: 981–1007.
[19] Poluzzi E, Raschi E, Piccinni C, de Ponti F. Data mining techniques in pharmacovigilance: analysis of the publicly accessible FDA adverse event reporting system. In: Karahoca A (ed). Data Mining Applications in Engineering and Medicine, InTech, 2012: 267–301.
[20] Ranjan J. Data mining in pharma sector: benefits, Int J Health Care Qual Assur 2009; 22(1): 82-92.
[21] Padhy N, Mishra D, Panigrahi R, et al. The survey of data mining applications and feature scope. [On-line] ArXiv Prepr. ArXiv12115723, 2012. Available at WWW: http://arxiv.org/abs/1211.5723
[22] Komenda M, Schwarz D. Visual Analytics in Environmental Research: A Survey on Challenges, Methods and Available Tools. In: Hřebíček J, Schimak G, Kubásek M, Rizzoli AE (eds). Environmental Software Systems: Fostering Information Sharing. Springer: Heidelberg 2013: 618–629.
[23] Saffer JD, Burnett VL, Chen G, van der Spek P. Visual analytics in the pharmaceutical industry. IEEE Comput Graph Appl 2004; 24: 10–15.
[24] Thomas J, Kielman J. Challenges for visual analytics. Inform Visual 2009; 8(4): 309–314.
[25] Sun GD, Liang RH, Liu SX. A survey of visual analytics techniques and applications: State-of-the-art research and future challenges. J Comput Sci Technol 2013; 28(5): 852–867.
[26] Parkinson J, Muthas D, Clark M, Boyer S, Valentin JP, Ewart L. Application of Data Mining and Visualization Techniques for the Prediction of Drug-Induced Nausea in Man. Toxicol Sci 2012; 126: 275–284.
[27] Coulter DM, Bate A, Meyboom RHB, Lindquist M, Edwards IR. Antipsychotic drugs and heart muscle disorder in international pharmacovigilance: data mining study. BMJ 2001;322: 1207–1209.
[28] Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel Data-Mining Methodologies for Adverse Drug Event Discovery and Analysis. Clin Pharmacol Ther 2012; 91: 1010–1021.
[29] Fisher B, Green TM, Arias-Hernández R. Visual Analytics as a Translational Cognitive Science. Top Cogn Sci 2011; 3: 609–625.
[30] Kayyali B, Knott D, van Kuiken S. The Big data revolution in US healthcare: Accelerating value and innovation. [On-line] 2013. Available at WWW: http://www.mckinsey.com/insights/health_systems_and_services/the_big-data_revolution_in_us_health_care
[31] Gupta M, Seiter S, von Allmen H. Jaffe H. The patient journey re-envisioned. The Magazine of Pharmaceutical Business and Marketing 2012; 31(4).
[32] Marr B. How Big Data Is Changing Healthcare. [On-line] 2015. Available at WWW: http://www.forbes.com/sites/bernardmarr/2015/04/21/how-big-data-is-changing-healthcare/
[33] Epstein RS, Sidorov J, Lehner JP, Salimi T. Integrating scientific and real-world evidence within and beyond the drug development process. Journal of Comparative Effectiveness Research 2012; 1(1s): 9–13.
[34] Kelly L. Defining Real World Data: An Interview with Toby Leete and Dr. Ekta Sood. [On-line] Available at WWW: http://social.eyeforpharma.com/market-access/defining-real-world-data-interview-toby-leete-and-dr-ekta-sood
[35] Taylor N. Special Report: Analyzing Real-World Data for Lifecycle Management. [On-line] Available at WWW: http://www.fiercebiotech.com/offer/analyzingdata

PRACTICAL FULLTEXT SEARCH IN MEDICAL RECORDS

Vít Volšička[1, 2*], Milan Blaha[1, 2]
[1] *Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Czech Republic*
[2] *Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic*
[*] *Corresponding author: vit.volsicka@uzis.cz*

ARTICLE HISTORY
Received 28 May 2015
Revised 23 June 2015
Accepted 24 June 2015
Available online 7 July 2015

KEYWORDS
fulltext
apache solr
search
medical records
sql
information retrieval

ABSTRACT — *Performing a search through previously existing documents, including medical reports, is an integral part of acquiring new information and educational processes. Unfortunately, finding relevant information is not always easy, since many documents are saved in free text formats, thereby making it difficult to search through them. A full-text search is a viable solution for searching through documents. The full-text search makes it possible to efficiently search through large numbers of documents and to find those that contain specific search phrases in a short time. All leading database systems currently offer full-text search, but some do not support the complex morphology of the Czech language. Apache Solr provides full support options and some full-text libraries. This programme provides the good support of the Czech language in the basic installation, and a wide range of settings and options for its deployment over any platform. The library had been satisfactorily tested using real data from the hospitals. Solr provided useful, fast, and accurate searches. However, there is still a need to make adjustments in order to receive effective search results, particularly by correcting typographical errors made not only in the text, but also when entering words in the search box and creating a list of frequently used abbreviations and synonyms for more accurate results.*

BODY

Performing a search through text documents is an integral part of the educational process and acquiring new information. Unfortunately, finding relevant information is not always easy, since many documents are saved in free text formats, thereby making it difficult to search through them.

We face the same problem when attempting to scan medical reports, which contain relevant information about the health conditions and treatment of certain patients. These reports are most often saved in free text formats, without any inner structure and parametric values that could facilitate a search. There may be good reason to, at least partially, parameterise and categorise textual medical records. For example, this type of information would help to find relevant groups of patients for clinical studies, analyse treatment methods for certain groups of patients, obtain larger statistics, or find real cases for educational purposes.

FULL-TEXT INDEX

Medical records are usually stored in relational databases, such as unstructured text strings. The basic method, an alias for a naive approach to search in text string databases, means combing through each record, word by word and comparing it precisely with the query. It is not the fastest and best solution, especially when a language is as morphologically complicated as the Czech language. The performance of such a procedure, implemented onto the database, depends on the number of records. With their growing number, the time needed to scan the records increases linearly, which is not very effective. Scanning a large volume of data can easily exceed the time for which the user is willing to wait for the result. There is a need to choose a faster and more effective method. This method is called a full-text search, or an inverted index. It stores words into an index with links to the documents in which these words occur. It is much faster due to a smaller volume of scanned data. Inverted index is also a favourable possibility for inputting complex queries containing more words, with an option to add specifying logical operators.

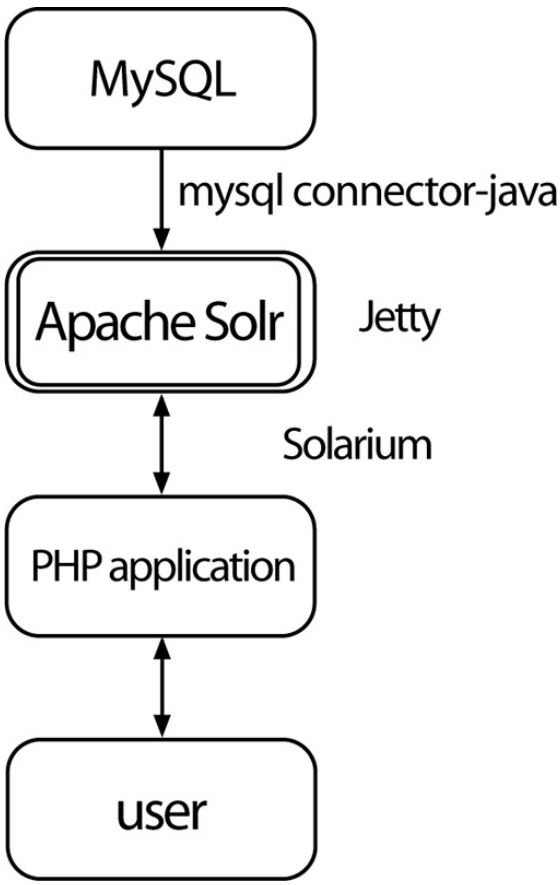

FIGURE 1 Scheme of connection

FIGURE 2 Frontpage

FIGURE 3 Results of search with highlighted keywords

MORPHOLOGY, NORMALISATION AND METHODS

We found out how to solve the performance problem, but not the problem with language morphology, which could be complex and tricky, especially with the Czech language. Words in the text might be found in various forms. Standard search looks for the same form of the word as if it was assigned in the query. If there is not a perfect match, the word is ignored. For example, when we insert the word "lékař" (doctor), we expect to receive records that contain additional words, such as "lékařův", "lékařovo", "nelékařský" etc., therefore finding all forms of word, including variants with various prefixes, suffixes, and plurals.

The next challenge is to deal with diacritical marks found in some languages. Some users write words without diacritical marks, and the computer does not recognise that it is the same word. It is possible to remove diacritical marks entirely, but this could increase the number of irrelevant results.

As we mentioned above, the scanned text should contain various forms of words, which further complicate searches. The solution to this problem is to transform words into their basic forms. This process is called normalisation and we apply it both to the records and to the given query. Consequently, the number of successfully returned words will increase, while the size of the index will decrease due to a smaller amount of indexed words. We will introduce several ways of normalisation.

Lemmatisation

Lemmatisation is the process of transforming a work to the grammatically correct basic form, lemma. For example, the word "vyléčil" (cured) should be transformed to "léčit" (to cure). Homonyms are a weak point in this method, however. Homonyms are words with the same form but with different meanings. If we want to choose the correct meaning of the word, we would have to make a semantic analysis of the whole sentence. But the cost of this process is very high, so it is rarely used. Therefore, the correct approach is to return to all forms, which yields an even lower accuracy.

Stemming

Stemming is the process of eliminating prefixes and suffixes, so the word is reduced to its basic form, or the root. The stem might not necessarily be the same as the morphological root, and might have no meaning. This is the one difference from lemmatisation, which returns a valid word. For example, the root of word "léčit" is "léč." According to the complexity of

the language, conversion into the stem of the word can be algorithmised with a certain quality. This has an advantage over lemmatisation, which needs footing in the base form of a specific vocabulary in order to function. We face two kinds of mistakes when applying the stem algorithm. The first problem is the over-stemming that occurs when two different words are transferred to the same stem. This is considered a false positive error. The second problem is understemming, which occurs when two words that should have the same stem actually have different stems. This is called a false negative error [1]. It has been proved that aggressive stemming algorithms reduce false positive errors, but at the expense of an increase in false-negative errors. Aggressive algorithms behave conversely [2].

N-Grams

An n-gram is a contiguous sequence of n items from a given word extracted from a text. The principle of this method consists in the fact that similar words have a large number of the same sequences. The length of a sequence is typically chosen as either 2 (bigrams) or 3 (trigrams). The value of n is appropriate to test for each language, because it may significantly affect the quality of the result. If n is too small, each word is divided into many n-grams and the query yields a large number of irrelevant results. However, when we increase the value of n, the size of the index will grow exponentially [3]. For example, we could divide the word "nemocnice" (hospital) by bigram into the sequence *N, NE, EM, MO, OC, CN, NI, IC, CE, E* and by trigram: **N, *NE, NEM, EMO, MOC, OCN, CNI, NIC, ICE, CE*, E**, where the asterisk means an empty space. The search query is decomposed into the same size of n-grams and afterwards the number of identical parts is compared. The advantage of this technique is that it provides language independence of the text.

Combination of different approaches

We combine several approaches to receive the best quality outcome. One suitable combination for the Czech language is the cooperation of stemming algorithms, Brute Force and Suffix Stripping. The Brute Force method is based on a "look-up" table containing a construction of state forms of words and their roots. The algorithm enters a query into the table and, in conformity, yields the root of the word. The Brute Force approach is criticised for its lower speed of data processing, as well as for its inability to cover the entire language. Due to the number of word forms in the Czech language, it is unrealistic to expect that all of their forms could be captured. Therefore, it is difficult to design the Brute Force algorithm, because we would need to record a large number of words to achieve an acceptable level of accuracy. However, continuously adding words only improves the accuracy. The advantage of this method is that there is zero risk of over-stemming or under-stemming.

An appropriate complement to this purely vocabulary-based method is the Suffix Stripping algorithm, which removes the Brute Force method's criticized properties. Suffix Stripping works on an application's simple rules for removing prefixes and suffixes to obtain the root word [4]. The pitfall of this approach is that it yields unusual situations, such as, in some cases, changing the word root.

RESULTS – FULL-TEXT SEARCH TOOL IMPLEMENTATION IN APACHE SOLR

The aforementioned theoretical solutions are implemented in a platform called Apache Solr, which was primarily intended for efficient full-text searches. It is an open-source application written in Java. This makes Solr a cross-platform, so it can be deployed on most systems. Other advantages include support in choosing the standardisation of words, removing diacritical marks, speed, and a broad range of settings. Solr needs to run its own server, though; fortunately, "virtual" servers (servlet container) – such as Jetty, for example – are an option.

From the above-mentioned methods, Suffix Stripping was chosen as the most appropriate method for the standardisation of Czech words [5]. It has already been implemented into Solr.

Firstly, for proper functioning we must connect Solr to a database or other repository, where a medical centre has stored its medical records. It is advisable to have the records stored in UTF-8 coding for a proper display and to retrieve characters with diacritical marks; however, Solr works with most databases. Interconnection is based on downloading the appropriate file *-connector-java.jar and entering access data into the configuration file.

The second part is to index records from the database. The appropriate configuration file will specify which database field would be indexed, and how it would be processed and normalised. Based on tests on which algorithm achieves the best results when processing a text written in Czech, we found that Built Stemmer was the best candidate [5]. Solr is natively able to process more than 33 languages. In case of the Czech language, it is worth to add a feature for the removal of diacritical marks. When we want to add new entries to the index, we might not start the whole process over again, but run a process called Delta import.

The last part of the process is to install a library called Solarium. Solarium is not absolutely necessary, but makes typing commands in PHP for communicating with Solr faster and easier. The PHP programming language is the most widely used in the web

scripting language, and as Solr, is not dependent on the platform. We use it to create a web application which will provide a user-friendly interface for searching through the indexed records. The appearance of the web application depends on specific requirements. Solr makes calculations based on proprietary algorithm relevance for each document of the given query, and determines the order of results accordingly. The way of ordering can be changed. Found words can also be highlighted in the text.

CONCLUSION

Performing a search in the given text documents, especially in medical reports, it is not a trivial matter. We encounter several issues that need to be resolved, depending on the morphological complexity of the given language. However, these problems can be solved, and the solutions have been already implemented in applications primarily designed for text searching. One of these applications is the Apache Solr, which includes a built-in support for more than 33 languages, including the Czech language. Additional languages can be added by using the library Hunspell, which provides support for 99 languages (http://hunspell.sourceforge.net/). Apache Solr thus offers a suitable solution for searching through medical records. Thanks to its cross-platform nature, the application can be deployed on existing solutions in medical facilities without the need for major interventions to their systems.

The tool has been implemented and successfully used for searching through medical records stored in a MySQL database. Records themselves might be located in different data stores and, due to the support, might be written in many languages. The application has been pilot-tested and used to scan 54,486 written records in Czech. Solr has dealt with the normalisation of words and with the size of databases very well. Solr provides useful, fast, and accurate searches. An example of a successful search is shown in Figure 3. Solr offers a wide range of options and customisations, so it can be easily modified to suit specific requirements.

The described solution for full-text searches supposes that there are syntactically correct medical records in the database, and that the query is correct. However, typographical errors are very common. Likewise, a mixture of Czech and Latin words and expressions, or common use of abbreviations is often used. This semantic part is rather problematic within computer text analysis, but very important for returning relevant records.

One possible solution to correcting typographical errors is to implement an algorithm called the Damerau–Levenshtein Distance. This algorithm counts the number of steps (insertion, deletion, or substitution of a single character, or a transposition of two adjacent characters) that are needed to retrieve another word. Damerau stated that the operations correspond to more than 80% of all human misspellings [6].

A solution for abbreviations and Latin words consists of creating a dictionary of synonyms and translations. Solr already offers this function. For our future work, it will be necessary to build such functionalities and implement the correction of typographical errors.

Vít Volšička

REFERENCES

[1] Jivani AG. A Comparative Study of Stemming Algorithms. Int J Comp Technol Appl 2011; 2(6): 1930-1938.
[2] Paice C. An evaluation method for stemming algorithms. Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval. Springer-Verlag: New York 1994: 43–50. ISBN 0-387-19889-X.
[3] Coetzee D. TinyLex: Static N-Gram Index Pruning with Perfect Recall. Proceedings of the 17th ACM Conference on Information and Knowledge Management. Association of Computation Machinery: New York 2008: 409–418. ISBN 978-1-59593-991-3.
[4] Kumar D, Rana P. Stemming of punjabi words by using brute force technique. Int J Eng Sci Technol 2011; 3(2): 1351–1358.
[5] Sikora R.. Vyhledávání v českých dokumentech pomocí Apache Solr. Masaryk University: Brno 2012. MSc thesis.
[6] Damerau FJ. A technique for computer detection and correction of spelling errors. Commun ACM 1964; 7 (3): 171–176.

SYNTHESIS OF EVIDENCE OF DIAGNOSTIC TESTS AND PREVENTIVE PROGRAMS IDENTIFYING PRE-DIABETES TYPE

Dagmar Tučková[1,2], Miloslav Klugar[1,2*], Jana Marečková[1,2], Jitka Klugarová[1,2]
[1] *Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Olomouc, Czech Republic*
[2] *The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Olomouc, Czech Republic*
[*] *Corresponding author: miloslav.klugar@upol.cz*

ARTICLE HISTORY
Received 1 June 2015
Revised 28 July 2015
Accepted 10 August 2015
Available online 18 August 2015

KEYWORDS
type 2 pre-diabetes
children
diagnostic test accuracy

ABSTRACT — ***Introduction:*** *Type 2 diabetes mellitus (T2D) has become the main type of diabetes in children and it is expected that in countries with high income diabetes it is projected to be one of the leading causes of death by 2030. Another fact is that programs and tests diagnosing pre-diabetes type 2 (T2P-DMC) are missing.*

Methods: *The aim of the paper is to present the steps for the synthesis of the evidence within the brand new type of the systematic review (SR): SR of diagnostic test accuracy (DTA). Using the acronym PIRD it was developed a review question, search strategy and inclusion and exclusion criteria.*

Results: *The initial search was done in two databases (MedLine and Cinahl) with 2 025 results. The second search after the improvement of the sensitivity and the specificity was done in 15 databases with 3 681 results.*

Conclusion: *This methodological paper introduces how to conduct the systematic review protocols of diagnostic test accuracy on the example of T2P-DMC.*

INTRODUCTION

Approximately 347 million of people around the world suffer from diabetes [1]. In 2004, an estimated 3.4 million of people died on the consequences related to fasting high blood sugar [2]. According to the research carried out in Europe [3], type 2 diabetes (T2D) prevalence in children is increasing. Although a very high prevalence of T2D has been observed in non-Caucasian groups (African Americans, Native Americans, Hispanics), T2D occurs in all races [4]. In the SEARCH study [5], the incidence rate (per 100,000 person-year) of T2D among children and adolescents varies greatly by ethnicity, with the highest rates observed among youths aged 15–19 years in minority populations. In particular, the reported incidence rate was 49.4 for Native Americans, 22.7 for Asian/Pacific Islanders, 19.4 for African Americans, 17 for Hispanics, and 5.6 for non-Hispanic whites. The increased prevalence of T2D in the obese paediatric population is paralleled by an increased prevalence of the prediabetes conditions. In particular, 25% of children and 21% adolescents with severe degree of obesity, irrespective of ethnicity, were found to have impaired glucose tolerance (IGT) [6]. The global rise of childhood obesity and physical inactivity is widely believed to play a crucial role. Healthy eating and lifestyle habits are a strong defence against the disease [7].

The effective pre-diabetes diagnostic tests, early diagnosis and preventive programs can help diabetes prevention development. The main problem in children is that there exist some recommendations regarding the diagnostic tests and preventive programs for pre-diabetes and diabetes. These have been made by the American Diabetes Association (ADA) [2], but formal screening is infrequent [3]. The diagnostic tests and preventive program options include fasting plasma glucose (FPG) and oral glucose tolerance tests (OGTTs), but both require fasting samples that the patients find inconvenient [4], and it is not clear what the best diagnostic strategy is—another reason why the screening is under-performed [5]. Childhood obesity epidemic brings a need to implement test practices for the paediatric population. The effective test practices for type 2 pre-diabetes mellitus in children (T2P-DMC) will allow us to deal with this disease in

the earliest stage of the ontogenetic development and to improve public healthcare in the developed countries around the world.

The main aim of the paper is to present the design of the synthesis of the diagnostic tests for the identification of T2P-DMC followed by intervention to prevention of subsequent onset of T2D. For better evidence we decided to focus on diagnostic test accuracy (DTA) which is a new developed methodology in the field of evidence synthesis [8]. It shows the propaedeutic of such a systematic review (SR) conducting DTA which are used to identify the presence, or the absence of a condition for the purpose of developing an appropriate treatment plan [9].

Before the initial search (the part of protocol development) the preliminary search was conducted in four databases (MedLine, Prospero, JBI Library and Cochrane library) to find out if there are any existing SR of this topic. There was not find any SR or a guideline related to the issue of DTA in T2P-DMC.

METHODS

The review question for our research is: Which diagnostic test is currently the most accurate in identifying T2P-DMC at the different stage of ontogenetic development?

PICO acronym was for SR of DTA replaced by more suitable acronym PIRD where P = population, problem, patient; I – index test; R – reference test; D – diagnosis of interest.

In primary studies of DTA, the test of interest (the 'index test') is compared to an existing diagnostic test (the 'reference test'), which is known to be the best test currently available for identifying accurately the presence or absence of the condition of interest [8]. This PIRD acronym is according to JBI methodological approach more suitable than PICO and it was developed peculiarly for DTA methodology. It covers all parts of what we want to find in the literature. As it was said in the introduction, the diagnostic tests for detection of T2P-DMC are not standardized. That was the reason the reference tests were not described within the PIRD which was determined [8,10].

Our inclusion/exclusion criteria in PIRD acronym:

Population – children with the risk of overweight, obesity, hypokinesis and metabolic syndrome at different stages of ontogenetic development. According to International Diabetes Federation [11] (IDF, 2014), the age criteria ontogeny for metabolic syndrome development is following:

- 6–10 years
- 10–16 years
- > 16 years

Index test – studies that evaluate any type of existing pre-diabetes diagnostic practices and programs, for example impaired fasting glucose level of 100–125 mg/dL, impaired glucose tolerance: A plasma glucose level (obtained 2 hours after a 75-g oral glucose challenge) > 140 mg/dL but < 200 mg/dL or haemoglobin A1c level of 5.7–6.4%. The minimum of 3 from 5 major criteria: (obesity determined by waist circumference, hypertension, low HDL levels, elevated triglyceride levels, and glucose intolerance)

Reference test – as we mentioned in the introduction, here does not exist any reference test for children.

Diagnosis of interest – studies that include the following diagnosis of interest: type 2 pre-diabetes mellitus.

For the needs of the initial search was developed search strategy which consists of the following terms/key words. The Table 1 shows the key words and terms which were used for MedLine database. This search strategy was adapted for each database.

The extensive systematic search strategy aims to find published and unpublished studies within the sources of both scientific literature and so called 'grey

TABLE 1. Key words used for initial search in MedLine database

1.	Children OR teenager* OR kids OR non adults OR early ontogenetic stages OR youngster, adolescent* OR youth
2.	Fasting glucose level OR fasting blood glucose level OR fasting plasma glucose OR FPG
3.	Impaired glucose tolerance: A plasma glucose level OR IGT
4.	H$moglobin A1c OR glycoh$moglobin A1c OR A1c h$moglobin OR Hb OR Hgb OR HbA1c OR HGBA1C
5.	Obesity OR overweight OR BMI + percentile
6.	Hypertension, high blood pressure
7.	Low HDL levels, elevated triglyceride levels, TG, triacylglycerol, TAG, or triacylglyceride
8.	Glucose intolerance
9.	Type 2 pre-diabetes mellitus OR T2P-DMC
10.	Metabolic syndrome OR metabolic syndrome X OR cardiometabolic syndrome OR syndrome X OR insulin resistance syndrome OR Reaven´s syndrome OR CHAOS
11.	Diagnosis OR detect OR accura* OR diagnostic accuracy test OR diagnostic accuracy clinical tests OR test of diagnostic accuracy OR diagnostic accuracy OR sensitivity OR specificity OR ROC
12.	Animal*
13.	1 AND (2 OR 3 OR 4) AND (5 OR 6 OR 7 OR 8) AND (9 OR 10) AND 11 NOT 12

TABLE 2. QUADAS-2 signalling questions [8]

Critical appraisal questions
Domain 1: Patient selection
• Was a consecutive or random sample of patients enrolled?
• Was a case-controldesign avoided?
• Did the study avoid inappropriate exclusions?
Domain 2: Index test
• Were the index test results interpreted without knowledge of the results of the reference standard?
• If a threshold was used, was it pre-specified?
Domain 3: Reference test
• Is the reference standard likely to correctly classify the target condition?
• Were the reference standard results interpreted without knowledge of the results of the index test?
Flow and timing
• Was there an appropriate interval between the index test and reference standard?
• Did all patients receive the same reference standard?
• Were all patients included in the analysis?

literature', government reports, dissertations, editorial etc. The search strategy used in this SR will include three steps according to JBI methodology [12]. Initial search has been done in MedLine and Cinahl database. This phase will be followed by two other steps. The next one will be to create the search strategy for each specific database included in the protocol. Each database has its own specific vocabulary/dictionary (engine). The third step will involve the review of the reference list of all studies that are retrieved for appraisal to search for additional studies.

Studies, which will be found, will be assessed by two independent reviewers in terms of relevance according to specified criteria. Then the standardized critical appraisal tools of JBI-DATARI will be used for critical appraisal done again by two independent reviewers. For the reviews of DTA we will use critical appraisal sheet using QUADAS-2 signalling questions [8].

The critical appraisal questions will be answered as "yes", "no", or "unclear". When this will be not applicable "not applicable". How much weight is placed on specific critical appraisal questions will vary between reviews and it is up to the reviewers to set what criteria, if any, will result in the exclusion of a study from the review. Many reviewers specify a set of questions which must be answered "yes" or the study will be excluded. It is important that these criteria will be applied consistently across studies [8]. Any disagreements that arise between the reviewers both during paper retrieval and during critical appraisal

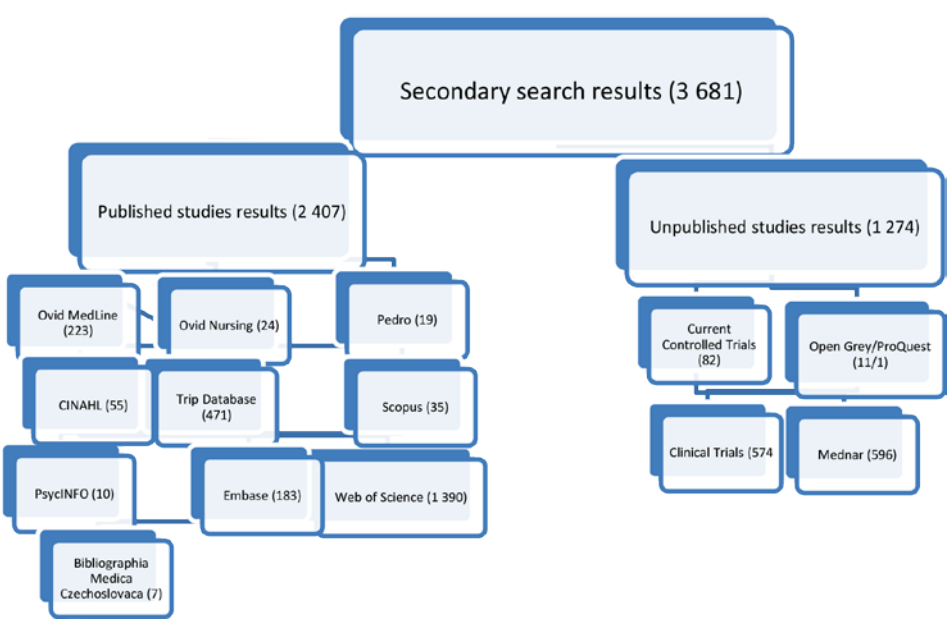

FIGURE 1. Second search results

will be resolved through a discussion, or with a third reviewer.

Next step will be to make data extraction. The decision threshold that was used to classify the results as positive or negative is an item of data unique to studies of DTA. For data extraction will be used a JBI tool called DATARI which is based on the standards for the reporting of diagnostic accuracy studies (STARD) checklist. All studies of DTA that comply with the STARD statement will include a 2×2 table with sensitivity, specificity, and predictive values that classify patient test results and disease status [8].

Data synthesis will be done after data extraction. The results of SR of DTA can be graphically represented through two different major ways. The first way is the use of forest plots. However, in order to present data on DTA "paired" forest plots must be used where two forest plots are presented side by side; one for sensitivity and the other for specificity. In this way these graphs display the means and confidence intervals for sensitivity/specificity for each of the selected primary studies [8].

RESULTS AND DISCUSSION

Results should bring the synthesis of findings of DTA of existing practices identifying T2P-DMC. So far the review question was developed by PIRD acronym. Specific criteria for inclusion and exclusion of studies were determined also by using of PIRD acronym. Key words for the search strategy were found and set into the searching strategy process. After we found out that there is not any existing SR for this topic (in preliminary search) we made initial an systematic literature search which was performed in MedLine and Cinahl databases. This search was done to verify whether the determined search strategy using determined key words and terms have balanced sensitivity and specificity. After this search, we got the initial search result of 2 025 studies – 958 from MedLine, and 1 067 from Cihnal. This was followed by secondary search. For this secondary search specificity was improved in search strategy and adjusted for every particular database. The second step of the search identified 3 681 studies in 15 databases. 2 407 published results and 1 274 unpublished results were found. As it is shown in the Figure 1, most published results were found in WoS (1 390), Trip Databas (471), Ovid MedLine (223), and Embase (182). Other results were found in Cinahl (55), Scopus (35), Ovid Nursing (24), Perdo (19), PsycINFO (10), and Bibliographica Medica Czechoslovaka (7). In unpublished databases we found 596 in MedNar, 574 in Clinical Trials, 82 in Current Controlled Trials, 11 in Open Grey, 1 in ProQuest. In COS Conference Papers was not found any result.

These are just the partial results. Another step will follow those in the process of SR development. This phase will be followed by selection of relevant studies, which will have two phases, a title and abstract screening and full texts screening. Then critical appraisal of relevant studies will be done, followed by data extraction of high quality studies and data synthesis.

CONCLUSION

The use of newly developed JBI methodology of DTA will help to analyse the situation in research field of interest. This approach will pool and synthesize relevant data which will be used for practice, healthcare policy and other stakeholder's information.

SR provides the highest level of scientific evidence. The issue of T2P-DM is currently one of the most discussed topics because of children's life style. And if there exist DTA helping to detect T2P-DM in adults this could work for children as well with the respect of their ontological development and metabolic rates. With early identification of T2P-DM using diagnostic tests it will be possible to prevent T2P-DM in children and teenagers. This could represent one of the solutions of this medical issue and become prevention for this disease. The consequences can further have the positive effect and the positive results could be expected in a long-term run in many fields of health and healthcare.

This paper was focused on presentation of brand new methodology and it shows the process of DTA SR development. It can be used as an initial inspiration for students of medical faculties doing their own research in the field of SR. As well as it can be pointed out that ability to develop the SR should be part of medical professional's skills. There should be a consideration to integrate subject aiming to SR development training into a curricula at the medical faculties.

Dagmar Tučková

ACKNOWLEDGEMENTS

The project was supported by the research grant IGA_LF_2015_024 (Faculty of Medicine and Dentistry, Palacký University in Olomouc).

REFERENCES

[1] Danaei G, Finucane MM, Lu Y, Singh, MM, Cowan MJ, Paciorek CJ, Stevens GA. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. Lancet 2011; 378 (9785): 31–40. ISSN 0140-6736
[2] WHO Global health risks: mortality and burden of disease attributable to selected major risks (1st ed.). World Health Organization: Geneva 2010.
[3] Haines L, Wan KC, Lynn R, Barrett TG, Schield JP. Rising incidence of type 2 diabetes in children in the U.K. Diabetes Care 2007; 30(5): 1097–1101.
[4] Rosenbloom AL, Silverstein JH, Amemiya S, Zeitler P, Klingensmith GJ. Type 2 diabetes in children and adolescents. Pediatr Diabetes 2009; 10(12): 17–32.
[5] Writing Group for the SEARCH for Diabetes in Youth Study Group, Dabelea D, Bell RA, D'Agostino RB Jr, Imperatore G, Johansen JM, Linder B, Liu LL, Loots B, Marcovina S, Mayer-Davis EJ, Pettitt DJ, Waitzfelder B. Incidence of diabetes in youth in the United States. JAMA 2007; 297(24): 2716–2724.
[6] Sinha R, Fisch G, Teague B, Tamborlane WV, Banyas B, Allen K, Savoye M, Rieger V, Taksali S, Barbetta G, Sherwin RS, Caprio S. Prevalence of impaired glucose tolerance among children and adolescents with marked obesity. New Engl J Med 2002; 346(22): 1756–1756.
[7] World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organization: Geneva 2010.
[8] Campbell J, Klugar M, Ding S, Carmody DP, Hakonsen, SJ, Jadotte YT, Munn Z. Diagnostic test accuracy systematic review: The Joanna Briggs Institut's approach. International Journal of Evidence-Based Healthcare. In press.
[9] White S, Schultz T, Enuameh Y. Synthesizing evidence of diagnostic accuracy. Lippincott Williams and Williams: Philadelphia 2011. ISBN 978-1-451163-88-9.
[10] Campbell J, Klugar M, Ding S, Carmody D, Hakonsen S, Jadotte Y, et al. The Systematic Review of studies of diagnostic test accuracy. In: JBI, editor. Joanna Briggs Institute Reviewers' Manual: 2014 edition/supplement. The University of Adelaide: Adelaide 2015. In press.
[11] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2013; 36(Suppl. 1): S67–S74.
[12] The JBI. Joanna Briggs Institute Reviewer´s Manual: 2014 edition. University of Adelaide: The JBI 2014. ISBN 978-1-920684-11-2

VIRTUAL CASES IN INTERNAL MEDICINE EDUCATION

Ilja Tachecí[1]*, Aleš Ryška[2]
[1] *2nd Department of Internal Medicine – Gastroenterology, University Hospital Hradec Králové and Faculty of Medicine in Hradec Králové, Charles University in Prague, Hradec Králové, Czech Republic*
[2] *Fingerland Department of Pathology, University Hospital Hradec Králové and Faculty of Medicine in Hradec Králové, Charles University in Prague, Hradec Králové, Czech Republic*
* *Corresponding author: tacheci@gmail.com*

ARTICLE HISTORY
Received 14 June 2015
Revised 11 August 2015
Accepted 11 August 2015
Available online 26 August 2015

KEYWORDS
MEFANET
gamification of medical education
serious games
virtual cases
problem-based learning
scenario-based learning
students' attitudes
evaluation
interactive training

ABSTRACT — *Virtual patients represent a useful tool in teaching students clinical reasoning skills. Virtual cases (www.e-kazuistiky.cz) represent a newly developed interactive problem-based learning system, drawing information from virtual clinics, covering different fields of internal medicine, generating sets of unique virtual patients according to user-predefined program settings (spectrum of diagnoses, number of patients and criteria for passing the course). Basic clinical information including personal data, medical history, symptoms, laboratory values, etc. is generated for each virtual patient. The main task for the student is to determine the optimal diagnostic algorithm (choose adequate diagnostic steps in the correct order), and to determine the correct diagnosis in each virtual patient. Results of diagnostics tests and clinical findings are presented utilising a multimedia presentation (images, video-sequences, audio-recordings). Evaluation of students includes not only assessment of correctly determined diagnosis, but also the diagnostic pathway, which led the user to the specific diagnosis. Thus, the system enables assessment of appropriateness of each test as well as reasonable sequencing of tests and also financial costs of all examinations.*

The program is now routinely used in the undergraduate curriculum at the Medical Faculty in Hradec Králové. User hands-on experience was evaluated through anonymous questionnaires. The most appreciated attribute of the system is the game-like involvement and multimedia-supporting environment (for students) as well as the possibility of a detailed analysis of each student's performance and clear identification of their weakest areas (for tutors).

The system is a useful tool for undergraduate medical education with positive feedback from both students and teachers. The main advantages are flexibility, potential for further growth and no restrictions regarding particular disease, clinical discipline, diagnostic procedure, etc.

INTRODUCTION

Undergraduate education in internal medicine is based on two main pillars – theoretical education during lectures and seminars and practical training at the patient's bedside. Teaching at the bedside remains an essential, integral part of the internal medicine curriculum and plays a crucial role in teaching and assessing clinical reasoning to medical students.

However, due to objective reasons, it may face a series of issues and limitations. The optimal organisation of practical education is focused on presentation of patients, revealing a wide spectrum of both common and rare symptoms, syndromes or diseases and their variants. The spectrum of demonstrated diagnoses is highly influenced by several factors in real educational practice – due to specialization of the given department, limited number of patients who are present at the department during the time of practical training and who agree to be examined by students and by the actual condition of the particular patient. Another important factor influencing the composition of presented cases is the teacher. Some teachers tend to present more often rare diseases or symptoms, with a limited or minimal impact on the future professional life of doctors outside tertiary centres. These teachers frequently underestimate the presentation of common (and thus from a practical point of view more important) cases. Presentation of patients is often limited by the duration of practical classes too. Thus, students are missing the experience regarding development of disease features from initial symptoms down to full manifestation.

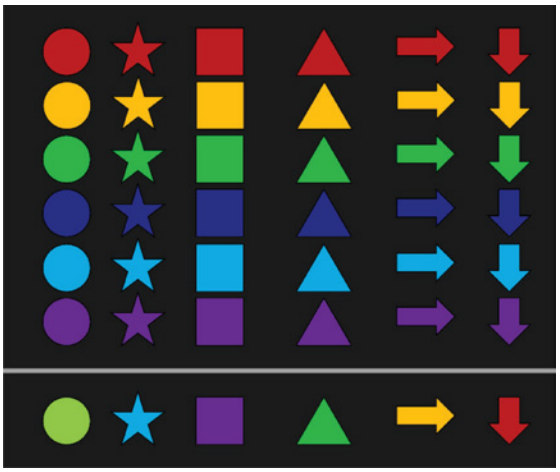

FIGURE 1. Pseudo-random combinations lead to a higher number of cases available for teaching/learning within the project

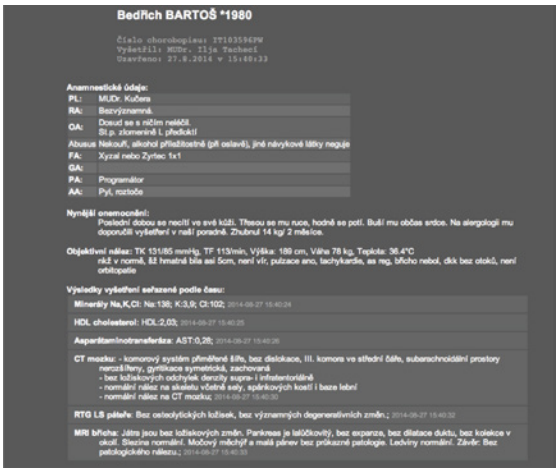

FIGURE 2. Patient history and clinical investigation are presented in the beginning of each case

Another important issue is the lack of diagnostic algorithm training (chaining of rational investigative methods to maximize the diagnostic yield and minimizing the risk of complications in different clinical situations) and the lack of economic knowledge on the part of students concerning costs of each test or diagnostic method.

A potential solution to all of these limitations in clinical reasoning training is a simulation of clinical scenarios utilising information technologies, including so-called serious games concept. Serious games could be defined here as "games" that are designed to entertain players as they educate, train or change behaviour [1]. Clark Abt first defined this term in the early 1970s for non-digital, math-related games by words [2]: "Games may be played seriously or casually. We are concerned with serious games in the sense that these games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement. This does not mean that serious games are not, or should not be, entertaining." The real advancement came with the expansion of computer technology in the last 20 years. In medical education, serious games are developed for non-entertainment purposes and designed primarily for clinical skills development [3]. They offer a good dose of realism together with the entertainment factor of a traditional game. Serious games can be used in medical education in many different ways. One of the most attractive and currently most realistic approaches in the serious games field is the use of so-called "virtual patients" or "virtual case studies". Virtual patients are computer-based scenarios used in simulation of real-life clinical situations in the context of a computer game [4–11].

AIM

The aim of our project was to develop a dynamic database of virtual patients in an outpatient clinic to supplement teaching of internal medicine within a general medicine study program. We wanted to test this environment subsequently during the internal medicine curriculum at the Charles University Medical Faculty in Hradec Králové.

Our newly developed serious games system was named "Virtual cases". Taking into account the fact that there is a substantial difference between computer simulation currently available (text plus multimedia) and the actual "patient simulator" (which should meet much higher requirements for audio-visual and tactile interaction between the student and computer simulator and which has not, according to the authors' knowledge, really been developed so far), we prefer the term "virtual cases" to "virtual patients" as this corresponds better with the technology available. The objective of computer simulation and environment of the program is to show students how to select the optimal diagnostic algorithm and establish diagnosis based on available clinical and laboratory data including a complete medical history and clinical examination results. The goal is thus not to replace the bedside teaching with direct contact with the patient by a computer simulation.

METHODS

The main requirements for the new system were to show clinical cases with as realistic features and real-life clinical data as possible. It was our intention to show a broad spectrum of diagnoses covering both frequent and rare disorders, to minimize the risk of repetition of particular cases, teaching optimal diagnostic algorithms taking into consideration also the economic background of the entire diagnostic procedure, to maximize flexibility of the system and establish a software environment capable of further development and expansion into other clinical disciplines

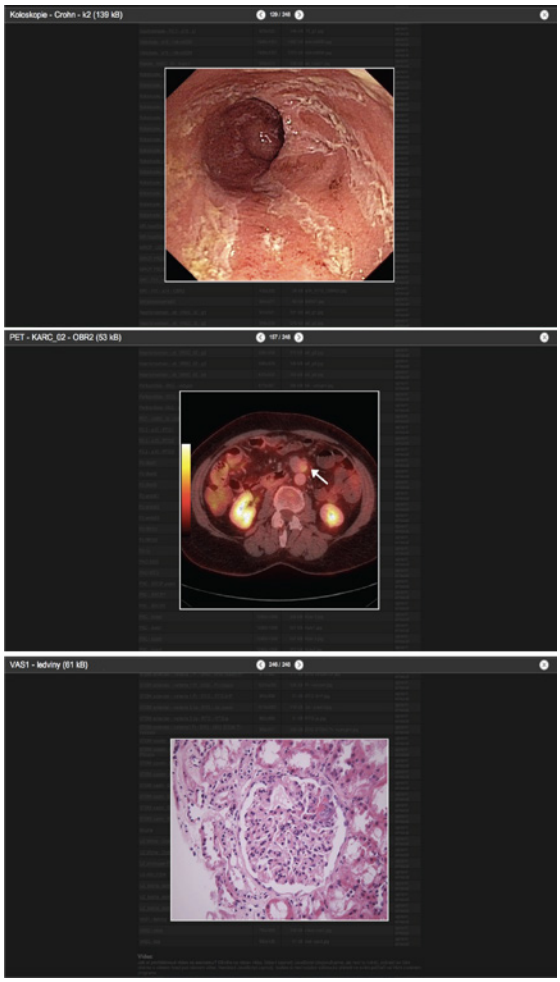

FIGURE 3. Results of diagnostic tests and biopsy are presented in a multimedia manner

without any restrictions regarding the discipline, type of disease or diagnostic procedure.

The main pedagogical objectives of the project were:

- Simulation of real-life diagnostic problems
- Training in rational use of diagnostic procedures
- Automatic evaluation of students' knowledge and skills
- Possibility of re-evaluating students' progress during subsequent analysis of individual cases with the teacher

Virtual cases were prepared as a web-based tool, because web technologies allow effective work with multimedia content and easy access from any computer. Due to the fact, that this is a device intended for the medical education (containing the anonymized data of real patients), access to the system is limited by the necessity of registration, and the approval of the operator.

Preparation of virtual cases from the author's point of view

Each virtual case was prepared in a similar way. The system allows the authors to choose the outpatient clinic (specialization), in which a given case generates, and select the correct diagnosis, which the student should select after completing the concrete case. There were 4–5 similar real patients with one concrete diagnosis identified (e.g. peptic duodenal ulcer) and all data was drawn from the hospital information system. Their anonymous medical case histories (segmented into history of present illness, past medical history including childhood diseases, family diseases, social history, regular medications, allergies and obstetric/gynaecological history in females), objective findings at physical examination and results of accompanying tests (such as laboratory, X-ray, computed tomography, ultrasound, endoscopic examinations, histological biopsy results, etc.) were put together and copied into the back office form (the same procedure for all virtual cases). Because the students have the possibility to select any diagnostic method or laboratory test presented in the system (not only the pathological one or the methods with a particular relationship to the case study), the authors should define the results for all of them.

The virtual cases generation system

The system was made to generate pseudorandom combinations of the above-mentioned segments (from all patients with the same diagnosis) – e.g. history of present illness from patient A, combined with family history of patient B, results of ultrasound of patient C, etc. (Figure 1). Naturally, it was essential to prevent mismatch of certain parameters (e.g. if a virtual case was a male patient, it is not possible to show chest X-ray results of a female patient, as the breasts visible on the X-ray would not match the clinical data). Mixing and matching individual segments of findings from several different patients prevents repetition of clinical histories and thus the user (student) is always facing a new (although substantially very similar) clinical scenario. With this feature, students are not distracted by their memory when going through repeated cases (such as "...a chemical engineer with symptoms starting on a holiday has a peptic ulcer...") and must always carefully read all clinical data and not skip any part of the case. The result of this pseudorandom combination is a unique clinical case (personal data, medical history and basic clinical investigation – current clinical picture for each diagnosis) presented to students (Figure 2).

FIGURE 4. Student results are presented with a conclusion, economy results, bonus and malus points reached in each concrete case

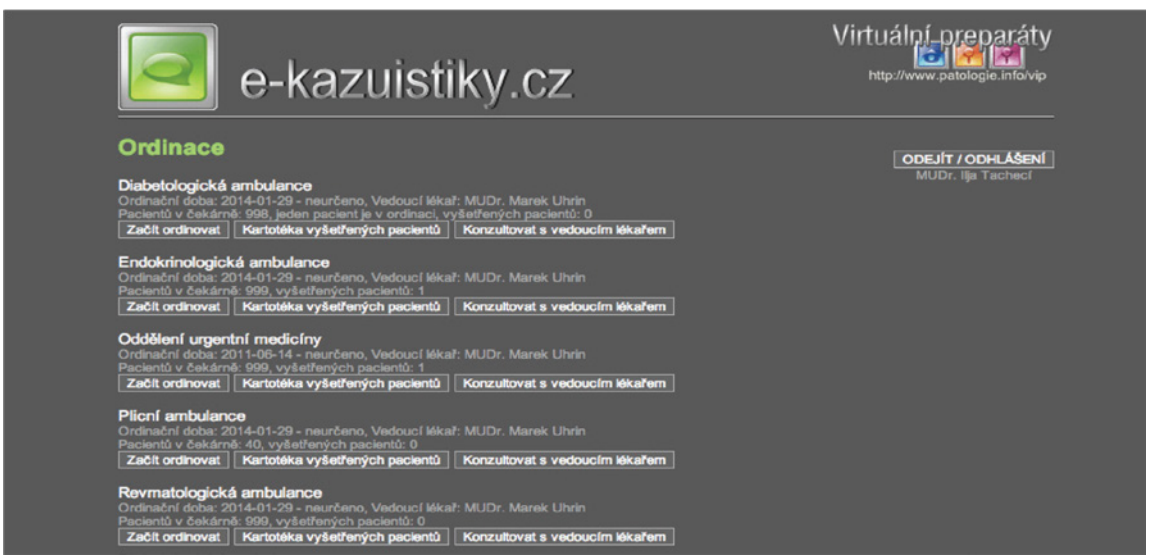

FIGURE 5. The set of outpatient clinics respects the main subspecialisations of internal medicine

User approach

After reading all initial clinical data (history, objective findings, complaints by the patient, etc.), students are requested to choose what tests should be performed. This is done in steps; therefore after seeing results of one test, the student may decide what the next examination will be. An extensive multimedia database enables presenting clinical findings and results of diagnostics in a fully multimedia manner – as text, photographs, videos, virtual histological slides, etc. (Figure 3). Examination results are realistic (we are using the real patients data) and present in addition to their benefits the common limitations of selected tests too (false positive and negative findings, poor quality of examinations, etc.).

Thus, students are expected to choose a diagnostic algorithm. After they gather sufficient data, they proceed with diagnosis. If the diagnosis is based on sufficient input information (the student is not just guessing, but the diagnosis is really evidence-based) and the suggested diagnosis is correct, the case is successfully completed.

Students evaluation system (bonus and malus points, cost and diagnostic methods sequence evaluation)

The unique multi-level system of evaluation of each student is based on scoring each individual step and their sequence in the diagnostic algorithm (Figure 4).

If the tests selected are beneficial for the diagnosis in a given situation, the student receives positive (so-called "bonus") points. If the investigation is without contribution or may even harm the patient, the student receives negative (so-called "malus") points. The optimal (minimal) level of "bonus" points (that must be made before it is possible to establish a definite diagnosis) is determined for each diagnosis by authors (otherwise the conclusion is considered speculative and not supported by a sufficient number of diagnostic tests performed). The maximum level of "malus" points is set for each diagnosis too; the solution of the case is considered unsuccessful and the case is terminated immediately after reaching the maximum malus points level.

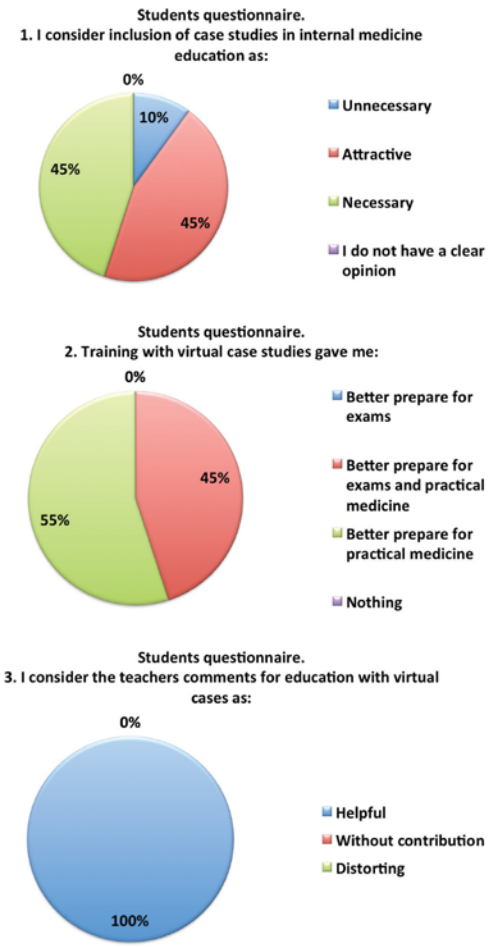

FIGURE 6. Student questionnaire

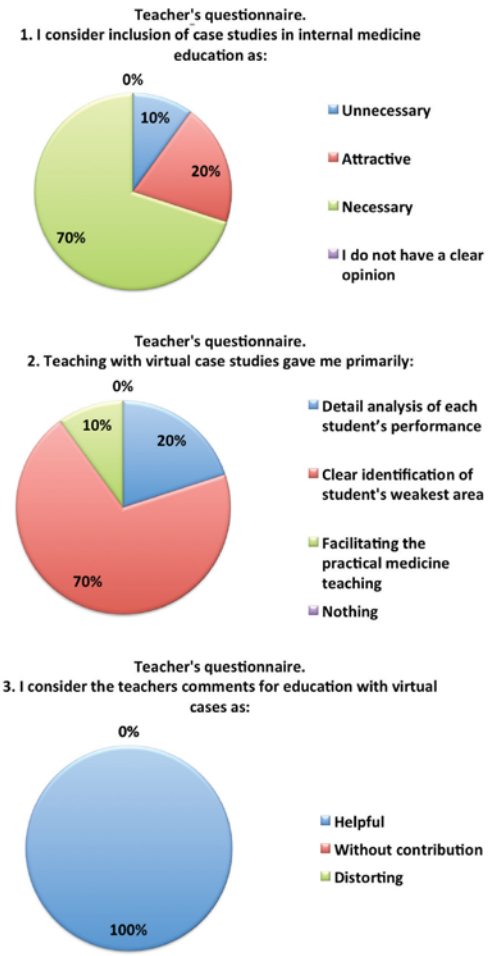

FIGURE 7. Teacher questionnaire

A very important part of Virtual cases is also the effort to show students the diagnostic process economic background. The actual current cost of each test method is included in the system and users have information about the selected diagnostic algorithm costs at every moment. After the maximum cost (determined by the author of the virtual case) is exceeded, the student's attempt is also considered unsuccessful and solution of the virtual case is terminated.

Because students can choose any diagnostic test available without any up-front limitation, it is essential that the system evaluate if the selected test contributes to reaching the final correct diagnosis. An additional key feature of the system is the original evaluation tool allowing assessment of optimal/acceptable/incorrect sequence of selected diagnostic methods.

Virtual cases were newly created in collaboration with teachers of clinical and preclinical disciplines, programmers experienced in clinical medicine and undergraduate medical students who tested the system. Thus, this multidisciplinary team of authors ensures a balance between the (often differing) expectations of teachers, physicians of clinical and preclinical disciplines, programmers and students regarding the ideal educational tool. The programme was used within the undergraduate curriculum at the Medical Faculty in Hradec Králové. The user (student and teacher) hands-on experience was evaluated in several anonymous questionnaires focused on different aspects of the project. Main topics covered in the questionnaires are: user ratings of virtual case studies, evaluation of case studies contribution from the perspective of teachers and students and evaluation of the teachers contribution to courses using virtual case studies.

RESULTS AND DISCUSSION

The team of authors created a basic set of 62 case studies covering the major specializations of internal medicine (gastroenterology, hepatology, cardiology,

endocrinology, hematology, diabetology, rheumatology, acute medicine and pulmonary medicine, as well as paediatrics) – Figure 5.

After the first 20 virtual cases have been published and tested, the system was put to use in the 4th year of undergraduate medical curriculum (internal medicine) at the Medical Faculty of Charles University in Hradec Králové. The quality and user-friendliness of the project as well as hands-on experience were evaluated through anonymous questionnaires completed by teachers and students, resulting in mostly positive feedback (Figures 6 and 7). The most appreciated attribute of the system from the students' perspective is the game-like involvement and multimedia-supporting environment; from the teachers' perspective, it is the possibility of detailed analysis of each student's performance and clear identification of the weakest areas of the tuition system. Teachers can objectively evaluate students' knowledge, their diagnostic approach and the strengths and weaknesses leading to the success or failure of each individual student. Each case solved by students is subsequently discussed at the seminar, where each student has to defend the algorithm of selected tests in front of his fellow students and the tutor, explain what they expected from each individual method and why they selected the particular diagnosis. The teacher explains what the other options were, how the approach could have been more straightforward and effective. Thus, each case does not represent only a benefit for the individual student, but for the entire learning group.

In the next phase of the project, additional 42 cases were added to broaden the spectrum of diagnoses. The system was enhanced by newly developed interactive authors' form, enabling a much easier input of new patients into the system without the need for detailed knowledge of the back-office system. It is expected that this form will help further increase the number of authors who will be willing to contribute with new clinical cases. One additional feature of the virtual cases is also a close link to the system of virtual microscopic slides, which are used in teaching histopathology in the 3rd year of the same medical curriculum. This has been proven over many years as an extremely useful method and has been a tool highly appreciated by students in preparation for both practical classes as well as for final exams in pathological anatomy (http://www.patologie.info/vip). The link to virtual cases further shows students how the knowledge of preclinical disciplines, such as pathology, may be of use in their further medical studies.

The Virtual cases project is specific from different points of view. The program is focused on teaching the formation of rational diagnostic algorithms and establishing the correct diagnosis. The main difference from the previously published tools using the virtual patients is the unique multilevel evaluation system of student's progress [10-11]. The aim of this evaluation is to achieve the maximum possible degree of realism. Automatic scoring of each student's step in the four aspects (1. benefits of the examination for determination of diagnosis – "bonus points", 2. burden of the patient by selected examination – "malus points", 3. economic burden of health care provider – the comparison of the price of the chosen algorithm in comparison with the "ideal" algorithm defined by the author of a case study and 4. evaluation of correctness of the diagnostic methods sequence in the chosen algorithm) allows the elimination of speculative conclusions and exclusion of dangerous, burdensome or unnecessary diagnostic methods from the algorithms. The system of virtual cases generation is innovative due to the use of pseudo-random combination of individual parts of the cases with the same diagnosis too. This approach leads to the higher number of virtual cases available, which is important to prevent false correlation, and allows focusing on all clinical data. Another important aspect of Virtual cases teaching, highly acclaimed by the users, is subsequent discussion of cases under the tutor supervision, allowing to obtain feedback for the whole group of students participating in seminar.

CONCLUSION

The Virtual cases is a newly developed educational system using the principles of so-called serious games for the purpose of undergraduate education of internal medicine. The 62 case studies (covering the main specializations of internal medicine) were created and the program has been successfully included into the undergraduate curriculum of internal medicine at the Charles University, Faculty of Medicine in Hradec Králové. Virtual cases were positively evaluated by all users (teachers and students). After adding new virtual cases, the original idea of having a system, which can fully simulate a spectrum of patients in everyday clinical practice, could be achieved.

Ilja Tachecí

ACKNOWLEDGEMENTS

The project was supported by Operational Programme EU: Education for Competitiveness (Modernizace výuky klinického rozhodování napříč pediatrickými obory lékařských fakult v síti MEFANET) – CZ.1.07/2.2.00/28.0038, and by research grant IGA NT11524-5/2010.

REFERENCES

[1] Stokes B. Video games have changed: time to consider "Serious Games"? Development Education Journal 2005; 11(3): 12.
[2] Abt C. Serious games. Viking Press: New York 1970.
[3] M. Graafland, J. M. Schraagen, M. P. Schijven. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012; 99 (10): 1322–1330.
[4] Berman N, Fall LH, Smith S, Levine DA, Maloney CG, Potts M, Siegel B, Foster-Johnson L. Integration strategies for using virtual patients in clinical clerkships. Acad Med 2009; 84(7): 942–949.
[5] Huang G, Reynolds R, Candler C. Virtual patient simulation at US and Canadian medical schools. Acad Med 2007; 82(5):446–451.
[6] Ellaway R, Poulton T, Fors U, McGee JB, Albright S. Building a virtual patient commons. Med Teach 2008; 30(2): 170–174.
[7] Ellaway R, Topps D, Lee S, Armson H. Virtual patient activity patterns for clinical learning. Clin Teach 2015; 12(4): 267–271.
[8] Kononowicz AA, Zary N, Edelbring S, Corral J, Hege I. Virtual patients – what are we talking about? A framework to classify the meanings of the term in healthcare education. BMC Med Educ 2015, 15: 11.
[9] Bloice MD, Simonic KM, Holzinger A. Casebook: a virtual patient iPad application for teaching decision-making through the use of electronic health records. BMC Med Inform Decis Mak 2014; 14: 66.
[10] Salminen H, Zary N, Björklund K, Toth-Pal E, Leanderson Ch. Virtual patients in primary care: developing a reusable model that fosters reflective practice and clinical reasoning. J Med Internet Res 2014; 16(1): e3.
[11] Štourač P, Harazim H, Schwarz D, Křikava I, Komenda M, Štoudek R, Smékalová O, Kosinová M, Hůlek R, Maláska J, Šustr R, Šnábl I, Dušek L, Gál R. Akutne.cz algorithms and sepsis-q scenarios as interactive tools for problem based learning sessions in medical education. Mefanet J 2013; 1(2): 61–73.

8TH YEAR OF THE MEFANET CONFERENCE: VIRTUAL REALITY IS BREAKING DOWN BARRIERS

Jakub Gregor, Martin Komenda, Daniel Schwarz*, Ladislav Dušek, Lenka Šnajdrová
Institute of Biostatistics and Analyses at the Faculty of Medicine and the Faculty of Science of the Masaryk University, Brno, Czech Republic
** Corresponding author: schwarz@iba.muni.cz*

ARTICLE HISTORY
Available online: 27 April 2014

KEYWORDS
MEFANET
medical informatics
e-learning
virtual patient
e-assessment

ABSTRACT — *With a slight overstatement, one could say that the 8th year of the MEFANET conference reached boundaries of the solar system, which can boast eight planets. In reality, ambitions of the Czech-Slovak cooperation in the area of medical education are more modest, but this year's programme has demonstrated that the international cooperation surely extends beyond the former federal boundary of Czechoslovakia.*

Teachers, students and experts interested in modern methods of medical education met as usual at the end of November in the Hotel International, not far away from two dominants of the City of Brno. The conference was ceremonially opened by Prof Jan Žaloudík, Chairman of the Committee on Health and Social Policy of the Senate of the Parliament of the Czech Republic, and recently the director of the Masaryk Memorial Cancer Institute in Brno, and by Prof Aleš Ryška from the Fingerland's Department of Pathology at the University Hospital Hradec Kralove.

A summary of key activities, recent successes and new challenges for the MEFANET network became a regular, keenly anticipated part of the conference opening. Similarly to the last year, Dr Daniel Schwarz (IBA, MU's Faculty of Medicine) took the floor, presenting a lot of positive information and good news. Namely, dr Schwarz introduced the new appearance of the MEFANET gateway, which transformed the current e-publishing platform into a "web-scale discovery" system enabling a more elaborate indexing, and a more precise search of articles. A recently published book entitled "Computer Applications, Systems and Netwoks for Medical Education. MEFANET: Czech and Slovak Medical Faculties Network" is another piece of good news, pointing to the impact of various forms of e-learning on the education process in different medical specialties, documenting the approaches to a simulation-based education, and last but not least, introducing the MEFANET network itself and its key role in the involvement of information and communication technologies in the education of physicians and other health professionals. Dr Schwarz also highlighted the importance of activities which attract and involve students in the education process, namely AKUTNE.CZ and WikiSkripta. He did not miss out other interesting national and international project linked to the MEFANET network, including the MEFANET Journal, which would complete its second year in December 2014. Several individual sessions of the conference were dedicated to selected activities of the MEFANET project, as mentioned further in this article.

But first, another integral part of the conference needs to be mentioned, i.e. keynote lectures of invited foreign guests. This year, Sokratis Nifakos and Christos Vaitsis (both from the Karolinska Institutet in Stockholm, Sweden) accepted the invitation. Sokratis Nifakos focused on mobile technology and its use in the education of both students and doctors. For example, he presented the so-called augmented reality, which is "a live direct or indirect view of a physical, real-world environment whose elements are augmented (or supplemented) by computer-generated sensory input such as sound, video, graphics or GPS data", as defined by Wikipedia. The prescription of antibiotics and three-dimensional models in cardiology were provided as examples. In the latter lecture, Christos Vaitsis dealt with the processing of big data, which needs to be done when preparing an education programme, teaching materials or electronic testing, as

FIGURE 1 Conference opening by Professor Jan Žaloudík

FIGURE 2 Christos Vaitsis (left) and Sokratis Nifakos participated not only as lecturers, but also as attendants, because English was the official language on the first conference day

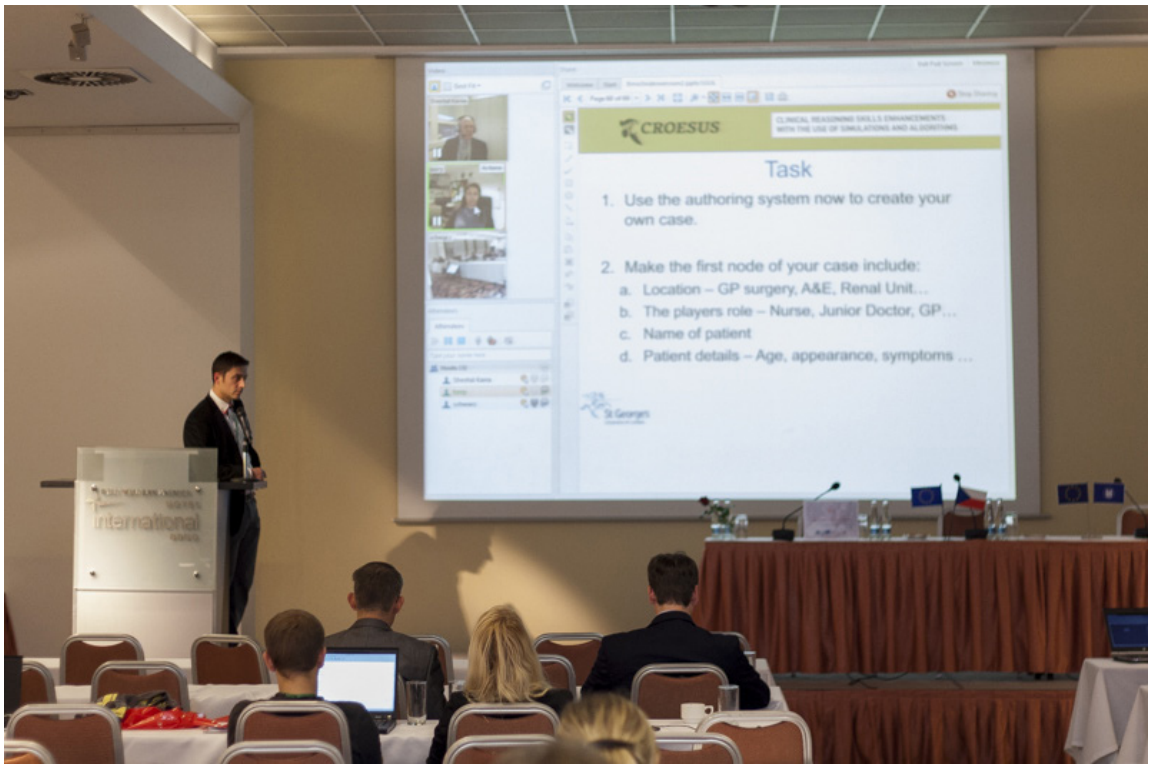

FIGURE 3 Workshop inspired by the CROESUS project, with a teleconference participation of colleagues from London

well as visualisation of such data, based on the so-called "visual analytics" approach. He emphasised several aspects, such as the necessity of providing a balanced amount of information (lack versus oversaturation), the correct definition of roles and targets, or putting oneself in the place of end users when developing any outputs. Development and optimisation of a study plan (whether it be the entire field of study, or a single course) was likened to a path on the map, which involves intermediate and final targets of the education process.

As mentioned above, two conference sections were dedicated to projects which are somehow linked to the MEFANET network. The OPTIMED project (http://opti.med.muni.cz/en) is one of them: this project, which is being solved at the Faculty of Medicine at Masaryk University (MU), aims to innovate the General Medicine field of study in its entirety: both compulsory and optional courses will be dealt with. This project is unique by the development of an entirely new dynamic system, which will ensure an effective communication between students and teachers on the levels of not only theoretical and pre-clinical courses, but also in practical courses. The system is based on the proposal of a parametric description of a medical curriculum, which can be searched quickly and precisely by the OPTIMED platform. Lectures in this session were intentionally designed in a manner which would introduce the motivation and the theoretical background of the entire project (a lecture given by Assoc Prof Julia Bienertová Vašků from the MU's Faculty of Medicine), and demonstrate its real use for the study of medicine (as explained by Dr Komenda from IBA, MU's Faculty of Medicine). The final lecture was presented by Jan Svancara, MSc (IBA, MU's Faculty of Medicine), who summarised the available data describing the education from the analytical point of view, and indicated that a properly designed database background would make it possible to view big data describing the curriculum as a whole picture, and to evaluate the education contents globally.

The conference programme continued with a workshop inspired by an international project called CROESUS, which involves the mutual cooperation of St George's, University of London (United Kingdom), Masaryk University (Czech Republic), and Pavol Jozef Šafárik University in Košice (Slovakia). Terry Poulton and Sheetal Kavia – two colleagues from St George's, University of London – participated in the workshop via teleconference. First of all, they informed workshop participants about their perspective and experience with the introduction of innovative methods into the education of medical and non-medical fields of study, such as the educational algorithms, virtual cases or virtual patients, which is the main topic of the CROESUS project. Workshop participants subsequently played an active role in the practical part: with the help of instructors, the participants created

FIGURE 4 Dr Petr Štourač opens the session dedicated to e-learning and simulations in acute medicine

FIGURE 5 Dr Daniel Schwarz was elected the new Chairman of the Coordinating Council of the MEFANET project

nodes of a virtual clinical case, trying out for themselves how the platform for administration of virtual patients works. The concluding discussion turned out to be very useful: both foreign guests answered many unusual questions concerned with the introduction of these modern methods into the education itself.

Despite the very successful social evening, the morning programme of the second conference day was not sleepy at all. Quite the opposite – expressions involving words such as "urgent" or "intensive" echoed throughout the lecture theatre, as interesting topics from acute medicine were on the programme, chaired by Dr Petr Stourac (Department of Anaesthesiology, Resuscitation and Intensive Care Medicine, University Hospital Brno). Virtual patients and simulations play an important role particularly in acute medicine, because the opportunities to involve students in real cases are obviously very limited. After all, simulations are perceived as a separate scientific discipline in the field of acute medicine, and many new tools and procedures are tested in simulations before being introduced into practice. This medical specialty is also rather specific from the point of view of a smooth running of the medical team, numerous interactions, and a limited time for making decisions – and all of these are obviously very difficult to encourage, teach and examine by ordinary methods of education. Electronic publishing and simulations also help the staff of non-university hospitals, who have rather limited opportunities to be in contact with leading experts in the field. Other conference sessions focusing on medical informatics, information science and generally ICT in the education of physicians and other health care workers also provided many interesting examples and plentiful experience, which were mutually shared by teachers interested in modern technology.

Meeting of the Coordinating Council of the MEFANET project concludes the conference each year, and this year was no exception. The discussion focused on principal points concerning the future of the entire educational MEFANET network, and on activities aimed at specific ways of cooperation among the involved institutions. Election of the Chairman of the Coordinating Council took place at the end of the meeting: Dr Daniel Schwarz (IBA, MU's Faculty of Medicine) replaced Assoc Prof Ladislav Dusek (from the same institution), who did not stand as a candidate any more. Dr Dusek was officially appointed honorary chairman by representatives of the involved faculties.

Jakub Gregor

SUPPLEMENTARY MATERIAL

Looking back on the MEFANET 2014 conference
https://youtu.be/kYGrIbX04Fw

MEFANET JOURNAL PROFILE

Aims and Scope

The journal is intended to present within a single forum all of the developments in the field of medical informatics, medical education, e-learning and thereby promote the synergism among these disciplines. The journal is the premier vehicle for disseminating information about MEdical FAculties NETwork, which covers all Czech and Slovak medical faculties.

The journal enables medical teachers and scientists to share and disseminate evidence demonstrating the actual practice in on-line education in medicine and healthcare sciences by focusing on:

- research in medical educational informatics and learning analytics
- applications of medical informatics into education
- design, usage and results of novel e-learning tools and innovative pedagogical methods in medical teaching and learning
- other interdisciplinary topics related to information and communication technology in medical education

In 2009–2012, MEFANET report was published as one volume per year and was printed in 1000 copies. Since 2013, MEFANET journal has been published biyearly.

Subjects of interest

- E-health and telemedicine
- E-learning
- Information science
- Innovative teaching methods
- Medical educational informatics and learning analytics
- Modeling and simulation
- Multimedia
- Social media pedagogy
- Evidence-based medicine in education

Indexing

MEFANET journal is indexed in:

- ERIH PLUS
- Directory of Open Access Journals
- Index Copernicus
- Google Scholar
- Bibliografia medica Čechoslovaca
- Bibliografia medica Slovaca

On-line access

All volumes are available in electronic version at http://mj.mefanet.cz

PUBLISHER

Facta Medica

The Facta Medica Ltd. publishing house, based in Brno, was founded in 2008 by Dr. Boris Skalka, Dr. Eliška Skalková, and Assoc. Prof. Zdeněk Susa. The publishing house was founded with the aim of focusing on the publication of specialized literature from the field of medicine and health care – both periodical and non-periodical, but also medicine-related literature of fact and that of fiction. Since 2009 the publishing house has been represented by B. Skalka and E. Skalková.

The scientific standards are guaranteed by the publishing house's Scientific Board, whose members are listed below:

- Prof. Josef Bednařík, MD, PhD, neurology
- Assoc. Prof. Igor Crha, MD, PhD, gynaecology and obstetrics
- Prof. Richard Češka, MD, PhD, internal medicine – metabolism
- Prof. Eva Češková, MD, PhD, psychiatry
- Prof. Petr Dítě, MD, PhD, gastroenterology
- Prof. Andrej Dukát, MD, PhD, internal medicine
- Prof. Jindřich Fínek, MD, PhD, oncology
- Prof. Petr Husa, MD, PhD, infectology
- Prof. Václav Monhart, MD, PhD, nephrology
- Prof. Marián Mokáň, MD, PhD, FRCP Edin diabetology
- Prof. Miroslav Penka, MD, PhD, haematology
- Prof. Miroslav Souček, MD, PhD, internal medicine
- Prof. Jindřich Špinar MD, PhD, FESC, cardiology
- Assoc. Prof. Petra Tesařová, MD, PhD, oncology

REVIEWERS

Papers published in the Mefanet Journal 3(1) were reviewed by:

Miroslav Dvořák
Jakub Gregor
Jiří Kratochvíl
Tomáš Kulhánek
Jaroslav Majerník
Jana Strenková
Petr Štourač
Michal Trnka
Christos Vaitsis

Editorial board and publisher greatly acknowledge reviewers' contribution to this Mefanet Journal issue.

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitivness
UNIVERSITAS MASARYKIANA BRUNENSIS
INVESTMENTS IN EDUCATION DEVELOPMENT

EDITORIAL BOARD

KEYWORDS INDEX

AUTHORS INDEX

CENTRAL GATEWAY OF THE MEFANET EDUCATIONAL NETWORK

mefanet | MEDICAL FACULTIES NETWORK

Search

medical disciplines linker

HTTP://PORTAL.MEFANET.CZ

MEFANET 2015 Conference

9th international conference of Czech and Slovak faculties of medicine, focused on e-learning and medical informatics in the education of medical disciplines

Specific field of medicine: Neurology — Psychiatry — Neuroscience

Brno, 25—26 November 2015