MUNI
FI

Empowering a Data-Driven Approach in Medical Education and Healthcare

Habilitation thesis

Martin Komenda

Declaration

I affirm that all work in this document is my original work. Further, I confirm that all writing is my own writing. All sources, references and literature used or adopted in the preparation of this thesis are properly cited with full reference to the relevant source.

Martin Komenda

Abstract

In an era where communication and information technologies, closely connected to the Internet, are part of our daily lives, determining the accuracy, truthfulness and objectivity of available information is a significant challenge. This thesis offers a comprehensive view of dependent outputs and activities in both academic and governmental sectors that concentrate on data collection, processing, analysis and visualisation. Recognising the importance of correctly managing, identifying, and disseminating accurate information and insights from the vast array of data of differing quality is essential. This necessity primarily drives the compilation of this summary thesis. It includes chapters that exemplify new findings and good practices in technology-enhanced learning and health informatics based on experiences from research and development projects, stakeholders' opinions and community requirements. Decision-making should always be grounded in data, but this is only effective if the data is accurately processed, understood and, most importantly, correctly interpreted. There are proven methodological and technological backgrounds, emphasised in this thesis, that provide a complete and robust framework for two unique and multidisciplinary domains of human interests and understanding: first, medical and healthcare education supported by the MEFANET network; second, health information and statistics describing mainly data from the National Health Information System. The chapters emphasise the entire and complex process based on used methodologies, technological frameworks, relevant related projects and data/analytical reporting tools. These topics are set in the context of achieved results applied in daily practice, sharing know-how, student involvement and community building. This work is multidisciplinary and summarises the results achieved, combining the fields of informatics, medical education, pedagogy, analysis, statistics and health literacy.

Keywords

data-driven approach, decision-making support, medical education, healthcare, CRISP-DM

Acknowledgements

I am deeply grateful to all my colleagues at Masaryk University and all the academic and government institutions for their significant contributions to the conception, design, and execution of the projects that form the basis of the case studies in this habilitation thesis. The productive, intense and long-term cooperation resulted in numerous innovative and valuable outputs, including creating open-minded communities driven by data. Furthermore, I must express my heartfelt appreciation to my family and close friends. Their unwavering support and understanding have given me the mental fortitude and time necessary to complete this summary thesis.

Abbreviations

CBL: Case-Based Learning

CRISP–DM: Cross-Industry Standard Process for Data Mining

DICOM: Digital Imaging and Communications in Medicine

EHDS: European Health Data Space

EVAMED: Evaluation of Medical Education

HIS: Health Information and Statistics

ICT: Information and Communication Technologies

IHIS: Institute of Health Information and Statistics

ISID: Information System of Infectious Diseases

LTS: Long-Term Support

MAMES: Monitoring, Analysis and Management of Epidemic Situations

MED MUNI: Faculty of Medicine of Masaryk University

MEDCIN: Medical Curriculum Innovations

MEFANET: Medical Faculties Network

MeSH: Medical Subject Headings

MHE: Medical and Healthcare Education

MOOCs: Massive Open Online Courses

MoH CR: Ministry of Health of the Czech Republic

NDI: National Data Infrastructure

NHIS: National Health Information System

NZIP: National Health Information Portal

OBE: Outcome-Based Education

OPTIMED: Optimized Medical Education

OSCE: Objective Structured Clinical Examination

RFA: Radiofrequency Ablation

RFO: Research Funding Organisation

RIs: Research Infrastructures

SAML: Security Assertion Markup Language

SBL: Scenario-Based Learning

SESSAD: Sharing Experience in Scientific Software and Applications Development

SIMU: Simulation Centre of MED MUNI

VP: Virtual Patient

WHO: World Health Organization

XML: eXtendable Markup Language

Content

Introduction

This section introduces the research topic of this habilitation thesis: the data-driven approach in medical education and healthcare. The individual chapters are intended to provide the necessary context and intersections of specific activities that resulted in implementing methodological and technological outputs into practice. These long-term activities logically fit into a conceptual solution by designing and building robust systems that will enable appropriate work with data not only in the creation (collection) itself but especially in their use in real situations as a tool for objective and qualified decision-making. Data should always serve as a basis for decision-making processes and mechanisms, but only if they are correctly processed, understood and interpreted. There are various ways to present results through descriptive statistics and data analysis, from summary tables to static graphs to interactive web visualisations. It is only possible to say which type and presentation format is best with additional information (such as the target audience or primary purpose of use) [1].

The thesis is based on two relevant domains: Medical and Healthcare Education (MHE) and Health Information and Statistics (HIS). What both have in common is the need for more comprehensive and sustainable tools in the selected areas, which are the main focus of this thesis; these tools need to be built on a solid and proven methodological background and supported by a strong user community. Chapter “Selected publications” briefly lists the author's most significant publication outputs in terms of journal articles and mentions a book representing the author's milestone outputs. Chapter “Research and development projects” briefly documents more than 25 projects that have addressed an area relevant to the topic of this thesis at the national or international level, where the author was the principal investigator or the research team leader. Chapter “Sharing know-how” continues with an introduction to well-established conferences that annually showcase new developments in their respective fields (MHE and HIS), provide a forum for sharing experiences from practice, and help to build a social community of interested users in a relaxed and enthusiastic way. Chapter “Student leadership and mentoring” focuses on conducting student work and projects in three complementary areas (computer science, natural science, and medicine). The scope of the whole thesis is interdisciplinary, combining knowledge from (i) computer science in the form of methodological, algorithmic, and technical approaches; (ii) natural science mainly in the form of a methodological framework for data mining, analysis and processing; (iii) medical and health education together with healthcare in the form of the specific expertise needed for validation and correct interpretation of the achieved outputs. Chapters “Results: Medical and healthcare education” and “Results: Health information and statistics” describe the results divided into methodological and technological levels, where mainly existing tools used in daily practice are presented. This includes a robust framework for health data sharing based on the national methodology [2]. Last, the societal overlap and impact on the selected communities are mentioned in Chapter “Societal impact”. The thesis concludes with a brief assessment of the primary and secondary outputs of the habilitation thesis, with a breakdown into MHE and HIS.

Medical and healthcare education

For the sphere of medical and healthcare education, it was crucial to design a comprehensive infrastructure for long-term sustainable sharing of not only educational materials but also data describing the entire curriculum that logically accompanies the process of teaching young medics. Below are described three pillars, the basic building blocks that have significantly supported the establishment and functioning of the Czech-Slovak network of medical and healthcare faculties called MEFANET (MEdical FAculties NETwork) [3]. The main goal of MEFANET is to support education and apply a paradigm of technology-enhanced learning in the medical and healthcare sciences field using modern information and communication technologies, which are always based on proven methodological backgrounds and concepts tested in practice. MEFANET provides a platform for sharing guaranteed educational resources, materials and methods among member institutions. These domains include multimedia teaching materials, online courses, video lectures and other digital resources, project outputs and building tools (i.e., educational software or complex web applications) to support everyday teaching. The aim is to improve the quality and efficiency of education in medicine and related fields. MEFANET also promotes the exchange of experiences and collaboration between teachers, students and professionals from different medical faculties. In this way, it strives to define or improve medical and healthcare education standards nationally and internationally. Considering the specifics of individual faculties, it is logical that adaptation to their needs and custom-made development according to specific requirements are desirable and necessary when trying to incorporate innovative trends into the regular curriculum. It is no different at the Faculty of Medicine of Masaryk University (MED MUNI) in Brno, where the MEFANET network has become a place for inspiration and transfer of good practice in pedagogical, medical and informatics terms.

This section mentions three main highlights, which appropriately combine methodological background, technical solution and impact on the academic community in medical and healthcare education. The path to these outputs included the implementation of more than ten national and international research and development projects, a comprehensive review of innovative approaches and subsequent adaptation to the MED MUNI environment, long-term sharing of good and bad practices, seeking inspiration at dedicated conferences and workshops, active collaboration with students and, last but not least, publications mapping partial and comprehensive results.

- **Collaborative framework for sharing medical and healthcare electronic educational resources**
- **Integration platform supporting medical and healthcare education using innovative pedagogical methods and technologies**
- **Implementing the proven methodology to support data-driven medical education and healthcare decision-making**

Health information and statistics

The domain of Czech healthcare is very closely linked to the global need for rational digitisation and maximum data availability, which can significantly help in making valid and objective decisions. The individual areas described in this thesis systematically work towards building a robust information service over healthcare data by current legislation and the needs of specific target groups. The core component is undoubtedly the National Health Information System (NHIS), which represents a unified nationwide public administration platform and contributes substantially to standardisation in reporting, evaluation and economic valuation of healthcare in the Czech Republic. The national concept of health data sharing, based on the priority strategic objectives of the Ministry of Health of the Czech Republic (MoH CR), is a superstructure of the available data from NHIS. It presents data sharing and accessibility as a uniform, systematic, sustainable and technically well-defined procedure for publishing data outputs for further manual or machine processing. At the beginning of this development, there was mainly a desire to make data transparent and accessible for management purposes. Still, with the progressive digitisation of health services and the increasing volume of centralised data, the objectives and requirements have expanded substantially. The concept outlines the key processes and tools necessary to build a sustainable system for sharing, publishing and secondary use of NHIS data. Shared experience, supported by communication with stakeholders from the public sector, academia, the business sphere and the data journalism domain, has emerged as a critical factor in designing a workable framework.

The three critical outputs listed below have been developed in collaboration with leading experts in Czech healthcare, especially from the Czech Medical Association of J. E. Purkyně. Based on the successful implementation of several national and international projects, it was possible to create robust tools supported by good practice and proven methodological procedures that realistically support data-driven decision-making in selected domains of Czech healthcare, such as information service over NHIS or inpatient care mapping. The National Health Information Portal (NZIP) has become the unified and overarching platform for comprehensive health literacy support in the Czech Republic; today, it is a platform for the general public and professionals.

- **National methodology for health data sharing in the Czech Republic**
- **Technological background for health literacy and data-driven decision-making improvement in Czech healthcare**
- **Real-time framework for effective management and data-driven support of intensive inpatient care, including complex information service during the pandemic**

Selected publications

The following selected publications are divided into two overviews according to the domain they fall under within the thesis: Medical and Healthcare Education (Table 1) and Healthcare Information and Statistics (Table 2). All of them are also annexed to this thesis. The last summary publication is a printed and electronic monography [1], which presents 18 case studies, again logically divided into the domains mentioned above.

Table 1 The author's most significant works related to the domain of Medical and healthcare education

No.	Publication	Author's contribution
1	Komenda et al. Medical faculties educational network: Multidimensional quality assessment. *Computer Methods and Programs in Biomedicine* 2012; 108(3): 900-909.	General idea (50%) Design and implementation (75%) Writing (90%)
2	Schwarz et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. *Journal of Medical Internet Research* 2013; 15(7): 298-311.	General idea (10%) Design and implementation (40%) Writing (10%)
3	Komenda et al. Curriculum Mapping with Academic Analytics in Medical and Healthcare Education. *PLoS One* 2015; 10(12): e0143748.	General idea (100%) Design and implementation (85%) Writing (90%)
4	Komenda et al. Practical use of medical terminology in curriculum mapping. *Computers in Biology and Medicine* 2015; 63: 74-82.	General idea (90%) Design and implementation (85%) Writing (90%)

Table 2 The author's most significant work on Healthcare information and statistics

No.	Publication	Author's contribution
5	Komenda et al. Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice. *Journal of Medical Internet Research* 2020; 22(5): e19367	General idea (75%) Design and implementation (75%) Writing (90%)
6	Krejčí et al. Development of the Czech Childhood Cancer Information System: Data Analysis and Interactive Visualization. *JMIR Public Health and Surveillance* 2021; 7(6): e23990	General idea (45%) Design and implementation (50%) Writing (50%)
7	Komenda et al. Sharing datasets of the COVID-19 epidemic in the Czech Republic. *PLoS One* 2022; 17(4): e0267397	General idea (90%) Design and implementation (75%) Writing (90%)
8	Komenda et al. Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic. *Journal of Medical Internet Research* 2022; 24(2): e33149	General idea (75%) Design and implementation (75%) Writing (90%)

Research and development projects

Below, in chronological order, are the annotations of all projects in which I was actively involved as principal investigator or co-investigator. The results of these activities have contributed significantly to developing both key areas of my research focus: Medical and Healthcare Education and Health Information and Statistics.

Standardisation and sharing of an educational platform between medical faculties within the MEFANET project[1]

The present project proposes to share digital educational content in the MEdical FAculties NETwork (MEFANET) using an interactive communication platform. For this purpose, a standardised environment for electronic publishing will be designed and set up using modern ICT and recommended methodological procedures for creating and sharing digital content in the medical faculty network will be developed. The target group, which consists of teachers and students of medical and healthcare disciplines, will be introduced to the nature and tools available in the MEFANET network through training seminars and conferences, to which experts from teaching hospitals and other specialised clinical departments will also be invited as active contributors. The project will lead to increased cooperation and the transfer of good practice and experience between institutions that naturally enter the process of training doctors and health professionals.

Modern trends in testing issues in medical education[2]

Modern methods of assessing students' knowledge are among the most discussed topics in the academic world. Standardised testing allows for fair, verifiable, objective assessment of student knowledge and skills. The project's primary goal is to strengthen the academic staff's pedagogical competencies in constructing and analysing test items across selected departments at MED MUNI through educational seminars.

Modernisation of teaching clinical decision-making across paediatric disciplines of medical faculties in the MEFANET network[3]

Clinical medicine education needs more contact between students and patients as hospital stays for patients decrease. Therefore, the project proposes to innovate the clinical phase of medical studies using virtual patients and other case-based learning tools. Virtual patients are recognised in the international medical education community as an effective method for developing clinical decision-making competence. Thanks to the natural support for training communication and psychotherapeutic skills, it is suitable, among others, for practical teaching of paediatric subspecialties of General Medicine, whose students and teachers are the project's target groups. An implementation team of experts in clinical medical education and ICT teaching

[1] https://www.muni.cz/en/research/projects/8423
[2] https://www.muni.cz/en/research/projects/26921
[3] https://www.muni.cz/en/research/projects/16663

applications from four medical faculties will technically define the platforms for virtual patients, case studies and interactive algorithms. It will also support digital content production and ensure know-how transfer from abroad.

Optimised teaching in general medicine: horizontal and vertical integration, innovation, and efficiency for practice[4]

The project aims to comprehensively innovate the system of teaching General Medicine at MED MUNI and to strengthen problem-solving-oriented teaching through the graduate's application in the clinical and academic fields. The core elements of the project are (i) horizontal innovation of all taught subjects using a tool in the form of a digital library with linked parametrically processed learning outcomes and teaching units and (ii) vertical linking of teaching on the axis: medical students' input knowledge – theoretical and preclinical knowledge – clinical knowledge and skills – medical graduates' skills after entering practice. Thus, OPTIMED relies on creating an innovative, sophisticated, and dynamic system that will facilitate students' and teachers' orientation in teaching and, in effect, make students' knowledge and skills for practice more effective. The critical parameter of the system is its dynamism, i.e., its ability to absorb new knowledge in medicine and link it rationally with patient-oriented teaching.

Testing in medicine education ON-LINE[5]

Standardised testing of knowledge will enable educators to assess knowledge and skills fairly, clearly, and objectively. Setting up uniform procedural procedures in the preparation and implementation of testing means reducing the time demands on teachers to develop and use tests. Therefore, the project aims to create an e-learning portal platform to systematically build and strengthen the pedagogical competencies of academic staff in test preparation at MED MUNI.

Evaluation of medical curriculum through advanced analytical methods[6]

The primary objective of the EVAMED project is to identify, in collaboration with two working groups, through newly proposed metrics, areas that need to be more consistent in their formalisation. Using the newly developed evaluation approaches and data mining methods, a follow-up step will be to effectively communicate graspable outputs to the authors in the form of recommendations generated from the working group representatives. Therefore, the main effort is to use modern approaches to analyse heterogeneous data sources to unify the content of the curriculum description database (designed during the OPTIMED project), which will increase the quality of the description of the whole curriculum. The proposed and implemented methods for quality assessment, data analysis, and visualisation have a high publication and citation potential due to their uniqueness, regardless of the specific pilot implementation over the OPTIMED database. The outputs of the submitted EVAMED project are aimed at

[4] https://www.muni.cz/en/research/projects/16523
[5] https://www.muni.cz/en/research/projects/29429
[6] https://www.muni.cz/en/research/projects/29268

disseminating know-how in medical curriculum optimisation and educational data mining, not only in the environment of medical faculties in the Czech Republic and Slovakia but internationally.

Medical curriculum innovations[7]

The project's objectives are as follows: (i) to standardise the medical curriculum management system (the OPTIMED platform) used at MED MUNI based on international recommendations and published specifications; (ii) to propose innovative methodological background based on approved analytical and data mining techniques to evaluate and map medical curricula; (iii) to build an original model using proper data mining and natural language processing techniques, enabling the systematic comparison of medical curricula at various medical schools; (iv) to foster increased awareness of existing educational standards produced by MedBiquitous and provide an exemplary implementation for disseminating best practices beyond the project partners.

OPTIMED portal: Implementation of new functions for efficient search[8]

The OPTIMED portal contains complete descriptions of the courses and has long provided one of the key references in General Medicine. One of the critical functionalities of the portal is the search, which allows the user to view the available content in the form of a list of unit details sorted by relevance based on the frequency of occurrence of words. The project aims to incorporate new search algorithms working with morphological forms of Czech words into OPTIMED and thus increase the accuracy of the results when searching the content.

Clinical reasoning skills enhancements with the use of simulations and algorithms[9]

The strategic partnership will support the development of a new online platform where medical and healthcare students and their teachers can learn, teach and cooperatively create the educational content. The platform will allow the implementation of innovative pedagogy approaches to clinical medicine education. The critical challenge of the project is to use the platform to modernise the teacher-based and classroom-oriented biomedical science component of the medicine and healthcare courses and, thus, to enrich the curriculum of partners' health education institutions with a focus on competence-based learning and assessment built around problem-based learning, virtual patients, simulations, and interactive algorithms.

[7] https://www.muni.cz/en/research/projects/31964
[8] https://www.muni.cz/en/research/projects/36430
[9] https://www.muni.cz/en/research/projects/28843

Detection of links between information systems for curriculum mapping and virtual patients[10]

The MERGER project aims at information science, applied informatics, and biomedical engineering. It proposes to develop methods for automatically detecting links between information systems to support clinically oriented stages of studies in medical and healthcare disciplines. For this purpose, the MERGER project intends to use large amounts of textual data collected in medical curriculum mapping systems and systems to create and playback virtual patients. The proposed machine learning methods for automatically detecting links between content entities in these systems will be implemented as a software toolbox. To quantitatively verify the accuracy of the developed algorithms, datasets extracted from OPTIMED (a curriculum mapping system at MED MUNI) and from OpenLabyrinth (a virtual patient system in the MEFANET network) will be used. Algorithms based on machine learning methods will be learned and validated using external expert knowledge supplied by senior MED MUNI students.

Widening access to virtual educational scenarios[11]

This project takes scenario-based learning (SBL) 'out-of-the-box' and combines the skill sets of both academic and enterprise partners to make SBL more accessible to a wide range of professions. Developments will include the ability to embed SBL activities directly into learning platforms and Massive Open Online Courses (MOOCs), adding renewed pedagogic value and ease of use to learners through improved integration, progress monitoring and the delivery of feedback. By embedding SBL, MOOCs will enable the development of real-world skills, competency and experience rather than knowledge. Providing training materials in multiple languages, including a 'How to create Virtual Scenarios' MOOC, will make VS more accessible for educators. The project will sustain and disseminate these activities through an extended partnership, the 'WAVES' network.

Training against medical errors[12]

TAME builds upon the recent introduction of interactive virtual patient (VP) cases for training in both eViP ePBLnet and Clinical Reasoning in CROESUS. It includes partners from more than 25 institutions who participated in several previous projects dealing with virtual patients. In many ways, these previous projects have taught the principle of clinical decision-making. But this project goes to the next level. This project takes VPs one crucial step further by targeting the most critical factor in clinical education, patient safety, and asks how we can best adapt VPs to maximise the possibilities of training against medical error. The process would be to use VPs in later parts of the course, specifically VPs that are more detailed in clinical attachments. The virtual patients should be able to challenge the students' thought processes built on the knowledge

[10] https://www.muni.cz/en/research/projects/36097

[11] https://www.muni.cz/en/research/projects/32844

[12] https://www.muni.cz/en/research/projects/32866

acquired during the early years of their degree. They would be constructed to allow students to make wrong decisions, decisions that seem entirely plausible and are usually based on real-world incidents; students would take actions based on those decisions and experience the consequences of those actions. Pedagogically, this project addresses a fundamental issue in clinical management: to safely train learners to make correct decisions without compromising patient safety and without requiring exposure to every clinical situation or disease the student may need to be familiar with.

Case-based learning and virtual cases to foster critical thinking skills of students[13]

Three institutions from three European countries (Romania, Czech Republic, Slovakia) want to create a strategic partnership to support transferring and sharing selected best educational practices between involved institutions. The mentioned best practices include case-based learning (CBL) in medicine, dentistry, and pharmacology, as well as the use of ICT platforms. The project duration is 36 months. The first skill set is connected to pedagogy methods from scenario-based learning (case-based learning, problem-based learning, team-based learning, hackathons, flipped classrooms, etc.). The second skill set is connected to various ICT platforms usable to support scenario-based learning didactic approaches – in medicine, of course, related to virtual patients (virtual cases). After analysing the needs performed in the fields of (i) pedagogy and (ii) medical informatics, we will train the teachers and use them as authoring forces supported by PhD students to create 28 interactive virtual cases.

INTENT[14]

The INTENT project aims to find solutions for innovative patient-centred cancer care. It targets and involves various types of actors: cancer care providers, patients, and policymakers. The aim is to work with these groups to understand better how to interpret the patient-centred approach and identify ways to improve cancer care in central Europe. Beyond the activities leading to an innovative cancer care model, the project will issue policy recommendations, provide an online benchmarking tool, identify priorities for improving existing local systems and create a virtual know-how centre.

Building curriculum infrastructure in medical education[15]

The project objectives are: (i) to share know-how in the field of curriculum design, innovating on and optimising the proven methodology of parametric description (curriculum building blocks) and guidelines for curriculum definition; (ii) to create new descriptions of selected curriculum parts in partner institutions per the local needs analysis and institutional requirements while fully complying with international standards; (iii) to define and solve new curriculum mapping research questions and issues and to explore hidden relations in described curricula using natural language processing, data/text mining, machine learning, analysis, and visualisation.

[13] https://www.muni.cz/en/research/projects/43366

[14] https://programme2014-20.interreg-central.eu/Content.Node/INTENT.html

[15] https://www.muni.cz/en/research/projects/43286

Online platform for real-time monitoring, analysis and management of epidemic situations[16]

This project is unique because it was created on the basis of an urgent need, i.e. without submitting a project proposal. An online analytical tool was created for managing and monitoring epidemic situations in real-time about the possibility of managing bed, equipment and personnel capacities and prospective modelling of the burden of the monitored population, including its structure determined by individual stages of the disease and other factors. The web tools enabled the analysis of the overall situation in the Czech Republic, the situation in particular regions and the localisation of outputs to individual healthcare facilities with authenticated data access.

Innovative partnership for action against cancer[17]

The project's general objective is to develop innovative approaches to advances in cancer control. The innovation that will be covered within the Joint Action consists of further development of cancer prevention, comprehensive approaches to the use of genomics in cancer control, cancer information and registries, improvements and challenges in cancer care, mapping of innovative cancer treatments and governance of integrated cancer control, including a new analysis of National Cancer Control Plans. The project's key focus will be implementation, reflected in the critical deliverable: the Roadmap on Implementation and Sustainability of Cancer Control Actions.

Masaryk University's strategic investment in education[18]

The SIMU+ project aims to provide strategic activities' infrastructure and material and technical needs. Its core part is constructing the complex Simulation centre (SIMU) implemented by the MED MUNI. The project also covers activities aimed at innovating the teaching of other MUNI study programmes. Implementing the project will require the acquisition of the necessary instrumentation and material equipment, significantly improving the quality of educational activities at MUNI. The project activities also aim to support people with specific needs.

Sharing experience in scientific software and applications development[19]

In natural sciences, medicine and healthcare, digital solutions enable innovations and scientific advancements more than ever. Scientific software systems and applications are the backbone of science, education, and knowledge transfer. Most prominently, since the beginning of the COVID-19 pandemic, software solutions have played a central role in collecting and analysing data for visualising and predicting infections and for developing, testing and distributing vaccines. Scientific software solutions are becoming increasingly large and complex, using the latest technologies such as artificial intelligence, big data, machine learning, image processing, simulation and visualisation.

---

[16] https://www.muni.cz/en/research/projects/55708

[17] https://www.ipaac.eu/

[18] https://www.muni.cz/en/research/projects/37264

[19] https://www.muni.cz/en/research/projects/65049

At the same time, these solutions must serve research with the highest standards regarding correctness, precision, performance, traceability, and explainability. A single person or organisation can no longer develop scientific software tools and applications. To fulfil all requirements, scientific software has to apply best software engineering practices and requires a joint effort across several teams and organisations and interdisciplinary collaboration. In both Interreg regions, research and innovation are essential for medical science, education and the economy of medical technologies and digital health.

Postgraduate medical education development programme[20]

The project addresses a common cross-border problem: the lack of pedagogically qualified medical and healthcare professionals who can act in their lecturing or mentoring practice according to the latest trends and methods of medical and healthcare education. The project aims to increase the relevance of educational content to the needs of the labour market and improve graduates' employability by improving the pedagogical competencies of postgraduate medical graduates in the cross-border region.

Raising awareness through the National Health Information Portal[21]

The project aims to improve the communication and credibility of information of the Ministry of Health of the Czech Republic, namely by providing a comprehensive methodological and professional support of the newly created National Health Information Portal (NZIP). Only fully guaranteed and methodologically standardised content is published on the portal; therefore, NZIP is intended to become a key communication channel in terms of healthcare for the state administration, especially for the general public and professionals.

New era in medical education[22]

During the spring of 2020, due to the advancing pandemic of COVID-19, there was a need to move teaching to the online environment quickly. This need affected all universities worldwide. Still, the position of medical faculties was specific in the fact that a large part of teaching generally takes place in the form of practical exercises and clinical practice, as well as in the fact that senior students – during the COVID-19 pandemic – were heavily involved in volunteering; this posed great demands on their time possibilities, which are limited even in regular times due to the complexity of the study. The project's primary goal was to rationalise online teaching at the participating medical faculties thanks to academics' unified methodological guidance in preparing syllabi and online study materials for students.

[20] https://www.muni.cz/en/research/projects/62148

[21] https://www.nzip.cz/o-projektu

[22] https://www.muni.cz/en/research/projects/59909

Introduction of objective structured clinical examination and related modifications to the teaching of intensive care medicine[23]

Intensive care medicine at MED MUNI is a crucial subject of undergraduate studies. In connection with the opening of the new simulation centre, a significant shift in the concept of the subject occurred with the introduction of teaching using high-fidelity simulations. Despite the considerable modification of education and testing, the next strategic step is the introduction of Objective Structured Clinical Examination (OSCE). In the context of these developments, there is a need to modify the existing teaching so that students experience situations in practical sessions that will be subsequently assessed. The current lesson plan and syllabus will be adapted to this. OSCE will be an excellent tool, and its implementation will complete the transformation of the "Intensive care medicine" course according to the modern format of education.

MUNI Self-Care Hub[24]

The MUNI Self-Care Hub will become a bilingual portal of targeted support and assistance for Czech and English students of MED MUNI, potentially expanding to the whole university environment. The vision of supporting students' mental health is based on the initiative of the MED MUNI Student Chamber with the professional guarantee of the Institute of Medical Psychology and Psychosomatics and the National Institute of Public Health. The content focuses on mental health, proper diet and appropriate exercise habits during studies.

Systematic education and increasing health literacy of the population through the National Health Information Portal[25]

Through the development and evolution of the educational content of the National Health Information Portal, target groups (people with health disadvantages, people with low health literacy and people with disabilities) are systematically educated in the area of health literacy, especially in the areas of prevention and healthy lifestyles, knowledge of certain diseases, reactions in various life situations and orientation in health data and information.

Methodology of creating plain language recommendations for knowledge translation to support evidence-based decision-making in healthcare in the Czech Republic[26]

The project aims to develop and test a methodology for producing plain language versions of priority recommendations for the public interest of adolescents, parents and the general adult population in the Czech Republic. The methodology will be based on the experience of developing lay recommendations in the context of COVID-19. The project is based on international collaboration and expertise with colleagues at

---

[23] https://www.muni.cz/en/research/projects/70813

[24] https://www.muni.cz/en/research/projects/70313

[25] https://www.uzis.cz/index.php?pg=o-nas--projekty&prid=35

[26] https://www.muni.cz/en/research/projects/70372

McMaster University, with whom we have developed a unique COVID-19 recommendation repository funded by the Canadian Institutes of Health Research.

National childhood obesity prevention programme BE FIT 24[27]

The subject of the project is the early detection of overweight people to prevent childhood obesity. Innovative tools for optimising health parameters in children will be tested, including preparing a universal health prevention program. The project includes the provision of methodological and technical background for collecting and evaluating essential predictors of health status in the child population, the optimisation of which has a significant preventative character in terms of health promotion.

[27] https://www.uzis.cz/index.php?pg=o-nas--projekty&prid=38

Sharing know-how

Sharing experiences across the professional community in multidisciplinary domains has significantly helped build new communities. Topics appropriately combining modern information and communication technologies, the specifics of medical and healthcare teaching and elements of pedagogy are the basis of the now traditional MEFANET conference. The target group, which includes teachers and students of all Czech and Slovak medical and healthcare faculties, provides a big potential for implementing the achieved results into practice. The epidemic associated with the uncontrolled spread of COVID-19 has increased interest in the availability of health data. However, the publication of these data is not only a technical task but also a matter of correct interpretation and regular communication that still needs to be set up. Thus, the opening of the National Health Information System in the autumn of 2019 established an entirely new tradition that continues the ongoing development of partnerships between entities actively interested in health data. Last but not least, the following is a national conference focused on e-health, where current topics are highlighted, and innovative approaches supporting the digitisation of healthcare are shared. I am a member of programme committees and organising committees for the conferences listed below.

Series of the MEFANET conferences

The traditional MEFANET conference (16 events between 2008 and 2023[28]) represents professional cooperation between Czech and Slovak medical and healthcare faculties with a focus on building and strengthening collaboration in the development of teaching using modern information and communication technologies. The primary aim is to share practical experience and systematically develop teaching support at the faculties.

Conferences on the National Health Information Portal & Health data opening

The conferences (seven events between 2019 and 2023[29]) are devoted to the topic of the National Health Information Portal, offering different target groups understandable and guaranteed information about health, the Czech healthcare system, health promotion, disease prevention and diseases themselves from various perspectives. An integral part of this is the accessibility and publication of data, as well as the collection and management of data, which is part of the agenda of many public institutions. The focus is, of course, on health data, but this has always been in the overall context of opening up data from public administration databases and registries.

[28] https://www.mefanet.cz/conference

[29] https://www.nzip.cz/kategorie/272-konference-nzip

Conferences on e-health and telemedicine

The annual conferences (two events in 2023 and 2024[30]) are dedicated to various aspects of e-health. The conference programme covers legislative, technological and implementation aspects of e-health services, all with a priority focus on the dynamically developing field of telemedicine technologies.

[30] www.telemedicina-brno.cz

Student leadership and mentoring

Below are the bachelor's and master's theses of students under my supervision across three specific fields of study: (i) computer science, Faculty of Informatics, Masaryk University (Table 3); (ii) computational biology, Faculty of Science, Masaryk University (Table 4); (iii) public health, Faculty of Medicine, Masaryk University (Table 5). The results helped propose solutions for selected partial problems with the primary motivation of verifying the functionality in a pilot application in practice.

Table 3 Final theses of students at the Faculty of Informatics, Masaryk University, supervised by the author

Title	Annotation
Analysis and visualisation of medical curriculum data[31] (T. Kolla)	This bachelor thesis will examine data visualisation and analysis within the SIMUportfolio platform. Questions the thesis will be answering are what is within the data in the SIMUportfolio platform, how we can visualise and analyse it, and what we can take away from it for the future. The basic information about the platform is summarised within the thesis, as well as the approach with which the visualisation and analysis of the data was created.
Comprehensive analysis and visualisation of open data in healthcare[32] (M. Kravec)	This master thesis aims to explain the concept of open data and describe its current usage in healthcare in the Czech Republic. Another goal is to create a web interface for effective and clear visualisation of provided data sets, specifically the National Registry of Health Services Providers and evaluate the quality of web portals of selected healthcare facilities. This thesis contains detailed descriptions of the entire process of creating such a web interface, from analysis through design and implementation to testing and description of procedures used to develop the web interface.
Design and development of a web application for archiving and publishing image documentation in orthopaedics[33] (A. Kollar)	This master thesis involves the design, development, testing and deployment of a new web application, OrtoDB, tailored for Masaryk University in medical education. The web application's primary purpose is to provide a system for storing medical data and images of patients' medical recordings that can be filtered and categorised. It will help teachers create a better education environment where students will have digital materials about actual cases, better preparing them for real-life issues. OrtoDB is deployed on MUNI servers.

[31] https://is.muni.cz/auth/th/nfd1w/
[32] https://is.muni.cz/auth/th/pto8m/
[33] https://is.muni.cz/auth/th/s94wl/

Title	Annotation
Interactive visualisation of the study trajectory of medical and healthcare programmes[34] (O. Kondek)	This master thesis deals with understanding, processing and visualising the provided data about the Faculty of Medicine of Masaryk University students. The thesis describes available tools and methodologies suitable for this purpose. Subsequently, it focuses on implementing the CRISP-DM methodology on the provided data using selected tools - mainly the Python language with the Pandas library and the Looker Studio business intelligence tool. As part of the implementation, the thesis describes the data format, data model, data and model preparation process, and visualisation.
MetaTrader – A trade investment platform[35] (T. Milo)	The bachelor thesis describes the concept of trading on stock markets and, subsequently, analyses the tools for algorithmic trading. The proposed thesis is based on work with the selected platform MetaTrader 4 and implementation in a programming language, MQL4. The practical meaning of the thesis is supported by outputs in the form of programmes or source code that demonstrate the selected functionality.
Open dataset publications in healthcare[36] (M. Uhliar)	The master thesis introduces the world of open data and defines their issues and certain specifics in healthcare. The introductory part describes the available data, their properties and status in the Czech Republic. The following section describes the processes for creating a dataset for publication. The last part demonstrates the process of dataset preparation together with metadata.
Phabricator extension design and implementation[37] (L. Jagoš)	The master thesis deals with project management tools. The theoretical part contains a description of the project and the project management. Selected software tools for project management are shown and compared to each other. The thesis focuses on the system Phabricator, which is described in detail. The practical part, the Phabricator extension, is designed based on needs assessment and requirements specifications. The main benefit of the work is that the module is implemented as an extension to the Phabricator, providing information about task difficulty and duration for better scheduling and team collaboration.
Using of reCAPTCHA service into the WikiLectures[38]	The aim of this bachelor thesis is to analyze available options of protection against automated attackers and to propose a solution of internet security for website of project WikiSkripta.

[34] https://is.muni.cz/auth/th/sccrd/
[35] https://is.muni.cz/auth/th/yxeli/
[36] https://is.muni.cz/auth/th/so0ge/
[37] https://is.muni.cz/auth/th/lm9zg/
[38] https://is.muni.cz/auth/th/w3dzi/

Title	Annotation
User behaviour analysis in the curriculum management system[39] (M. Karolyi)	The master thesis focuses on data mining, processing and visualisation techniques, which can be used to analyse user behaviour. Concepts such as exploratory data analysis, explanatory data analysis, data storytelling, web-based application design and implementation are presented here. The primary purpose of the thesis is to understand the audience of the curriculum management system. The further objective is to find out how to effectively present datasets to a typical user and propose improvements in the existing curriculum management system based on the knowledge obtained.

Table 4 Final theses of students at the Faculty of Science, Masaryk University, supervised by the author

Title	Annotation
Algorithm design and implementation for medical education data analysis and comparison[40] (K. Ježová)	This master thesis concerns text data analysis, namely automated processing of medical curriculum descriptions. High-quality medical care is dependent on doctors receiving an appropriate medical education. However, given its complexity and magnitude, designing and maintaining a well-balanced curriculum takes time and effort. This thesis aims to optimise data input and subsequently create a tool that provides analysis of the individual parts of a curriculum and helps to discover overlaps or gaps in content by comparing it to a current curriculum or an entry supplied by a user.
Health data classification using neural networks[41] (Hoa Vu Thu)	This master thesis is focused on designing and implementing the classification of textual health data. The aim is to find a suitable solution for comparing documents from two website portals of the National Health Information System, namely Reporting and CZ-DRG, and to identify their semantic similarity. The practical output is an algorithm with a public interface that can be used with external applications. It compares two documents and classifies them as similar or not.

[39] https://is.muni.cz/auth/th/m9sd8/
[40] https://is.muni.cz/auth/th/amqxe/
[41] https://is.muni.cz/auth/th/lcu84/

Title	Annotation
Knowledge discovery algorithm design and implementation on medical education data[42] (A. Bóna)	The script written in R programming language for filtering and visualising data from implementing objective structured clinical examination into first aid courses at Faculty of Medicine of Masaryk University is introduced in this thesis. The output of the suggested script consists of files with filtered data and its visualisation as a group of plots. With a focus on OSCE, the data mining method CRISP-DM (of which two phases, data preparation and modelling, are performed by suggested script) and its alternatives, a short discussion precedes the practical part of this bachelor thesis.
Textual data comparison from medical and healthcare fields[43] (P. Růžičková)	This master thesis focuses on text preprocessing and similarity analysis of textual data, represented by the curriculum description and study materials from the medical and healthcare fields. This thesis aims to provide an in-depth analysis of the similarity of the text and discuss the possibilities of recommending systems. Several filtering steps are used to prepare the data, and the central part of the preparation is devoted to the lemmatisation phase, which is used to aggregate various word forms. The thesis output is an algorithm used to compare learning units describing the medical curriculum in the context of the MeSH classification.
Tool for hospital records classification in the CZ-DRG system[44] (M. Drábková)	The master thesis topic is a tool for classifying hospital records in the CZ-DRG system. The requirements of the classification tool are summarised, and based on them, the implementation using the Predictive Model Markup Language standard is chosen. Finally, the created prototype is tested and analysed.

Table 5 Final theses of students at the Faculty of Medicine, Masaryk University, supervised by the author

Title	Annotation
Comprehensive research on platforms for promoting and developing health literacy[45] (K. Valentová)	This master thesis uses the selected methodology to comprehensively search available online information tools for providing long-term support in health literacy, primarily aimed at the lay or informed public. The work will include the following steps: (i) identification of current projects and activities based on a defined set of keywords characterising health literacy; (ii) detailed analysis of the information architecture and supporting functionalities of the systems; (iii) summary comparison and overview report; (iv) recommended conclusions about further cultivation and development of the nzip.cz platform.

[42] https://is.muni.cz/auth/th/vv76p/
[43] https://is.muni.cz/auth/th/t7yf5/
[44] https://is.muni.cz/auth/th/lmxs4/
[45] Work in progress

Title	Annotation
Using business intelligence tools for interactive visualisation of healthcare data[45] (O. Machálka)	The master thesis aims to design and implement a comprehensive online report on a selected healthcare domain, following the proven CRISP-DM methodology for data mining. The first and essential step is to choose the appropriate domain (topic), the goals of the resulting visualisation and the analysis of the availability of data sources. Subsequently, all six steps of the methodology will be discussed in detail, and the necessary steps will be carried out practically, leading to the final implementation of the online solution that will be usable in practice. The report on the available data will be descriptive statistics and will not require advanced knowledge of analytical methods.

Results: Medical and healthcare education

The ICT methodologies behind

This chapter involves a detailed description of the research findings covering data collection, processing, analysis, web application design, development and implementation in medical and healthcare education. The presentation of the research results is divided into sections based on different pillars, which document specific outputs.

Standards implementation

Technical standards in medical and healthcare education are usually developed and applied to systematically enhance continuous improvements of teaching/learning processes [4]. Standardisation in the curriculum domain addresses the need for a structured data format describing individual curriculum building blocks and their characteristics and relations. A set of proven data formats used in practice at MED MUNI is required for standards-compliant curriculum innovations. One of the most productive associations, the MedBiquitous Consortium[46], covers various medical and healthcare professional associations, universities, commercial organisations and governmental organisations. The selected technical standards (primarily used for curriculum development, management and mapping), the Competency framework and the Curriculum inventory have been successfully implemented into the MEDCIN (Medical Curriculum Innovations) project and the SIMUportoflio integration platform [5]. The benefits include faculty management, policymakers and decision-makers being able to better evaluate and measure teaching against the required outcomes, institutions performing structured analyses and being able to compare their curricula. At the same time, students can better understand their intended learning [6].

The “medical disciplines linker” is one of the fundamental elements of the MEFANET portal platform, which enables basic categorisation of the educational contents and covers all learning areas among medical faculties. It is based on a varied list of medical specialties (containing 48 medical specialties) adopted from significant medical publishing resources. It can be modified only with the agreement of the MEFANET Coordinating Council [7].

[46] https://www.medbiq.org/

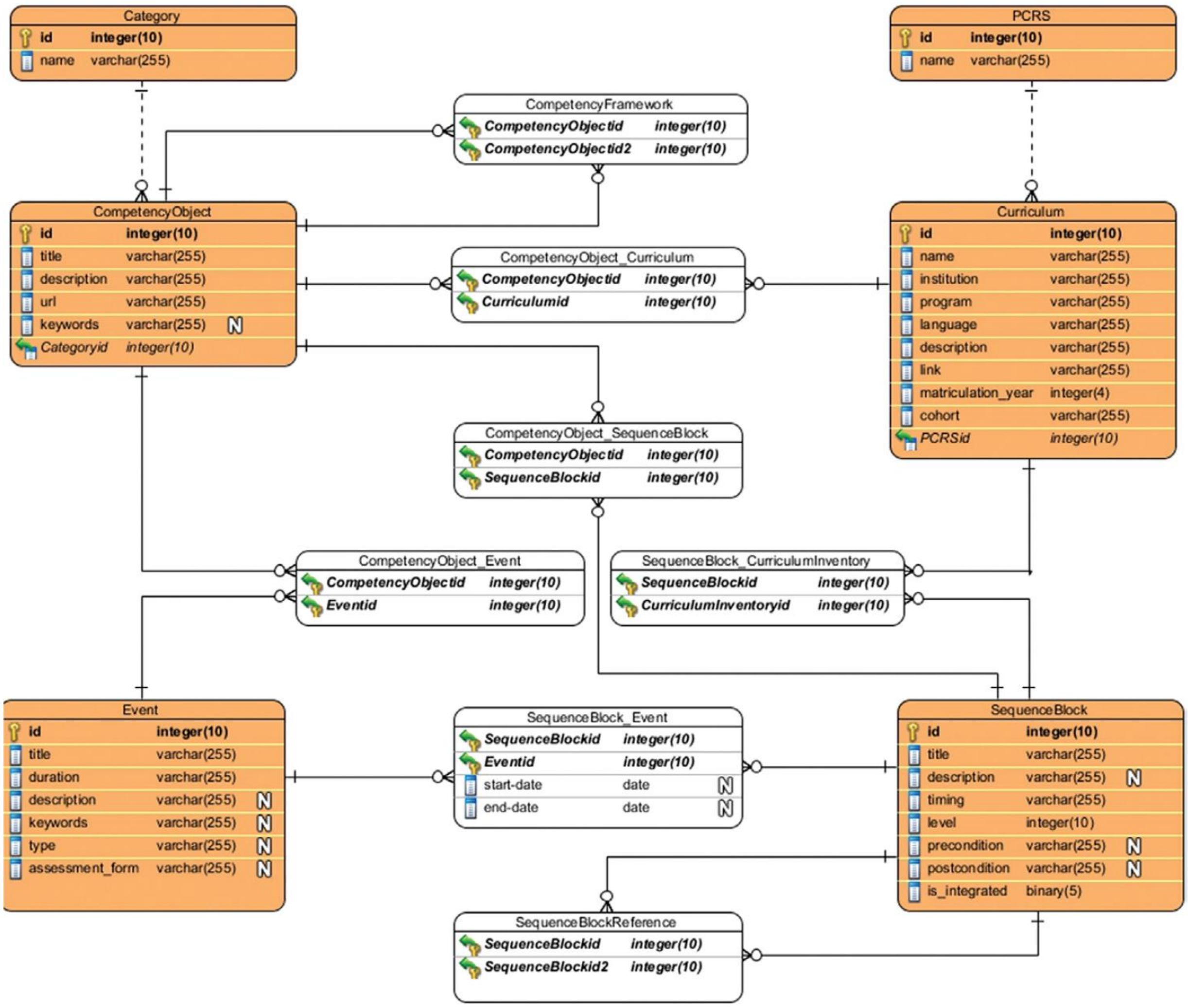


Figure 1 The MEDCIN database model supporting the MedBiquitous standards

Medical terminology adoption

While many references, surveys, and books on medical coding and classification for healthcare informatics applications exist, much less can be found in standardised vocabularies as regards medical and healthcare education. Based on a deep review [8], the MeSH biomedical dictionary, whose main objective is to classify learning activities appearing in the curriculum, has been implemented in the OPTIMED platform. One of the main requirements for standardised dictionary integration was regular updates of the Czech version, which MeSH fulfils as the only available solution. No other language mutations were foreseen, but a possible change should bring only a few complications regarding the proposed structure.

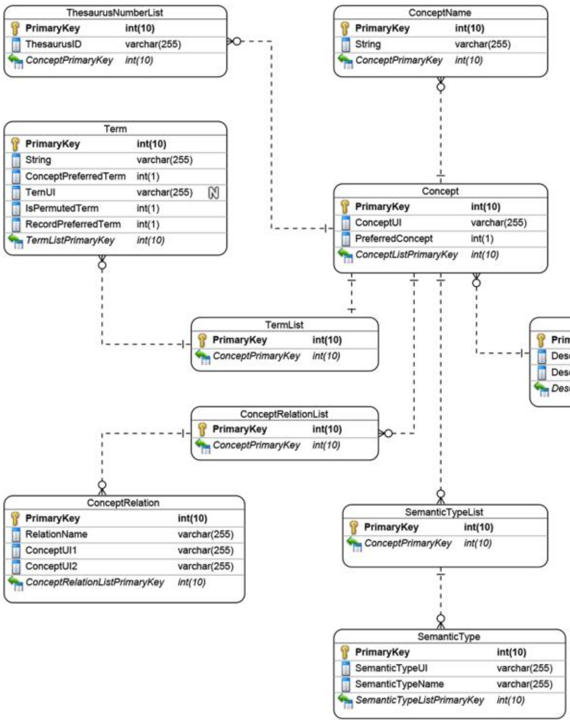


Figure 2 Complete data model of MeSH entities in the OPTIMED database

Multidimensional quality assessment

The 4-D quality assessment (Figure 3) represents a process which stands on four independent principles (review, typological classification, level of the target group, user self-study score), enabling an easy classification as well as a complex online review workflow with the use of XML template forms in the MEFANET portal platform.

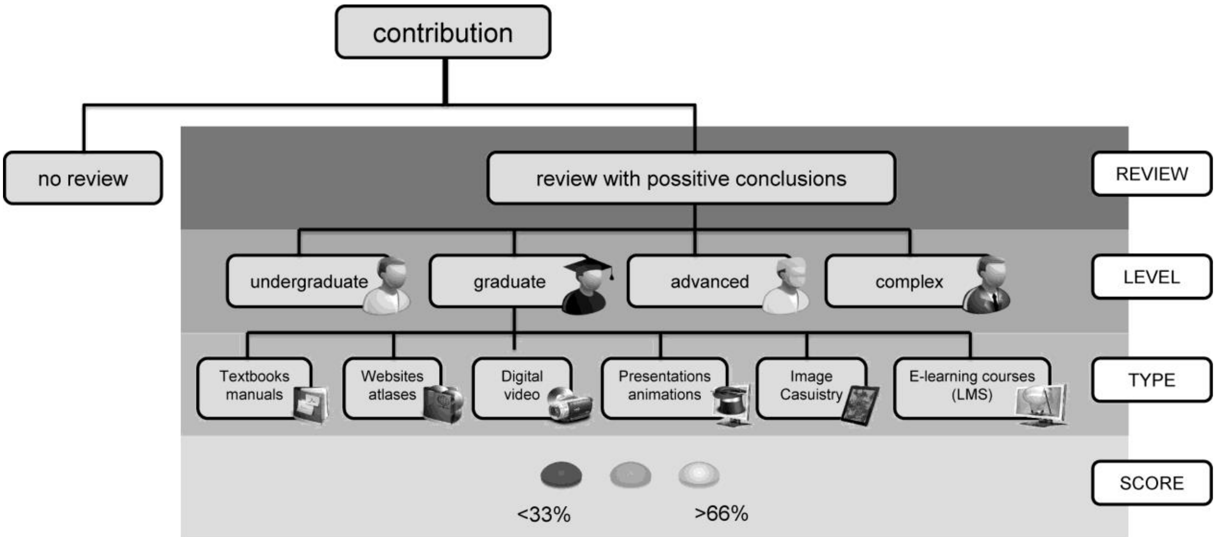


Figure 3 The MEFANET 4-D quality assessment mechanism

Quality criteria model

The MEFANET quality criteria model (Table 6) was inspired by (i) several studies published in the area of quality content assessment and (ii) the MEFANET stakeholders' experiences; as a result, the most significant properties were recommended. The model comprises website quality (design, development, performance, security, and privacy), content quality and service quality indicators.

Table 6 Quality criteria of the educational portal platform

Website quality	Web content quality	Web services quality
Screen resolution	Varied formats	Multi-level navigation
Corresponding icons and symbols	Content quality assessment	Searching facilities
Simple and easy navigation	Transparent categorisation	Communication
Readability of the information	Target audience	Printable version of content
Accessibility of the sites	No anonymous content	Technical support
Consistent representation	Updated information	Authors support
Site aesthetics, colour harmony	Object description	Administration background
Code and links validity	Amount of data	Support of online inquiries
Authorisation	Content uniqueness	
SEO		
Optimisation, standardisation		
Website feedback monitoring		
Page loading time		
Response time		
Failure restriction		
Rights management		

MEFANET user's identity framework

The MEFANET authentication mechanism is based on the SAML (Security Assertion Markup Language) concept, which permits the secured transfer of authentication data between independent organisations with an established trust relationship. The eduID.cz federation provides a single web-based sign-in approach that ensures users access multiple applications using just one password, which users generate and use within their home organisation. The educational content is comprised of contributions which are accessible to anyone who searches the portal. Every attachment (file or link) has its access rights (see below), which strictly define users who can access it:

- non-registered anonymous user,
- registered anonymous user who accepts the terms of use upon his registration,
- user of the MEFANET network, i.e., student or teacher from any Czech or Slovak medical faculty,
- user from a local university whose affiliation to that university has been verified at the portal via the regional information system of that university,
- user from a local medical faculty whose affiliation to that faculty has been verified at the portal via the regional information system of the respective university or faculty,
- user to whom attachments are made available only with the author's explicit consent.

Data exploration methodology

Many different data mining methodologies have been developed. They are well established in systematically approaching and solving tasks similar to medical and healthcare curriculum data analysis, modelling and deeper understanding. The CRISP-DM (Cross-industry standard process for data mining) methodology (Figure 4) has been adopted as a model that completely fits the curriculum exploration, which, among other things, fully supports the following crucial steps: (i) the understanding of the curriculum innovation objectives from the academic perspective, which are defined by research questions; (ii) understanding initial curriculum data arrangement, including the identification of data quality problems; (iii) medical experts' evaluation and refinement of mined results are more practical to higher education institutions [1].

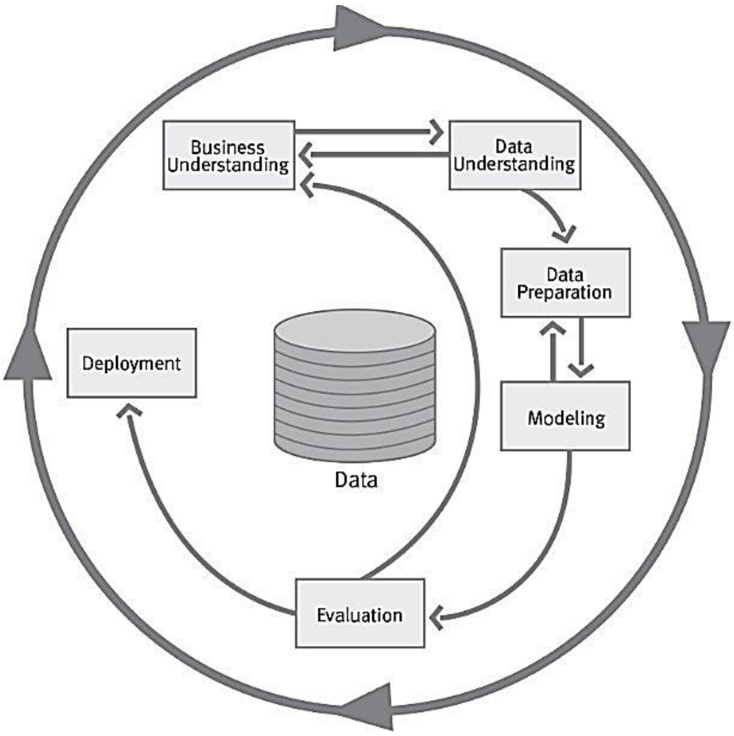


Figure 4 Diagram of the CRISP-DM stages

Technology-enhanced assessment

Testing of students' knowledge is a critical moment in the educational process. Appropriately composed tests also play a significant role in individual study courses and the whole educational process, as the form and content of a test often determine which areas students will focus on and which skills they will master. Knowledge assessment methods were described in detail, including written and computer-based testing possibilities, with particular attention to medical and healthcare education. The whole cycle of test preparation includes setting teaching goals, test blueprinting, appropriate item construction, item review and test piloting [9].

Exam preparation, running and evaluation based on the OSCE (Objective Structured Clinical Examination) method is complex. Running must be practical and as little as possible time-consuming for the OSCE guarantor (i.e. the person planning the agenda, managing the technical team and overseeing the smooth running), for the teacher (i.e. an expert preparing the scenario for a given station and acting as the observer/evaluator) and for students themselves. From a methodological point of view, the following stages were identified:

- creating a station, including allocating a room and assigning observers, simulated patients, and students;
- carrying out the examination, with evaluation recorded online into predefined forms in real time;
- evaluation and analysis of results according to stations attended by students to provide instant feedback to students, observers and the OSCE guarantor.

For this reason, OSCEs at MED MUNI are divided into three main stages (a dedicated submodule in the SIMUportfolio is shown in brackets): (i) design and planning (the “Sketch” module), (ii) running of the examination itself (the “Execute” module), (iii) evaluation of the examination and providing results (the “Report” module) [10].

Curriculum mapping

A correctly compiled and balanced curriculum across medical fields is essential for medical education. According to the conclusions of the Bologna Process, an all-in-one web-oriented platform covering all elements linked to global curriculum management was designed. It provides transparent information about what knowledge shall be obtained over the six years of medical study programmes, what topics will be on the agenda, which areas will be covered repeatedly and which courses are linked to the studies [11]. Curriculum mapping is designing a multidimensional model of an educational programme for a complete, more transparent, and better-integrated learning experience. Based on a state-of-the-art narrative review [12], a list of 12 critical characteristics expected to be implemented in the planned technical curriculum mapping infrastructure were identified in four thematic categories that gathered the focal points of interest of the involved community (Table 7) .

Table 7 Thematic categories in descriptions of technical infrastructure for curriculum mapping

Category	Characteristic
Visualisations	Visual overview of the curriculum
	Visual relations between various components of the curriculum
	Identification of redundancies in learning objectives
Text-based analytic functions	Complex reporting based on available curriculum building blocks
	Possibility to search by keywords
Outcome-based approach	Evaluation of learning objectives
	Outcome-based education compatibility
	Export of curricula by course, study field, department, and faculty
Adaptability	Possibility to modify reports and outputs according to institutional requirements
	Integration of international recommendations
	Integration of different user roles
	Available online

The curriculum planning model (Figure 5), supported by the curriculum management system, helps academics reengineer their curriculum activities. The planning process is delineated into three significant stages and is followed by the final evaluation process.

This model represents an original methodology for creating a new structured set of curriculum building blocks designed to assess and adjust to real medical and healthcare education [13].

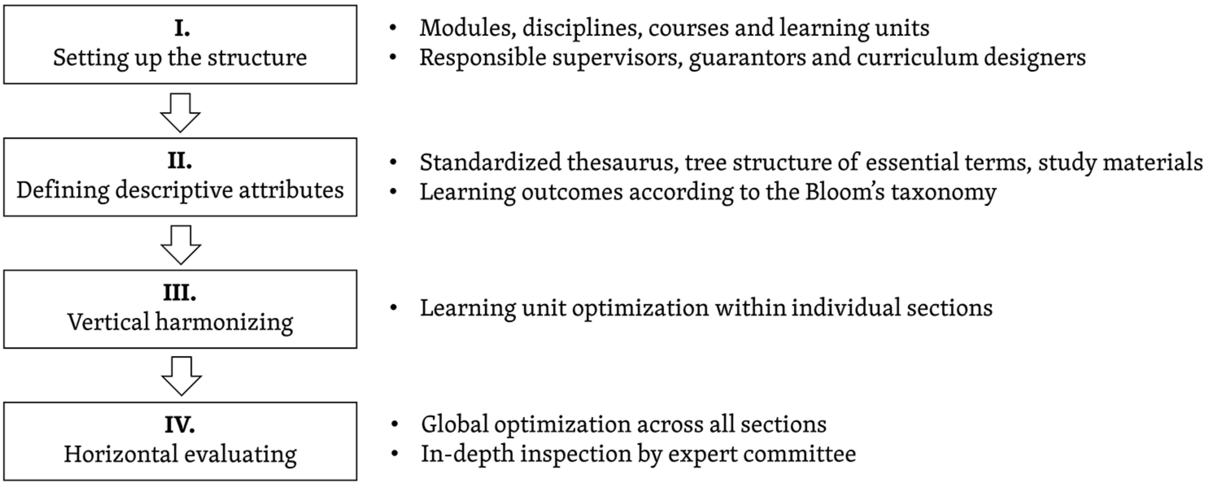


Figure 5 The planning model introducing curriculum innovations

The proven CRISP-DM methodology was chosen to increase the success and accuracy of solving individual tasks. One of the analytical tasks was to verify whether the medical curriculum of undergraduate training at MED MUNI addresses the issue of non-healing wounds in different years of study in both preclinical and clinical medical disciplines. Natural language processing techniques were used to map critical concepts onto the curriculum. The results in summary tables and graphs showed where to focus across the curriculum of each discipline to improve the overall quality of teaching as regards pressure ulcers and other non-healing (chronic) wounds [14]. Another task identified overlaps among medical or healthcare disciplines using term similarity. Close attention was focused on Rehabilitation and Physical Medicine and its role in the General Medicine study field at MED MUNI. Achieved results have demonstrated overlapping areas among the defined disciplines in the explored curriculum. The resulting comprehensive analytical report presented the term occurrence in figures and tables, which were thoroughly evaluated by experts familiar with the curriculum design process [15]. Another topic was the practical use of the selected thesaurus in a curriculum mapping application. Our research showed how the curriculum was consistent with the MeSH thesaurus and how the MeSH thesaurus can be used to demonstrate the interconnectivity of the curriculum through association analysis [8].

A proof of concept in online data crawling and scraping was carried out in a student project. Accurate data from a guaranteed online source on dental care were automatically mined and processed by a set of Python algorithms and stored in a relational database running on our servers. The challenging task of complex mapping between data describing the distribution of dental practices was successfully solved. It provided an original overview of data stored in the portal of the Czech Dental Chamber because it revealed hidden relations between graduates' and dentists' distribution across the Czech Republic [16]. Another case study was focused on natural language processing techniques and morphological analysis of medical and healthcare curricula. The main idea was based on adopting a morphological analyser to enhance the medical curriculum management system and the ability to identify features of morphological analysis that make the search engine more effective (i.e. more accurate and faster) [17]. An automated similarity detection between study materials and curriculum descriptions in medical and healthcare education was also challenging. The following research questions were formulated to define and solve a particular research problem: What is the relation between the results achieved in the form of detected similarities done by computer and an evaluation by users (medical students and teachers)? Which similarity detection approach can be effectively implemented in a particular domain of medical education? Based on statistical calculations, similarities and standard features/terms were found between interactive algorithms in the AKUTNE.cz platform and learning units in the OPTIMED platform. This study did not find significant similarity between our calculated results of Cosine, extended Jaccard and Dice similarity and ratings of medical students and teachers [18]. Another related task was to measure similarities (based on keyword extraction) between particular virtual patients and specific curriculum building blocks. Here, the string-based similarity of terms (the frequency-based approach) was used with the normalised Pearson correlation coefficient as the similarity measure [19]. For illustration, Figure 6 shows the most relevant curriculum building blocks for a particular virtual patient.

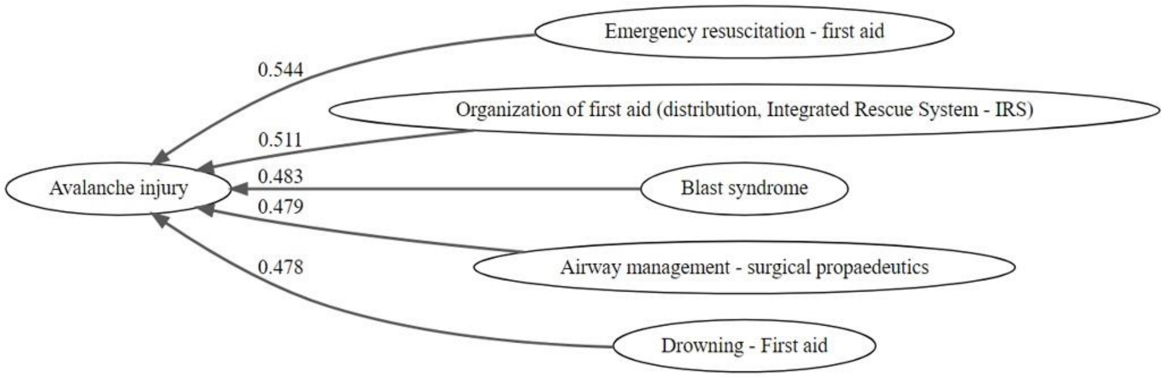


Figure 6 Similarity graphs show the five most relevant learning units to a virtual patient called "Avalanche injury"

Moreover, the following research problem was applying data mining and analytical methods effectively for a crucial keyword-based exploration of medical curriculum data and which graphic component fits the best for a web-based presentation of achieved results. A pilot R package analysis, including web-based visualisations per a proven data-mining methodology, led to actual results in practice [20]. One of the advanced tasks was focused on the kind of centrality to identify the most accurate set of keywords in the curriculum metadata description. A novel method for identifying critical terms from a free text covering the medical curriculum has been implemented. This method is based on computing node centralities in a similarity graph of terms in a given medical and healthcare discipline description, whereas a popular word2vec model provides the degree of similarity [21]. The overlapping areas and gaps in medical and healthcare curricula were explored using the social network analysis [22].

Assessment data analysis and reporting

Thanks to the experience gained at the national and especially international level, the MED MUNI team has understood and fully adopted the comprehensive OSCE workflow.

It is essential to appropriately combine the management requirements, the time, technical and spatial possibilities of the workplace, and the methodological background with recommendations for good practice. SIMUportfolio fully covers all the complicated OSCE process features using the newly developed dedicated OSCE module, where primary and advanced overviews describing the progress of the OSCE assessment in an understandable way are available [23]. A unique script for transforming the MeSH tree structure to a linear representation of individual terms made it possible to simplify the links between the keywords and the particular curriculum building block. For further analysis, we merged the data (Table 8) about learning units (unitName, unitSection, unitClass) with data describing MeSH keywords.

Table 8 Attributes describing the medical curriculum used for case study analysis

Name of attribute	PostgreSQL data type	Sample value
unitName	character varying(255)	Skin anatomy
unitSection	character varying(255)	Internal medicine
unitClass	character varying(100)	Dermatovenerology
meshName	character varying(255)	Epidermis
meshPreferred	Integer	Yes
meshSemanticType	character varying(50)	Tissue
meshTreeId	character varying(100)	A10.272.497
meshTreeLevel	Integer	3
meshCategory	character varying(50)	Anatomy
meshSubcategory	character varying(50)	Tissues

Sharing experience in software development

The primary goal was to establish a method for fostering collaboration among cross-border development teams engaged in creating scientific software solutions for academic and non-profit organisations, particularly in medical and simulation education. Designed and published guidelines are based on the evidence of best practices and recommendations in the field of design, development and implementation of web applications, mainly in medical education and science. A set of descriptive interactive overviews [24] of the available data brings a report keyword overview, clustering based on application domains, and a summary of the Likert scale quantitative evaluations. An interregional collaboration served as a foundational model for cooperation within the Czech-Austrian network of software development institutions. The implementation of the project increased the competitiveness of employees and institutions in the participating regions and was of national and international importance. It encouraged collaboration in developing web presentations and applications tailored to the precise requirements of medical faculties and simulation centres.

Technological frameworks

The MEFANET's e-publishing ecosystem

Practical, open, and long-term cooperation among all Czech and Slovak medical faculties in the medical and healthcare education fields has been covered by the MEFANET network. The original MEFANET educational web portal platforms represent a complex e-publishing system consisting of 12 standalone portal instances (i.e. the e-publishing portal of MED MUNI [25]) and the central gateway [26]. It serves as a robust e-publishing tool, which provides the following benefits in particular to the users at engaged medical faculties: (i) easy data sharing and capturing; (ii) ease of use and access; (iii) common authentication framework; (iv) quality assessment of published contents. Each instance of the portal platform is composed of two major sections (Figure 7): FrontOffice (non-secured interface to present the published contents to users) and BackOffice (secured interface to manage a selected area of functionality used by editors and administrators).

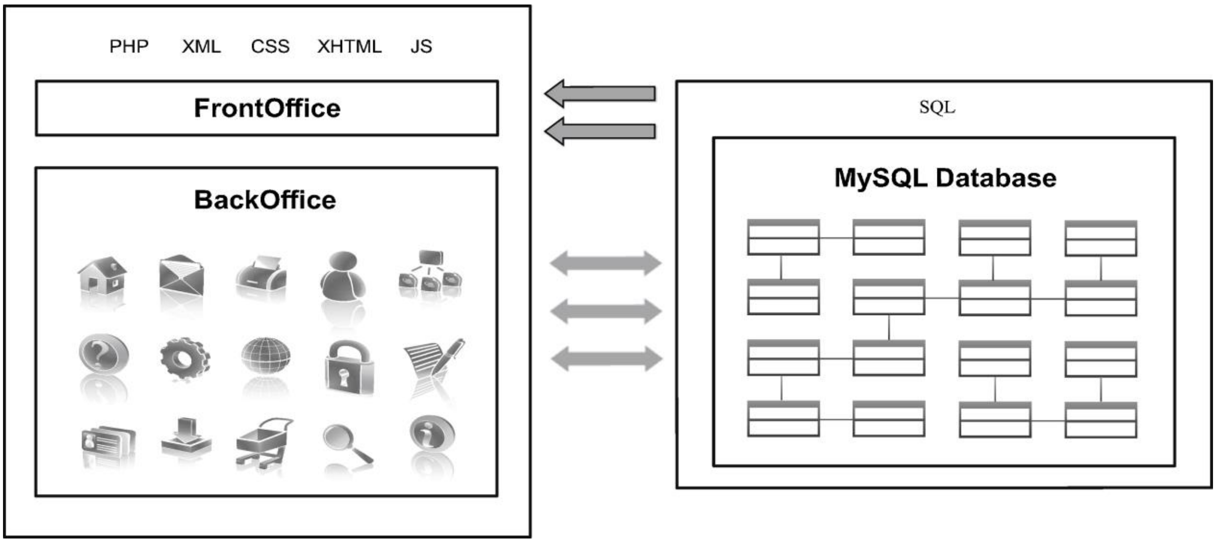


Figure 7 Global view of the structure of the MEFANET web portal instance

Each system module is built on the most widespread software architecture for web applications called Model-View-Controller (Figure 8). This model is responsible for managing the data and storing and retrieving entities an application uses. Three types of views are implemented in the MEFANET portal platform: FrontOffice view, BackOffice view, and Central Gateway view. The controller handles the model and view layers together, receives requests from clients, invokes the model to perform the requested operations, and sends data to the view [27].

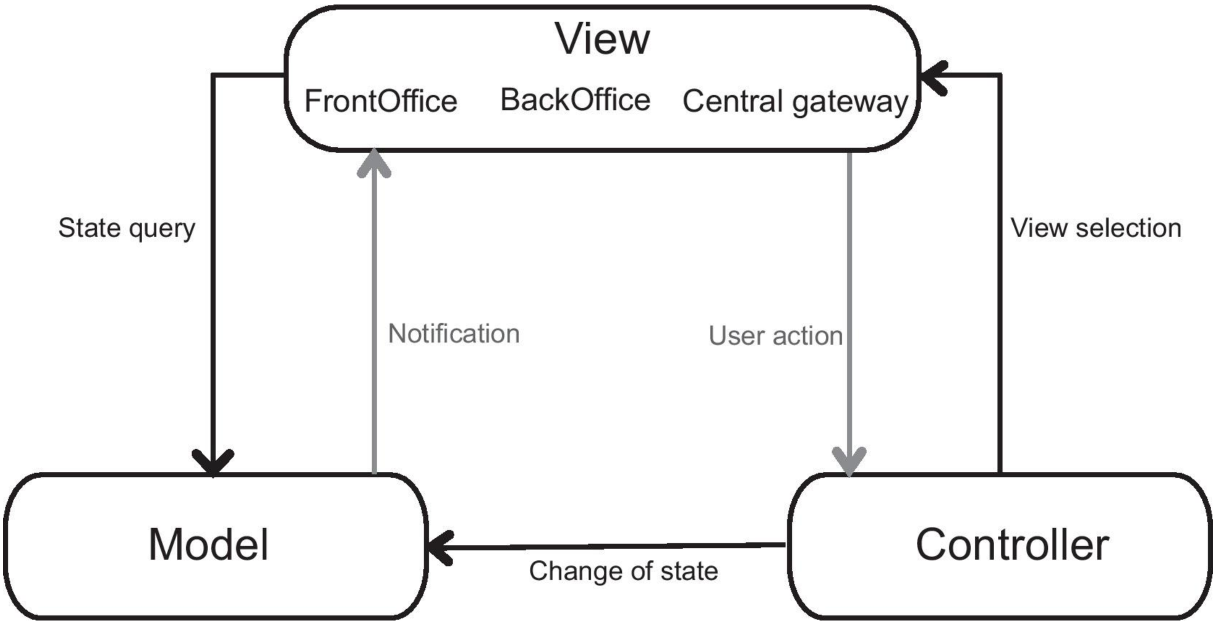


Figure 8 The description of relationships between components in the MVC architecture

AKUTNE.CZ: An interactive tool for simulation-based learning sessions

The education portal AKUTNE.CZ [28] is integral to the MEFANET's content, focusing primarily on simulation-based tools for teaching and learning acute medicine issues. Interactive algorithms take the form of content-rich virtual cases as they link together process flowcharts and multimedia. The portal properly combines algorithms development methodology, including a peer-review process and technologies supporting simulation-based learning objects in daily medical and healthcare education. The unique advantage of interactive algorithms AKUTNE.CZ is the possibility to create complex and branching scenarios in compliance with the proven methodology for incorporation in the learning and teaching of acute medicine [29].

Curriculum management platforms: OPTIMED and EDUportfolio

The number and diversity of publications presenting software solutions in curriculum mapping show high interest in electronic support for curriculum management and mapping; however, hardly any of those support international design and harmonisation of curricula in higher education. Based on the identified requirements by the needs analysis (Figure 9) among curriculum designers, teachers and managers on national and international levels and evidence published in the literature on curriculum development, web-based tools were developed to help manage, map and analyse curricula in the medical and healthcare study fields [30].

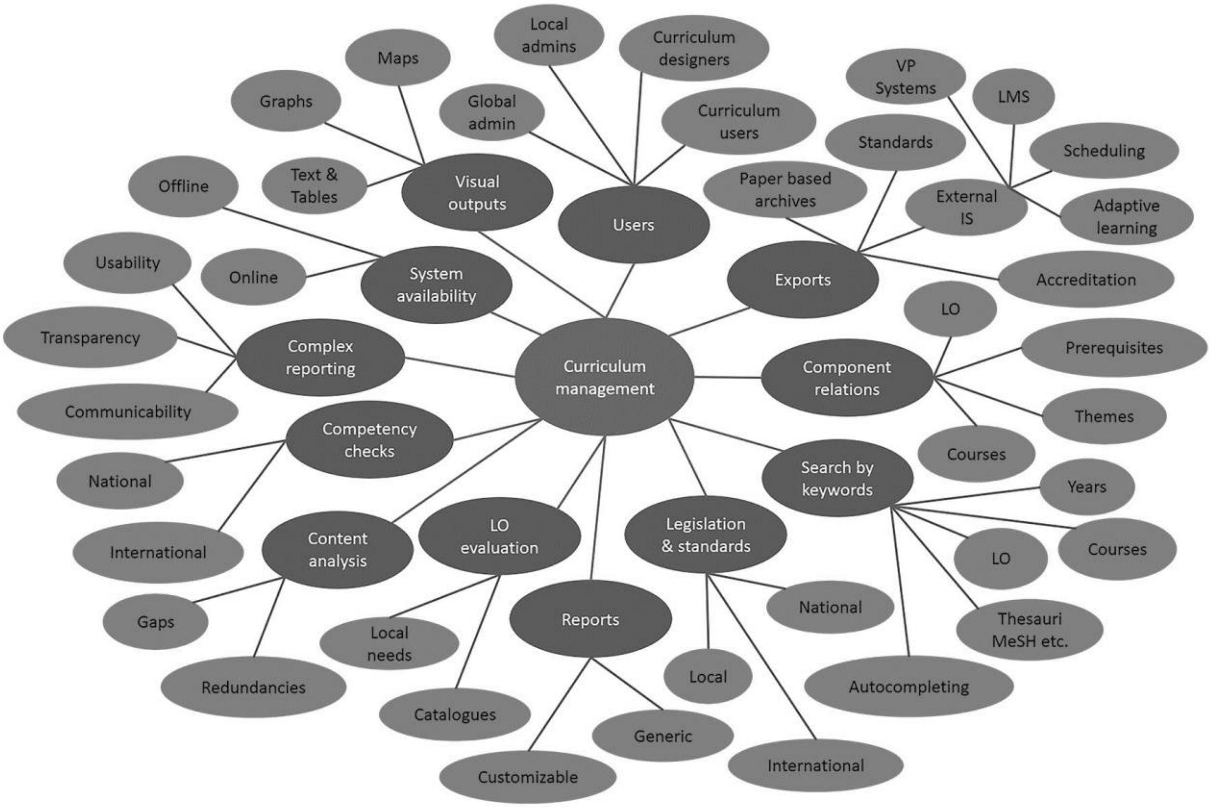


Figure 9 A mind map of features to consider while developing and implementing a curriculum mapping system

By approving the paradigm of learning outcomes as a fundamental building block of education [31], an original methodological model opens the possibility of reforming the curriculum structure effectively, as all elements are available in the form of parametric descriptions that have been designed and implemented [32]. A robust framework, the OPTIMED platform, demonstrated the purpose of outcome-based medical education through adequate management. The methodology supported by the platform has helped academics in their curriculum reorganisation efforts by providing a transparent overview of the curriculum structure [33].

The new curriculum management platform, EDUportfolio, with curriculum mapping functions, was developed as part of the BCIME project[47]. It is based on a compounded curriculum model, which makes it possible to define a hierarchy of teaching blocks and thus characterise it in a structured, parametric way at different levels of detail (e.g., for the study programme, medical discipline, course, learning unit and learning outcome) by international standards provided by the MedBiquitous association [34].

SIMUportfolio: Complex platform enhancing medical education

SIMUportfolio [35] represents an integration platform being developed at MED MUNI, which makes orientation in the teaching and learning processes more accessible for both students and teachers and, consequentially, should make students' knowledge and practical skills more effective. It contains an extensive description of the entire curriculum and metadata on recommended study materials. A predefined data structure

[47] https://www.muni.cz/vyzkum/projekty/43286

to store all interactions between users and the system, combined with the above-described data on the teaching process itself, presents exciting challenges from the points of view of advanced data analysis, data mining from extensive texts, recommending relevant contents to users, curriculum mapping and comprehensive online reporting [36].

The ESPO platform: Effective searching for project opportunities

ESPO [37] is a web application that supports an effective search of project opportunities and collaboration. The application uses data from project opportunities from the Crowdhelix website and projects from MED MUNI. Using the data available, the application provides an easy way of searching for collaboration opportunities based on specific researchers, individual projects or even a simple phrase [38].

The ORTODB platform: Orthopedic images presentation

The OrtoDB platform [39] is a web-based application that enables the creation and searching of orthopaedic examination records of actual patients at MED MUNI. Each record includes image documentation and essential patient information such as age, sex and details on the performed surgery or other procedure. One of the crucial features is the support of the most commonly used file formats in medical education, like DICOM (Digital Imaging and Communications in Medicine), together with the standard image formats.

Results: Health information and statistics

The ICT methodologies behind

To design appropriate concepts based on proven methodological foundations, combining experience from health information, knowledge extraction from databases, and specific domains of medicine (such as epidemiology or oncology) was crucial. A detailed understanding of the stakeholders' needs, knowledge of currently used technologies, and mapping of clearly defined evidence-based medicine recommendations helped design several methodological outputs successfully implemented in practice.

National methodology of health data sharing

Comprehensive management of access to and sharing of centralised data involves a wide range of steps, from standardising primary data content from NHIS, ensuring efficient collection, follow-up and editing of data to publishing it with guaranteed updates and considering user feedback. The primary aim is to implement tools that fully cover all stages of this process. The centralisation of any data potentially affects the entire healthcare system, i.e. tens of thousands of healthcare providers, their founders, health insurance companies, specialised institutes etc. The system must, therefore, be efficient, secure, easy to operate and control.

The motivation was to describe the methodological framework, including the essential technical background for publishing datasets from NHIS. Basic and advanced epidemiological characteristics related to the COVID-19 epidemic in the Czech Republic were logically selected as a pilot example given the current needs in 2020. All regimes of the data provision from NHIS strictly require a certain degree of legislative regulation and must fully meet the criteria set for NHIS by the Czech legislation, particularly Act No. 372/2011 Coll., on Health Services and Their Provision. In other words, publishing open data cannot be misinterpreted as the publication of primary records without any regulation and standardisation; the term “open data” does not necessarily describe primary database records (the data may be aggregated, statistically processed etc.). The dataset design, preparation and publishing should respect the algorithm of dataset preparation shown below, which always respects several principal rules: (i) individual natural persons must not be identifiable; (ii) individual legal persons must not be identifiable unless expressly stated by the law; (iii) secondary processing must lead to the pseudonymisation of the dataset; (iv) the purpose of the dataset publication must correspond to the NHIS purpose; and (v) the standardised process of approval and publishing must be adhered to (Table 9). The essential requirement of a comprehensive process, which this approach meets, is to ensure the necessary completeness, validity and overall data quality [40].

Table 9 A chart of dataset production and publication

Step	Description	Details
1	Concept design	Proposed by state administration, external subjects (health insurance companies, expert societies, research institutions)
2	Concept evaluation	Purpose, data availability, feasibility, legal perspective
3	Feasibility analysis	Data extraction, processing, analysis, validation
4	Dataset production	Structure, methods of production, metadata description
5	Review	Personal data protection, factual content, IT solution
6	Publication	National Health Information Portal (https://www.nzip.cz/)

The dataset publication consists of six steps describing the critical preparation and implementation phases.

- **Step 1**: Proposal of a dataset concept in the form of a short description, which can be raised by any entity, usually a state institution, a health insurance company, a professional (medical) society, a research institution or an academic institution.
- **Step 2**: The delivered concept is reviewed from the perspective of data availability, export feasibility, design and personal data regulations.
- **Step 3**: After approval, a methodology for the dataset processing is proposed (data export from central registries, data pre-processing and cleaning, analytical adjustments and validation mechanisms).
- **Step 4**: The dataset is generated in an open data standardised format according to the predefined scheme, including an obligatory description with metadata.
- **Step 5**: Reviewing and validating the dataset and its content is a mandatory procedure before publication (a manual approach is always needed, as well as technical control to follow open data standards and best practices).
- **Step 6**: Publication of the final dataset in the National Health Information Portal.

Legal background

The database for centralised data sharing is NHIS, the components of which have a clear legislative mandate under the acts listed below:

- Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision;
- Act No. 325/2021 Coll., on the Digitisation of Healthcare;
- Act No. 89/1995 Coll., on the National Statistical Service;
- Act No. 285/2002 Coll., on the Donation, Procurement and Transplantations of Tissues and Organs;
- Act No. 258/2000 Coll. on Protection of Public Health and Amendment to Some Related Acts.

The forthcoming proposal for a regulation on European Health Data Space (EHDS) will have implications for NHIS in both primary data processing and secondary data use. In primary data management, implementation of the EHDS can be expected to emphasise standardising electronic documentation content directly with health services providers and health insurance companies. These processes will increase the information value of data and its usability. The national methodology already implements the main principles of secondary use of health data, and several conditions resulting from the EHDS have already been met within the national legal regulation [2].

All regimes of the data provision from NHIS strictly require a certain degree of legislative regulation and must fully meet the criteria set for NHIS by the Czech legislation, particularly Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. In other words, publishing open data cannot be misinterpreted as the publication of primary records without any regulation and standardisation; the term "open data" does not necessarily describe prior database records (the data may be aggregated, statistically processed etc.) [1].

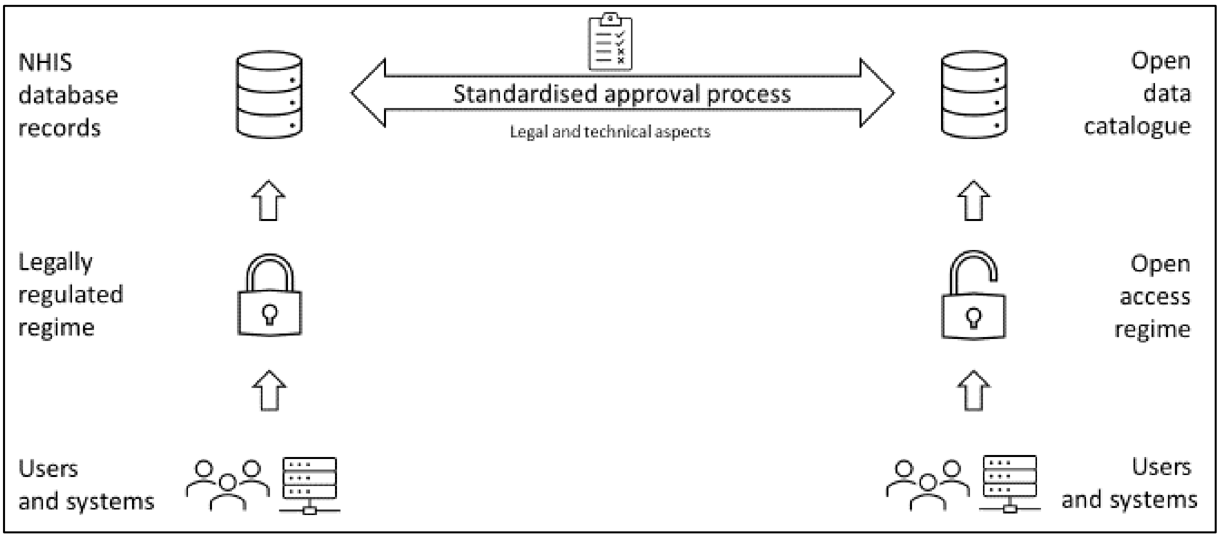


Figure 10 National Health Information System: Data provision schema

Main approaches to health data publication

Below are various forms of output of published NHIS data from sources that are non-public by law and are therefore always made available after necessary modifications: (i) aggregated statistical publications, NHIS reports and yearbooks; (ii) aggregated datasets; (iii) thematic open datasets; (iv) univariate open datasets; (v) comprehensive datasets published with synthetic data. As for the individual target groups, each form has a clear added value and, at the same time, clearly defined descriptive attributes to minimise misinterpretation of the published data.

An example is a thematic open dataset, where the range of parameters used need only follow the specified theme and is finite once the thematic dataset is published. If the topic is required to extend further, a new set must be prepared, again narrowly focused on the given topic. Chaining and concatenation are not possible with this format. For disaggregated data, the statistical units are identified by a unique identifier that is not linked to other open data. In the case of aggregation, the modification of the data before publication must group the statistical units in question (patients, procedures, hospital admissions etc.) into groups whose information scope meets the conditions for sufficient protection of personal data (categorisation of continuous variables, such as age, level of territorial subdivision of the Czech Republic or region only, groups of diagnoses instead of individual disease codes etc.). A defined minimum frequency rule (threshold of 10) is applied if necessary, but this is not a strict rule. An unlimited number of datasets focusing on different sub-areas can be created for a single data source or topic. Each thematic dataset must be accompanied by more structured metadata that aims to provide more information than the established open data standard (title, basic description, primary objective, interpretation and information value of the data, examples of use etc.)

Target groups

An essential part of processing various data outputs in healthcare is connected with a proper definition of target groups and stakeholders. A universal and comprehensive methodological description of the process of design, preparation, validation and publication of data is essential information with which the broad community (composed of several general target groups) dealing with data in healthcare must be familiar to re-use the datasets. Ensuring clear and consistent communication about open data in healthcare is essential to a functional system. User profiling should consider communication goals and information, health and data literacy. The outcome of such thinking is a division into three basic levels, which may overlap in preference of output format (Figure 11).

Figure 11 Communication of health data

Clarity, level of detail, guarantee of validity, and correct interpretation are essential in terms of minimum requirements and demands. Such a concept requires close collaboration between experts across the necessary expertise (theoretical physician or clinician, member of an expert medical society, healthcare expert, computer scientist, analytical guarantor, guarantor of data visualisation or open data guarantor). Depending on the format of the output, these people are actively involved in the actual process of communicating with the selected target group. In particular, data explorers and experts have an excellent opportunity to disseminate the results correctly among stakeholders mandated to reach the general public, i.e. data novices [1].

The Czech national cancer registry: Data processing, analysis and visualisation

Within the processing of various medical disciplines falling under the individual NHIS components and agendas, several comprehensive methodological materials have been created, which appropriately combine health informatics, legislative aspects and the given speciality, including their specifics. The field of oncology is presented in this thesis for illustration. The methodology covers several procedures on available data from the Czech National Cancer Registry; more specifically, data describing childhood cancers. These data were validated using the clinical database of childhood cancer patients and subsequently combined with data from the National Registry of Hospitalised Patients and death certificates. Records on cancer patients were subsequently classified into 12 main groups, according to the International Classification of Childhood Cancer (3rd Edition) [40]. Regarding statistical analyses, the authors focused on three epidemiological indicators: incidence, mortality and survival. The design and development of the web portal equipped with an online data browser was driven by the modern and practice-proven Symfony PHP framework. The data repository was built on PostgreSQL – an open-source object-oriented system which is routinely used to organise more complex data structures. All individual steps, i.e. (i) the validation of theoretical background, which describes the basic terms, (ii) the methodology of cancer classification, and (iii) static analytical reports, were performed internally; that is, in cooperation with the analytical team and a group of senior doctors who are specialists in childhood cancers and have many years of experience with the methodology of childhood cancer classification [41].

Another outcome introduces the processing, modelling, analysis, and visualisation of cancer epidemiology and care data in compliance with a proven and validated methodology. The authors aim to provide online access to unique cancer care and epidemiology data, including an interactive visualisation of various analytical reports to provide relevant information to the general public and experts, such as healthcare managers, environmental experts and risk assessors. With the proven CRISP-DM methodology, the challenge of performing efficient modelling, storage and visualisation of data on cancer epidemiology from a territory of the Czech hospital network was solved. Figure 12 illustrates the data warehousing concept, which makes it possible to integrate information from heterogeneous databases and to query extensive databases efficiently. The main idea of staging, operational data storage, primary data warehouse and data mart was successfully implemented [42].

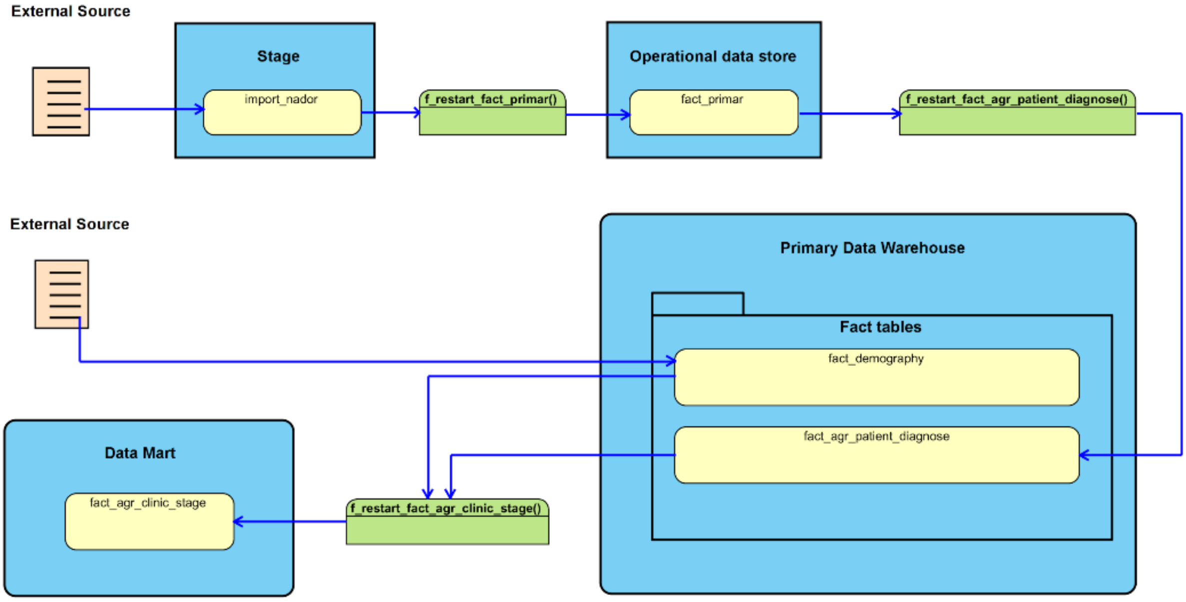


Figure 12 Data flow in data warehouse subset

Information service on COVID-19 epidemic data: Background

Due to an urgent need for a tool during a pandemic situation that would provide essential reports based on valid data sources, a team of government experts and researchers focused on the design and development of a web application intended to provide a regularly updated overview of COVID-19 epidemiology in the Czech Republic to the general population. Figure 10 describes the complete workflow of how the patient suspected of being SARS-CoV-2 positive is referred by his general practitioner or regional public health authority for biological sampling at the testing point, how the patient record flows through the system and, finally, how the record is further processed and stored in the Information System of Infectious Diseases (ISID) [44].

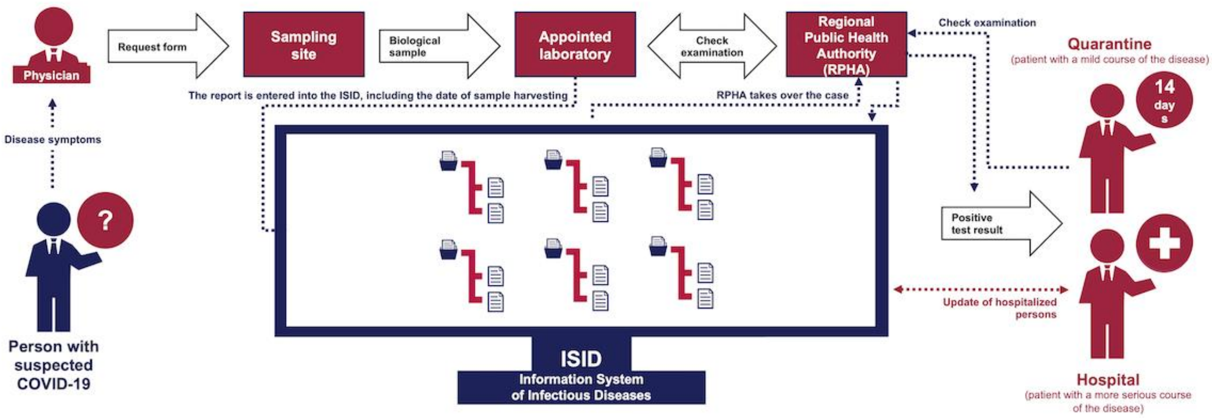


Figure 13 Simplified schema of newly identified cases of COVID-19

The entering, processing, and subsequent visualisation of data describing the current epidemiological situation about the new coronavirus SARS-CoV-2 required meticulous validation. Some datasets might have been extended by additional attributes, or their structure might have been changed after publishing (for example, data describing demography or region/district coding, which provide an extra point of view regarding population distribution in this particular case). The datasets were available in all versions using the application programming interface (API) to maintain backwards compatibility. After careful data design, processing, preparation, validation and standardisation, the data were published manually in the local catalogue using a graphical user interface, which was accessible only to authorised members of the open data team at the Ministry of Health of the Czech Republic. The entire process is captured in Figure 14 [40].

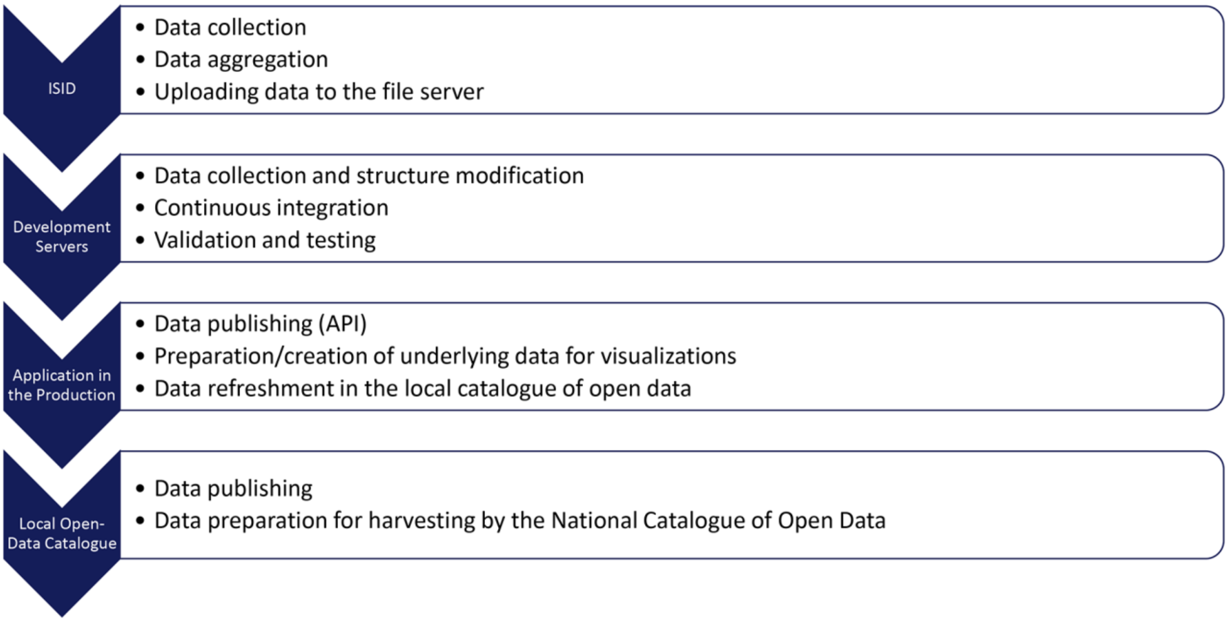


Figure 14 The process of data processing and publishing

Several data sources must be validated thoroughly and transformed into a form suitable for computer processing and subsequent visualisation as a daily-based online information service. In particular, the designed model (Figure 15) involves a set of control mechanisms ensuring the completeness of individual records. It involves data cleaning (incomplete records are detected, corrected, or removed), construction of particular data views (selected attributes with primary identification are derived and merged) and data integration (final data sets are automatically generated).

The two years of the COVID-19 pandemic have also made it possible to study and evaluate the phenomenon of infodemic in the countries that were among the most affected in Europe. An effective combination of good journalistic practice, medical expertise and modern information and communication technologies made it possible to identify selected forms of information distortion in various media. These included, for example, inadequate or incorrect work of journalists with data, wrong choice of reference examples and comparisons, or the need for more knowledge of medical terminology. The main issue was identifying the essential characteristics influencing a digital infodemic in the Czech Republic, including a detailed understanding of media communication and the performance of key stakeholders on national and regional levels. These characteristics are based on the information “cake” model [45], which illustrates the very first broad roadmap to fight an infodemic. A newly designed information matrix of COVID-19 infodemic management in the Czech Republic extended this approach and properly combined four fundamental pillars (facilitate accurate knowledge translation; knowledge refinement, filtering, and fact-checking; build e-health literacy; monitoring, infodemiology, infoveillance, and social listening), involved stakeholders and tools (experts and researchers; government and the Ministry of Health; media) driving publishing information about the COVID-19 epidemic [46].

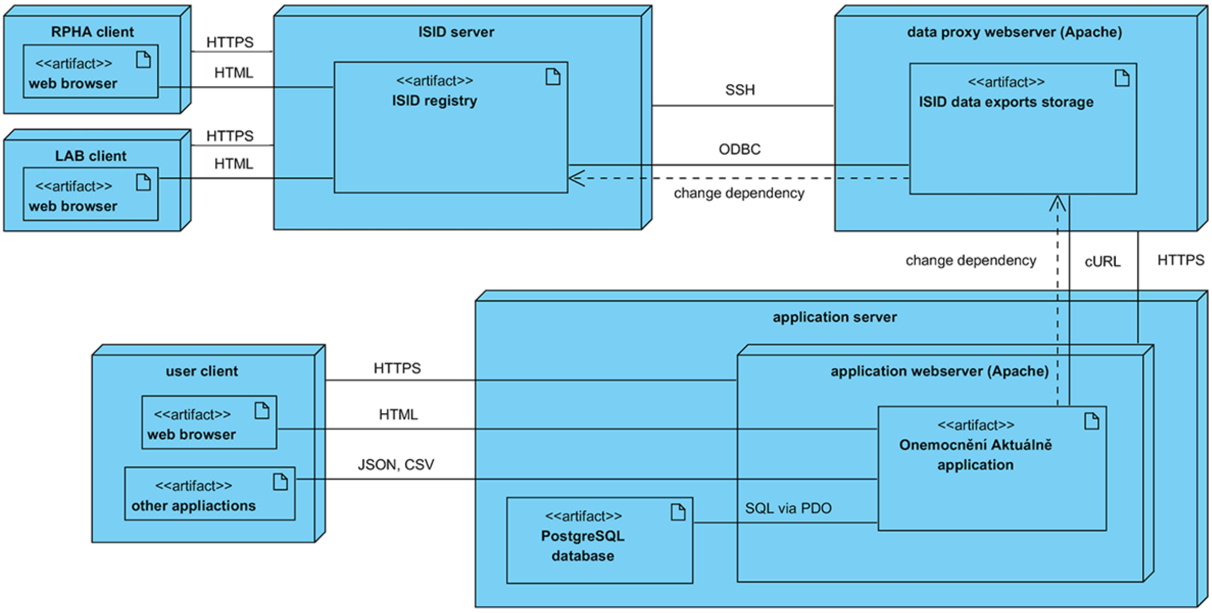


Figure 15 The diagram of an implementation of individual components, showing the communication among three independent servers transmitting data sources.

Real-time control centre for intensive care: Background

During the COVID-19 epidemic in the Czech Republic, strategic data-based and organisational support for individual regions and acute care providers at the nationwide level was highly needed. Control Centre for Intensive Care as a web application was developed and made available to meet the needs of systematic online technical support for managing intensive inpatient care. By the defined methodology, this system was brought into operation in April 2020 to enable the continuous monitoring and daily reporting of available data on the occupancy rate of inpatient beds.

Use of the application (available in the Czech language only) relied on several fundamental principles: (i) user authentication and authorisation; (ii) definition of user roles and rights to edit and read; (iii) the leading representative approves each new user registration; and finally, (iv) private and free email addresses were not allowed. All users were representatives of individual healthcare facilities (management, coordinators of intensive care, physicians, and nurses), emergency medical service workers, pharmacy workers, and Integrated Central Management Team members. The use case diagram in Figure 16 provides an overview of the entire system, including critical functions assigned to users in specific roles. The main objective of this diagram is to depict the various actors as well as interactions that are available to these actors. Use cases provide insights into the basic structure of functional requirements. Based on the modelling of individual use cases, which are associated with the primary actors, authors can see individual parts of the system so that it is possible to decompose and divide the whole system into separate submodules and continue with phases of design and development [47].

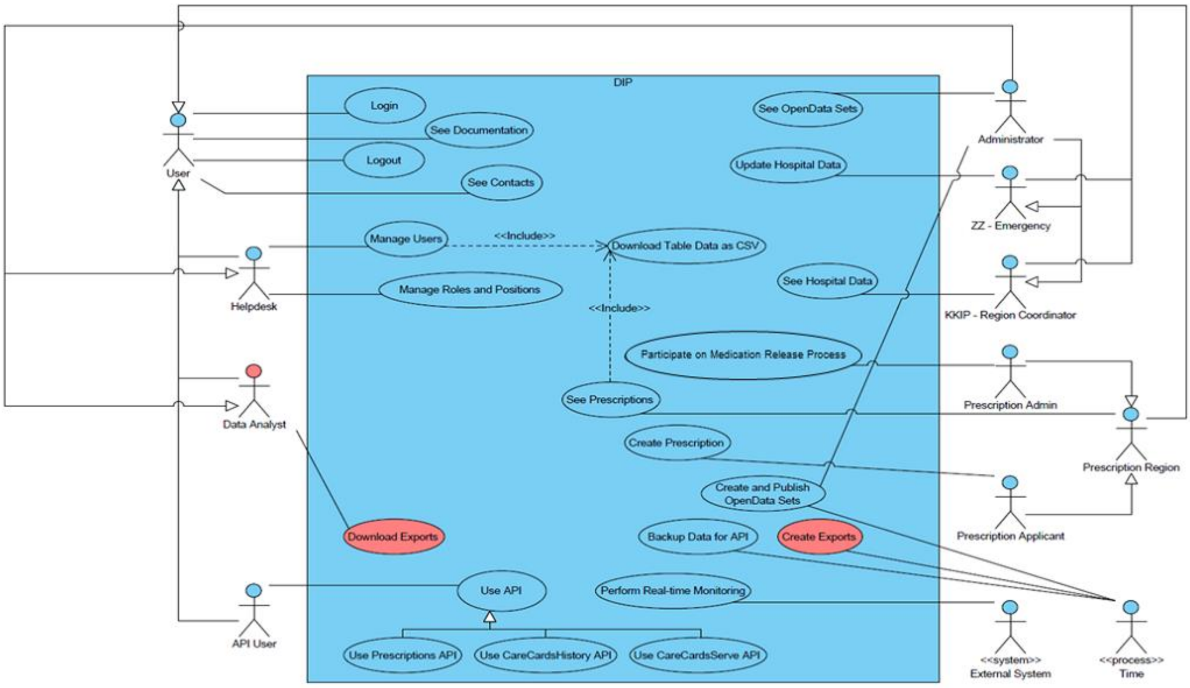


Figure 16 Use case diagram for the Control Centre for the Intensive Care system

Technological frameworks

National Health Information Portal supporting health literacy

The current state of the presentation of health information on the Internet in the Czech Republic needs improvement. It is characterised by a high degree of fragmentation and low credibility of information. There are dozens of specialised portals, partly published by government authorities and partly by various entities which do not indicate the authors, guarantors or relevant links to sources for the information presented. To unify communication channels and establish a centrally managed online tool, the National Health Information Portal (NZIP) [48] was created to aggregate and publish health information in a fully authorised form and understandable way to the target groups. Its main objective is to improve the communication and credibility of information provided by the Ministry of Health of the Czech Republic, both among public administration representatives providing healthcare and towards the general public. NZIP helps to promote the development of health literacy through comprehensible content. The primary division is divided into main sections (specific situations in life, prevention and healthy lifestyle, information on diseases, recommended websites and data reporting), covering essential information in articles, explanatory index terms, and additional external sources [49]. Interactivity as a tool for promoting health literacy is represented by incorporating the concept of virtual patients or clinical teaching scenarios, which have been adapted to the general public as educational quizzes and story games.

Data reporting on the National Health Information System

Health data and its correct and precise interpretation also play a crucial role in healthcare, especially in supporting objective decision-making. As an essential part of NHIS, nationwide registries must recognise their substantial contribution to current medical research, mainly in highly progressive segments of care, where fast innovations make a long-term follow-up of stable cohorts of patients impossible. These registries also play a crucial role in mapping heterogeneous trajectories of patients exposed to a wide variety of interventions, e.g. in the field of palliative care. In such a fluid situation, clinical studies requiring sufficient size and homogeneity of compared samples are almost unfeasible. On the other hand, modern registries have proven their utility by producing valuable data that allows reliable disease surveillance in such complicated circumstances. In principle, we distinguish three principle levels of clinical data registration: (i) central registries that create an identity space for e-health; (ii) nationwide registries that centralise the primary sources of administrative data and enable population-representative analyses of health status, access to care, consumption of health services and outcomes of care; (iii) sub-diagnostic and clinical registries that complement nationwide administrative data with detailed diagnostic indicators and predictive and prognostic markers or monitor the longer-term evolution of patients' health status. Linking the administrative data of providers and health insurers significantly reduces the data that would otherwise have to be collected in disease-specific sub-registries. Administrative and operational data are then parametrically supplemented by additional records on diagnosis, treatment and disease progression as computerisation progresses, which gradually shifts the narrowly focused clinical registries to the position of IT background for research projects (Figure 18) [50,51].

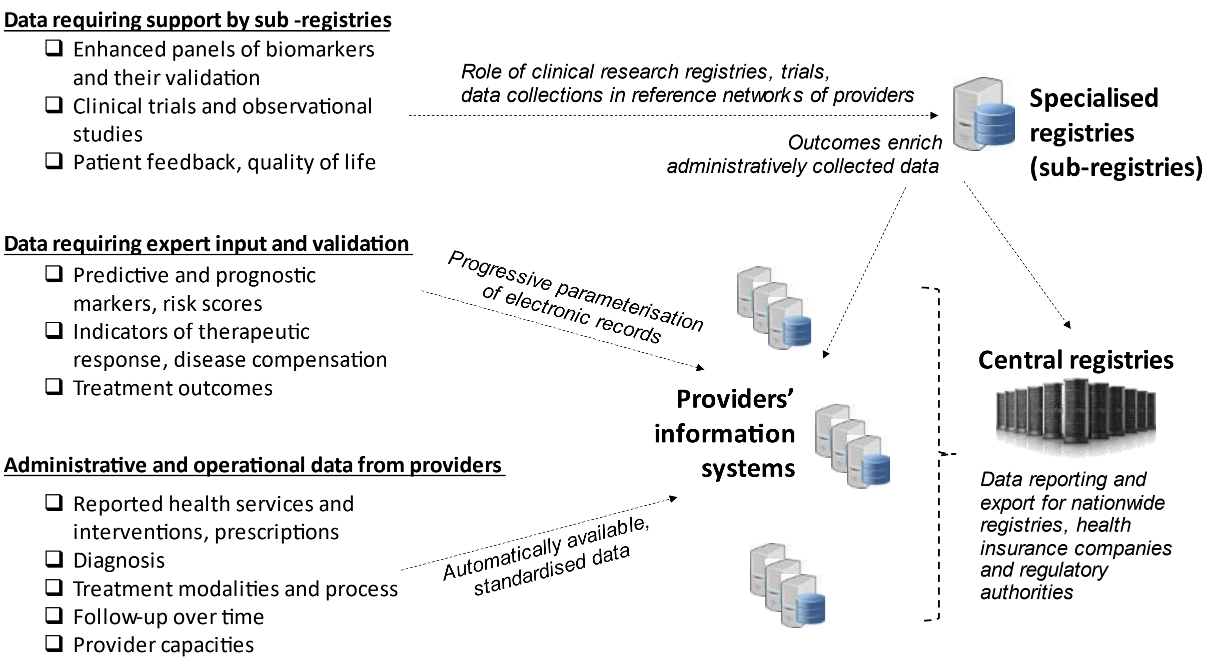


Figure 17 Different data sources for the evolving computerisation and standardisation of providers' medical records

The newly built information system based on health data (NHIS) provides various analytical outputs as valid and correctly interpreted material for strategic decision-making. The overarching platform for publishing and communicating these outputs is data reporting, a comprehensive information service and an integral part of the National Health Information Portal [48]. It covers all the individual components and agendas of NHIS in various formats (data summaries, analytical studies, yearbooks, open data etc.), which are always designed and published in collaboration with professional societies. The primary aim is to provide a comprehensive overview of specific disciplines in Czech healthcare. Selected topics of Czech healthcare have been prepared in close cooperation with representatives of professional societies as a comprehensive information service, such as cardiology [52], vaccinations [53] or reproductive health [54]. As of 16 April 2024, a total of 147 open datasets have been published in the Data reporting module by the national methodology, together with 59 data summaries, 10 analytical studies and dozens of interactive visualisations in the form of stand-alone data viewers or infographics. The design and publication of various NHIS outputs represent an innovative and sustainable resource guaranteed by experts and contain extended and correct descriptions, including interpretations. The primary objective is to provide up-to-date data for further processing by proper citation standards (the mentioned data outputs have legitimate authors and their use is to be cited similarly to a peer-reviewed publication). The current overview of citations shows that the professional, scientific community already works with the NHIS outputs above in this way, and hundreds of correct citations appear across published outputs.

Interactive tools on health data

As a result of the projects implemented and the cooperation with national and international partner institutions, including the active involvement of students, dozens of specific tools have been created that publish NHIS data outputs in various forms. These are briefly described below, including relevant links and references.

Czech childhood cancer information system

This web portal [55] provides information on childhood cancer epidemiology in the Czech Republic. It also includes an interactive tool for analytical reporting, which provides information on the following essential topics in graphs or tables: incidence, mortality and overall survival. The primary aim is to improve healthcare professionals' awareness of childhood cancers and promote education. The target groups involve paediatricians, non-specialised paediatric departments of healthcare facilities, professionals in basic and applied research and civic associations of patients and their family members. The interactive browser, the principal component of the portal, contains predefined analytical tools that make it possible for the user to look into epidemiological data from different points of view, both in graphical and tabular representation. From the user's viewpoint, this is how the interactive browser is used: (i) selection of the main module (incidence, mortality, survival); (ii) selection of analysis type; (iii) selection of the analysed group of patients, setting the analysis outputs [42].

Comprehensive cancer care network: Interactive viewer

The developed data viewer [56], which follows the outcomes of the CANCON Joint Action[48], allows the user to investigate general epidemiological trends for a particular diagnostic group. The user has an overview of all provided data views in the current module and can choose the one that interests them. The remaining part of the navigation page contains a selection of a particular diagnosis. The diagnoses are divided into twelve major cancer groups defined in the domain and data understanding

[48] https://cancercontrol.eu/

section. The viewer is divided into three modules, each containing a set of specific analytical reports visualised by interactive graphs and maps (Table 10) [43].

Table 10 Overview of available reports

Viewer module	Analytical report
Cancer epidemiology	Incidence and mortality trends over time Prevalence trend over time The age structure of incidence and mortality Age-specific incidence and mortality rates
Regional benchmarking	Incidence rates in regions Mortality rates in regions
Cancer diagnostics	Distribution of clinical stages Incidence trends by clinical stages Distribution of clinical stages by age Comparison of incidence of clinical stages by age

Reporting from the tuberculosis registry

Health registries at the national level allow for systematic and continuous control of the state of healthcare, the digitisation of certain services and the overall improvement of the state of care. As a result of multidisciplinary cooperation, which included three critical groups – experts in the field of tuberculosis and other mycobacteriosis, data analysts, and software developers – a freely available tool displaying interactive visualisations [57] of data from the nationwide Tuberculosis Registry (covering all cases diagnosed in the Czech Republic) has been developed. It has a great potential to benefit both professionals and the general public. The tool is based on modern technologies, making future long-term sustainability, scalability and extensibility much easier [58].

Analytical overview of the mental health care system

It is necessary to understand how the mental health care system works to systematically support the reform of mental health care services through information and communication technologies. The overarching platform, the Czech National Portal of Mental Health Care [59], presents mental health care primarily in data processing, analysis, and visualisation. All analytical outputs were prepared with the target group in mind so their supporting information was as practical, easy to understand and straightforward as possible. The combination of an interactive browser, infographics, open data and static reports took into account users with different IT knowledge and skills. The portal significantly helped to improve the mapping of the quality of care for people with mental health conditions based on valid materials for data-driven decision support [60].

Adverse events reporting system

Adverse events reporting is essential for increasing the quality of care for patients. The complex National Adverse Events Reporting System [61], in the form of an online portal, has been developed for data collection among all involved/participating healthcare providers, especially for methodological support. It includes static analytical reports, dynamic and interactive data visualisations and benchmarking among healthcare providers in the Czech Republic. Moreover, the methodological team ensures and coordinates strategic analyses to evaluate the quality of healthcare. Examples involve (i) analysis of data with the possibility of "benchmarking" between the different categories of healthcare providers – internally available data – internal data visualisation; (ii) analysis of personal and staff capacity indicators of the healthcare system of the Czech Republic; (iii) further cooperation on the analysis of NHIS data with a focus on the promotion and protection of public health with a particular emphasis on the evaluation of the quality of care; (iv) identification of errors in the so-called “quality of care”; and (v) the analysis of the quality of care [62].

Information service on COVID-19 epidemic data

On 11 March 2020, the first version of the official Czech web portal, named “COVID-19: The overview of the current situation in the Czech Republic” [63], was released. It provided a set of outputs in tables, graphs, and maps intended for the general public and the media. Its primary objective was to provide a well-arranged visualisation and clear explanation of basic information included in the overview of COVID-19 epidemiology in the Czech Republic [44]. All presented overviews were based on open data sets linked to records in the National Catalogue of Open Data[49]. Great emphasis was put on (i) the understanding and correct implementation of all six steps of the proven CRISP-DM methodology (business understanding, data understanding, data preparation, modelling, evaluation and deployment), including the infrastructure of a nationwide information system; (ii) the methodological setting of communication channels between all involved stakeholders; and (iii) data collection, processing, analysis, validation, and visualisation.

The Dashboard [64], a management tool for the Integrated Central Management Team, was developed to provide a real-time analytical overview and data sharing. This tool offered complex analytical reporting for the government of the Czech Republic and the supporting workgroups and predictive models for the mathematical determination of various epidemic scenarios. The application allowed the user to change the graph parameters and to show data based on visual filters (regions, districts, districts of Prague, age groups, etc.). Access to this application was restricted to government members, top management of the Ministry of Health of the Czech Republic and workgroups participating in crisis management, representatives of the Czech Armed Forces and other institutions involved. Finally, a simple epidemiological model [65] was based on a published dataset describing the current epidemiological situation in the Czech Republic and developed to help decision-makers understand the course of the epidemics (including the estimation of the adequate reproduction number) and to facilitate short-term predictions.

Finally, a robust web-based application, the Management Tool for Epidemic Monitoring, Analysis, and Visualisation [66] was designed to integrate different data sources, including real-time analysis, and provide transparent interactive reporting of current status over available data in a standardised format. Thanks to the experience of the teams involved, the understanding phases were relatively quick; all the more attention was paid to working with the selected models, which were tested in practice. The result was an interactive application with the possibility of custom data import and display, which was very positively assessed during the project evaluation. The R Shiny

[49] https://data.gov.cz/datov%C3%A9-sady

technology ensured the application's interactivity, which also made it possible to set many variables concerning the current or predicted state of the epidemic [67].

Real-time control centre for intensive care

This web-based application [68] enabled continuous monitoring and decision-making during the highest pressure on critical care. The Control Centre for Intensive Care has become an indispensable part of a set of online tools employed regularly for crisis management during the COVID-19 pandemic. This platform provided two tools of crucial importance to offer practical support for crisis management teams: (i) a module for regular online updates and overall monitoring of currently free capacities in hospitals (healthcare technology/medical devices, beds, staff) in real-time, and (ii) a module for online entering and overall record-keeping of requirements on medications for COVID-19 patients. No patient personal data or medical records were stored in this application; thus, no anonymisation or pseudonymisation algorithms had to be used. Data on free capacities were available and exported as an open data set in the National Catalogue of Open Data. The data set contained overviews of machine equipment (extracorporeal membrane oxygenation, mechanical ventilation, continuous renal replacement therapy, intermittent haemodialysis, ventilators—portable and those in operating theatres) and occupancy rate of inpatient beds (only beds in the emergency department and intensive care units for adults and standard beds with oxygen in the entire hospital of interest) [47].

Societal impact

From my perspective (i.e. as the author of this thesis), I consider the most significant contribution to be that a robust community has been built in both areas – medical education and health information – that is interested in continuously cultivating the environment with innovations. The successful implementation of dozens of institutional, national and international projects has brought new know-how, experience and partnerships that form the basis for strategic and conceptual changes. Close links between experts (physicians, representatives of professional societies, management of academic institutions and public sector entities, educators and students) and informatics and data-driven teams bring a long-needed degree of rationalisation to the proposed and newly implemented approaches, technologies and platforms. A detailed analysis of the input user and technical requirements of all interested target groups, the adoption of methodological procedures proven in practice and their subsequent adaptation to the real needs of the Czech environment seems to be crucial.

Several key factors should be mentioned in successfully building a community that works with data and effectively uses modern and data-related communication technologies (information systems, web presentations and portals, interactive data visualisation). This combination of factors helps create a strong and sustainable community that can use data appropriately as a supportive input for objective decision-making and achieving common goals.

- **Sharing knowledge and experience:** The community can share its knowledge and experience, whether through workshops, lectures, online resources or newsletters.
- **Technological background**: For a stable technical background in the development, operation and further improvement of information systems, educational tools and communication portals, it is necessary to secure the development team regarding personnel and finances. The team must be able to coordinate with the faculty representatives regarding leadership, requirements specification, design of methodologies and algorithms, and the coding itself.
- **Access to resources**: The community can access relevant datasets, data analysis tools and development platforms. This access allows for practical experience and innovation.
- **Networking and collaboration**: Fostering relationships and cooperation within the community increases its value and helps find mentors and collaborators on scientific research topics.
- **Diversity**: An important aspect is to include experts from different disciplines (computer science, natural sciences, medicine and health, education etc.). Diversity increases the creativity and innovation capacity of the community.
- **Leadership support**: Active involvement and visible support from key institutional leaders provide the community with the necessary resources, legitimate framework and motivation.
- **Regular activities and events**: regularly organised formal and informal events such as hackathons, workshops and conferences help keep community members active and engaged.
- **Openness and transparency**: Sharing information and opening up data, methodologies and concepts promotes collaboration, increases trust, and

allows community members to influence the outcome and learn from each other actively.

- **Supporting learning and development**: Providing resources for learning, professional and personal development helps those involved to improve their skills and keep up to date with current trends in the field.

The individual partial outputs presented in the sections above in the form of successfully implemented projects, the establishment of partnerships, the establishment of regular conferences, the involvement of students and the research and development publications have led to the emergence of two major communities that are very recognisable in their domain and thus have the opportunity to influence the future direction and critical decisions.

The MEFANET community

MEFANET (Medical Faculties NETwork) [3] is an initiative that unites medical and healthcare faculties primarily across the Czech Republic and Slovakia. It encourages the use of digital media in teaching and learning, supports the development of e-learning tools, and promotes best practices through workshops, conferences and joint projects. By doing so, MEFANET strives to enhance the quality of medical education and to prepare future healthcare professionals more effectively for their careers. This activity is certainly not meant to affect or control teaching activities at individual faculties: all objectives of the MEFANET project fully respect the independence of medical and health sciences faculties. Since 2007, the MEFANET has covered all Czech Republic and Slovakia medical faculties. It is driven by the MEFANET Coordinating Council, which comprises representatives of all participating members. In 2012, health sciences faculties joined the MEFANET educational network. From an IT perspective, implementing innovative approaches and technologies in the MEFANET network through centrally managed activities and developed tools has been successful in the long term. Uniformity at the data and application level provides numerous benefits for each network member and enables sustainable and systematic development. The network has thousands of active students, academics and faculty management representatives. Each year, one traditional MEFANET conference is organised under the auspices of the MEFANET Coordination Council, and highly specialised workshops and seminars on selected topics (e.g. OSCE, portal platform, virtual clinical scenarios etc.) are organised.

Health data community

The creation of a community around the topic of health data dates back to February 2020, the period associated with the turbulent spread of COVID-19. At the same time, the dangers and risks of misinterpreting the data (intentionally or unintentionally) became fully apparent. Since then, there has been an ever-expanding professional initiative based on the new communication channel of established contacts in the COVID-19 reporting. The national methodology [2] is a universal and comprehensive description of health data's design, preparation, validation and publication process. This document is being consulted and opposed by a range of experts and will also be presented to the broader health data community after final approval by MoH management. Ensuring clear and consistent communication of health data is an essential component of a functional system. The lay and professional public must know the current status, publication plan and overall vision. Communication tools that will be continuously used not only to communicate the current status but also to gather suggestions, requests and feedback include:

- national and regional conferences, webinars;
- discussion panels and educational seminars;
- individual consultations;
- National Health Information Portal as a primary source of guaranteed, valid, and verified information for the general public;
- data reporting tool for up-to-date information and news based on health data;
- local catalogue of open data in the health sector;
- social media communication;
- press conferences and official stakeholders' opinions;
- publications (papers, books, handouts).

Communication across experts at international level to share and transfer experiences is also an integral part of the efforts described above, especially in the specific domain of open data in healthcare. NZIP plays a crucial role in long-term sustainability and efficient use of resources.

- NZIP is a structured publication medium that includes sections for the general public and complementary content for NHIS specialist news. It is the portal on which all NHIS data and statistical outputs are published – publication on a single platform contributes significantly to the clarity and transparency of outputs.
- NZIP serves as a gateway for submitting requests for NHIS outputs, including authors' proposals for new datasets or summaries (or proposals for modifying already published outputs).
- NZIP links published data and statistical outputs to other key, data-driven agendas such as published clinical guideline practices, quality of care indicators, prevention programme evaluations, etc.
- The content of NZIP provides a basis for retrospective evaluation of the impact of already published outputs, whether by citation, traffic or download metrics or by tracking subsequent use and interpretation.
- Monitoring the traffic, topic searches and interests of professional and lay users of NZIP facilitates the prediction of needs for NHIS outputs and enables effective targeting of future outputs.

Summary

The factors above that are crucial for building a strong community are briefly described in Table 11 from the perspective of both domains of this thesis. The multidisciplinary and centralised coordination of the supporting activities has provided a strong background for the strategic development of community networks. It has also improved technical and methodological background to support data-driven decision-making over real-world issues.

Table 11 Key factors influencing community building in both domains of this thesis

Factor	Medical and healthcare education	Health information and statistics
Sharing knowledge and experience	The MEFANET annual conferences and other related conferences	The NHIS Open annual conferences and other related conferences
Technological background	The MEFANET e-publishing platforms, including the central gateway	National Health Information System, National Health Information Portal, National Catalogue of Open Data
Access to resources	Education content exports, user behaviour statistics (both on faculty level)	Complex data reporting service over the National Health Information System
Networking and collaboration	Strong national and international background in academic partnership	Established inter-ministerial cooperation (e.g. Ministry of Health, Ministry of Interior, Institute of Health Information and Statistics, Digital Information Agency etc.)
Diversity	Cooperation among teams from different faculties (teachers, students, faculty management, computer scientists, data analysts etc.)	Close connections between technologists, data analysts and representatives of the Czech Medical Association of J. E. Purkyne and other relevant stakeholders
Leadership support	Long-term support on the university management level from all involved academic subjects	Support on the relevant ministry level (regardless of the political balance of power)
Regular activities and events	Organisation of specialised workshops and discussion panels focusing on dedicated topics	Organisation of specialised workshops and discussion panels focusing on dedicated topics
Openness and transparency	Data sharing on the local level of individual faculty, methodology sharing among the MEFANET network	Open data first strategy and sharing outputs in various data formats
Supporting learning and development	Professional and personal development of individual team members	Professional and personal development of NZIP team members based on specialised trainings and workshops

Summary and main contribution

The following paragraphs summarise key outcomes that demonstrate practically the application of sound and proven practices in the methodological areas of informatics, data work, healthcare and pedagogy. A crucial aspect of the applicability and long-term sustainability of the results achieved was the consideration of the requirements of key stakeholders, the demands of the target groups and the technological and staffing capacity for the actual implementation and operation.

Medical and healthcare education

Collaborative framework for sharing medical and healthcare electronic educational resources

A community of all medical faculties provides a large target audience covering numerous academic staff (in the order of tens of thousands of potential users). Logically, there was a need for a transparent and unified source of teaching content. The primary purposes of the MEFANET educational portal are:

- to help teachers with publishing their educational works;
- to help authors of multimedia teaching and learning tools to communicate with the users of these tools;
- to help people involved in the lifelong learning of physicians and healthcare professionals;
- to help students of medicine and healthcare disciplines navigate various tools for electronic support of their education.

The educational web portal (a central gateway) [26] is the official publication platform for the MEFANET network, hosting electronic versions of academic resources and multimedia teaching materials. Its primary function is to provide access to all types of electronic instructional materials created throughout the MEFANET network. It includes eleven portal platforms connected by this gateway, facilitating efficient searches across all available content. Local portal instances complement, rather than replace, the established LMS systems at medical and health sciences faculties. Unlike traditional learning management systems, this educational portal offers additional e-publishing tools and established guidelines for quality assessment. It also supports federated authentication, eliminating the need for users to maintain separate accounts for each portal instance. Figure 18 documents the number of visits from April 2023 to March 2024 on the MED MUNI portal instance, which is one from 11 local e-publishing platforms covered by the MEFANET central grateway.

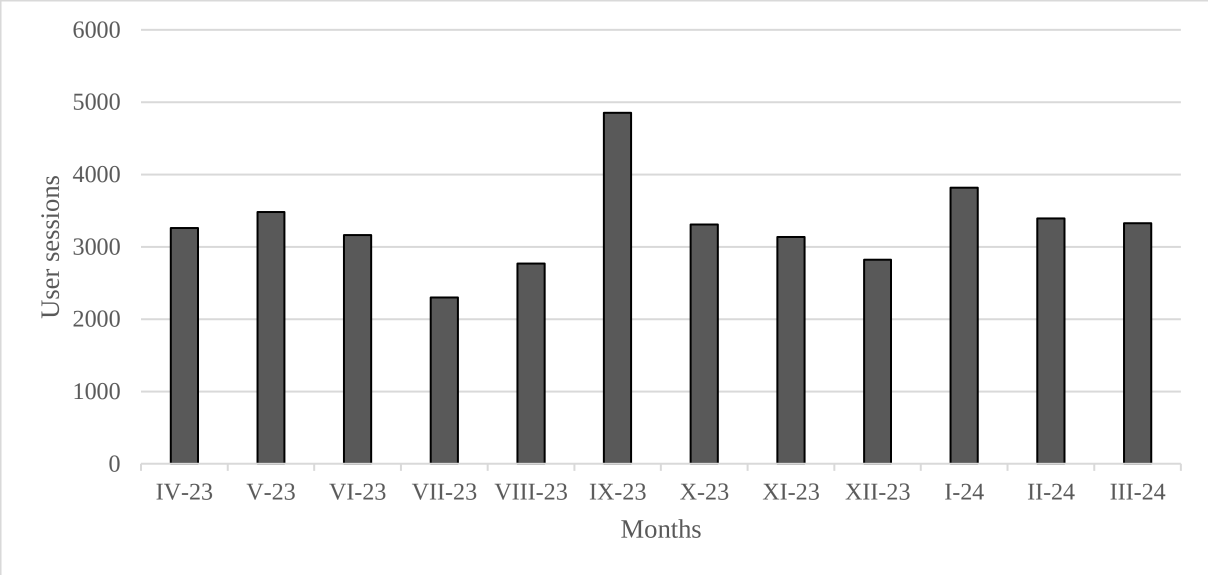


Figure 18 List of visits at https://portal.med.muni.cz/ from April 2023 to March 2024

Related publications

— Komenda, M., Schwarz, D., Feberová, J., Štípek, S., Mihál, V., Dušek, L. Medical faculties educational network: Multidimensional quality assessment. Computer Methods and Programs in Biomedicine 2012; 108(3): 900-909. Q3
— Schwarz, D., Komenda, M., Dušek, L. The system of e-publishing education materials and information discovery service in MEFANET. In Daniel Schwarz, Ladislav Dušek, editors. Computer Applications Systems and Networks for Medical Education: MEFANET: Czech and Slovak Medical Faculties Network. 1st ed. Brno: Masaryk University, Institute of Biostatistics and Analyses and Facta Medica, 2014: 95-109. ISBN 978-80-904731-9-5.

Integration platform supporting medical and healthcare education using innovative pedagogical methods and technologies

The SIMUportfolio platform [35] is an integrated online web application developed from the insights gained across more than ten previous international and national projects. It aims to introduce innovative and advanced features crucial for educational activities at the Faculty of Medicine of Masaryk University. Thanks to the strong support of the faculty management and the close collaboration with all stakeholders involved (including guarantors, teachers, technicians, students and the development team), the platform can continually adapt to shifting priorities and needs, both technologically and methodologically. SIMUportfolio has been evolving and maintained for nearly six years, during which it has become a valuable and indispensable tool not just for student assessment. The key benefits of using SIMUportfolio include:

— online accessibility: the platform is available 24/7 with user access control management;
— time efficiency: its automation allows educators to concentrate more on teaching and less on administrative duties;
— adaptability: it supports various assessment formats, catering to different courses;

- data-driven decision-making: data from SIMUportfolio provides critical insights into student performance.

Related publications

- Komenda M, Karolyi M, Woodham L, Vaitsis C. The role of technical standards in healthcare education. In: Konstantinidis ST, Bamidis PD, Zary N (eds). Digital innovations in Healthcare Education and Training. London: Academic Press; 2020. p. 47–59. ISBN 978-0-12-813144-2.
- Karolyi M, Ščavnický J, Růžičková P, Šnajdrová L, Komenda M. Design and Management of an Objective Structured Clinical Examination using the SIMU portfolio Platform. In: Lane HC, Zvacek S, Uhomoibhi J (eds). Proceedings of the 12th International Conference on Computer Supported Education (CSEDU) – Vol 1. Setúbal: SciTePress; 2020. p. 269–76.
- Komenda, M., Karolyi, M. SIMUportfolio: Complex all-in-one web-based platform enhancing medical education. In: Proceedings of the 2019 International Symposium on Educational Technology (ISET). Piscataway: Institute of Electrical and Electronics Engineers; 2019. p. 70–3. ISBN 978-1-7281-3389-8.

Implementing the proven methodology to support data-driven medical education and healthcare decision-making

Data-driven decision-making has proven to be effective in developing medical and healthcare curricula. Various projects and activities have been launched in recent years at the Faculty of Medicine of Masaryk University, focusing on the systematic utilisation of data as a reliable tool for making informed decisions. According to faculty management, course guarantors and teachers, employing data systematically facilitates the gradual introduction of new teaching methods and technologies. All case studies which have been successfully finalised and published utilise the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. They demonstrate the outcomes and insights gained from partner cooperation, which include domains like curriculum mapping, virtual clinical scenarios, student assessment etc. The CRISP-DM model is applied to cover all critical stages of data handling (Business understanding, Data understanding, Data preparation, Modelling, Evaluation, Deployment), emphasising the often overlooked importance of domain understanding, which suggests further research.

Related publications

- Komenda M, Víta M, Vaitsis C, et al. Curriculum mapping with academic analytics in medical and healthcare education. PLoS One. 2015;10(12):e0143748.
- Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. 390 p. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.
- Komenda, M., Ščavnický, J., Pokorná, A. Mapování dekubitů a nehojících se ran v medicínském kurikulu. Česká a slovenská neurologie a neurochirurgie. 2019;82(S1):S33-S36.

Health information and statistics

National methodology for health data sharing in the Czech Republic

Without a strategic document that methodically describes the different areas of health data sharing, it would be impossible to build technical tools and communication channels for the target groups. Thanks to the newly developed concept for NHIS data sharing [2], in which stakeholders from the public sector and academia have been involved as reviewers, it is feasible to establish a preliminary roadmap for deliverables and gradually open up health data in compliance with the current legislation. The concept works with different forms of data publication, making a fundamental distinction between data sources that are legally public and can be released in a fully open format without modification, and data sources in primary (legally non-public) formats, which need to be modified before publication or can be shared remotely only to legally designated institutions. At the same time, the decision-making and control processes to which any release of NHIS data must be subject are defined.

Related publications

- Komenda M, Klimeš D, Zvolský M, et al. A Concept of Sharing and Secondary Mining of Data from the National Health Information System [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2023 [cited 16 Feb 2024]. Available from: https://share.uzis.cz/s/zTBm7EQMSW3GoGq.
- Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. 390 p. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.

Technological background for health literacy and data-driven decision-making improvement in Czech healthcare

For a unified technological background and online accessible platform with the ambition of long-term development, it was necessary to implement a centrally controlled environment under the auspices of the Ministry of Health. This is what the National Health Information Portal (NZIP) became in 2020. It is an essential online source of valid and guaranteed information and recommended tools to improve the population's health literacy. At the same time, it provides access to a wide range of information on healthcare, disease prevention, healthy lifestyles and current measures adopted or recommended by health authorities. NZIP's main contribution is to improve citizens' awareness of health issues, which helps to improve understanding and more effective use of health services. As a result, individuals can make better decisions about their health and actively contribute to the prevention and management of potential health problems.

NZIP also promotes transparency and accessibility of health data using a dedicated data reporting module, which is essential for strengthening public trust in the health system. Structured metadata descriptions clarify the need for data outputs, define all attributes, correct interpretation and information limits, and provide use examples. Importantly, this is intended to be a single place on the Internet where users can find information about different areas of the Czech healthcare system. Data outputs published in this way, e.g. in the form of open data or data aggregates, are the basis for advanced analyses to provide factual and highly valid material for subsequent decision-making on the strategic direction of the Czech healthcare system (e.g. a study which followed all Czech patients with implanted pacemakers [69]). An overview of traffic to the NZIP platform over the last year documents its use and further potential for improvement (Figure 19).

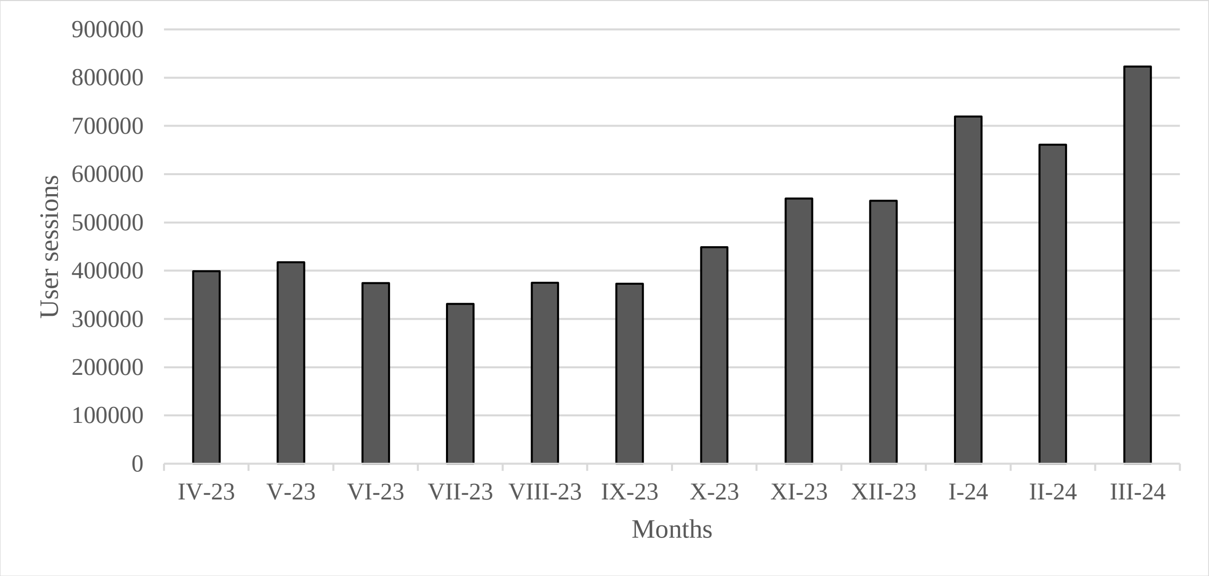


Figure 19 List of visits at https://www.nzip.cz/ from April 2023 to March 2024

Related publications

— National Health Information Portal [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2020 [cited 23 Apr 2024]. Available from: https://www.nzip.cz. ISSN 2695-0340.
— Komenda M, Klimeš D, Zvolský M, et al. A Concept of Sharing and Secondary Mining of Data from the National Health Information System [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2023 [cited 16 Feb 2024]. Available from: https://share.uzis.cz/s/zTBm7EQMSW3GoGq.

Real-time framework for effective management and data-driven support of intensive inpatient care, including complex information service during the pandemic

The situation during the COVID-19 epidemic in the Czech Republic has shown the urgent need to map the occupancy of inpatient care and, at the same time, the importance of near real-time data availability. A comprehensive methodology supported by a technical solution to define, monitor and evaluate hospital capacity was developed in collaboration with leading experts. The "Control Centre for Intensive Care", as a web-based platform, has provided systematic online 24/7 support for managing intensive inpatient care throughout the Czech Republic. It includes two essential tools as part of the national methodology: one module for regular online updates and real-time monitoring of available intensive care capacities and another for entering and maintaining records of medication requirements for COVID-19 patients. The whole solution has become part of the legislation and has been adopted by the relevant departments of inpatient care providers. Thanks to its general design and wide configuration options, the Czech system is ready to use this platform and monitor the necessary attributes in real time for any other unexpected situations of similar cases.

Epidemiological data were an essential part of the crisis management in the Czech Republic during the epidemic. Within a few weeks, a technological solution was developed based on the open data first principle and complemented with interactive visualisations. A comprehensive methodology covering data collection, processing, publication and correct interpretation for further sharing and communication with the general public was an integral part of the project. The experience of the COVID-19 epidemic provided the basis for a global approach to health data sharing, taking into account the diverse requirements of the actors who subsequently work with the data.

- Komenda M, Panoška P, Bulhart V, et al. COVID-19: Overview of current situation in the Czech Republic. Disease at the Moment [Internet]. Ministry of Health of the Czech Republic; 2020 [cited 23 Apr 2024]. Available from: https://onemocneni-aktualne.mzcr.cz/covid-19. ISSN 2694-9423.
- COVID-19: Online Intensive Care Dispatching [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2020 [cited 2 Apr 2024]. Available from: https://dip.mzcr.cz/.
- Komenda M., Černý V., Šnajdárek P., et al. Control Centre for Intensive Care as a tool for effective coordination, real-time monitoring, and strategic planning during the COVID-19 pandemic. J Med Internet Res 2022;24(2):e33149.
- Komenda M, Bulhart V, Karolyi M, et al. Complex reporting of coronavirus disease (COVID-19) epidemic in the Czech Republic: use of interactive web-based application in practice. J Med Internet Res. 2020;22(5):e19367.
- Komenda M., Jarkovský J., Klimeš D., et al. Sharing datasets of the COVID-19 epidemic in the Czech Republic. PLOS ONE 2022;17(4):e0267397.

Conclusion

From a methodological and technical perspective, the habilitation thesis describes building long-term functional and operational concepts supported by strong communities of interested users. In practice, the importance of systems enabling the collection and presentation of available data is demonstrated. The actual importance and societal impact are documented using the solutions described by the professional and lay public. The thesis summarises the years 2008 to 2024, when the author was involved in a wide range of research and development, education, coordination and management activities in the academic and public sectors. The combination of multiple specialties, including clinical and theoretical medical disciplines, with informatics and analytical experts, allowed me not only to design but, above all, to realistically implement the required solutions that met the specifications and expectations of the competitive tender. Thanks to the appropriately designed architecture, it is possible to work efficiently with the available data and prepare valid documents for strategic and conceptual decision-making in medical education and Czech healthcare. Thus, it distinguishes between subjective and objective perceptions, feelings and impressions, and conclusions directly derived from data.

The story in the form of individual outcomes and concrete results that fit together only documents the need and necessity of each. The benefit of newly gained experience, learning and adopting new approaches or technologies, and knowing that a given path is not the right one, is unquestionable. The collaboration of several often isolated teams has been crucial to achieving the published results, to whom many thanks are due. Without their proactive approach, the implementation of the intended solutions – sometimes in record time – would be virtually impossible. The societal impact and added value of presented results are documented by user statistics of individual online platforms and published papers in established scientific multidisciplinary journals and by positive feedback and evaluation by experts from practice. Thanks to the established communities, it is possible to continuously discuss solutions, gain new knowledge and inspiration and further cultivate the whole system.

Thanks to the functionality of the proposed solutions, the long-term support of stakeholders and the enshrinement in legislation, it will be possible to continue to improve the quality of medical education and the Czech healthcare system through the effective use of modern methodologies and technologies, but in accordance with the views of representatives of medical societies.

References

[1] Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. 390 p. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.

[2] Komenda M, Klimeš D, Zvolský M, et al. A Concept of Sharing and Secondary Mining of Data from the National Health Information System [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2023 [cited 16 Feb 2024]. Available from: https://share.uzis.cz/s/zTBm7EQMSW3GoGq.

[3] Medical Faculties Network [Internet]. Masaryk University; 2007. [cited 1 Mar 2024]. Available from: https://www.mefanet.cz/index-en.php.

[4] Komenda M, Karolyi M, Woodham L, Vaitsis C. The role of technical standards in healthcare education. In: Konstantinidis ST, Bamidis PD, Zary N (eds). Digital innovations in Healthcare Education and Training. London: Academic Press; 2020. p. 47–59. ISBN 978-0-12-813144-2.

[5] Karolyi M, Komenda M, Woodham L, Ščavnický J, Vaitsis C, Spachos D. Medical curriculum standards: Towards relational database transformation. Health Technol. 2020;10(3):759–66.

[6] Komenda M, Karolyi M, Vaitsis C, Spachos D, Woodham L. A Pilot Medical Curriculum Analysis and Visualization According to Medbiquitous Standards. In: Konstantinidis ST, Bamidis PD, Rodrigues PP (eds). IEEE 30th International Symposium on Computer-Based Medical Systems. Los Alamitos: IEEE Computer Society; 2017. p. 144–9. ISBN 978-1-5386-1710-6.

[7] Komenda M, Schwarz D, Feberová J, Štípek S, Mihál V, Dušek L. Medical faculties educational network: Multidimensional quality assessment. Comput Methods Programs Biomed. 2012;108(3):900–9.

[8] Komenda M, Schwarz D, Švancara J, Vaitsis C, Zary N, Dušek L. Practical use of medical terminology in curriculum mapping. Comput Biol Med. 2015;63:74–82.

[9] Štuka Č, Martinková P, Vejražka M, Trnka J, Komenda M. Testing in Medical Education. Test Construction and Analysis in Medical Faculties. Prague: Karolinum; 2013. 156 p. ISBN 978-80-246-2369-6.

[10] Karolyi M, Ščavnický J, Růžičková P, Šnajdrová L, Komenda M. Design and Management of an Objective Structured Clinical Examination using the SIMU portfolio Platform. In: Lane HC, Zvacek S, Uhomoibhi J (eds). Proceedings of the 12th International Conference on Computer Supported Education (CSEDU) – Vol 1. Setúbal: SciTePress; 2020. p. 269–76.

[11] Komenda M, Schwarz D, Dušek L. Towards a System of Enhanced Transparency of Medical Curriculum. Eur J Biomed Inform. 2013;9(3):9–16.

[12] Kononowicz AA, Balcerzak Ł, Kocurek A, et al. Technical infrastructure for curriculum mapping in medical education: a narrative review. Bio-Algorithms Med-Syst. 2020;16(2):20200026.

[13] Komenda M, Víta M, Vaitsis C, et al. Curriculum mapping with academic analytics in medical and healthcare education. PLoS One. 2015;10(12):e0143748.

[14] Komenda, M., Ščavnický, J., Pokorná, A. Mapping of pressure ulcers and non-healing wounds in medical curriculum. Česká a slovenská neurologie a neurochirurgie. 2019;82(S1):S33–6.
[15] Karolyi M, Komenda M, Janoušová R, Víta M, Schwarz D. Finding overlapping terms in medical and health care curriculum using text mining methods: rehabilitation representation – a proof of concept. Mefanet J. 2016;4(2):71–7.
[16] Karolyi M, Ščavnický J, Buďa J, Jurková T, Mazalová M, Komenda M. Mapping of Dental Care in the Czech Republic: Case Study of Graduates Distribution in Practice. In: Ganzha M, Maciaszek L, Paprzycki M (eds). Proceedings of the 2019 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2019. p. 599-603. ISBN 978-83-955416-0-5.
[17] Karolyi M, Ščavnický J, Komenda M. First Step Towards Enhancement of Searching Within Medical Curriculum in Czech Language using Morphological Analysis. In: McLaren BM, Reilly R, Zvacek S, Uhomoibhi J (eds). Proceedings of the 10th International Conference on Computer Supported Education. SciTePress; 2018. p. 288–93.
[18] Ščavnický J, Karolyi M, Růžičková P, et al. Pitfalls in users' evaluation of algorithms for text-based similarity detection in medical education. In: Ganzha M, Maciaszek L, Paprzycki M (eds). Proceedings of the 2018 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2018. p. 109–16. ISBN 978-83-949419-7-0
[19] Komenda M, Ščavnický J, Růžičková P, Karolyi M, Štourač P, Schwarz D. Similarity Detection Between Virtual Patients and Medical Curriculum Using R. Stud Health Technol Inform. 2018;255:222–6.
[20] Komenda M, Karolyi M, Vyškovský R, et al. Towards a keyword extraction in medical and healthcare education. In: Ganzha M, Maciaszek L, Paprzycki M (eds). Proceedings of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2017. p. 173–6. ISBN 978-83-946253-9-9.
[21] Komenda M, Víta M, Karolyi M, Kríž V, Pokorná A. Automatic Keyword Extraction from Medical and Healthcare Curriculum. Proceedings of the 2016 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2016. p. 287–90. ISBN 978-83-60810-90-3.
[22] Víta M, Komenda M, Pokorná A. Exploring medical curricula using social network analysis methods. In: Ganzha M, Maciaszek L, Paprzycki M (eds). Proceedings of the 2015 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2015. p. 297–302. ISBN 978-83-60810-66-8.
[23] Komenda M, Ramler R, Růžičková P, Hohensigner D, Kerberová M, Pichler M. Visualisation of text-based data describing best practices in software development. In: Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. p. 373–90. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.

[24] SESSAD Guidelines [Internet]. Masaryk University; 2023. [cited 3 Mar 2024]. Available from: https://www.med.muni.cz/sessad/guidelines.
[25] Schwarz D, Komenda M, Šnábl I, Dušek L. Multimedia support in the education of clinical and health care disciplines at Faculty of Medicine, Masaryk University [Internet]. Masaryk University; 2021 [cited 13 Mar 2024]. Available from WWW: http://portal.med.muni.cz. ISSN 1801-6103.
[26] Schwarz D, Komenda M, Šnábl I, Dušek L. A uniform platform for electronic publishing and sharing educational content in the network of medical faculties MEFANET [Internet]. Masaryk University; 2019 [cited 13 Mar 2024]. Available from WWW: http://portal.mefanet.cz.
[27] Majerník J, Schwarz D, Komenda M, Dušek, L. A Unified Educational Platform of Multimedia Support in Education at Medical Faculties of MEFANET Project. Procedia Social and Behavioral Sciences. 2012;55:476-84.
[28] The educational portal AKUTNĚ.CZ [Internet]. Masaryk University; 2015 [cited 13 Apr 2024]. Available from: https://www.akutne.cz/en/. ISSN 1803-179X.
[29] Schwarz D, Štourač P, Komenda M, et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. J Med Internet Res. 2013;15(7):e135.
[30] Majerník J, Kačmariková A, Komenda M, et al. Development and implementation of an online platform for curriculum mapping in medical education. Bio-Algorithms Med-Syst. 2022;18(1): 1–11.
[31] Harden RM, Crosby JR, Davis MH. AMEE Guide No. 14: Outcome-based education: Part 1. An introduction to outcome-based education. Med Teach. 1999;21:7–14.
[32] Komenda M, Schwarz D, Vaitsis C, Zary N, Štěrba J, Dušek L. OPTIMED Platform: Curriculum Harmonisation System for Medical and Healthcare Education. In: Cornet R, Stoicu-Tivadar L, Hörbst A, Parra Calderón CL, Andersen SK, Hercigonja-Szekeres M (eds). Digital Healthcare Empowering Europeans: Proceedings of MIE2015. Amsterdam: IOS Press; 2015. p. 511–5. ISBN 978-1-61499-511-1.
[33] Komenda M, Schwarz D, Hřebíček J, Holčík J, Dušek L. A Framework for Curriculum Management – The Use of Outcome-based Approach in Practice. In: Zvacek S, Restivo MT, Uhomoibhi J, Helfert M (eds). Proceedings of the 6th International Conference on Computer Supported Education (CSEDU) – Volume 2. Setúbal: SciTePress; 2014. p. 473–8. ISBN 978-989-758-020-8.
[34] Karolyi M, Ščavnický J, Bulhart V, Růžičková P, Komenda M. EDUportfolio: Complex Platform for Curriculum Management and Mapping. In: Lane H, Zvacek S, Uhomoibhi J (eds). Proceedings of the 11th International Conference on Computer Supported Education (CSEDU) – Volume 2. Setúbal: SciTePress; 2019: 352–8. ISBN 978-989-758-367-4.
[35] SIMUportfolio – a platform for enhancing the use of advanced elements of simulation medicine in teaching [Internet]. Masaryk University; 2021 [cited 22 Feb 2024]. Available from: https://portfolio.med.muni.cz/.
[36] Komenda, M., Karolyi, M. SIMUportfolio: Complex all-in-one web-based platform enhancing medical education. In: Proceedings of the 2019 International Symposium on Educational Technology (ISET). Piscataway: Institute of Electrical and Electronics Engineers; 2019. p. 70–3. ISBN 978-1-7281-3389-8.

[37] Sasín J, Komenda M. ESPO – a web application that supports an effective search of project opportunities and collaboration [Internet]. Masaryk University; 2023 [cited 18 Feb 2024]. Available from: https://espo.iba.muni.cz/.
[38] Sasín J. Web application supporting effective search of project opportunities and collaboration [M.Sc. Thesis]. Brno: Masaryk University, Faculty of Informatics; 2023. Available from: https://is.muni.cz/th/sruye/.
[39] Kolár A., KomendaM. OrtoDB: Application for structured description of orthopaedic image documentation [Internet]. Masaryk University; 2023 [cited 18 Feb 2024]. Available from https://ortodb.med.muni.cz/.
[40] Komenda M, Jarkovský J, Klimeš D, et al. Sharing datasets of the COVID-19 epidemic in the Czech Republic. PLoS One. 2022;17(4):e0267397.
[41] Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. Cancer. 2005;103(7):1457–67.
[42] Krejčí D, Karolyi M, Pehalová L, et al. Development of the Czech Childhood Cancer Information System: Data Analysis and Interactive Visualization. JMIR Public Health Surveill. 2021;7(6):e23990.
[43] Dušek L, Mužík J, Karolyi M, Šalko M, Krejčí D, Komenda M. A pilot interactive data viewer for cancer screening. In: Hřebíček J, Denzer R, Schimak G, Pitner T (eds). Environmental Software Systems: Computer Science for Environmental Protection. IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing AG; 2017. p. 173–83. ISBN 978-3-319-89934-3.
[44] Komenda M, Bulhart V, Karolyi M, et al. Complex reporting of coronavirus disease (COVID-19) epidemic in the Czech Republic: use of interactive web-based application in practice. J Med Internet Res. 2020;22(5):e19367.
[45] Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res. 2020;22(6):e21820.
[46] Moravec V, Dušek L, Gregor J, et al. COVID-19 Infodemic. Prague: Academia Publishing; 2022. 437 p. ISBN 978-80-200-3434-2.
[47] Komenda M, Černý V, Šnajdárek P, et al. Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic. J Med Internet Res. 2022;24(2):e33149.
[48] National Health Information Portal [Internet].Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2020 [cited 23 Apr 2024]. Available from: https://www.nzip.cz. ISSN 2695-0340.
[49] Komenda M, Svačina Š, Seifert B, et al. National Health Information Portal: Text-based open data. In: Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. p. 309–20. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.
[50] Dušek L, Komenda M, Těšitelová V. Central and clinical registries, importance of data interpretation for clinical medicine. In: Táborský M (ed). Digital Medicine 2022. Prague: EEZY Publishing; 2022. p. 176–90. Medicína. ISBN 978-80-908638-8-0.

[51] Dušek L, Krejčí D, Jarkovský J, et al. Overview of available population-based data in the Czech Republic with a focus on elderly. In: Gürlich R, Teplan V, Adámková V (eds). Surgery in Elderly and Their Internal Preparation for an Operation. Prague: Maxdorf; 2022: 252–82. ISBN 978-80-7345-745-7.
[52] Jarkovský J, Komenda M, Vráblík M, et al. National Cardiology Information System. National Health Information Portal [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2023. Available from: https://www.nzip.cz/nkis.
[53] Chlíbek R, Cabrnochová H, Prymula R, et al. Central Vaccination Records in the Czech Republic. National Health Information Portal [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2023. Available from: https://www.nzip.cz/ceo.
[54] Jírová J, Komenda M, Pařízek A, et al. National Registry of Reproductive Health. National Health Information Portal [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2023. Available from: https://www.nzip.cz/nrrz.
[55] Krejčí D, Ščavnický J, Zapletalová M, et al. Czech Childhood Cancer Information System [Internet]. Institute of Health Information and Statistics of the Czech Republic and Masaryk University; 2018 [cited 16 Apr 2024]. Available from: https://ccc-is.uzis.cz.
[56] Comprehensive Cancer Care Network – a pilot model [Internet]. Masaryk University; 2017 [cited 23 Apr 2024]. Available from: https://cccn-viz.onconet.cz/.
[57] Nováková T, Wallenfels J, Jarkovský J, et al. Interactive data viewer of the Registry of Tuberculosis [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2021 [cited 17 Apr 2024]. Available from: https://tbc.uzis.cz.
[58] Karolyi M, Nováková T, Panoška P, Wallenfels J, Komenda M, Jarkovský J. Interactive Online Reporting of Registry of Tuberculosis. In: Ganzha M, Maciaszek L, Paprzycki M (eds). Proceedings of the 2020 Federated Conference on Computer Science and Information Systems (FedCSIS). Warsaw: Polish Information Processing Society; 2020. p. 599–603. ISBN 978-83-955416-0-5.
[59] Soukupová J, Melicharová H, Šanca O, Bartůněk V, Jarkovský J, Komenda M. National Portal of Mental Health Care [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2022 [cited 17 Apr 2024]. Available from: https://psychiatrie.uzis.cz/.
[60] Komenda M, Soukupová J, Melicharová H, et al. Data and analytical basis for a modern mental health care system in the Czech Republic. In: Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. p. 231–46. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.
[61] Pokorná A, Štrombachová V, Mužík J, et al. National Adverse Events Reporting System [Internet]. Institute of Health Information and Statistics of the Czech Republic; 2016 [cited 12 Apr 2024]. Available from: https://shnu.uzis.cz.

[62] Pokorná A, Dolanová D, Štrombachová V, et al. Adverse Events Reporting System: Development of the platform for healthcare providers and the lay public. In: Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. p. 287–306. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.

[63] Komenda M, Panoška P, Bulhart V, et al. COVID-19: Overview of current situation in the Czech Republic. Disease at the Moment [Internet]. Ministry of Health of the Czech Republic; 2020 [cited 23 Apr 2024]. Available from: https://onemocneni-aktualne.mzcr.cz/covid-19. ISSN 2694-9423.

[64] COVID-19 Dashboard [Internet]. Ministry of Health of the Czech Republic; 2020 [Cited 14 Jan 2024]. Available from: https://dashboard.uzis.cz.

[65] Májek O, Ngo O, Jarkovský J, et al. Modelling the first wave of the COVID-19 epidemic in the Czech Republic and the role of government interventions [Internet]. medRxiv; 2020 [Cited 23 Jan 2024]. Available from: https://doi.org/10.1101/2020.09.10.20192070.

[66] Pavlík T, Uher M, Komenda M, et al. MAMES – Monitoring, Analysis and Management of Epidemic Situations [Internet]. Masaryk University; 2021 [Cited 23 Jan 2024]. Available from: https://webstudio.shinyapps.io/mames.

[67] Pavlík T, Přibylová L, Eclerová V, et al. Management tool for epidemic monitoring, analysis, and visualisation. In: Komenda M, et al. Data-Driven Decision-Making in Medical Education and Healthcare. Brno: Masaryk University and Institute of Health Information and Statistics of the Czech Republic; 2023. p. 355–71. ISBN 978-80-280-0392-0 and 978-80-7472-198-4.

[68] COVID-19: Online Intensive Care Dispatching [Internet]. Ministry of Health of the Czech Republic and Institute of Health Information and Statistics of the Czech Republic; 2020 [cited 2 Apr 2024]. Available from: https://dip.mzcr.cz/.

[69] Táborský, M., Skála, T., Dušek, L., et al. Clinical characteristics and mortality in all Czech patients after pacemaker implantation in the last decade. Frontiers in cardiovascular medicine, 2023. Available from: https://doi.org/10.3389/fcvm.2023.1248145.

Attachments

List of the author's most significant works related to this thesis.

[1] Komenda et al. Medical faculties educational network: Multidimensional quality assessment. Computer Methods and Programs in Biomedicine 2012; 108(3): 900-909.

[2] Schwarz et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. Journal of Medical Internet Research 2013; 15(7): 298-311.

[3] Komenda et al. Curriculum Mapping with Academic Analytics in Medical and Healthcare Education. PLoS One 2015; 10(12): e0143748.

[4] Komenda et al. Practical use of medical terminology in curriculum mapping. Computers in Biology and Medicine 2015; 63: 74-82.

[5] Komenda et al. Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice. Journal of Medical Internet Research 2020; 22(5): e19367

[6] Krejčí et al. Development of the Czech Childhood Cancer Information System: Data Analysis and Interactive Visualization. JMIR Public Health and Surveillance 2021; 7(6): e23990

[7] Komenda et al. Sharing datasets of the COVID-19 epidemic in the Czech Republic. PLoS One 2022; 17(4): e0267397

[8] Komenda et al. Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic. Journal of Medical Internet Research 2022; 24(2): e33149

computer methods and programs in biomedicine 108 (2012) 900–909

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Medical faculties educational network: Multidimensional quality assessment

Martin Komenda[a,*], Daniel Schwarz[a], Jitka Feberová[b], Stanislav Štípek[c], Vladimír Mihál[d], Ladislav Dušek[a]

[a] Masaryk University, Institute of Biostatistics and Analyses, Czech Republic
[b] Charles University in Prague, Second Faculty of Medicine, Czech Republic
[c] Charles University in Prague, First Faculty of Medicine, Czech Republic
[d] Palacky University in Olomouc, Faculty of Medicine and Dentistry, Czech Republic

ARTICLE INFO

Article history:
Received 10 May 2011
Received in revised form
23 January 2012
Accepted 1 May 2012

Keywords:
Educational portal
E-learning
Computer communication network
E-publishing
Quality assessment tools
Medical education

ABSTRACT

Today, World Wide Web technology provides many opportunities in the disclosure of electronic learning and teaching content. The MEFANET project (MEdical FAculties NETwork) has initiated international, effective and open cooperation among all Czech and Slovak medical faculties in the medical education fields. This paper introduces the original MEFANET educational web portal platform. Its main aim is to present the unique collaborative environment, which combines the sharing of electronic educational resources with the use tools for their quality evaluation. It is in fact a complex e-publishing system, which consists of ten standalone portal instances and one central gateway. The fundamental principles of the developed system and used technologies are reported here, as well as procedures of a new multidimensional quality assessment.

1. Introduction

There are a lot of experts who see the medical education field in different ways and with various attitudes. It is widely known that the progress of modern information technologies is also an important part of medical and health care education.

Things have progressed dramatically in the past decade. Information and telecommunications technology offers an opportunity to revolutionize the way we provide education [1]. Technological developments, in particular the rapid growth of the Internet and e-learning, are altering the nature of medical educational environment, and the possibilities offered are proving impossible to resist. Internet continuing education is gaining popularity, and most participants are satisfied with the experience and find it to be an effective learning format [2].

In medicine as in other fields, there has been a rapid expansion of e-learning. Almost every student in a medical school spends a part of their day or their week online searching using Google or some other search engine for information on a topic, communicating with a colleague or teacher, or studying a unit, module or course developed in their institution or elsewhere. E-learning has grown in popularity because of its convenience and flexibility and because of the increasing availability of computers and students' familiarity with them. A legitimate aspiration of e-learning is to make existing approaches to teaching and learning more effective and efficient. If this is to

* Corresponding author. Tel.: +420 549494469.
E-mail address: komenda@iba.muni.cz (M. Komenda).
0169-2607/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cmpb.2012.05.002

be achieved, e-learning implementation has to be monitored and managed by staff with the necessary education, technology and content expertise. There is no doubt, that if delivered appropriately, e-learning can be effective and can enrich the learning experience for the student while at the same time allowing the teacher to take on more productive and rewarding roles [3].

Learning is facilitated by a curriculum map, guided-learning resources, "ask-the-expert" opportunities, and collaborative or peer-to-peer learning. The educational philosophy is "just-for-you" learning (learning customized to the content, educational strategy, and distribution needs of the individual physician) and "just-in-time" learning (learning resources available to physicians when they are required). Implications of the new learning technologies are profound. E-learning provides a bridge between the cutting edge of education and training and outdated procedures embedded in institutions and professional organizations. There are important implications, too, for globalization in medical education, for interdisciplinary education, and for the continuous education from undergraduate to postgraduate and further to continuing education [4].

There have been many educational web portals designed in recent years. However, they usually, with a few exceptions, represent either storage of diverse study/teaching materials or huge data warehouses without tools for organizing and guaranteeing learning objects. Such digital libraries or large reusable learning objects [5] repositories like Jorum [6] and Wisconsin Online Resource Center [7] provide learning content for educators mostly for free without any registration or fees. On the other side, medical educational repositories appear that use various review processes and publish mainly peer-reviewed learning content, such as MedEdPortal [8], MERLOT [9], HEAL [10].

The need and importance for a well-designed collaborative environment as a source of information for students and academic staff has been recently growing. We explored how to most effectively use available technologies to advance the training of medical students, such that they can make the best possible use of published teaching materials via the Internet. The MEFANET project (MEdical FAculties NETwork) represents the cooperation between eight Czech and three Slovak medical faculties. It was decided to develop an original educational web portal platform for publishing and sharing teaching content among them. The objective of this paper is to describe our development experience and to present in detail a unique set of procedures for multidimensional quality assessment.

The development was driven by three major motivating factors, which in combination represent a demand for a novel approach to the effective on-line delivery of educational content. All of them are consequences of rapid ICT expansion, innovative teaching progress and the requirements of today's students. The first factor follows the web accessibility rules, which are defined by an international community called The World Web Consortium (W3C). It's an organization, where members, a full-time staff and the public work together to develop Web standards, technical specifications, guidelines and strategies [11]. The second motivation is to ensure high-level service quality of the platform for its users, including searching facilities, user support, communication elements, personalization and powerful administration background. The third motivating factor is the need for a quality assessment of published contents, in order to distinguish the MEFANET portal platform from usual anonymous web educational environments, which do not always provide updated and correct information. In our opinion, all the published information should be relevant, up-to-date and accurate for the needs of the audience.

2. Digital teaching contents

2.1. *Progress of e-learning*

The Internet is a very large scale hypertext information space where different types of users can search and find information in various domains. There is now access to a wealth of freely accessible online educational materials in the educational/instructional domain. Tutors can use these resources in their teaching, and students can use them to support their learning. A student is likely to first browse the web looking for relevant open access resources (tutorials, portals, example code, courses, assessments, experts, institutions, etc.). This scenario is true across all subject domains; the Internet has become an important tool within education [12]. Using the Internet is the norm for today's university students and a lot of reasons exist justifying the use of this phenomenon as a source for learning. According to Geueke and Stausberg [13], the Internet promises many advantages: Access from all over the world with a low-cost technical infrastructure; independence from proprietary solutions in hardware and software by means of a common browser as front-end; independence from time restrictions because the material is available 24 h a day, 7 days a week without any technical necessity for downtime.

2.2. *Quality of digital on-line content*

Since one of the roles of educators is to assess and select learning resources based on curriculum goals and student needs, the development of standardized methods for educational content evaluation becomes vital. To the learner, it is important for reviews of the quality of the resource to be readily available. Unfortunately, systematic evaluation of computer-based education in all its various forms (including integrated learning systems, interactive multimedia, interactive learning environments, and microworlds) often lags behind development efforts [14].

Guaranteeing the correctness of the published content seems to be a long-term problem and is often related to system design, development and offered services. Various types of indicators, which should generally organize the quality of web portals and published information, were already described. Experts have been presented with many effective mechanisms to ensure and guarantee published content for many years.

Yang et al. [15] have published an instrument which measures user-perceived service quality of web portals. They proposed and validated a five-dimension service quality toolset involving: usability, usefulness of content, adequacy of information, accessibility, and interaction. These tools provide a useful scale for researchers, who wish to measure the service quality of web portals.

Bottentuit and Coutinho [16] have selected a set of indicators that, in their point of view, should be necessarily integrated on a portal dedicated to educational issues:ease of use, services, communication, content, performance and information. It is necessary to take these issues of quality into account in the construction and/or management of the educational portal, if one wants a successful portal with a larger traffic of site users and promoters.

Caro et al. [17] have presented a portal data quality model (PDQM), focused on the data consumer's perspective. To produce the PDQM model, they defined a four-phase process: 1. Identification of web data quality attributes and definition of a Classification Matrix; 2. Classification of data quality attributes into the Matrix; 3. Validation; 4. PDQM.

During the first phase, they recompiled 41 web data quality attributes, which they believe should therefore be applicable to web portals. Caro et al. [18] also did an extensive web data quality revision, and identified relevant quality attributes. From this revision they captured several data quality attributes. The most considered are Accuracy, Completeness, Timeliness, Conciseness, Consistency, Currentness, Interpretability, Relevance and Security.

Elissavet and Economides [19] have described a very sophisticated evaluation framework for hypermedia courseware, which is concerned with social and practical acceptability. The term social acceptability is related to the social basis of an educational system (student-, teacher- or patient-centered). The practical acceptability is examined through the evaluation of the following four sectors: (a) content, (b) presentation and organization of the content, (c) technical support and update processes and finally, (d) the evaluation of learning.

2.3. *MEFANET's portal platform and central gateway*

The MEFANET project aims to develop and strengthen the cooperation among medical faculties. It focuses on the progress in education of medical and health care disciplines using modern information and communication technologies by means of a common platform for sharing educational digital content. The primary objective of the MEFANET portal platform is to ensure a horizontal accessibility of electronic teaching tools for both teachers and students with full respect for the independence of the individual faculty [20–22].

A community of ten medical faculties provides a large target audience, which covers numerous academic staff – over 30,000 potential users. Logically, there appeared a need for a transparent and unified source of teaching contents. It was decided to design and implement a robust e-publishing platform, which provides the following benefits in particular to the users at engaged medical faculties:

- Easy data sharing and capturing;
- Ease of use and access;
- Common authentication framework;
- Quality assessment of published contents.

The main purposes of the MEFANET educational portal are:

- To help teachers with publishing their educational works;
- To help authors of multimedia teaching and learning tools to communicate with the users of these tools;
- To help people involved in the lifelong learning of physicians and health care professionals;
- To help students of medicine and health care disciplines with navigating through various tools for electronic support of their education.

The educational web portal of the Medical Faculty at Masaryk University has been accepted as the uniform groundwork for the solution of the common portal platform [23]. Every member of the network provides a standalone instance of the platform, which differs only in a local configuration, graphic template and various ISSN (International Standard Serial Number). All these independent web-based applications are covered by a central gateway [24]. This crucial part of the e-publishing system enables the effective gathering and monitoring all metadata about the contents published with the use of portal instances. A complete image of the available digital contents across the whole network is constructed in this way. Users can freely browse the objects sorted by medical discipline, author, faculty, or quality assessment criteria. The gateway, as well as each instance of the portal platform, includes a sophisticated tool for advanced searching in the entire MEFANET database and in external resources such as Google, Ariadne and Globe databases. Titles, keywords, other metadata and full-text indexes are searched and the results are given to the user in order of relevance.

3. E-publishing platform

3.1. *System description*

This subchapter describes the running environment and technologies used in the development process. The MEFANET portal platform runs on the most-used and widespread web servers – either an Apache server or a Microsoft Internet Information Server (IIS). The latest statistics show [25] that these two web server technologies host over 80% of the top 10,000 websites on the Internet (included Apache, IIS 5, IIS 6, IIS 7). We use Linux/Ubuntu and Windows Server operating systems for optimal performance. All technologies related to programming and system construction are PHP, XHTML, XML, CSS 2, JavaScript (JS), AJAX, MySQL.

We have also taken advantage of third party frameworks, such as jQuery (javascript library), Google Analytics (monitoring and analytics tool) and CKEditor (WYSIWYG text and HTML editor). The third party server application is used for the federative authentication process, which establishes user identity trust and authentication between universities. The Czech academic identity federation eduID.cz [26] fully provides means for inter-organizational identity management and access control into the MEFANET. The eduID.cz federation is based on one of the available tools for web single sign-on processes named Shibboleth [27]. All of the portal development has adhered to recommended rules and declared standards.

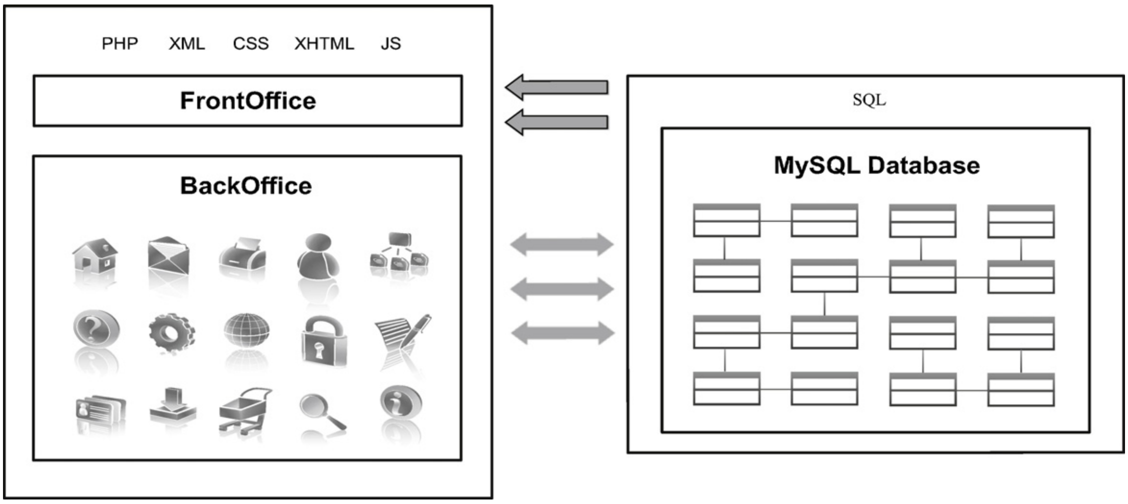


Fig. 1 – Global view of the structure of the MEFANET web portal instance.

Each instance of the portal platform is composed of two major sections: FrontOffice and BackOffice. The FrontOffice interface is used to present the published contents to users. It is a non-secured and freely available on-line part of the educational portal. Each visitor can browse and search digital teaching materials here. But not every user has free access to all published educational content. There are defined restrictions in access rights, which are described in detail in the subchapter MEFANET user's identity. The BackOffice is the second, secured part of the system, which is used by editors and administrators. This section is divided into separate units called modules. Each module serves to manage a selected area of functionality (Guest book, Users, Contributions, Submitted contributions, Comments, Editing, File Manager, Links, Sections, Authors, Departments, Courses, Metadata, Full pages, Languages, Import), see Fig. 1.

Each system module is built on the most widespread software architecture for today's world web applications called Model-View-Controller (MVC). This concept was described by Reenskaug [28] and it was first implemented by J. Althoff for the Smalltalk-80 class library. The MVC pattern was conceived as a general solution to the problem of users controlling a large and complex data set [28]. The main idea consists in separating an application into three parts, which represent core functionalities: model, view and controller. Fig. 2 shows all direct associations and notifications.

The model is responsible for managing the data, storing and retrieving entities used by an application, usually from a database, and contains the logic implemented by the application [29]. The view is responsible for displaying the data provided by the model in a specific format. It has a similar usage with the template modules present in some popular web applications, such as Wordpress [30] or Joomla [31]. Three types of view are implemented in the MEFANET portal platform – FrontOffice view, BackOffice view and Central Gateway view. The controller handles the model and view layers to work together and receives a request from the client, invoke the model to perform the requested operations and send the data to the view. The view then formats and presents the data to the user in a web application as an html output [32]. Models do not directly call views. Instead, each view registers itself with its model, and the model notifies all registered objects whenever it is updated. This lets a developer add or change views without altering the model. It also ensures that views are synchronous because each view reflects the same model state [33].

3.2. *Fundamental elements of the MEFANET portal platform*

Scalability and extensive customizations are important and desired properties of the MEFANET educational web portal. On the other hand, there are also several legitimate requests for particular common conventions which should be followed on the parts of local administrators. With this in mind, we established three fundamental elements in the MEFANET e-publishing platform, which are common and obligatory for all portal instances and the Central Gateway [34]:

1. Medical disciplines linker – to sort and categorize the published items. 2. Authentication framework – to provide an effective and easy authentication mechanism for user identification. 3. Multidimensional quality assessment – to ensure the publishing of quality data content.

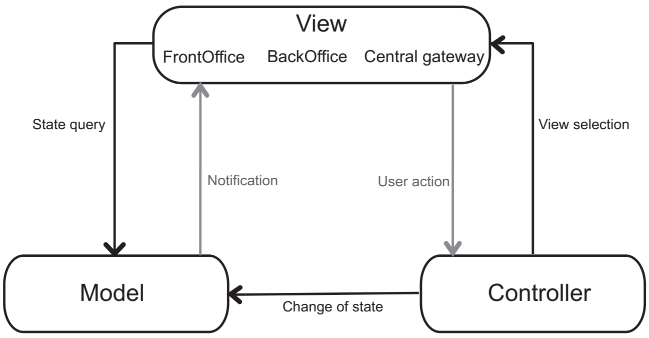


Fig. 2 – The description of relationships between components in the MVC architecture.

Table 1 – The list of medical disciplines for the categorization of educational content.

Anaesthesiology and Intensive Care Medicine	Anatomy
Biology	Biophysics
Cardiology, Angiology	Dentistry
Dermatology	Diabetology, Dietetics
Emergency Medicine	Endocrinology, Metabolism
Epidemiology, Preventive Medicine, Hygiene	Gastroenterology and Hepatology
General Practice Medicine	Genetics
Geriatrics	Hematology
Health Care and Nursing	Histology, Embryology
Immunology, Allergology	Infectology
Internal Medicine	Laboratory Diagnostics
Medical Ethics and Law	Medical Chemistry and Biochemistry
Medical Informatics	Microbiology
Nephrology	Neurology
Neurosurgery	Nuclear Medicine
Obstetrics, Gynecology	Occupational Medicine and Toxicology
Oncology, Radiation Therapy	Ophthalmology and Optometry
Other	Otorhinolaryngology
Pediatrics, Neonatology	Pathology and Forensic Medicine
Pharmacology	Physiology and Pathophysiology
Psychiatry, Psychology, Sexology	Public Health Care, Social Medicine
Radiology and Imaging	Rehabilitation, Physiotherapy, Ergotherapy

3.2.1. *Medical disciplines linker*

This feature enables basic categorization of the educational contents and covers all learning areas among medical faculties. In the beginning, a single-level or multilevel list of medical specializations was considered, as well as the possibility to adapt an existing scheme from the National Library of the Czech Republic, which is based on the standard Conspectus method [35]. However, the medical disciplines mapping according to the Conspectus method showed to be inapplicable for MEFANET purposes. The medical disciplines linker is based on a varied list of medical fields adopted from significant medical publishing resources and can be modified only with agreement of the MEFANET Coordinating Council. As of February 2011, the list contains forty-eight medical fields, see Table 1.

3.2.2. *Authentication framework*

The objective of this part is to describe the MEFANET authentication framework implementation, which is provided by the eduID.cz federation. The primary goal consists in the application of web single sign-on (SSO), which ensures users access to multiple applications using just one password, which is generated and used by users within their home organization. Federated authentication and authorization services are key requirements in creating a secure and trusted technology infrastructure for the sharing and use of information by individuals at multiple locations [36]. The authentication represents a verification process that someone or something is, in fact, who or what it is declared to be ("I am who I say I am"). The authorization is used to find out if the person, once identified, is permitted to have the resource or to do the operation ("what I can see or do based on what my role is").

The eduID.cz federation is based on technologies which supported a standard-compliant SAML implementation. The SAML, developed by the Organization for the Advancement of Structured Information Standards (OASIS), is an XML-based framework for communicating user authentication, entitlement, and attribute information. It is a flexible and extensible protocol designed to be used – and customized, if necessary – by other standards. It enables web SSO through the communication of an authentication assertion from the first site to the second, which, if confident of the origin of the assertion, can choose to log in the user as if they had authenticated directly. SAML is used by several other standards groups to provide a security and identity underpinning for their work, such as Liberty Alliance, XACML, WS-Security, WS-* and Shibboleth, which is used in the MEFANET portal platform [37].

The Shibboleth is a standard-based and open-source architecture for supporting inter-institutional sharing of resources that are subject to access control. It is an attempt to solve the federated administration of authentication/authorization by allowing a user's home institution to retain control of the authentication process while releasing enough information to the remote resource so it is able to make authorization decisions and preserve the user's privacy to the maximum extent possible [27].

3.2.2.1. *MEFANET user's identity.* Users' identities are verified without need of previous registration and users are only required to enter their existing login and password to confirm affiliation to their institution. The MEFANET authentication mechanism is based on SAML concept, which permits secured transfer of authentication data between independent organizations that have an established trust relationship. As noted above, the SAML has been adopted for use with several other standard frameworks. Its specification defines the structure and content of both the assertions and the protocol messages used to transfer requested information. The following process explains the simplified workflow of SAML-based SSO service. There are three involved parts: User (access via web browser), Identity Provider (IdP – creates, maintains, and manages user identity), Service Provider (SP – controls access to services and resources). 1. The user attempts to access a resource on the MEFANET portal platform. 2. The web server determines that it has no credentials about the user (not even who the user is or what institutional affiliation it has). 3. The SP generates an SAML authentication request and sends a redirect response to the user's browser. 4. The browser redirects the uniform resource locator and includes the encoded SAML authentication request that should be submitted to IdP SSO service. 5. The SSO service determines whether the user has an existing logon security context at the IdP that meets authentication policy requirements. The IdP SSO generates an SAML assertion and returns the encoded response message placed in an HTML form to the browser. This message contains the authenticated user's data provided by the IdP to the SP within the MEFANET network, which included controlled attributes (for example GivenName, Surname, Mail, Organization) and the local attribute named http://www.mefanet.cz/mefaperson. This parameter was especially designed for MEFANET purposes and its value ("true" or "false") differs in dependence on user's medical faculty affiliation. 6. The browser, due either to a user action or execution of an "auto-submit" script, issues an HTTP POST request to send the form to the SP's Assertion Consumer Service (ACS). The ACS obtains, processes and verifies the message from the HTML form. If the response is successfully verified, ACS redirects the user to the destination URL. 7. The user has sufficient authorization to access the resources on the MEFANET educational portal [26,27,37].

The access to the portal itself is not restricted anyhow. The educational content is comprised of contributions which are accessible for anyone who searches the portal. A particular contribution consists of a title, a short abstract, keywords, categorization, associated files and hypertext links. Every attachment (file or link) has its own access rights, which strictly define users who can access this material. Authors can choose from following user groups, in order to permit/deny access to their materials:

- non-registered anonymous user,
- registered anonymous user, who accepts the terms of use within his registration,
- user of the MEFANET network, i.e. student or teacher from any Czech or Slovak medical faculty university member,
- user of a local university, whose affiliation to that university has been verified at the portal via the local information system of that university,
- user of a local medical faculty, whose affiliation to that faculty has been verified at the portal via the local information system of the respective university or faculty,
- user to whom attachments are made available only with the author's explicit consent.

The most widespread and preferred is "user of the MEFANET network", so it means that all students and academic staff from the MEFANET network can freely access and view the offer of electronic study materials.

Every attachment has special parameters, which determine clinically sensitive material. These attachments consist of any materials, which could only be made after the patient gave his informed consent and will not be available to students of non-medical faculties. A ranking of "clinically sensitive" depends on the decision of the author of the respective contribution(s). Clinically sensitive materials can include various media, such as video records from multimedia atlases in which patients talk about their condition, or photos and videos from surgical procedures which reveal a patient's identity.

3.2.3. *Multidimensional quality assessment*

After individual portal instances had been implemented, the central gateway was developed and consequently included into the MEFANET e-publishing platform. Since the start of the inter-university educational portal in March 2008, the number of published contribution has rapidly grown, see Fig. 5. We started to place major emphasis on quality instead of the quantity of learning content and the development of a new effective classification and review mechanism began.

Two different review approaches inspired the creation of procedures for multidimensional quality assessment:

1. An opponency action – in the case of a pedagogy work which meets basic criteria set by an academic board of a faculty, the work is reviewed and then published with a symbol of an editorial committee or another authority of that faculty.
2. Guarantee signatures – there are guarantees defined for each of the medical disciplines in the linker described in Table 1. They are allowed to express their opinions and objections for each contribution assigned to their disciplines. Their final decision can be one of the following meanings:

a. Accept – the contribution will be signed by a positive icon with an alternate text identifying the guarantor;
b. Reject – the contribution will be signed by a negative icon with an alternate text identifying the guarantor. The portal's administrator should discard this contribution from the medical discipline.
c. Undefined state – the contribution is not signed by any icon. Either it has not been read by the guarantor or there are some insignificant objectives which do not imply discarding the contribution.

We decided to combine the above described approaches with new ideas. The final set of procedures is called the 4-D quality assessment. The whole process stands on four independent principles, which enable easy classification as well as a complex on-line review workflow with the use of XML template forms in one of the BackOffice modules. The four dimensions consist of the following parts, see Fig. 3:

- review,
- typological classification,
- level of the target group,
- user self-study score.

3.2.3.1. Review. Each contribution published on a portal instance includes the teaching materials as well as additional metadata, including information about authors and their affiliations. Anonymous or social collaborative forms of publishing of educational resources are impossible. The first dimension represents the editorial review process, which ensures an easy and effective control mechanism for publishing educational content under the guarantee supervision. The guarantor is a selected expert from medical society who expresses the opinion to newly created contributions, establishes the review process and nominates reviewers. Each reviewer consequently fills in the form with binary and open questions in the BackOffice of the portal platform. Only the contributions with positive review conclusions and guarantor approval can be published on the central gateway.

3.2.3.2. Typological classification. Each contribution needs to classify to one or more of the following categories, which cover all possible media types of teaching content which might be published:

- Textbooks and manuals
- Educational websites and atlases

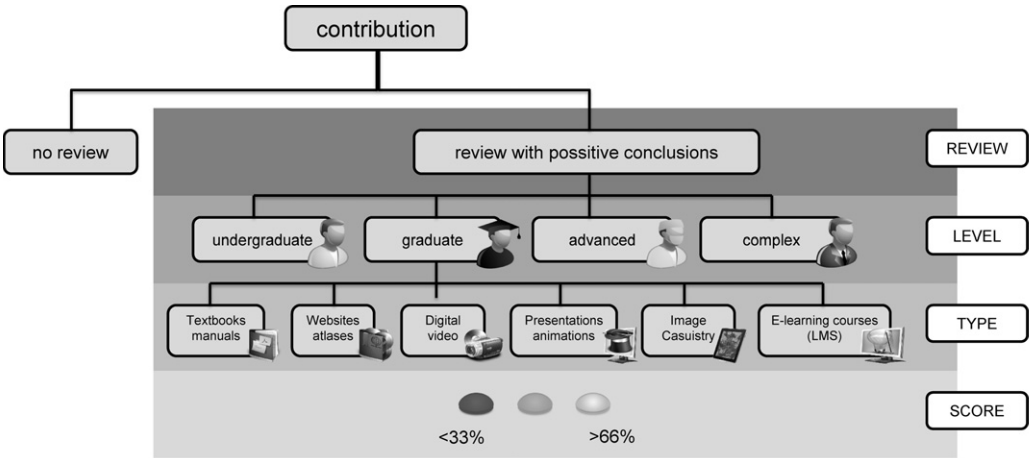


Fig. 3 – The 4-D quality assessment mechanism.

- Digital video
- Presentations and animations
- Casuistics in images
- E-learning courses (Learning Management System)

The list of typological classification options is fixed and approved by the MEFANET Coordinating Council. Every media type has its own graphic representation which is highlighted when selected.

3.2.3.3. *Level of the target group.* The following list of user groups represents the scale of levels of user formal education.

- Undergraduate level – teaching materials advisable for bachelor and master degree.
- Graduate – study materials advisable for graduates and advanced graduates.
- Advanced Graduate – study materials advisable for deeply interested graduates.
- Complex – study materials that cover all the previous levels comprehensively.

Authors and reviewers select one or more options for each contribution and their preferences are logically unified and graphically reproduced in the 4-D quality assessment section of the contribution when displayed to users.

3.2.3.4. *Users' self-study score.* The users anonymously express their opinions on the suitability of provided teaching materials for their self-study. They select a scalar value on a scale from 0 to 100 points and further add information about their education level: (a) student, (b) graduate, (c) postgraduate. The final score is computed as a weighted average of all votes:

$$Sc = \frac{\sum_{i=1}^{n}(S_i \times w_i)}{\sum_{i=1}^{n} w_i} \quad (1)$$

where *Sc* is final self-study score, s_i is individual user's score (range from 0 to 100), w_i is individual weight of vote (calculated via Table 2) and *n* is total users' votes amount.

3.3. *Quality criteria model*

The proposed criteria model was inspired by several studies published in the quality content assessment area, as well as in the research of educational portals, as referred to previously by Butcher [38], Bottentuit and Coutinho [16], Yang et al. [15], Caro [17,18], Elissavet and Economides [19] and Karlsson [39]. We used our experiences and recommended the most important and significant properties. The final model is divided into three appropriate parts (see Fig. 4). There appears diversity in comparison with traditional commercial web portals, it is recommended to use described indicators primary for educational portals.The quality criteria model consist of website quality (design, development, performance, security and privacy), web content quality and web services quality indicators (see Table 3).

4. Discussion

All ten instances of the portal platform and central gateway are available on-line and open to everyone in Czech/Slovak and English versions. Published contributions have no access restriction, but related attachments always have strictly defined access rights, which define users who can view this study material. There are no specific browser requirements, visitors can use a standard web browser extended by common

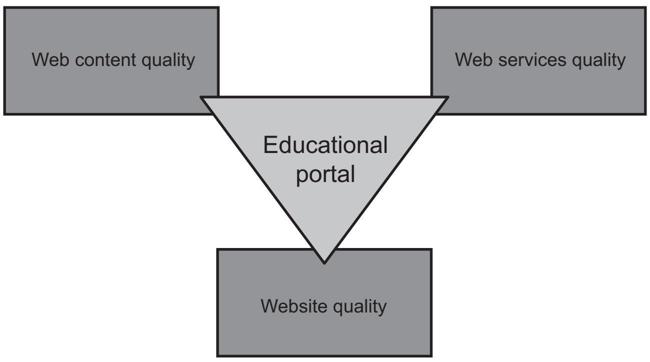


Fig. 4 – Quality criteria of educational portal platform.

Table 2 – Weight of votes into the self-study qualification mechanism.

Level of the target group	Undergraduate	Graduate	Advanced graduate	Complex
User's education level				
Student	2	1	1	1
Graduate	2	2	1	1
Postgraduate	2	2	2	2

plug-ins. The latest release marked as version 1.8 contains the 4-D quality assessment implementation and many other upgrades. All members of the educational network can use its tools for quality evaluation of their published electronic teaching materials. Every updated version contains users' documentation including UML (Unified Modeling Language) standardized models. It is the set of diagrams of a system, which represents static and dynamic structure (for example entity-relations diagrams, use-case diagrams and activity diagrams). We can recommend an application of quality criteria during the development process, which are described in detail in Section 4.2. It is effective and useful from both the user's and administration's point of view.

The number of contributions published at the central gateway [24] has been growing since the launch of the central gateway (see Fig. 5). Apparently, the graph of published contributions implies three local rapid changes – June 2008, June 2010 and September 2010. These significant risings relate to the launching of new instances into the portal network. In April 2011, the central gateway contained 2035 contributions, which are categorized according to well-arranged medical fields.

The multidimensional quality assessment enables effective quantifying of digital content in the entire MEFANET network. It is an original and innovative approach, because other freely accessible medical repositories only exceptionally use various peer-reviewed mechanisms for quality guarantee of publishing resources (e.g. MedEdPortal, MERLOT, HEAL). Probably the largest educational content publication service MedEdPortal (provided by the Association of American Medical Colleges) clearly marked all learning materials with the appropriate review logo based on the evaluation approach (Peer Reviewed, Editorial Reviewed, Special Collection Reviewed) [8]. In essence, the primary objective is common for all mentioned systems: provide relevant, correct and useful educational data to medical learners and educators. However, the MEFANET solution promises a more complex functionality and wider range of tools for organizing the published contents as well as a possibility to present the contents completed by comments from tutors selected from expert medical societies. The graphic visualization of designed model gives users all appropriate information related to the learning material. We believe in success of the multidimensional assessment in the medical faculties network and hope it is the best suited classification and review mechanisms for all involved participants, which are directed by local editorial committees or other rules.

The preparation of educational content intended for medical disciplines differs from other fields. Authors have to consider higher claims and special rules related to the publishing of electronic medical learning materials. It is quite common to find clinically sensitive attachments including various media, such as videos from hypertext multimedia atlases in which patients talk about their condition, or videos from surgical procedures which reveal a patient's identity. Ranking the attachment as clinically sensitive depends on the author's decision and its content is then unavailable to students of non-medical faculties. The use of clinically significant materials form an integral and very important part of the matriculation oath of all medical students. In general, producers of medical-based learning content have to be very careful. All mistakes and incorrect statements contained in teaching materials have dramatic impact on the knowledge and skills of future physicians. All these published inaccuracies cause incorrect habits and can result in malpractice during the real human treatment. One of the MEFANET portal platform benefits for students and teachers is the complex classification and assessment mechanisms, which enable a gradual

Table 3 – Quality criteria of educational portal.

Website quality	Web content quality	Web services quality
Screen resolution	Varied formats	Multi-level navigation
Corresponding icons and symbols	Content quality assessment	Searching facilities
Simple and easy navigation	Transparent categorization	Communication
Readability of the information	Target audience	Printable version of content
Accessibility of the sites	No anonymous content	Technical support
Consistent representation	Updated information	Authors support
Site esthetics, color harmony	Currency of information	Administration background
Code and links validity	Amount of data	Support of online inquiries
Authorization	Content uniqueness	
SEO	Object description	
Optimization, standardization		
Website feedback monitoring		
Page loading time		
Response time		
Failure restriction		
Rights management		

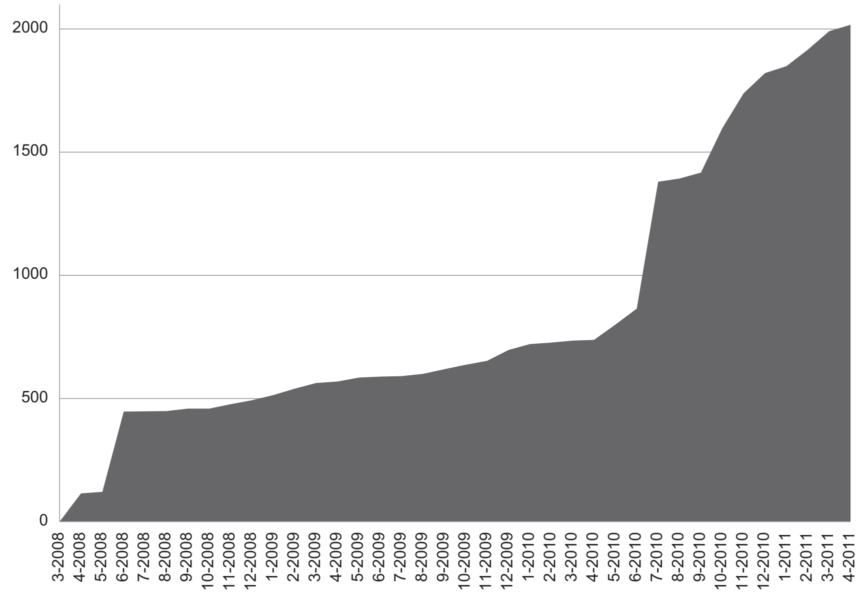


Fig. 5 – Progress of published contributions (May 2008–April 2011).

improvement in the quality of the content throughout the medical faculty network.

5. Conclusions

In this paper, the MEFANET portal platform and its advantages in the areas of biomedical and health care education with the use of modern information and communication technologies were presented. Students – about 16,500 potential users and academic staff – about 3900 potential users from all Czech and Slovak medical faculties can find their e-learning content via ten standalone faculties' instances of the educational portal and central gateway, which makes it possible to collect all metadata from instances in the network and provides a complete image of available digital content. This web-oriented e-publishing platform offers a well-organized and guaranteed collection of many different types of study materials. The cardinal importance represents the implementation of the multidimensional quality assessment, which should solve long-term problem of unreviewed content publishing.

Many other interesting challenges now exist in the field of medical education, such as virtual patients (VPs) and mobile technology for learning. VPs technology covers computer simulation and virtual reality educational tools and offers several advanced methods for improving the standardized patients learning. There are enough primary content resources available through the MEFANET e-publishing platform, which could be used to create wide range of very complex VPs as well as other objects for training of clinical reasoning. Today a lot of web applications for VPs authoring are available and, from the MEFANET perspective, the key task rests in appropriate selection according to compatibility options. The MEFANET network is advantageous for VP technology deployment, because of the already existing infrastructure, including secure user authentication and powerful tools for monitoring the quality of published contributions. The same benefits can then be applied to the introduction of new interfaces for portable devices so that the innovative learning approaches called m-learning (mobile learning) could be made available to students and academics in the network.

Conflict of interest statement

The authors disclose any financial and personal relationships with other people or organizations that could inappropriately influence our work.

Acknowledgements

This work was supported by the grant project GACR no. 102/09/H083 and by the project CZ.1.07/2.4.00/12.0050 funded from European Social Fund and the state budget of the Czech Republic.

REFERENCES

[1] B. Barnes, C. Friedman, Using technology incontinuing professional development The Continuing Professional Development of Physicians, vol. 2003, The Continuing Professional Development of Physicians, Chicago, 2003, pp. 205–220.
[2] S.C. Cobb, Internet continuing education for health care professionals: an integrative review, Journal of Continuing Education in the Health Professions 24 (3) (2004) 171–180.
[3] R.M. Harden, E-learning-caged bird or soaring eagle? Medical Teacher 30 (1) (2008) 1–4.
[4] R.M. Harden, A new vision for distance learning and continuing medical education, Journal of Continuing Education in the Health Professions 25 (1) (2005) 43–51.
[5] Cisco Systems Reusable Information Object Strategy: Definition, Creation Overview and Guidelines (Online). Available: http://www.cisco.com/warp/public/779/ibs/

solutions/learning/whitepapers/el_cisco_rio.pdf/ (accessed 18.03.11).

[6] Jorum – Learning to Share (Online) Available: http://www.jorum.ac.uk/(accessed 24.03.11).

[7] Wisc-Online (Online) Available: http://www.wisc-online.com/ (accessed 24.03.11).

[8] MedEdPortal (Online) Available: https://www.mededportal.org/ (accessed 18.01.12).

[9] Multimedia Educational Resources for Learning and Online Teaching (Online). Available: http://www.merlot.org/merlot/index.htm (accessed 25.01.12).

[10] Health Education Assets Library (Online) Available: http://www.healcentral.org/ (accessed 24.01.12).

[11] World Wide Web Consortium (W3C) (Online). Available: http://www.w3.org/ (accessed 25.03.11).

[12] A. Woukeu, et al., Ontological hypermedia in education: a framework for building web-based educational portals, in: Proceedings of ED-MEDIA, 2003, pp. 23–28.

[13] M. Geueke, J. Stausberg, A meta-data-based learning resource server for medicine, Computer Methods and Programs in Biomedicine 72 (3) (2003) 197–208.

[14] B. Flagg, Formative evaluation for educational technologies, vol. 1990, Harvard University/Lawrence Erlbaum Associates Publishers, 1990.

[15] Z. Yang, S. Cai, Z. Zhou, N. Zhou, Development and validation of an instrument to measure user perceived service quality of information presenting Web portals, Information & Management 42 (4) (2005) 575–589.

[16] J.B. Bottentuit Jr., C.P. Coutinho, The Conception of a Rubric to Evaluate Educational Portals on the Web (Online) Available: http://repositorium.sdum.uminho.pt/handle/1822/7765 (accessed 18.03.11).

[17] A. Caro, C. Calero, I. Caballero, M. Piattini, Defining a data quality model for web portals Web Information Systems – WISE, vol. 42, Springer, Berlin, Heidelberg, 2006, pp. 363–374.

[18] A. Caro, C. Calero, I. Caballero, M. Piattini, Towards a data quality model for web portals Web Information Systems and Technologies, vol. 1, Springer, Berlin Heidelberg, 2007, pp. 228–237.

[19] G. Elissavet, A. Economides, An Evaluation Instrument for Hypermedia Courseware Educational Technology & Society, vol. 2003, 2003, pp. 31–44.

[20] D. Schwarz, L. Dušek, The MEFANET Project (Online] Available: http://www.mefanet.cz/index-en.php (accessed 18.03.11).

[21] D. Schwarz, M. Komenda, R. Šustr, I. Šnábl and L. Dušek, MEFANET after four years of progressing: 4-D model for digital content quality assessment, Presented at the MEFANET 2010 Conference, Brno, 2010.

[22] D. Schwarz, S. Štípek, V. Mihál, L. Dušek, MEFANET: Educational Network of Medical Faculties for Developing and Sharing Digital Content on a Common Platform, Presented at The SCO 2009 Conference, Brno, 2009.

[23] Educational Portal of Masaryk University's Faculty of Medicine (Online) Available http://portal.med.muni.cz/index-en.php (accessed 18.03.11).

[24] MEFANET Central Gate (Online) Available: http://portal.mefanet.cz/index-en.php (accessed 18.03.11).

[25] Web Server Technology Web Usage Statistics (Online) Available: http://trends.builtwith.com/Web-Server (accessed 18.03.11).

[26] Czech Academic Identity Federation eduID.cz (Online) Available: http://www.eduid.cz/wiki/en/eduid/index (accessed 18.03.11).

[27] M. Needleman, The shibboleth authentication/authorization system, Serials Review 30 (3) (2004) 252–253.

[28] T. Reenskaug, MVC – XEROX PARC 1978–79 (Online) Available http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html (accessed 18.03.11).

[29] E. Curry, P. Grace, Flexible self-management using the model-view-controller pattern, IEEE 25 (3) (2008) 84–90.

[30] WordPress: Blog Tool and Publishing Platform (Online) Available: http://wordpress.org/ (accessed 24.03.11).

[31] Joomla! (Online) Available: http://www.joomla.org/ (accessed 24.03.11).

[32] Model View Controller (MVC) in PHP (Online) Available: http://php-html.net/tutorials/model-view-controller-in-php/ (accessed 24.03.11).

[33] M.J. Mahemoff, L.J. Johnston, Handling multiple domain objects with model-view-controller, in: Proceedings Technology of Object-Oriented Languages and Systems, TOOLS 32, 1999, pp. 28–39.

[34] M. Komenda, D. Schwarz, I. Šnábl, L. Dušek, S. Štípek, V. Mihál, The common portal platform in the MEFANET, Presented at the AMEE 2010, Glasgow, 2010.

[35] B. Stoklasová, Conspectus method used for collections mapping and structuring of portals in Czech libraries, Finnish Research Library Association Journal (2006), Available: http://info.jib.cz/dokumenty/helsinkistoklasovatext.pdf/ (accessed 12.03.11) (Online).

[36] H.D. Lambert, C.F. Leonhardt, Federated authentication to support information sharing: Shibboleth in a bio-surveillance information grid, International Congress Series 1268 (2004) 135–140.

[37] SAML V2.0 Executive Overview, OASIS, 2005, Available: http://www.oasis-open.org/committees/download.php/13525/sstc-saml-exec-overview-2.0-cd-01-2col.pdf/ (accessed 02.03.11) (online).

[38] N. Butcher, Best Practice in Education Portals: Research Document Prepared for The Commonwealth of Learning and SchoolNet Africa, Braamfontein, 2002, Available: http://www.col.org/SiteCollectionDocuments/02EducationPortals_Report.pdf (accessed 20.03.11) (Online).

[39] K. Karlsson, J. Olsson, University Web Portals Quality Assurance: User Perspective (Online). Available http://gupea.ub.gu.se/handle/2077/10544 (accessed 18.03.11).

Original Paper

Interactive Algorithms for Teaching and Learning Acute Medicine in the Network of Medical Faculties MEFANET

Daniel Schwarz[1], MSc, PhD; Petr Štourač[1,2], MD, PhD; Martin Komenda[1], MSc, Dr; Hana Harazim[1,2], MD; Martina Kosinová[2], MD; Jakub Gregor[1], MSc, PhD; Richard Hůlek[1], BSc; Olga Smékalová[2], MD; Ivo Křikava[2], MD; Roman Štoudek[2], MD; Ladislav Dušek[1], MSc, Dr, PhD

[1]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[2]Department of Anesthesiology and Intensive Care Medicine, University Hospital Brno, Faculty of Medicine, Masaryk University, Brno, Czech Republic

Corresponding Author:
Daniel Schwarz, MSc, PhD
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Kamenice 5
Brno, 625 00
Czech Republic
Phone: 420 549492854
Fax: 420 549492855
Email: schwarz@iba.muni.cz

Abstract

Background: Medical Faculties Network (MEFANET) has established itself as the authority for setting standards for medical educators in the Czech Republic and Slovakia, 2 independent countries with similar languages that once comprised a federation and that still retain the same curricular structure for medical education. One of the basic goals of the network is to advance medical teaching and learning with the use of modern information and communication technologies.

Objective: We present the education portal AKUTNE.CZ as an important part of the MEFANET's content. Our focus is primarily on simulation-based tools for teaching and learning acute medicine issues.

Methods: Three fundamental elements of the MEFANET e-publishing system are described: (1) medical disciplines linker, (2) authentication/authorization framework, and (3) multidimensional quality assessment. A new set of tools for technology-enhanced learning have been introduced recently: Sandbox (works in progress), WikiLectures (collaborative content authoring), Moodle-MEFANET (central learning management system), and Serious Games (virtual casuistics and interactive algorithms). The latest development in MEFANET is designed for indexing metadata about simulation-based learning objects, also known as electronic virtual patients or virtual clinical cases. The simulations assume the form of interactive algorithms for teaching and learning acute medicine. An anonymous questionnaire of 10 items was used to explore students' attitudes and interests in using the interactive algorithms as part of their medical or health care studies. Data collection was conducted over 10 days in February 2013.

Results: In total, 25 interactive algorithms in the Czech and English languages have been developed and published on the AKUTNE.CZ education portal to allow the users to test and improve their knowledge and skills in the field of acute medicine. In the feedback survey, 62 participants completed the online questionnaire (13.5%) from the total 460 addressed. Positive attitudes toward the interactive algorithms outnumbered negative trends.

Conclusions: The peer-reviewed algorithms were used for conducting problem-based learning sessions in general medicine (first aid, anesthesiology and pain management, emergency medicine) and in nursing (emergency medicine for midwives, obstetric analgesia, and anesthesia for midwifes). The feedback from the survey suggests that the students found the interactive algorithms as effective learning tools, facilitating enhanced knowledge in the field of acute medicine. The interactive algorithms, as a software platform, are open to academic use worldwide. The existing algorithms, in the form of simulation-based learning objects, can be incorporated into any educational website (subject to the approval of the authors).

(J Med Internet Res 2013;15(7):e135) doi:10.2196/jmir.2590

KEYWORDS
medical education; patient simulation; algorithms; students; community networks; problem-based learning; serious games; survey

Introduction

Medical education is constantly evolving by gradually, but significantly, shifting from traditional methods (eg, textbooks, lectures, bedside teaching) to a more comprehensive approach that also employs modern information and communication technology (ICT) tools (eg, e-learning, interactive algorithms, computer simulations, virtual patients). Such approaches have been demonstrated to enhance and improve the learning skills of medical students and residents in comparison to traditional methods [1-3]. Several ancillary factors in medicine and medical education have also contributed significantly to these trends; in particular, the rapid development of new technologies and the generally preferred shorter hospital stays, which reduces the student's exposure to a given case or diagnosis. The economic efficiencies of Web-based education and traditional face-to-face education approaches were compared under randomized controlled trial conditions in Maloney et al [4] and it was shown that the Web-based education approach was clearly more efficient from the perspective of the education provider.

Although most of the modern interactive tools are intended for extending and supplementing the traditional methods rather than replacing them, they have undoubtedly brought a number of advantages, such as equal and easy access for the students to all diagnoses, simulation of a variety of real-life situations, comprehensive interdisciplinary learning, and a higher level of comfort for hospitalized patients. Simulation-based learning also provides the unique opportunity of practicing knowledge application in a manner that mimics real-time patient care without posing a risk to the patient [5,6]. On the other hand, developing simulations and e-learning materials requires investment of the time of skilled professionals (eg, physicians, teachers, programmers); therefore, it is necessary to ensure that the time and resources expended is justified by the educational impact [7]. Furthermore, the developed tools are often accepted uncritically and with emphasis on technological sophistication at the expense of the underlying psychopedagogical theories [1].

Improved efficiency in the development of digital teaching and learning materials, as well as their higher quality, can be achieved by sharing the educational content and by initiating collaborative multi-institutional authoring teams together with joint efforts in establishing the methods for quality evaluation. The management of multisource content among academic institutions brings the necessity of correct indexing, metadescription, and proper categorization [8], as well as reimbursement [9-11] for the created resources. The idea of the medical faculties in the Czech Republic and Slovakia sharing their educational digital contents surfaced in 2006 for the first time. Soon after, in 2007, all 7 Czech medical faculties as well as all 3 Slovak medical faculties formally joined the new network. In 2012, representatives of the Czech and Slovak health care institutions joined the Medical Faculties Network (MEFANET) education network. The MEFANET project [12] aims to develop cooperation among the medical faculties to further the education of medical and health care disciplines using modern ICT via a common platform for sharing digital education content, as well as for assessing their quality through a multidimensional approach [13].

Most of the digital teaching described in recent literature has been prepared as Web-based works because Web technologies allow for easy incorporation of multimedia objects, interactive algorithms, animated simulations, etc. The work may then be easily accessed from any computer and by a defined target audience (eg, students of a particular medical school or course). The developed tools and simulations cover a wide range of medical disciplines, such as critical care [14,15], cardiology [3], hematology [1], neurology [16], surgery [17], metabolic disorders, imaging methods [18,19], and cytogenetics [20].

Acute medicine is a dynamic environment with high demands on team communication and leadership, requiring correct clinical reasoning and quick decision making under time pressure. Simulation offers a good and interesting platform for training multidisciplinary medical teams, facilitating interaction among the team members and enabling the team to function in an effective and coordinated manner [6]. Internet education resources for intensive care medicine have recently been reviewed by Kleinpell et al [14], who demonstrated that most of them are electronic forms of textbooks and articles rather than interactive algorithms and dynamic simulations. Davids et al [7] described an interactive Web-based simulation in which the user treats patients with electrolyte and acid-base disorders, selects the therapies and doses, and can immediately see the treatment results.

In this paper, we present the education portal AKUTNE.CZ [21] as an important part of the MEFANET's contents. It aims to be a comprehensive source of information and education materials covering all aspects of acute medicine for undergraduate and postgraduate students of the medical and health professions. We focus here primarily on the simulation-based tools for teaching and learning algorithms for acute patient care that form the backbone of AKUTNE.CZ. The simulations take the form of interactive algorithms and represent the basis for a new extension of MEFANET's activities incorporating focus on serious games.

Methods

Overview

MEFANET [12] has established itself as the standard-setting body for medical educators in the Czech Republic and Slovakia, 2 independent countries that once comprised a federation, have similar languages, and still retain the same curricular structure for medical education. One of the basic goals of the network is to advance medical teaching and learning with the use of modern information and communication technologies. As an instrument, MEFANET has decided to develop an original and uniform solution for educational Web portals that are used, together with a central gateway, to offer and share digital education content.

Students—approximately 16,500 potential users and academic staff and approximately 3900 potential users from all Czech and Slovak medical faculties—can find their e-learning materials at 11 standalone faculties' instances of an educational portal with the use of the indexing and searching engine, MEFANET Central Gateway [22].

MEFANET e-Publishing System

The idea of a shared e-publishing system is based on a set of standalone Web portals rather than on a centralized application hosted for all medical schools, which might be an inflexible and more vulnerable alternative solution. Each portal instance represents an independent publication media with its own International Standard Serial Number (ISSN) code and an editorial board. Local metadata describing the digital educational contents are replicated regularly to the central gateway (see metadata harvesting in Figure 1). There are 3 fundamental elements that have to be rigidly maintained on the part of local administrators: (1) the medical disciplines linker, (2) the authentication/authorization framework, and (3) multidimensional quality assessment. The other features, properties, and functionalities can be adapted or localized to meet the needs of the particular institution. A detailed description of the 3 fundamental elements is as follows. See [13] for full and comprehensive information.

The medical disciplines linker represents the main taxonomy of contributions within the frame of the network. With its single-level list of 56 medical specializations, it forms the only obligatory structure of a portal instance. Any change to its content is subject to approval of the MEFANET Coordinating Committee.

The authors of the shared teaching materials can choose from the following user groups to permit or deny access to their materials: (1) nonregistered anonymous users, (2) registered anonymous users who accept the terms of use within their registration, (3) users of the MEFANET network, that is, a student or teacher from any Czech or Slovak medical school (MEFAPERSON), (4) users from a local university whose affiliation to that university has been verified at the portal via the local information system of that university, (5) users to whom attachments are made available only after the author's explicit consent. Services of the Czech academic identity federation, eduID.cz [23], are used to check the affiliations of the users of the portal instances. This federation uses the Shibboleth technology, which is one of the several authentication frameworks allowing the sharing of Web resources among institutions using the Security Assertion Markup Language (SAML) protocol standard. The portal instances behave like service providers in this federation, whereas the information systems of the involved schools act as identity providers.

There are 4 dimensions of critical importance when evaluating the quality of electronic teaching materials: (1) expert review, (2) education level of target users, (3) classification by type, and (4) self-study score. The review includes binary questions as well as open questions. The structure of the review form can be localized by modifying an extensible markup language (XML) template file. The second dimension is represented by the education level of the target group of the teaching material, which is a useful piece of information for the users and the reviewers. The next dimension is represented by a multiple-choice classification according to the types of attachments—the enumerated scale includes static files for Web-based learning and interactive e-learning courses encapsulated in the learning management systems. The last dimension—a self-study score—indicates what users think about the usability of a particular contribution in their self-studies. The values of the first 3 dimensions of the 4D assessment are composed by authors, guarantors, and reviewers. Their activities and the workflow of a contribution are explained in Figure 2. In addition to the 4D quality assessment, all contributions submitted to the central gateway undergo an additional editorial process called *mentally active monitoring*. It focuses on the following issues: (1) metadata is filled in properly, (2) granularity of the attachments is suitable, and (3) all attached documents and the links are accessible for at least MEFAPERSON users. The monitoring of these 3 important issues is done not only at the syntax level, but also semantically; therefore, it is carried out by a team of editors in cooperation with the editors responsible for the local Web portals.

Recently, new tools for technology-enhanced learning have been introduced to the MEFANET network in addition to the common e-publishing portal platform. These new tools complement the portal platform suitably because they provide a higher level of interactivity for students during their self-study process. Figure 1 shows how the new 4 tools—Sandbox [24], WikiLectures [25], Moodle-MEFANET [26], and Serious Games [27]—are related to the already established and standardized MEFANET Central Gateway.

The Serious Games extension is the latest development in MEFANET and it is designed for indexing metadata about simulation-based learning objects, also known as *electronic virtual patients* or *virtual clinical cases*. The first comprehensive set of such interactive learning objects is composed by algorithms for acute patient care published at the AKUTNE.CZ educational portal [21] together with other digital education materials covering a wide range of acute medicine topics.

Figure 1. MEFANET involves all medical schools in the Czech Republic and Slovakia. They share one another's digital teaching and learning materials by using an e-publishing system that consists of 11 educational Web portals and a central gateway. The extensions of the MEFANET e-publishing system appear as standalone platforms for their users. However, all teaching or learning materials indexed by the MEFANET Central Gateway undergo the same procedures of multidimensional quality assessment.

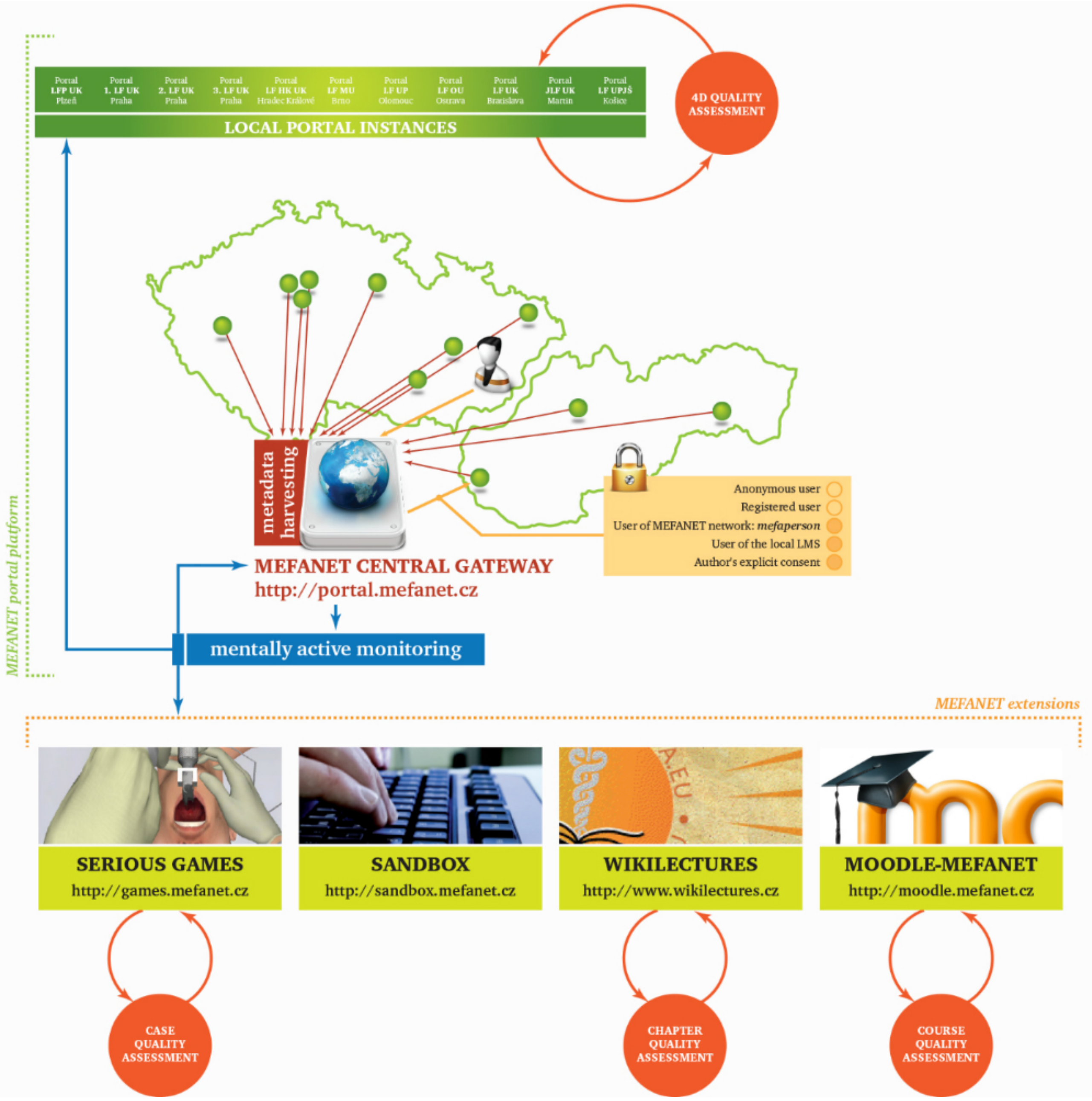
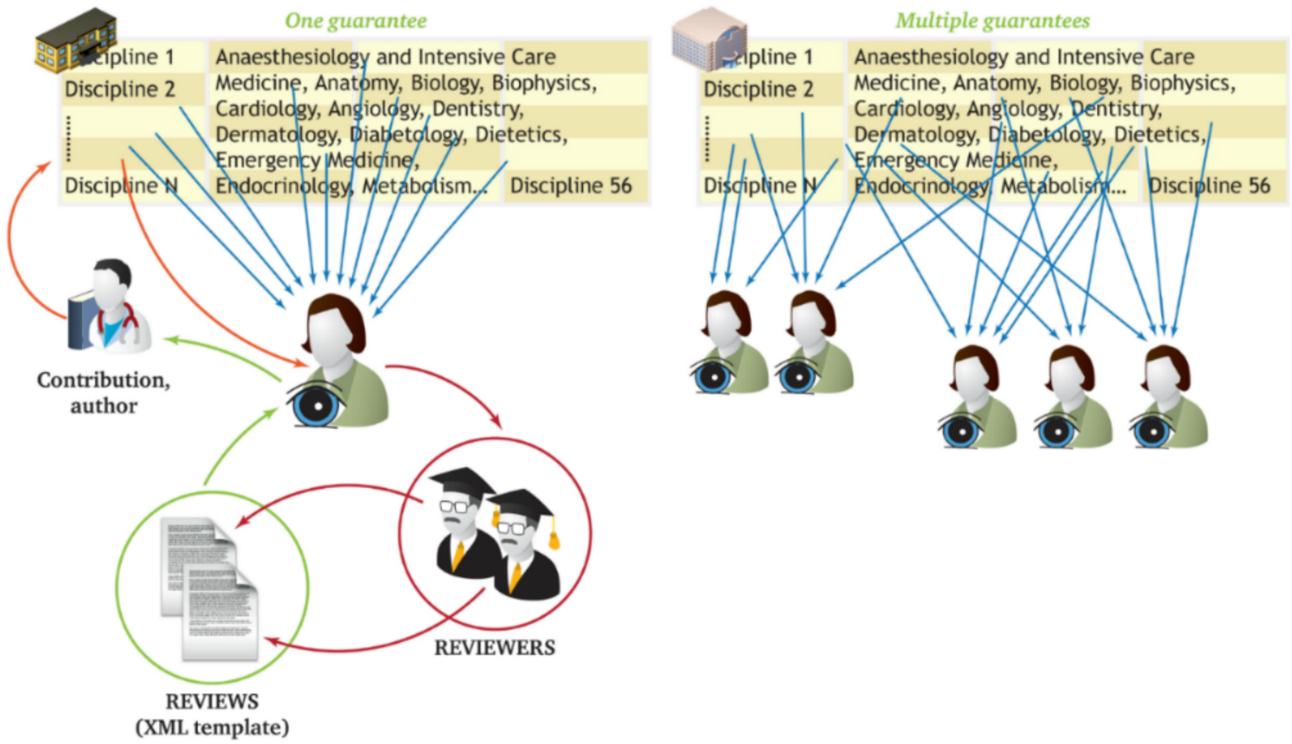


Figure 2. The contribution workflow scheme: (1) the author and technical editor finishes the contribution, (2) the guarantor, who is associated with a particular medical discipline, is notified about a new contribution to his/her field of interest, (3) the guarantor, either alone or with the help of the faculty's editorial committee, invites 2 reviewers to present their reviews online with the use of template-generated forms.



Interactive Algorithms for Teaching and Learning Acute Medicine

Each physician dealing with acute patients needs algorithmic thinking and correct clinical reasoning. Our interactive algorithms take the form of content-rich virtual cases because they link together process flowcharts and multimedia. Creating such algorithms or electronic virtual patients is laborious, time-consuming, and often accompanied by ambiguities and hesitations. Following the principles of student-centered learning, our authoring teams consisted of medical students in the final years of their studies, supervised by an experienced clinician. The complete workflow of the authoring process is outlined in Figure 3.

It takes 10 to 50 hours of active work to produce 1 interactive algorithm. The time of the team members is spent on collaborative work, meetings, and on self-studying. Student–authors consult their problems and reservations with a supervisor assigned to them and the resulting product is then submitted to an external reviewer, usually an experienced clinician or an academic staff from another workplace. After the incorporation of all reviewers' comments, the algorithm is completed by metadata to be published on the AKUTNE.CZ educational portal. Finally, sets of algorithms are compiled together with their metadata into a contribution to be published and indexed on the MEFANET Central Gateway. These contributions with a wider scope than individual algorithms are subjected to the multidimensional quality assessment described previously. Finished and published algorithms are used by other students either as outlines for problem-based learning (PBL) sessions or as supplementary materials for training and adopting correct clinical reasoning.

The interactive algorithms are authored with the use of a Web-based (PHP/MySQL) BackOffice application that provides the student–authors the following functionalities through its online forms and drag and drop control: (1) node-based scenario design, (2) description of the situation in each node, including the intervals of parameter values of physical examinations, intervals of laboratory values, and multimedia, (3) description of the correct answers as well as distractors with the option to repeat or end in a fatality, and (4) data export for each finished algorithm into an XML document. The XML documents are then rendered into a Flash object resembling a serious game. A student–player uses the game or this simulation-based learning object by moving between the nodes, which may be of different types, as shown in the sample algorithm in Figure 4. Each move causes a shift in the timeline as a side effect of the student–player's action, lending authenticity to the scenario and creating a stress effect, which is pronounced in real-life situations when dealing with acute patients. Continuous change of various numerical parameters reflecting the development of patient's clinical status and vital functions in time (eg, blood pressure, pulse, oxygen saturation) is also available (see the example of a node of a selected algorithm in Figure 5).

Students' Feedback on the Interactive Algorithms

We asked students about their attitudes and interest in using the interactive algorithms as part of their medical or health care studies. The purpose was to ascertain how the students perceived our efforts on authoring and implementing simulation-based learning tools that are so demanding to create. An anonymous questionnaire of 10 items (see Table 1 for complete overview of questions and answer options) was created and presented via SurveyMonkey [28], a free online survey software. Data collection lasted for 10 days in February 2013. The students who enrolled at 1 of the educational workshops or a conference organized by the group around the AKUTNE.CZ portal were asked to complete the survey. The first 4 questions were aimed at obtaining basic data about the respondents, so that the ones who did not study any field of medicine or health care could be filtered out as well as the ones who did know about our interactive algorithms at all. Further questions were answered with a 5-point Likert scale and 1 binary question was aimed at seeking feedback on the use of our interactive algorithms in the studies of acute medicine topics.

Figure 3. The authoring workflow of an interactive algorithm from choosing the topic through a review process to deployment to teaching in the form of a moderated problem-based learning session.

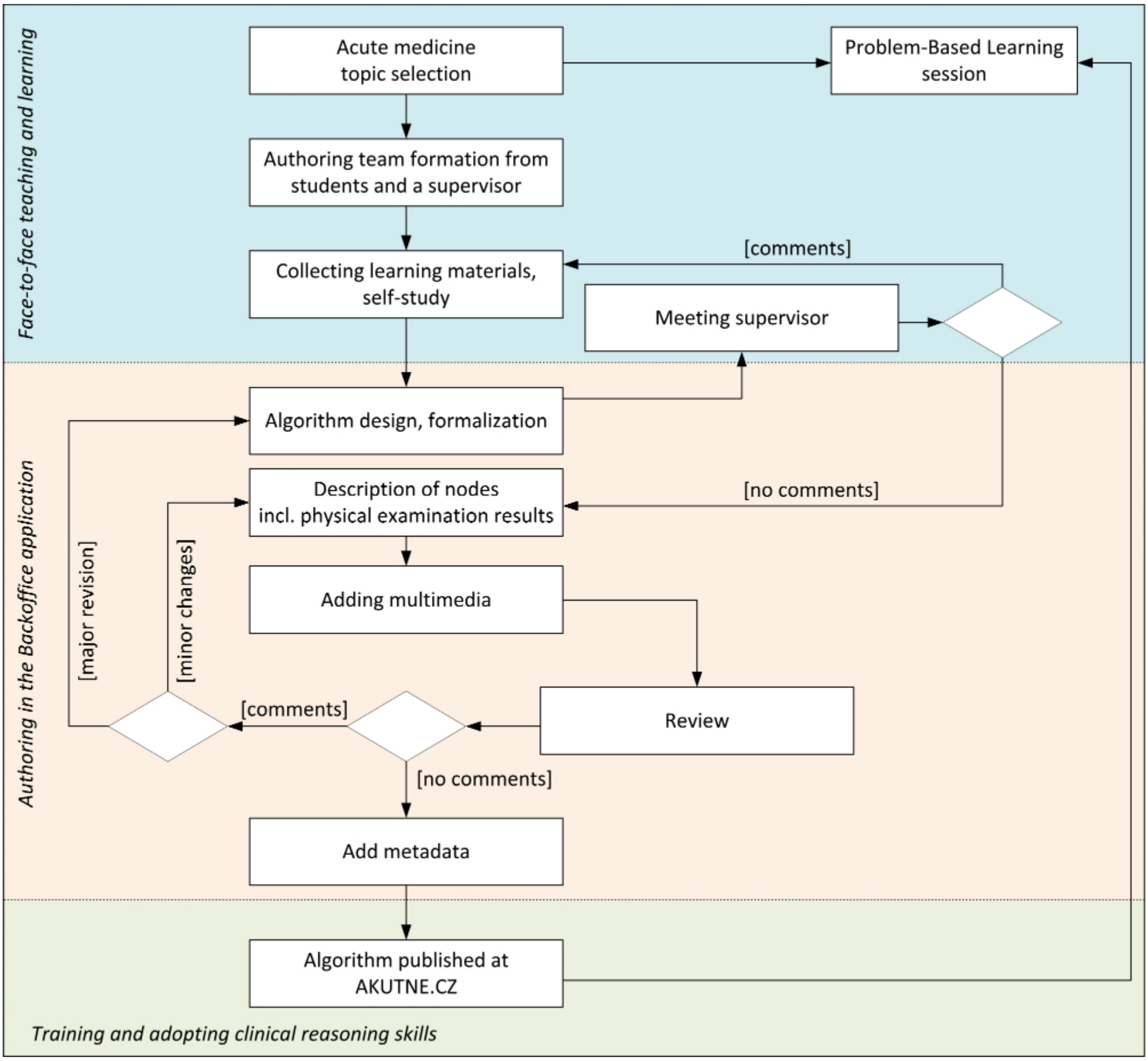


Figure 4. Various types of nodes and options/answers that may be used for authoring an interactive algorithm.

NODE PATH	RIGHT OPTION	WRONG OPTION, TO REPAIR	WRONG OPTION, FATAL	ALTERNATIVE OPTION	TERMINATING NODE
1A	•				
1B		•			
1C			•		
2A	•				
2B		•			
2C		•			
3A	•				
3B			•		
3C			•		
3D				•	
4A	•				
4B				•	
5A	•				
5B		•			
6A	•				
6B		•			
6C		•			
7					•
8					•
9					•

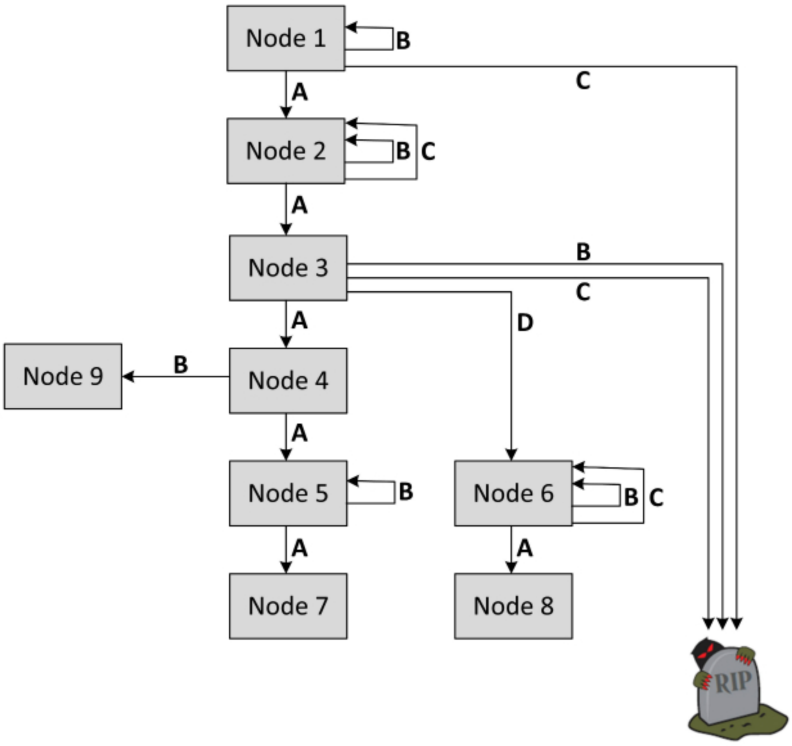


Figure 5. An explained screenshot for 1 node of an algorithm for training clinical reasoning skills in acute coronary syndrome.

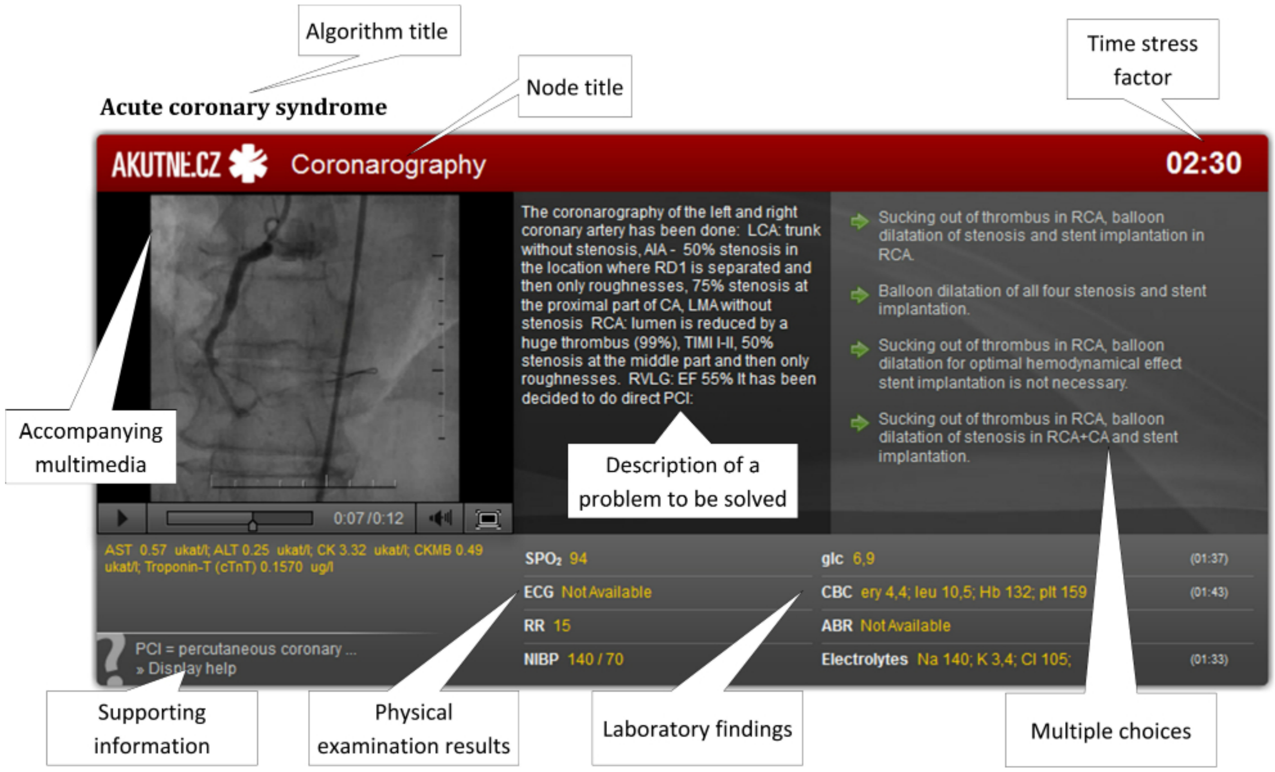


Table 1. Questionnaire for collecting the students' feedback on the interactive algorithms.

#	Question	Answer options
1	State your gender.	Male or female
2	What is your field of study?	General medicine
		Dentistry
		Health care specializations (MSc)
		Health care specializations (BSc)
		Midwifery (BSc)
		Postgraduate doctoral program. another (specify, please)
3	What is your attitude toward the interactive algorithms AKUTNE.CZ?	I do not know what they are
		I know what they are, but I have never used them
		I tried to solve at least 1 interactive algorithm
		I am an author or a coauthor of at least 1 interactive algorithm
4	Have you ever used for your studies a serious game (simulation of real situations for teaching and learning) or-any other interactive algorithm AKUTNE.CZ?	I have not used any at all, not even any interactive algorithm
		No. I have used only the interactive algorithms
		Yes. I have used also...(specify which):
5	The interactive algorithms AKUTNE.CZ are an effective tool for my learning.	5-point Likert scale from strongly disagree to strongly agree
6	The use of the interactive algorithms AKUTNE.CZ improved my knowledge in the field of acute medicine.	5-point Likert scale from strongly disagree to strongly agree
7	The use of the interactive algorithms AKUTNE.CZ represents for me a better way to study than static textbooks.	5-point Likert scale from strongly disagree to strongly agree
8	I like playing the interactive algorithms AKUTNE.CZ not only at home, but also at school under the supervision of teachers, together with consulting possible answers as well as with discussion on all issues related to the topic.	5-point Likert scale from strongly disagree to strongly agree
9	Multimedia accompanying the decision nodes together with the time stressor evokes an authentic atmosphere of clinical reasoning and decision making.	5-point Likert scale from strongly disagree to strongly agree
10	Would you recommend the interactive algorithms AKUTNE.CZ to your friends?	Yes or no

Results

Over 5 years, almost 25 interactive algorithms in the Czech and English languages have been developed and published on the AKUTNE.CZ educational portal to allow the users to test and improve their knowledge and skills in the field of acute medicine. Another 5 algorithms will be finished during 2013. They cover a wide range of acute medicine topics in the following 5 packages:

Basic Life Support and Advanced Life Support

Algorithms cover many basic life support (BLS) and advanced life support (ALS) procedures described in the current European Resuscitation Council guidelines. We developed a BLS for adults algorithm, ALS for bradycardia, BLS for choking children, and a foreign-body airway obstruction in adults algorithm.

Emergency Medicine

Emergency medicine is a very specific type of care in exceptional conditions. We tried to create an ambience of a real car accident in the interactive algorithm. Further topics of emergency medicine are algorithms for water rescue, severe hypothermia in the mountains in winter, out-of-hospital craniocerebral injury, and syncope.

Critical Care Medicine

Critical care medicine (CCM) is the flagship of medicine in general. It is no coincidence that the most demanding and complex algorithms are from this field. The surviving sepsis algorithm is based on the surviving sepsis guidelines of the Society of Critical Care Medicine (SCCM). The acute coronary syndrome algorithm provides a complete decision tree for a patient with acute myocardial stroke. The algorithm for diabetes mellitus deals with sudden loss of consciousness in a diabetic patient.

Anesthesiology

These algorithms cover both interesting acute and propaedeutic situations during anesthesia. We developed an algorithm describing the correct approach to the parturient with postdural puncture headache after epidural labor analgesia. Another acute situation is described in the algorithm for toxic reaction to anesthetic agents. Propaedeutic skills are represented by algorithms introducing the insertion of central venous catheter or the choosing of venous entry routes.

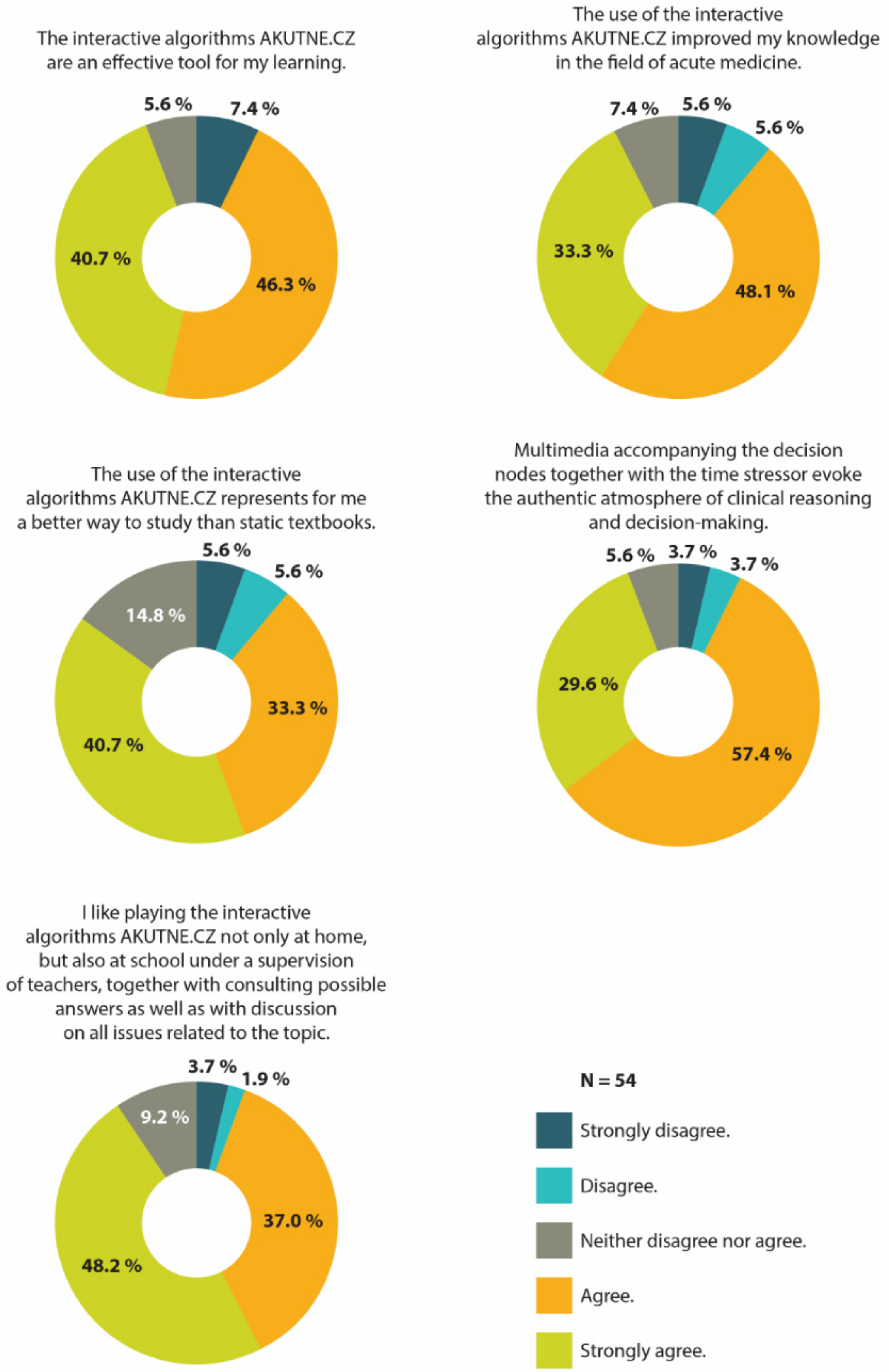
Pain Management

Providing good analgesia for acute and chronic pain is a global issue. We cover these issues with an acute postoperative pain algorithm and by algorithms with correct approach to analgesia in a general practitioner's and a dentist's surgery/clinic.

User's attendance to the interactive algorithms was analyzed with the use of Google Analytics in context of the whole website AKUTNE.CZ within a 1-month period (January 15 to February 14, 2013). In this period, 3342 unique users visited the website (5452 visits in total, 176 visits per day, SD 53.1). All interactive algorithms together had 816 unique users. Of 816 users, 297 (36.4%) accessed the algorithms from Brno and were, therefore, identified as students of the Faculty of Medicine in Brno. Other large groups of visitors were from Prague (99/816, 12.1%) and Bratislava (26/816, 3.2%), both major cities with established medical education facilities. On the other hand, 259 accesses (31.7%) were from places where no faculty of medicine exists. Although we are aware of the limited information value of such analysis (eg, not all visits from Brno are performed at school, or a visitor from a small village could be a student from the Brno faculty of medicine), these results document that the interactive algorithms have been used within the whole MEFANET network and a significant proportion of students use them in places outside of the school (ie, in their homes and during leisure time). The most frequently played algorithms were the diabetes mellitus (94/816 unique users, 11.5%), hypothermia (89 unique users, 10.9%), and surviving sepsis (52 unique users, 6.4%).

In the feedback survey, 62 participants (13.5%) completed the online questionnaire out of the overall 460 asked to participate. Of all respondents, 66.1% were women and 33.9% were men. After filtering out the participants who were not students of any medical or health care program, and those who did not know about the availability of the interactive algorithms AKUTNE.CZ, the resulting responses from 54 participants were analyzed (see Figure 6). The participants were asked whether the interactive algorithms served as an effective tool for their learning. Four responses were negative or very negative (7.4%), 3 responses were neutral (5.6%), and 47 responses were positive or very positive (87.0%). The participants were further asked whether the interactive algorithms improved their knowledge of acute medicine. Six responses were negative or very negative (11.2%), 4 responses were neutral (7.4%), and 44 responses were positive or very positive (81.4%). In all, 40 participants agreed or strongly agreed (74.0%) that the interactive algorithms represented for them a better study method in comparison to static textbooks, whereas 6 participants disagreed or strongly disagreed (11.2%), and a further 8 respondents neither agreed nor disagreed (14.8%). The participants' attitude toward interactive algorithms as a tool for face-to-face teaching and learning was positive or very positive in 46 responses (85.2%), negative or very negative in 3 responses (5.6%), and neutral in 5 responses (9.2%). Most participants agreed or strongly agreed (47/54, 87.0%) that multimedia and the time-stress factor provided an authentic atmosphere for pertinent clinical reasoning, whereas 4 participants disagreed or strongly disagreed (7.4%) with this fact and 3 were unsure (5.6%). All participants (100%) stated that the interactive algorithms were worth recommending to their friends.

Figure 6. Attitudes and interests of students about using the interactive algorithms as part of their medical or health care studies.



Discussion

Principal Findings

High-quality digital education content production has become a matter of prestige at medical schools in the Czech Republic and Slovakia, and the volume of teaching and learning materials available is growing rapidly thanks to the MEFANET project and its ICT platforms, which have been continuously developed and adopted to the needs of the MEFANET community during the past 6 years. Four new extensions, which complement the e-publishing portal platform standardized in MEFANET, are usable independently; however, their complex application in conjunction with the portal platform as a tool for final e-publishing will allow more effective repurposing of the materials created with the use of the extensions, as well as broader integration of the digital education contents among the MEFANET community. Further development aims to encourage the publication of materials for the teaching of clinical reasoning based on the concept of interactive algorithms or virtual patients. Such simulation-based learning objects are aimed to help the student in developing the much-needed confidence to manage acute conditions, to react accurately, and to avoid distraction by secondary issues.

The unique advantage of interactive algorithms AKUTNE.CZ is the possibility to create complex and branching scenarios. Nevertheless, real-life medical emergencies offer little or no extra options; in many cases, there is only 1 correct course of action. Unfortunately, this feature has not been adopted on a wide scale. The reason could be the characteristics of real-time acute medicine situations that are often linear with no space for branching. On our part, we have complied as much as possible with the guidelines of medical societies. Any deviation from the approved procedures may lead to deteriorating outcomes in real clinical situations. This is the reason why we prefer creating simplified and linear algorithms. An algorithm that approaches realistic simulation (nonlinear or open format) could be more attractive for the students, but we believe that to happen at the expense of didacticism. We also prefer topics that are endorsed and processed by the guidelines or recommendations of the European medical societies (ie, European Resuscitation Council, SCCM, European Society of Regional Anaesthesia and Pain Therapy) and/or national medical societies (ie, Czech Society of Anaesthesiology and Intensive Care Medicine, Czech Society of Intensive Care Medicine, Czech Society of Hematology, Czech Society of Cardiology, Czech Gynecological and Obstetrical Society, and Czech Pain Society). The linear scenarios help to maintain a didactic focus of the interactive algorithms. This mechanistic approach may, however, be detrimental to the students' understanding of the underlying physiological processes. In order to overcome this limitation, we prefer to use the interactive algorithms for teaching in the form of moderated PBL sessions. Inspired by several works in the field of advanced physiological simulators with a mathematical background [29-31], we will focus our future developments toward a technology mashup, which would allow to incorporate time-dependent, complex physiological simulation of multiple variables and their response to perturbations into the multimedia part of the interactive algorithms.

We cover a wide range of acute medicine topics through the AKUTNE.CZ algorithms. Of course, there is room for additional themes, for example, the widely publicized case of methanol poisoning in 2012 in the Czech Republic, which led to fatalities. Other topics under consideration include selected amyotrophic lateral sclerosis scenarios and out-of-hospital medical emergencies. Interactive algorithms are also used during obstetric anesthesia and analgesia lessons for the midwives—severe peripartal bleeding, amniotic fluid embolism, and out-of-hospital delivery algorithms. The primary aim is to achieve a situation whereby each acute medicine teaching unit has at least 1 interactive algorithm for PBL.

Although the algorithms were tailored to the teaching and learning of acute medicine issues, it is possible to use them for education in other medical and health care disciplines as well. The selection of the parameters from physical examination results and laboratory tests can be changed easily and, thus, adopting the tool for use elsewhere. In comparison with other examples of simulation-based learning objects, such as virtual patients [32], we have a different approach to handling the selected physical examination results and laboratory findings. We follow real-world scenarios and provide the possibility to record these parameters as they are recorded during management of real acute patients too. Each measurement is linked to an increase of the time-stress factor. Thus, students not only learn about dynamics of these characteristics, but also about the unpleasant price in terms of time spent for unnecessary measurements.

A major problem with any medical issue is topicality. AKUTNE.CZ algorithms overcome such problems by ensuring regular updates through the combined efforts of medical students and the authors, in addition to holding regular meetings on time-scheduled updated topics. The algorithms truly reflect on the current medical recommendations and guidelines of the medical societies.

In general, our survey points to a fairly strong preference for the AKUTNE.CZ interactive algorithms by the students as part of their medical or health care studies, although it is notable that the participants were only just aware of the interactive algorithms—a small proportion (9%) reported using other serious games or simulation-based learning objects for their studies. Nevertheless, positive attitudes toward the interactive algorithms outnumbered negative responses. Confirming our expectations, one of the strongest positive answers concerned the participants' desire to use the interactive algorithms not only for their self-studies during leisure time, but also in face-to-face teaching and learning. Based on our several preliminary attempts at implementing the PBL principles into our teaching, we are fully confident about PBL-like sessions conducted on the node-based scenarios of selected interactive algorithms as the appropriate way to fulfill that wish. The medical and health care institutions in the Czech Republic and Slovakia involved in MEFANET are currently, however, in the very preliminary phases of implementing PBL into their curriculum. Hopefully, the use of interactive algorithms in the process of PBL implementation shall pave the way toward increased attractiveness of our teaching, as well as deeper interest on the part of the students not only in acute medicine issues.

Limitations

A limitation of the study is that we did not collect data to observe effects of the use of algorithms on expected improvements of participants' knowledge or on their reactions in real situations. We can only guess about the positive impacts of the interactive algorithms from the fact that most of the student–authors did not have any difficulties launching their professional careers in acute medicine. Another improvement indicator can be inferred from the repeated successes of student–authors and student–players in international competitions of medical rescue teams.

Conclusions

The methodological aspects of our interactive algorithms for incorporation in the learning and teaching of acute medicine were presented. These interactive algorithms comprise the main part of the educational content of the AKUTNE.CZ portal and recently became the basis for a new extension for MEFANET, the education network of all medical faculties in the Czech Republic and Slovakia.

There are 25 algorithms in the Czech/Slovak and English languages, published online and covering a wide range of topics in acute medicine. The peer-reviewed algorithms were used for conducting PBL-like sessions in general medicine (first aid, anesthesiology and pain management, emergency medicine) as well as in nursing (emergency medicine for midwives, obstetric analgesia and anesthesia for midwives).

We investigated the students' perception of our interactive algorithms as an adjuvant to their medical and health care studies, especially in relation to clinical reasoning. The feedback from the survey among the AKUTNE.CZ users suggests that the students identify the interactive algorithms as an effective learning tool, serving to enhance their knowledge in the field of acute medicine. In addition, they expressed their keen desire to apply them not only in their leisure time, but also during face-to-face contact with their teachers at school or during clinical practice in the university hospital.

The AKUTNE.CZ interactive algorithms, as a software platform, are open to academic use worldwide. The already created and peer-reviewed algorithms, as simulation-based learning objects, can be included easily into any education website (subject to approval of the authors).

Acknowledgments

The grant project MEFANET clinical reasoning reg no: CZ.1.07/2.2.00/28.0038 is supported by the European Social Fund and the state budget of the Czech Republic.

Conflicts of Interest

Daniel Schwarz is the principal investigator of the MEFANET clinical reasoning grant project, which funded development of the MEFANET e-publishing system as well as the interactive algorithms for teaching and learning acute medicine. The grant project also paid Daniel Schwarz, Petr Štourač, Martin Komenda, and Hana Harazim a small portion of their salaries at Masaryk University.

References

1. Holzinger A, Kickmeier-Rust MD, Wassertheurer S, Hessinger M. Learning performance with interactive simulations in medical education: Lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. Computers & Education 2009 Feb;52(2):292-301 [FREE Full text] [doi: 10.1016/j.compedu.2008.08.008]
2. Wong RW, Lochnan HA. A web-based simulation of a longitudinal clinic used in a 4-week ambulatory rotation: a cohort study. BMC Med Educ 2009;9:8 [FREE Full text] [doi: 10.1186/1472-6920-9-8] [Medline: 19187554]
3. Subramanian A, Timberlake M, Mittakanti H, Lara M, Brandt ML. Novel educational approach for medical students: improved retention rates using interactive medical software compared with traditional lecture-based format. J Surg Educ 2012;69(4):449-452. [doi: 10.1016/j.jsurg.2012.05.013] [Medline: 22677580]
4. Maloney S, Haas R, Keating JL, Molloy E, Jolly B, Sims J, et al. Breakeven, cost benefit, cost effectiveness, and willingness to pay for web-based versus face-to-face education delivery for health professionals. J Med Internet Res 2012;14(2):e47 [FREE Full text] [doi: 10.2196/jmir.2040] [Medline: 22469659]
5. Brett-Fleegler MB, Vinci RJ, Weiner DL, Harris SK, Shih MC, Kleinman ME. A simulator-based tool that assesses pediatric resident resuscitation competency. Pediatrics 2008 Mar;121(3):e597-e603 [FREE Full text] [doi: 10.1542/peds.2005-1259] [Medline: 18283069]
6. Pothiawala S, Lateef F. Hong Kong Journal of Emergency Medicine. 2012 Jan. Simulation training in emergency medicine (STEM): an integral component of residency curriculum URL: http://www.hkcem.com/html/publications/Journal/2012-1%20Jan/p41-45.pdf [accessed 2013-02-27] [WebCite Cache ID 6Ek9Md4ig]
7. Davids MR, Chikte UM, Halperin ML. Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. Adv Physiol Educ 2011 Sep;35(3):295-306 [FREE Full text] [doi: 10.1152/advan.00127.2010] [Medline: 21908840]
8. Beux PL, Fieschi M. Virtual biomedical universities and e-learning. Int J Med Inform 2007;76(5-6):331-335. [doi: 10.1016/S1386-5056(07)00060-3] [Medline: 17407747]

9. Bamidis P, Kaldoudi E, Pattichis C. mEducator: A best practice network for repurposing and sharing medical educational multi-type content. In: Camarinha-Matos L, Paraskakis P, Afsarmanesh H, editors. Leveraging Knowledge for Innovation in Collaborative Networks. Berlin Heidelberg: Springer; 2009:769-776.
10. Kaldoudi E, Dovrolis N, Konstantinidis S, Bamidis P. Social networking for learning object repurposing in medical education. The Journal on Information Technology in Healthcare 2009;7(4):233-243 [FREE Full text] [WebCite Cache]
11. Kaldoudi E, Dovrolis N, Konstantinidis ST, Bamidis PD. Depicting educational content repurposing context and inheritance. IEEE Trans Inf Technol Biomed 2011 Jan;15(1):164-170. [doi: 10.1109/TITB.2010.2092442] [Medline: 21134819]
12. Schwarz D, Dušek L. The MEFANET Project. URL: http://www.mefanet.cz/index-en.php [accessed 2013-02-27] [WebCite Cache ID 6EkA1pn4M]
13. Komenda M, Schwarz D, Feberová J, Stípek S, Mihál V, Dušek L. Medical faculties educational network: multidimensional quality assessment. Comput Methods Programs Biomed 2012 Dec;108(3):900-909. [doi: 10.1016/j.cmpb.2012.05.002] [Medline: 22640818]
14. Kleinpell R, Ely EW, Williams G, Liolios A, Ward N, Tisherman SA. Web-based resources for critical care education. Crit Care Med 2011 Mar;39(3):541-553. [doi: 10.1097/CCM.0b013e318206b5b5] [Medline: 21169819]
15. Creutzfeldt J, Hedman L, Heinrichs L, Youngblood P, Felländer-Tsai L. Cardiopulmonary resuscitation training in high school using avatars in virtual worlds: an international feasibility study. J Med Internet Res 2013;15(1):e9 [FREE Full text] [doi: 10.2196/jmir.1715] [Medline: 23318253]
16. Nathoo AN, Goldhoff P, Quattrochi JJ. Evaluation of an Interactive Case-based Online Network (ICON) in a problem based learning environment. Adv Health Sci Educ Theory Pract 2005 Aug;10(3):215-230. [doi: 10.1007/s10459-005-7851-3] [Medline: 16193402]
17. Servais EL, Lamorte WW, Agarwal S, Moschetti W, Mallipattu SK, Moulton SL. Teaching surgical decision-making: an interactive, web-based approach. J Surg Res 2006 Jul;134(1):102-106. [doi: 10.1016/j.jss.2005.11.583] [Medline: 16488433]
18. Roubidoux MA, Chapman CM, Piontek ME. Development and evaluation of an interactive Web-based breast imaging game for medical students. Acad Radiol 2002 Oct;9(10):1169-1178. [Medline: 12385511]
19. Dikshit A, Wu D, Wu C, Zhao W. An online interactive simulation system for medical imaging education. Comput Med Imaging Graph 2005 Sep;29(6):395-404. [doi: 10.1016/j.compmedimag.2005.02.001] [Medline: 15996851]
20. Holzinger A, Emberger W, Wassertheurer S, Neal L. Design, development and evaluation of online interactive simulation software for learning human genetics. Elektrotech. Inftech 2008 May;125(5):190-196. [doi: 10.1007/s00502-008-0537-9]
21. Štourač P, Křikava I, Štoudek R, Harazim H, Smékalová O, Kosinová M. AKUTNE.CZ educational portal. URL: http://www.akutne.cz/index-en.php [accessed 2013-02-27] [WebCite Cache ID 6EkA9tNl4]
22. MEFANET Central Gateway. URL: http://portal.mefanet.cz/index-en.php [accessed 2013-02-27] [WebCite Cache ID 6EkAEOHoF]
23. Czech academic identity federation eduID.cz. URL: http://www.eduid.cz/wiki/en/eduid/index [accessed 2013-02-27] [WebCite Cache ID 6EkAIeBd5]
24. MEFANET SANDBOX. URL: http://sandbox.mefanet.cz/index-en.php [accessed 2013-02-27] [WebCite Cache ID 6EkAM8VLQ]
25. WikiLectures. URL: http://www.wikilectures.eu/index.php/Main_Page [accessed 2013-02-27] [WebCite Cache ID 6EkAPMgtf]
26. Moodle-MEFANET e-learning courses. URL: https://moodle.mefanet.cz/ [accessed 2013-02-27] [WebCite Cache ID 6EkATjZqo]
27. MEFANET GAMES. URL: http://games.mefanet.cz/index-en.php [accessed 2013-02-27] [WebCite Cache ID 6EkAfPUZp]
28. SurveyMonkey. URL: http://www.surveymonkey.com/ [accessed 2013-02-27] [WebCite Cache ID 6EkAiwxIh]
29. Kofranek J, Matousek S, Rusz J, Stodulka P, Privitzer P, Matejak M, et al. The Atlas of Physiology and Pathophysiology: Web-based multimedia enabled interactive simulations. Comput Methods Programs Biomed 2011 Nov;104(2):143-153. [doi: 10.1016/j.cmpb.2010.12.007] [Medline: 21232813]
30. Thomas SR, Baconnier P, Fontecave J, Françoise JP, Guillaud F, Hannaert P, et al. SAPHIR: a physiome core model of body fluid homeostasis and blood pressure regulation. Philos Trans A Math Phys Eng Sci 2008 Sep 13;366(1878):3175-3197. [doi: 10.1098/rsta.2008.0079] [Medline: 18565814]
31. Iliescu R, Lohmeier TE. Lowering of blood pressure during chronic suppression of central sympathetic outflow: insight from computer simulations. Clin Exp Pharmacol Physiol 2010 Feb;37(2):e24-e33. [doi: 10.1111/j.1440-1681.2009.05291.x] [Medline: 19769610]
32. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Med Educ 2009 Apr;43(4):303-311. [doi: 10.1111/j.1365-2923.2008.03286.x] [Medline: 19335571]

Abbreviations

ALS: advanced life support
BLS: basic life support
CCM: critical care medicine

ICT: information and communication technology
ISSN: International Standard Serial Number
MEFANET: Medical Faculties Network
PBL: problem-based learning
SAML: Security Assertion Markup Language
SCCM: Society of Critical Care Medicine
XML: extensible markup language

Edited by P Bamidis, D Giordano, C Pattichis, N Zary; submitted 01.03.13; peer-reviewed by L Woodham, R Iliescu, J Majernik, I Provaznik; comments to author 19.03.13; accepted 05.04.13; published 08.07.13

Please cite as:
Schwarz D, Štourač P, Komenda M, Harazim H, Kosinová M, Gregor J, Hůlek R, Smékalová O, Křikava I, Štoudek R, Dušek L
Interactive Algorithms for Teaching and Learning Acute Medicine in the Network of Medical Faculties MEFANET
J Med Internet Res 2013;15(7):e135
URL: http://www.jmir.org/2013/7/e135/
doi:10.2196/jmir.2590
PMID:23835586

PLOS | ONE

RESEARCH ARTICLE

Curriculum Mapping with Academic Analytics in Medical and Healthcare Education

Martin Komenda[1,2]*, Martin Víta[2], Christos Vaitsis[3], Daniel Schwarz[1], Andrea Pokorná[1], Nabil Zary[3], Ladislav Dušek[1]

1 Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **2** Faculty of Informatics, Masaryk University, Brno, Czech Republic, **3** Department of Learning, Informatics Management and Ethics, Karolinska Institutet, Stockholm, Sweden

* komenda@iba.muni.cz

OPEN ACCESS

Citation: Komenda M, Víta M, Vaitsis C, Schwarz D, Pokorná A, Zary N, et al. (2015) Curriculum Mapping with Academic Analytics in Medical and Healthcare Education. PLoS ONE 10(12): e0143748. doi:10.1371/journal.pone.0143748

Editor: Kent E. Vrana, Penn State College of Medicine, UNITED STATES

Received: August 10, 2015

Accepted: November 9, 2015

Published: December 1, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors were supported from the following grant projects: (i) OPTIMED - optimized medical education: horizontal and vertical connections, innovations and efficiency in practice (http://opti.med.muni.cz/en) – reg. no.: CZ.1.07/2.2.00/28.0042, which is funded by the European Social Fund and the National (state) budget of the Czech Republic; (ii) CROESUS – clinical reasoning skills enhancements with the use of simulations and algorithms (http://www.croesus.eu) – reg. no.: 2014-1-CZ01-KA203-002002, which is funded by the

Abstract

Background

No universal solution, based on an approved pedagogical approach, exists to parametrically describe, effectively manage, and clearly visualize a higher education institution's curriculum, including tools for unveiling relationships inside curricular datasets.

Objective

We aim to solve the issue of medical curriculum mapping to improve understanding of the complex structure and content of medical education programs. Our effort is based on the long-term development and implementation of an original web-based platform, which supports an outcomes-based approach to medical and healthcare education and is suitable for repeated updates and adoption to curriculum innovations.

Methods

We adopted data exploration and visualization approaches in the context of medical curriculum innovations in higher education institutions domain. We have developed a robust platform, covering detailed formal metadata specifications down to the level of learning units, interconnections, and learning outcomes, in accordance with Bloom's taxonomy and direct links to a particular biomedical nomenclature. Furthermore, we used selected modeling techniques and data mining methods to generate academic analytics reports from medical curriculum mapping datasets.

Results

We present a solution that allows users to effectively optimize a curriculum structure that is described with appropriate metadata, such as course attributes, learning units and outcomes, a standardized vocabulary nomenclature, and a tree structure of essential terms. We present a case study implementation that includes effective support for curriculum reengineering efforts of academics through a comprehensive overview of the General Medicine study program. Moreover, we introduce deep content analysis of a dataset that was

European Commission ERASMUS+ program; and (iii) EVAMED – evaluation of medical education (http://eva.med.muni.cz) – reg. no.: MUNI/A/1512/2014, which is funded by the Specific University Research Grant, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2015. MK is the principal investigator of the OPTIMED medical innovation grant project, which funded development of the OPTIMED curriculum management system as well as the EVAMED research study project focusing on the systematic assessment of described medical curriculum. DS is the principal investigator of the CROESUS clinical reasoning grant project, which supported implementation of the ICT platform for scenario-based learning including methodological preparation and training of virtual patients in clinical medicine. The grant projects also paid MK, DS, and AP a small portion of their salaries at Masaryk University. The authors disclose any financial and personal relationships with other people or organizations that could inappropriately influence their work.

Competing Interests: The authors have declared that no competing interests exist.

captured with the use of the curriculum mapping platform; this may assist in detecting any potentially problematic areas, and hence it may help to construct a comprehensive overview for the subsequent global in-depth medical curriculum inspection.

Conclusions

We have proposed, developed, and implemented an original framework for medical and healthcare curriculum innovations and harmonization, including: planning model, mapping model, and selected academic analytics extracted with the use of data mining.

Introduction

A correctly compiled and balanced curriculum that covers all theoretical and clinical fields of medicine is an essential prerequisite for medical doctor education [1]. The keyword "innovation" is perceived as an activity that enables higher education institutions (HEIs) to make their curricula more up-to-date, while maintaining transparency and a high degree of structure. The innovations, if performed in teaching domains formalized with the use of an unambiguous parametric description, and with entities adopted from the outcomes-based concept of education (LUs: learning units, LOs: learning outcomes), will enhance the transparency and continuity of the environment in which both teachers and students work daily [2,3]. In this manner, an innovation is more likely to have positive and practically applicable effects on higher medical education, if it is based on a balanced and well-structured medical curriculum.

Such efforts have recently demonstrated how a medical curriculum can be used as a tool to drive innovation for the sake of understanding and perceiving the complexity of medical education. This ability has proved that it is possible to verify the existence of the constructive alignment within the medical curriculum [4], and that performing a high-level gap analysis of the relationships between major curriculum components such as LOs, teaching activities, educational material, and assessment methods may eventually support critical decisions concerning quality improvement in the medical education. Different approaches like visual analytics [5–8], web-based learning objective databases [9], simple visualizations of academic analytics applied to a medical curriculum [10], and the results of our previous work [1,11–14], have all been used for this purpose. The aim of those studies was to improve the traditional method for aligning the medical curriculum and to strengthen the complicated decision making process, which is usually performed by medical education stakeholders such as teachers, tutors, curriculum designers, or institutional managers. That goal has previously been successfully if only partially addressed, as these studies investigated possible novel solutions at an exploratory level. A solution concerning the medical curriculum as a whole, which could enable its effective encapsulation, description, and management together with ability to demonstrate its most important aspects, is still missing.

1.1 Background of curriculum mapping

The advent of information technology in various educational domains has led to large volumes of data that is stored in various formats, including students' data, teachers' data, alumni data, resource data, curriculum data, etc. A multidisciplinary area called Knowledge Discovery in Databases (KDD) covers several methodologies for extracting useful information from data, including database design, statistics, pattern recognition, machine learning, and data visualization. A set of proper KDD methods is required to uncover knowledge from these large data repositories, in order to improve understanding and decision making [15]. By adopting these proven methodologies for extracting useful information from the large databases of curriculum management systems, we are able to create visual representations of the curricula that are based on real time information; this is one way to increase collaboration and collegiality in HEIs. This technique is called curriculum mapping. It introduces two main objectives: (i) to make the curriculum more transparent to all stakeholders; and (ii) to demonstrate the links and relationships between the various components of the curriculum. In general, curriculum maps can identify whether the intended material is actually being taught and what students actually learn, and demonstrate the links among the different components of the curriculum. The key to a really effective integrated curriculum is to get teachers to exchange information about what is being taught and to coordinate this so that it reflects the overall goals of the institution [16–19]. Curriculum mapping is about spatially representing the different components of the curriculum so that the whole picture and the relationships and connections between the parts of the map are easily seen. There are several case studies that have been successfully implemented in practice [5,20–24]. One of the main objectives of this paper is to introduce the use of mapping for the effective evaluation of medical curriculum, and to provide automatic tasks that could be used to build well-balanced courses that are both theoretically- focused and clinically-based.

Methods

2.1 Planning model for curriculum innovations

Our research is concentrated on proposing a way to channel clear communication between stakeholders (i.e., curriculum designers, supervisors, guarantors, faculty management, and students). On the grounds of the long-term and iterative form of curriculum innovations, we have already designed and implemented an in-house curriculum management system [25], which offers a wide range of web-based tools suitable to implement the innovations efficiently. The curriculum planning model (see Fig 1) that is supported by the curriculum management system helps academics reengineer their curriculum activities. The planning process is delineated in three major stages and is followed by the final evaluation process. (i) Setting up the medical curriculum structure. The study field is divided into individual sections, medical disciplines, courses, and LUs, including the responsible supervisors and guarantors. (ii) Defining the descriptive attributes of the LUs, linked MeSH vocabulary keywords and essential terms, and associated Los, according to Bloom's taxonomy. The outcomes typically consist of a noun or noun phrase (i.e., the subject matter content) and a verb or verb phrase (i.e. the cognitive process). In this case, each LO defines what students are expected to know, understand, and/or be able to demonstrate at the end of the learning period, typically as a graduate. This concept has been already applied by a number of academic institutions, especially in medical education [26,27]. (iii) Vertical harmonizing, i.e., LU optimization and further discussion within individual sections, under the supervision of responsible guarantors. (iv) Horizontal evaluating, i.e., follow-up discussions across all sections under the management of supervisors, including an in-depth inspection in collaboration with an established expert committee, which can logically influence the whole structure of the defined curricula. The model represents an original methodology for the creation of a new structured set of LOs designed to assess and adjust to real education.

2.2 Curriculum mapping model

Following the formal description of the medical curriculum [28], we decided to use selected KDD methods to uncover novel and potentially useful information from the medical curriculum

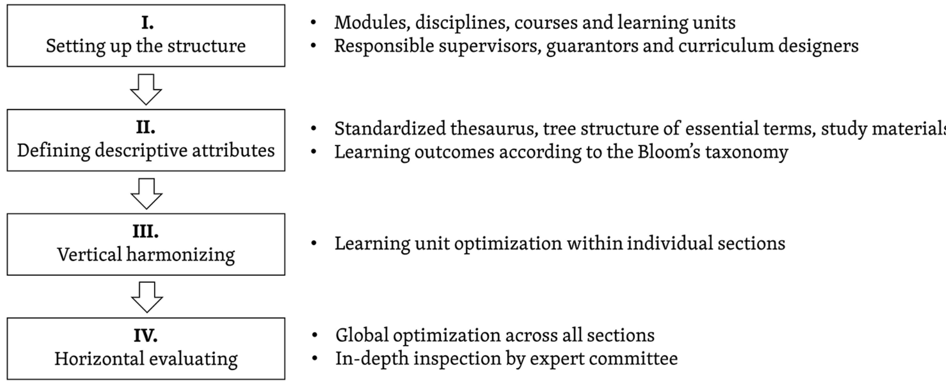


Fig 1. The planning model introducing innovations into a curriculum.

doi:10.1371/journal.pone.0143748.g001

data [29]. To increase the success and accuracy of solving individual KDD tasks, the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology was chosen. It defines a non-rigid sequence of six phases, which allow researchers to build and implement various data mining models to be used in a real environment (see Fig 2).

From an informatics perspective, pre-processing and analysis of available data stored in a curriculum management system is challenging. We have already prepared the normalized database schema based on the formal description of curriculum, which enables easy extraction of information. We have identified many KDD tasks in a form advanced analysis of medical curriculum, which can significantly help curriculum designers with the construction and maintenance of a well-balanced curriculum. These are selected questions that the application of KDD algorithms may help to answer:

- Which educational domains (medical disciplines) do not conform to an expected form within the curriculum data?
- Can we identify the most outlying areas and undesirable overlaps across the medical curriculum?
- Which communities/clusters of medical disciplines are we able to identify in the curriculum?
- Which courses belong to the most important parts of the curriculum (with respect to "content distance" to other courses)?

The proposed set of steps, in accordance with the six-stage-sequence of the CRISP-DM process, facilitates the search for potentially problematic areas and the construction of comprehensive reports for subsequent global, in-depth inspection.

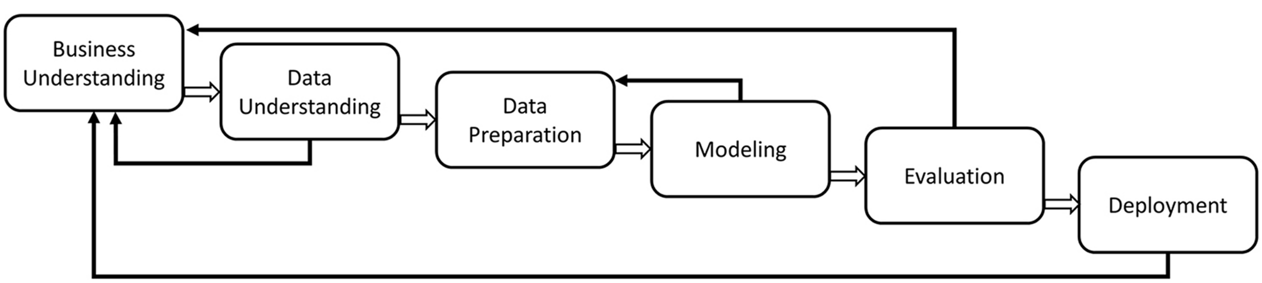


Fig 2. The six phases of the CRISP-DM reference model [30].

doi:10.1371/journal.pone.0143748.g002

2.2.1 Business understanding. This initial phase is focused on understanding the objectives and problem definition in terms of medical curriculum innovation. The goal is to detect outlying and overlapping areas in the General Medicine study program, using data/text mining techniques. The results of the data mining might then be the subject of an evaluation provided by senior curriculum experts. The mentioned goal can be reformulated in terms of data mining issues, such that the task is to investigate the properties of a graph obtained from a dissimilarity matrix over a collection of documents representing individual parts of the curriculum. Our approach is based on Social Network Analysis (SNA), and it was chosen since it can naturally deal with graph-like/network-like data and encompasses centrality notions. First of all, we generate the entire similarity graph based on the similarities of individual medical disciplines as represented by their textual descriptions; specifically, the cosine distance is used. Next, we compute various centrality measures to model the interconnection and remoteness of the nodes (i.e., medical disciplines, in our particular case) in the graph. We are interested in uncovering densely connected subgraphs called communities, as these might represent the most crucial and important parts of a curriculum, in terms of its global harmonization. Based on the community detection, we are able to explore novel and useful information about the structure of the General Medicine study program. We can identify the nodes with extreme (i.e., the highest/the lowest) values of selected centrality measures, which may indicate that these medical disciplines belong to an essential part of the curriculum, or, conversely, it may cause undesirable preference of mentioned disciplines that in fact cannot be identified by manual human inspection. A similar approach has been proposed by Trigo and Brazdil on mining affinity relationships within groups of researchers [31].

2.2.2 Data understanding. We have already built the database for collecting medical curriculum data in a parametric format. It covers all elements pertaining to global curriculum harmonization, including detailed metadata specification down to the level of LUs, linkages to Los, and standardized biomedical vocabulary. The organization of the metadata and its linkages is provided in the curriculum model, which can be implemented without any restrictions within any relational database technology. In order to effectively process curricular metadata stored in our system, we used aggregation and filtering to get the final dataset from the primary data, followed by application of particular KDD methods for exploring new knowledge (see Fig 3).

2.2.3 Data preparation. This step represents an important preprocessing step, which then allows the use of any of the data mining algorithms. All the stated content was created in the Czech and English languages. For the purpose of in-depth analysis of the present pilot, the English version was used, due to easier preprocessing. We identified a set of information-rich attributes, in collaboration with curriculum designers and guarantors, which are suitable for exploring the medical curricula. In our case, for each discipline a single plaintext file was generated, so that it contained merged contents of several attributes (see Table 1) describing LO and its affiliation to curriculum (i.e. link to LU, course, medical discipline, and section). For the purpose of analytical pilots, we processed only the attributes that were related to the particular LOs (i.e., sentences with action verbs, according to Bloom taxonomy) and to all of the descriptive indexes (i.e., grouped outcome, primary index, secondary index). For illustration, the input dataset could have been quantified by approximately 636 normative pages (containing 1,800 characters of text) or with 118,798 words.

2.2.4 Modeling. This phase uses modeling techniques (i.e., analytic methods, particularly data mining methods) on the dataset prepared in the previous phase. The entire text preprocessing was implemented with the use of the R software and its standard packages tm (a widely used package for text mining that allows us to deal with collections of textual data in a convenient way) and lsa (a standard package for latent semantic analysis). The collection of plaintext

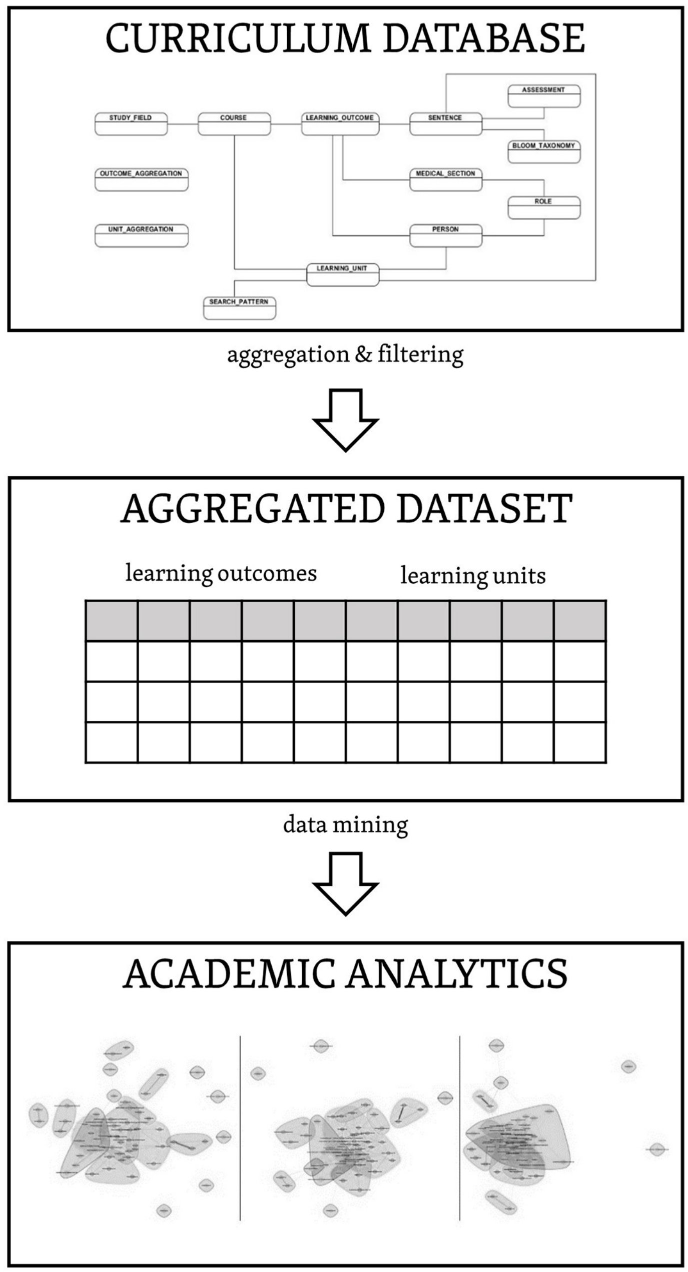


Fig 3. Phases of the CRISP-DM reference model.

doi:10.1371/journal.pone.0143748.g003

files that represent individual disciplines was loaded as a corpus (from tm package) in order to prepare their bag-of-word representation. A usual sequence of pre-processing issues was applied for each document in the corpus after tokenization: transformation to lowercase, stemming (using the Snowball system), punctuation removal, numbers removal, stop words removal, and whitespace stripping. The text preprocessing step was followed by the computation of two matrices. First, a document-term-matrix (DTM) based on *tf-idf* weighting [32] was generated. Consequently, a dissimilarity matrix was computed on the base of the cosine similarity over the vectors in the DTM. Values were rounded to two decimal digits, and values

Table 1. Attributes of final dataset describing the medical curriculum.

Attribute	Sample
LO	
Group outcome	Hypothalamic-pituitary portal system, liberins, statins
Primary index	Endocrinology
Secondary index	Hypothalamus
Sentence	Student is able to assign functional characteristics of different hormones to the structure of the hypothalamic-pituitary system.
Assessment form	Final exam—oral
LU	
Name	Endocrine system
Courses	Physiology I—practice, Physiology I—lecture
Medical discipline	Physiology
Section	Theoretical sciences
Importance	The aim of this teaching unit is to introduce students to the basics of the physiological functions of the endocrine system. The students will be able to compare basic types of intercellular communication (endocrine, paracrine, autocrine)….
Description	The purpose of this teaching unit is to introduce students to the basics of endocrine control of the body. First, it broadly defines the general principles of regulation and the role of the chemical structure of signaling molecules in the pathway of effector response activation…
MeSH keywords	Insulin, Endocrine gland
Significant terms	Glandula suprarenalis, Facies, Hilus, Pancreas, …

doi:10.1371/journal.pone.0143748.t001

lower than a certain threshold were replaced by zeros, since extremely low similarities were considered irrelevant. This dissimilarity matrix naturally corresponds to an undirected graph with weighted edges. The dissimilarity matrix can be viewed as an adjacency matrix of this graph (where the nodes are the disciplines, the edges represent similarities between pairs of documents, and the weight of an edge corresponds with the similarity value); thus, we are able to use various SNA techniques to work with graph representation of data.

In our case study, we have concentrated on the problem of community finding and estimating the importance of each node by various centrality measures, which is an easy way to explore the real composition and intersections of individual disciplines in the curriculum. For community finding, the WalkTrap algorithm [33] was chosen, due to its low computing complexity and the fact that it is implemented within the igraph package. The algorithm is based on short random walks in the graph. Its main premise is that short random walks tend to stay in the same community [34]. Here, we experimented with different values of WalkTrap parameters, mainly with the length of the random walk. After uncovering the communities, the following three centrality measures were computed: closeness centrality, betweenness centrality, and Eigenvector centrality. The closeness centrality of a node is defined by the inverse of the average length of the shortest paths to all of the remaining nodes in the given graph. The (vertex/node) betweenness centrality is, in the simplest case, defined as the number of shortest paths that go through a given node. Eigenvector centrality is one way of computing the approximate importance of a given node. The idea behind this measure is that the centrality of each node is the sum of the centrality values of its neighboring nodes. More precisely, the Eigenvector centrality values correspond to the values of the first eigenvector of the adjacency matrix. In all cases, we will use versions of these measures for graphs with weighted edges, and all values are normalized. The short description of these measures is given below, and more details can be found in [35,36].

Closeness centrality: The closeness centrality of a node v in a graph G is defined by the inverse of the sum of the lengths of the shortest paths to/from all the other nodes in the graph G:

$$c(v) = \frac{1}{\sum_{i \in V(G),\, i \neq v} d(i, v)}, \tag{1}$$

Where $d(i, v)$ is the length of the shortest path from node i to node v. If there is no path between a pair of nodes, then the total number of nodes in the graph is used instead of the path length. Through the mentioned calculation, we obtain the so-called raw closeness of the node. To obtain the normalized closeness of the node v, we multiply the raw closeness by $(n - 1)$, where $|V(G)| = n$, and we use this normalized version in subsequent calculations. Nodes (disciplines) with low values of closeness are those disciplines whose content is distant from other ones; thus, roughly said, they are independent of the others.

Betweenness centrality: In the simplest case (i.e., without edge weighting), the raw betweenness centrality of a node v corresponds with the number of shortest paths from all nodes to all others that go through the node in question:

$$b(v) = \sum_{i,j,v \in V(G), i \neq j, i \neq v, j \neq v} \frac{g_{ivj}}{g_{ij}}, \tag{2}$$

Where g_{ij} is the total number of shortest paths going from node i to j, and g_{ivj} is the total number of all the shortest paths from node i to node j going through v. To get the normalized betweenness $b_n(v)$ of the node v, we calculate $b_n(v) = \frac{2b(v)}{(n-1)(n-2)}$, where $|V(G)| = n$. This definition can be extended for weighted networks. The nodes (disciplines) with high betweenness centrality are the ones that are the best for joining the students' knowledge from different collections of disciplines (i.e., through intersubject transfer of knowledge).

Eigenvector centrality: Eigenvector centrality is one of the methods of computing approximate importance, and is a measure of the influence of a given node. The idea behind this measure is that the centrality of each node is the sum of the centrality values of its neighbor nodes. More preciously, the eigenvector centrality values correspond to the values of the first eigenvector of the adjacency matrix. The eigenvector centrality, in our case, models identification of important disciplines in the curriculum. All three of these measures were implemented within the igraph package.

2.2.5 Evaluation. In this phase, a checking procedure was performed in order to assess whether the approach we used was appropriate, in terms of its usability in real curriculum management. The obtained results were organized and presented in a clear way for the curriculum designers and the committee experts. From the perspective of medical teachers, who are often authors of LUs and LOs, several clear benefits were identified: (i) transparent overview over the medical education in a particular HEI; (ii) detection of any outlying LUs; and (iii) continuous monitoring of the consistency of a curriculum in the never-ending process of innovation. An example evaluation of the achieved analytical results obtained from senior curriculum experts is given in 3.1.2.

2.2.6 Deployment. Creation of the models is generally not the end of the project. The knowledge that was gained needs to be organized and presented in a way that the academics can use on a long-term basis. It often involves applying live models within a HEI's decision making processes. Depending on the requirements, the deployment phase can be as simple as generating an analytical report or as complex as implementing a repeatable data mining process across the HEI. We have systematically tailored our generic process model, in accordance with a pre-defined context of medical curriculum exploration. In general, we aim to analyze and consolidate the experiences of this single pilot project towards a follow-up mapping issues in form advanced analysis of medical curriculum for future usage in comparable contexts. Basically, a global strategy for mapping consists of five steps: (i) analyze a specific field of study; (ii) remove any details not applicable to given field; (iii) add any details specific to given field; (iv) specialize (or instantiate) generic contents according to concrete characteristics of given field; and (v) possibly rename generic content to provide more explicit meaning in a given context, for the sake of clarity [37].

Results

3.1 Curriculum innovations in practice

The developed curriculum management system and relating planning model have been implemented into the practice in the Optimization of Medical Education (OPTIMED) project by the Faculty of Medicine of Masaryk University. Since 01 April 2014, OPTIMED has been fully open to all medical students and teachers. Depending on the various user roles, academics are able to use individual modules for global curriculum overview and also for managing the LUs and related LOs. Google Analytics monitoring was integrated into the system, thus up-to-date statistics, reports, and analyses based on site traffic and visitor's behavior is available. Figs 4 and 5 demonstrate the latest Google Analytics summary from 01 April 2014 to 01 August 2015, where audience behavior is reported.

Our approach to curriculum innovation offers an appealing way to effectively reform medical education, where the emphasis is on the product—what sort of graduates should be produced—rather than on the educational process itself. The primary effort in this project was a comprehensive curriculum innovation of the General Medicine study program. The innovations are driven towards a smoother continuity between the theoretical and clinical phases of the study, and by need to deliver graduates with 21st century skills. The key point of the project is the use of information and communication tools in order to achieve a horizontally-innovated structure of compulsory and compulsory-optional subjects. The objective is not a radical change in learning or teaching, but rather it is an exploratory mapping of the current state of the general medicine curriculum, with a prospect for innovations that produce more transparent educational environment [28]. The institutional management designated curriculum experts across medical disciplines, which interacted with the study harmonization and streamlining process in different roles (i.e., curriculum designers, guarantors, coordinators, reviewers). They proposed a set of fundamental knowledge and skills known as the Global Minimum Essential Requirements (GMER), in accordance with the proposed curriculum planning model, which is based on the outcomes-based paradigm. A total of 385 teachers have been interacting with the study harmonization and streamlining process for two years. Together, they have produced and evaluated the huge amount of metadata records that define the General Medicine study program, which are categorized to a pre-defined structure of medical curriculum (for illustration, it takes more than 2,500 normative pages of text).

The entire study field (General Medicine) is split into four individual sections (i.e., diagnostic and neurosciences, internal medicine, theoretical sciences, and surgical sciences), and includes details about the responsible supervisors. Each section contains a set of courses, which are divided into particular LUs. This categorization provides easy organization of the metadata about the education. Each LU covers LOs, which represent the basic requirements on the graduate from the selected field. All the curriculum metadata were collected using specialized modules within the developed platform, called LO & LU registers. They serve as online tools for

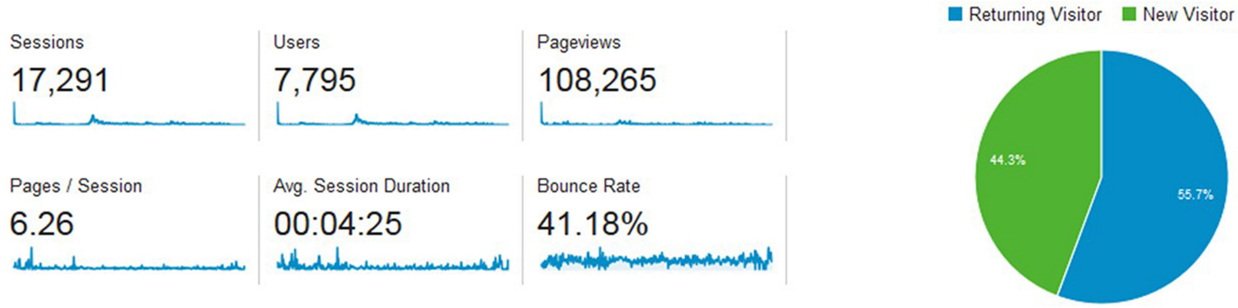


Fig 4. Google Analytics global overview (01 April 2014–01 August 2015).

doi:10.1371/journal.pone.0143748.g004

medical curriculum management and provide detailed metadata specifications, down to the level of LUs and connections to the LOs. Table 2 gives the complete summary in numbers. In accordance with the proposed planning model, the vertical harmonization represents the verification and further discussion within the individual section, under supervision of the responsible guarantor. Finally, the horizontal evaluating brings the process of horizontal harmonization, which consists of follow-up discussions across all sections under the management of supervisors, including in-depth inspection in collaboration with an established expert committee.

3.2 Curriculum mapping in practice

In accordance with CRISP-DM, we used the methods described above for solving the issue of medical curriculum mapping, which will enable medical curriculum designers and managers

Count of Sessions	Sessions	Pageviews
1	7,660	41,145
2	2,479	14,921
3	1,378	9,011
4	883	6,517
5	638	5,007
6	479	3,026
7	354	2,627
8	298	2,710
9-14	951	6,921
15-25	710	4,994
26-50	516	3,645
51-100	330	2,295
101-200	190	1,295
201+	425	4,151

Fig 5. Google Analytics detailed overview of sessions (01 April 2014–01 August 2015).

doi:10.1371/journal.pone.0143748.g005

Table 2. Summary of the descriptive attributes defining the medical curriculum.

Medical curriculum domain	Total number
Sections	4
Medical disciplines	44
Courses	144
LUs	1,347
LOs	6,974
Curriculum experts	385
Students	over 2,000

doi:10.1371/journal.pone.0143748.t002

to better understand the multidimensional structure and complex content of the General Medicine study program. By applying the Walktrap algorithm to the collection of 59 documents (medical disciplines), we obtained a similarity graph, labelled with various colors, that defined a number of communities (Fig 6). The presented graph was computed with the length of the random walk set to $k = 4$; this setting provided the best results in terms of expert evaluation.

To obtain an overall insight into the curriculum, the three centrality measures described in section 2.2.4 were used. Table 3 presents values of the centrality measures in alphabetical order for one particular community, which is shown as the enlarged curriculum map in Fig 7. Extremely low centrality values can indicate an inappropriate description of the discipline (e.g., missing important parts of the description). On the other side, the higher or extremely high values of the proposed measures can identify the core and most important parts of the curriculum.

Discussion

4.1 Principal results

4.1.1 Medical curriculum innovations. The Bergen ministerial conference of the Bologna Process in May 2005 discussed reforms to degree structures, credit transfer, quality assurance, and curricular development, which are transforming the European Higher Education Area. An outcomes-based approach is arguably best viewed as a fundamental building block of the Bologna education reforms, and brings greater transparency to higher education systems [38]. The process of medical curriculum innovation supported by a robust planning model, which is described here, is based on the well-established concept of a standardized definition of learning outcomes. One of its key features is its dynamism, namely the ability to easily manage any domain closely related to the medical curriculum through upgrades, and to absorb and incorporate all changes into the educational process. For teachers and faculty management, such an approach enables them to effectively administer education processes, to clearly see who teaches what and to what extent, to determine whether the lectures of individual teachers overlap thematically, to assess whether the overall teaching schedule is appropriate, or to decide whether a certain reform of the teaching schedule would be convenient and helpful. And for school managers, the presented tools will provide a practical view of the real content teaching. Further, it will also provide clear and comprehensible data about who teaches what and in what context, as well as information on the deficiencies and overlaps present in the curriculum. It is currently used by hundreds of senior teachers, curriculum designers, and professional guarantors, who use it for content definition and further inspection of the medical curriculum.

4.1.2 Medical curriculum mapping. With respect to the achieved results, we decided to evaluate the curriculum in co-operation with senior guarantors (senior experts) of medical education at our university. The checking procedure is a mandatory phase of the CRISP-DM

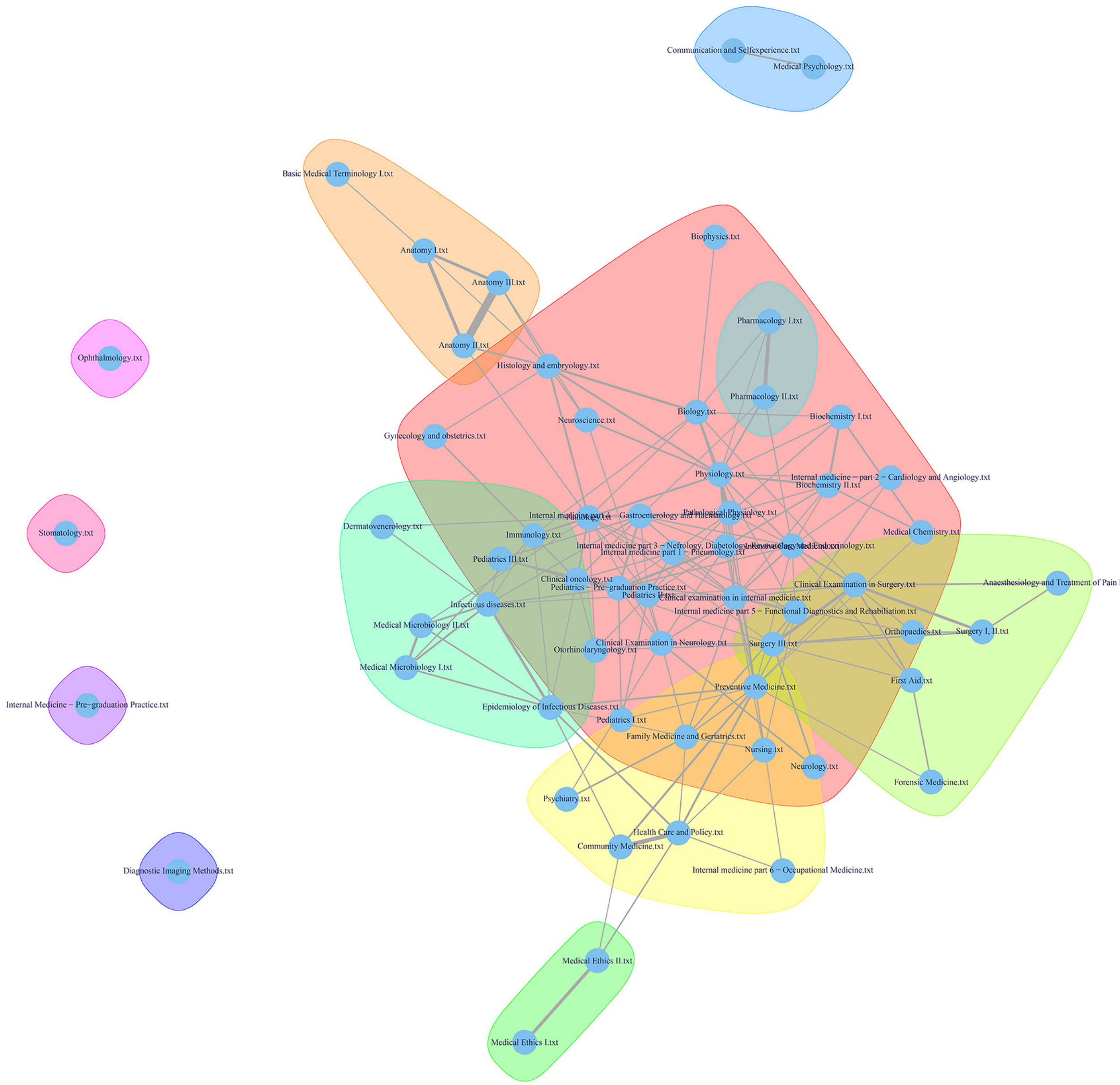


Fig 6. Overview of the medical curriculum, labelled with the communities that were uncovered using the Walktrap algorithm.

doi:10.1371/journal.pone.0143748.g006

methodology, in order to make sure that the given objectives are fulfilled. Expert opinion helps to indicate whether the medical disciplines are balanced or unbalanced, and to identify weak points and shortcomings in terms of inconvenient interdisciplinary relations. In total, three experts were chosen, based on specific criteria: at least 10 years of experience in medical education (educational and scientific background), at least 15 years of professional clinical experience (following the award of the diploma), hold a managerial position in the medical faculty at a scientific level (i.e., at minimum, associate professor level), proven professional experience at a strategic level in medical education, thorough knowledge and proven experience in the principles of developing, applying, and implementing curriculum in medical education, and experience in managing educational projects/program concerning medical pre gradual education. All

Table 3. The values of betweenness, closeness, and Eigenvector centrality for one particular subset within the analyzed medical curriculum.

Medical discipline	Closeness centrality	Betweenness centrality	Eigenvector centrality
Clinical oncology	0.3084	0.1041	0.7991
Dermatovenerology	0.3042	0.0019	0.2070
Epidemiology of Infectious Diseases	0.3072	0.0309	0.6931
Immunology	0.3037	0.0047	0.3995
Infectious diseases	0.3071	0.0232	0.6795
Medical Microbiology I.	0.2975	0.0000	0.2588
Medical Microbiology II.	0.3001	0.0000	0.2400
Otorhinolaryngology	0.3052	0.0016	0.3744
Pediatrics III.	0.3048	0.0037	0.4866

doi:10.1371/journal.pone.0143748.t003

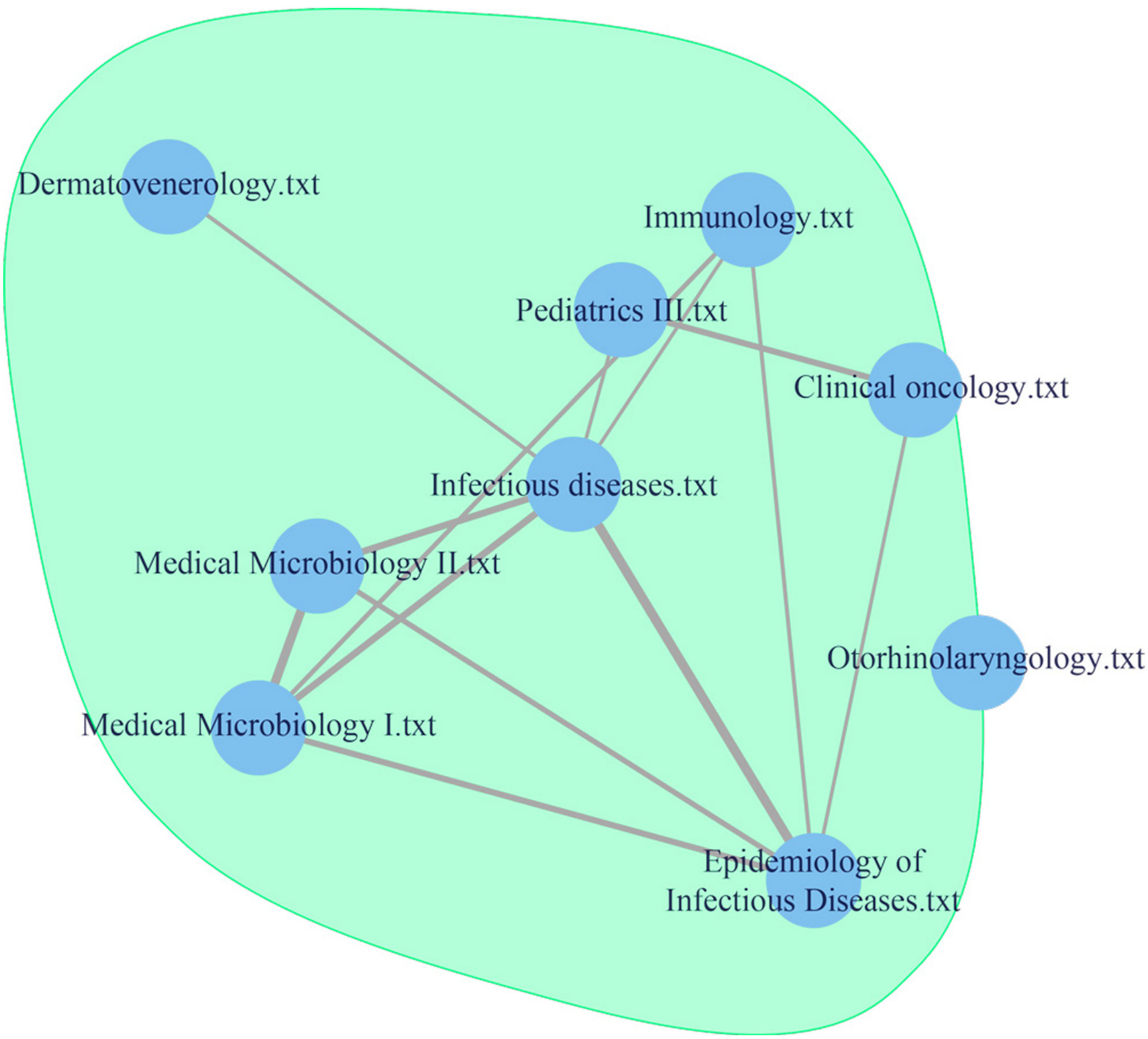


Fig 7. Detailed view of the interconnection between Microbiology and Infectious Diseases in different specialties.

doi:10.1371/journal.pone.0143748.g007

senior guarantors were familiar with the project activities, but were not deeply involved in the technical evaluation of the curriculum. All of them evaluated the generated graphs (see Figs 6 and 7) of the curriculum, and reviewed the pictures that described possible interconnections among each study subjects independently. The consensus of their review, based on their negotiations and discussion, is presented below. The reviewers were already familiar with the description of the technical evaluation process for evaluating interconnections (in terms of similarities, according the key words mentioned in the curriculum), and were introduced to some technical aspects of the evaluation, but without any guidance, or suggestions, or evocative recommendations, as this could influence the results of their review. As mentioned above, all reviewers held very similar positions, a fact that was due to set criteria mentioned above, and their views did not differ. All reviewers used a simple template to provide feedback, namely a checklist that evaluated the structure of the visualized curriculum. In addition, they described an evaluation in their own words, and discussed it afterwards all three experts together during an unstructured interview. This the method was used to analyze the qualitative feedback (simple semantic analyses and subsequent synthesis of conclusions was done).

Below, the short evaluation of similarity graphs is stated. Below, the short evaluation (a summary of comments) of similarity graphs is stated. The colored groupings of individual communities with related contents across the curriculum makes viewing the particular medical curriculum simpler and easier to understand. The courses that fall into the preclinical disciplines show suitable overlap and are centered together. The interconnection among the surgical disciplines, Anesthesiology, and First aid is excellent, implying good consistency. The distant location of the Ethics course is logical. However, this course is not interconnected with either Communication or Psychology; instead, it stands on the opposite side of the curriculum. It is possible that Ethics, Communication, and Psychology are not sufficiently represented in the preclinical and clinical disciplines, which unfortunately is a prevalent trend in medical education at Czech medical faculties [39]. Another explanation might lie in their different locations within the curriculum, in terms of the year of study (i.e., Ethics is taught in the second year of study). It is very difficult to interconnect specific ethical cases on which the education of ethics should be based, if students have not yet participated in clinical practice or internships. This problem might be addressed by extending the education of ethics to include the internship period.

An interconnection between Microbiology and Infectious Diseases (see Fig 5) can also be rationally explained: this confirms clear interdisciplinary relations. In this context, Dermatology and Venereology are also suitably located, because they can be regarded as specialties focused on the most common infectious diseases. the distant location of the Stomatology course (see Fig 6) is understandable, as Dentistry is a stand-alone program of study in the Czech Republic, and students in General Medicine are merely introduced to this subject. Ophthalmology does not occupy large space within the medical curriculum and that it also has a low perceived importance. In fact, the distant location of preclinical practice in the course of Internal Medicine can be justified with regards to the involvement of different disciplines of internal medicine; otherwise, a high consistency of these disciplines (i.e., Diabetology, Nephrology, and Gastroenterology in the center of the schema [light red color]) is evident. A similar consistency can be seen in the location of courses focusing on community care and public health, including Clinical Practice in Community Medicine (marked in yellow). A high degree of content similarity (according to the location) can also be seen in the preclinical courses (i.e., Basics of Medical Terminology, Anatomy 1, 2, and 3), with an overlap with Histology and relationships with other clinical courses (marked in light red). The distant location of the Imaging Methods course is troubling, in terms of the curriculum consistency. On the one hand, one can argue that imaging techniques in this LU are rather specific, in that it teaches about uncommon examinations, as the common diagnostic methods are shown in individual specialized courses. On the other hand, relative serious shortcoming of this course's remoteness is that it has no interconnection with other clinical disciplines; each clinical field/course/LU must certainly mention some examination techniques including imaging techniques, as mentioned above, and therefore it is essential to respect both horizontal and vertical interdisciplinary connections and relationships, which are based on a mutual sharing of requirements by individual teachers. The evaluation is also essential for the identification of the extent to which knowledge and skills have been acquired in individual courses. The general view of the curriculum is consistent, involving sporadic distant points that represent courses that focus on soft skills (i.e., Ethics, Communication, and Psychology) on the one hand, and courses that focus on specialized fields of medicine (i.e., Stomatology, Ophthalmology) on the other hand.

4.2 Limitations

Regarding medical curriculum innovations, we are investigating how to further optimize our curriculum innovations platform towards interoperability powered by broadly accepted standards and recommendations. The MedBiquitous organization has developed American National Standards Institute accredited technical standards for the health professions education (http://www.medbiq.org/), and for different purposes such as measuring students' performance, describing health competency frameworks, learning activities and content, and others. In the near future, we will adopt and implement two MedBiquitous standards, Competency Framework and Curriculum Inventory Standard [40] into our curriculum management system. The former allows LO to be used as the backbone of curriculum management systems, and enables users to search for resources that address a specific competency and determine where competencies are addressed in the curriculum. The latter is used to define curriculum data within a specific health profession education program, to facilitate their exchange, and to shift curriculum mapping and reporting from a disjointed and institution-specific undertaking to something that is shared among multiple medical schools and across whole medical education systems. We aim to be able systematically construct the complete profile of a graduate, which will be based on all the available descriptive metadata attributes. The broad support for the mentioned standards gives us a challenging opportunity to determine how to compare various characteristics of students across standardized curriculum management systems.

4.3 Comparison with prior work

Prior published works have introduced powerful existing virtual learning environments that focused on the curriculum management and relating planning activities but only from a certain perspective. Usually, these systems generate a huge amount of data, but they unfortunately differ in their levels of detail and their description style. Logically, a lack of any kind of standardization, unification, or common parameterization hampers global transparency and comprehensibility. As a result, it is very difficult to view the whole study field from a broader perspective, and diminishes the possibility of searching easily across the curriculum. Therefore, we decided to build a platform that covers all elements associated with global curriculum harmonization, including a detailed parametric description down to the level of the LUs and LOs. Thanks to this platform, we were able to classify learning activities appearing in the medical curriculum, in accordance with a specialized biomedical dictionary that took the form of aggregated tabular and graphic outputs, including interactive visualizations [41]. For further progress in in-depth curriculum analysis and advanced reporting, this unification on standards level is an essential step forward.

Conclusions

An original method of fostering curriculum innovations and harmonization was introduced in this paper. The presented solution can be used for the potentially perpetual process of specification and subsequent upgrades in a curriculum at a higher educational institution, and can provide tools to describe the educational process as effectively and easily as possible. Our endeavor was to describe a solution to curriculum mapping issues, and to show how it might make a particular study program more transparent for a broad academic community. In comparison with previously published approaches, we present a comprehensive system that includes a proven planning model based on the outcomes-oriented paradigm, which was adopted in practice by hundreds of senior curriculum designers and innovators as well as by thousands of students in the general medicine study program at Masaryk University. Furthermore, a pilot exploration of a medical curriculum that used social network analysis was performed, and outlying and overlapping medical disciplines were identified.

Regarding reproducibility, the implementation of the presented solution is fully independent of the particular information and communication technologies, as well as the particular field of study to be innovated and harmonized. The basic preconditions for running similar projects in other faculty are: (i) definition of curriculum in accordance with a proposed planning model; (ii) use of a pre-defined parametric structure of the study fields, providing an easy way how to extract required information from database; (iii) engagement of enthusiastic and reliable authoring and reviewing teams of curriculum experts; and (iv) specification of particular curriculum mapping issues, which identify information rich data relations and offer a clear and transparent overview, and in turn creates simpler understandings of the curriculum structure. This paper is oriented to a community interested in a curriculum harmonization process provided by HEIs, especially teachers, tutors, curriculum designers, institutional management, and anyone who might be engaged in the innovation of educational processes. The target group of the proposed platform includes the broader academic community: students, teachers, and the management of HEIs. In our future work, we will initiate the innovation of e-learning content linked to the LUs that have already been restructured by the curriculum designers/innovators. The methodical development and careful implementation of our solution integrates state of the art technologies for effective data handling. Moreover, it establishes a structured environment for further innovation, such as another module (e.g., the reporting tools, which will be developed), in order to provide a real time visual analytics service. We investigate how this tool could both represent the complex curriculum to enable easier analysis and processing, and also elicit scientific value currently hidden in curricula data.

Acknowledgments

The authors were supported from the following grant projects: (i) OPTIMED—optimized medical education: horizontal and vertical connections, innovations and efficiency in practice —reg. no.: CZ.1.07/2.2.00/28.0042, which is funded by the European Social Fund and the National (state) budget of the Czech Republic; (ii) CROESUS—clinical reasoning skills enhancements with the use of simulations and algorithms—reg. no.: 2014-1-CZ01-KA203-002002, which is funded by the European Commission ERASMUS+ program; and (iii) EVAMED—evaluation of medical education—reg. no.: MUNI/A/1512/2014, which is funded by the specific research and rector's programme of Masaryk University.

Author Contributions

Conceived and designed the experiments: MK MV. Performed the experiments: MK MV. Analyzed the data: MK MV. Contributed reagents/materials/analysis tools: MK MV CV DS NZ.

Wrote the paper: MK MV CV DS AP NZ LD. Critical interal review: MK MV CV DS AP NZ LD.

References

1. Komenda M, Schwarz D, Dušek L Building platform for optimization of medical education, EFMI STC Shar. Knowl. Tools Decis. Support Biomed. Health Care. (2013).
2. Adam S. An introduction to learning outcomes, Consid. Nat. (2006). Available: http://ccs.dcu.ie/afi/docs/bologna/a_consideration_of_the_nature_function.pdf. Accessed 11 July 2012.
3. Harden RM, Crosby JR, Davis MH. AMEE Guide No. 14: Outcome-based education: Part 1ÐAn introduction to outcome-based education, Med. Teach. 21 (1999) 7–14.
4. Biggs J, Tang C. Teaching for quality learning at university: What the student does, McGraw-Hill Education (UK), 2011. Available: https://www.google.com/books?hl = cs&lr=&id=VC1FBgAAQBAJ&oi = fnd&pg=PP1&dq=Teaching+for+quality+learning+at+universit+&ots=E6yLjA6APp&sig = gufHqAdbsNsakd1-M3ndL0E7RAk. Accessed 18 May 2015.
5. Vaitsis C, Nilsson G, Zary N. Visual analytics in healthcare education: exploring novel ways to analyze and represent big data in undergraduate medical education, PeerJ. 2 (2014) e683. doi: 10.7717/peerj.683 PMID: 25469323
6. Vaitsis C, Nilsson G, Zary N. Big Data in Medical Informatics: Improving Education Through Visual Analytics, Stud. Health Technol. Inform. 205 (2014) 1163. PMID: 25160372
7. Vaitsis C, Nilsson G, Zary N. Visual Analytics in Medical Education: Impacting Analytical Reasoning and Decision Making for Quality Improvement, Stud. Health Technol. Inform. 210 (2015) 95–99. PMID: 25991109
8. Nifakos S, Vaitsis C, Zary N. AUVA -Augmented Reality Empowers Visual Analytics to explore Medical Curriculum Data, Stud. Health Technol. Inform. 210 (2015) 494–498. PMID: 25991196
9. Hege I, Nowak D, Kolb S, Fischer MR, Radon K. Developing and analysing a curriculum map in Occupational-and Environmental Medicine, BMC Med. Educ. 10 (2010) 60. doi: 10.1186/1472-6920-10-60 PMID: 20840737
10. Olmos M, Corrin L. Academic analytics in a medical curriculum: Enabling educational excellence, Australas. J. Educ. Technol. 28 (2012).
11. Komenda M. Advanced system for optimization of medical curriculum, (2013). http://is.muni.cz/th/98951/fi_r/thesis-proposal-komenda.pdf (accessed June 26, 2013).
12. Komenda M, Schwarz D, Hřebíček J, Holčík J, Dušek L. A Framework for Curriculum Management—The Use of Outcome-based Approach in Practice, in: Proc. 6th Int. Conf. Comput. Support. Educ., SCITEPRESS, Barcelona, 2014. 10.5220/0004948104730478.
13. Komenda M, Schwarz D, Dušek L. The role of information technologies in medical curriculum harmonisation, PeerJ. (2015). doi: 10.7287/peerj.preprints.1093v1
14. Komenda D S M. OPTIMED Platform: Curriculum Harmonisation System for Medical and Healthcare Education, Stud. Health Technol. Inform. 210 (2015).
15. Mannila H. Data mining: machine learning, statistics, and databases, in: Sci. Stat. Database Manag. Int. Conf. On, IEEE Computer Society, 1996: pp. 2–2. 10.1109/SSDM.1996.505910.
16. Uchiyama KP, Radin JL. Curriculum mapping in higher education: A vehicle for collaboration, Innov. High. Educ. 33 (2009) 271–280.
17. Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning, Med. Teach. 23 (2001) 123–137. PMID: 11371288
18. Wachtler C, Troein M. A hidden curriculum: mapping cultural competency in a medical programme, Med. Educ. 37 (2003) 861–868. PMID: 12974839
19. Plaza CM, Draugalis JR, Slack MK, Skrepnek GH, Sauer KA. Curriculum Mapping in Program Assessment and Evaluation, Am. J. Pharm. Educ. 71 (2007). Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858603/. Accessed 16 August 2015.
20. Ball S. Dynamic Learning Maps—Final Project Report, (2009). Available: https://learning-maps.ncl.ac.uk/.
21. Horner P, Cotterill S, Peterson J, Skelly G. Dynamic Learning Maps, in: Learn. Forum Lond. 2010-Internet Subj. Forum EPortfolio-Key Competencies-Identity Lond. 5–7 July 2010 Savoy Place Publ. EIfEL ISBN 2-9524576-9-7, n.d.: p. 98. Available: http://2012.epforum.eu/proceedings/ePortfolio%202010.pdf#page=98. Accessed 14 July 2012.

22. SKabicher S, Derntl M. Visual Modelling for Design and Implementation of Modular Curricula, Z. Für Hochschulentwicklung. (2008). Available: http://www.zfhe.at/index.php/zfhe/article/view/64. Accessed 19 July 2012.
23. McDanlel E, Roth B, Miller M. Concept mapping as a tool for curriculum design, Informing Sci. Int. J. Emerg. Transdiscipl. 2 (2005) 505–513.
24. Aldrich PR. The curriculum prerequisite network: Modeling the curriculum as a complex system, Biochem. Mol. Biol. Educ. 43 (2015) 168–180. doi: 10.1002/bmb.20861 PMID: 25708224
25. Schwarz D, Dušek L, Bareš M, Bienertová Vašků J, Feberová J, Kofránek J, et al. Computer applications systems and networks for medical education: MEFANET: Czech and Slovak Medical Faculties Network, Facta Medica, 2014. Available: http://www.muni.cz/research/publications/1211459. Accessed 16 July 2015.
26. Schwarz MR, Wojtczak A. Global minimum essential requirements: a road towards competence-oriented medical education, Med. Teach. 24 (2002) 125–129. PMID: 12098430
27. Zhang X-I, Li X, Wan X, Zeng C, Li Z. A survey of Chinese medical students: Attitudes toward the global minimum essential requirements in medical education [J], China High. Med. Educ. 5 (2002) 1–4.
28. M. Komenda, D. Schwarz, L. Dušek, Towards a System of Enhanced Transparency of Medical Curriculum, (2013). Available: http://www.muni.cz/research/publications/1131373. Accessed 26 March 2014.
29. Witten IH, Frank E, Trigg LE, Hall MA, Holmes G, Cunningham SJ. Weka: Practical machine learning tools and techniques with Java implementations, (1999). http://researchcommons.waikato.ac.nz/handle/10289/1040 (accessed May 31, 2014).
30. Azevedo AIRL. KDD, SEMMA and CRISP-DM: a parallel overview, (2008). Available: http://recipp.ipp.pt/handle/10400.22/136. Accessed 9 July 9 2013.
31. Trigo L, Brazdil P. Affinity Analysis between Researchers using Text Mining and Differential Analysis of Graphs, (2011). Available: https://phdsession-ecmlpkdd2014.greyc.fr/sites/phdsession-ecmlpkdd2014.greyc.fr/files/proceedingsPhDsession_ECML_PKDD_2014final.pdf#page=174. Accessed 2 May 2015.
32. Feldman R, Sanger J. The text mining handbook: advanced approaches in analyzing unstructured data, Cambridge University Press, 2007. Available: https://www.google.com/books?hl = cs&lr=&id=U3EA_zX3ZwEC&oi = fnd&pg=PR1&dq=The+text+mining+handbook:+advanced+approaches+in+analyzing+unstructured+data&ots=2MzEMbEwID&sig = nXIOh6zfcXJtnpWm67lJ3Zw-COM. Accessed 25 May 2015.
33. Pons P, Latapy M. Computing Communities in Large Networks Using Random Walks, J. Graph Algorithms Appl. 10 (2006) 191–218. doi: 10.7155/jgaa.00124
34. R igraph manual pages, (2015). Available: http://igraph.org/r/doc/walktrap.community.htm. Accessed 22 May 2015.
35. Bonacich P. Power and centrality: A family of measures, Am. J. Sociol. (1987) 1170–1182.
36. Freeman LC. Centrality in social networks conceptual clarification, Soc. Netw. 1 (1979) 215–239.
37. Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, et al. CRISP-DM 1.0 Step-by-step data mining guide, (2000). Available: http://www.citeulike.org/group/1598/article/1025172. Accessed 9 July 2014.
38. Keeling R. The Bologna Process and the Lisbon Research Agenda: the European Commission's expanding role in higher education discourse, Eur. J. Educ. 41 (2006) 203–223. doi: 10.1111/j.1465-3435.2006.00256.x
39. Majernik J, Schwarz D, Komenda M, Dušek L. A Unified Educational Platform of Multimedia Support in Education at Medical Faculties of MEFANET Project, Procedia—Soc. Behav. Sci. 55 (2012) 476–484. doi: 10.1016/j.sbspro.2012.09.527
40. Ellaway RH, Albright S, Smothers V, Cameron T, Willett T. Curriculum inventory: Modeling, sharing and comparing medical education programs, Med. Teach. 36 (2014) 208–215. doi: 10.3109/0142159X.2014.874552 PMID: 24559305
41. Komenda M, Schwarz D, Švancara J, Vaitsis C, Zary N, Dušek L. Practical use of medical terminology in curriculum mapping, Comput. Biol. Med. (2015). doi: 10.1016/j.compbiomed.2015.05.006

Computers in Biology and Medicine 63 (2015) 74–82

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Practical use of medical terminology in curriculum mapping

Martin Komenda [a], Daniel Schwarz [a,*], Jan Švancara [a], Christos Vaitsis [b], Nabil Zary [b], Ladislav Dušek [a]

[a] Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Czech Republic
[b] Department of Learning, Informatics Management and Ethics, Karolinska Institutet, Sweden

ARTICLE INFO

Article history:
Received 5 March 2015
Accepted 7 May 2015

Keywords:
Medical informatics
Medical coding
MeSH
Medical education
Curriculum management
Visual analytics
Association analysis

ABSTRACT

Background: Various information systems for medical curriculum mapping and harmonization have been developed and successfully applied to date. However, the methods for exploiting the datasets captured inside the systems are rather lacking.
Method: We reviewed the existing medical terminologies, nomenclatures, coding and classification systems in order to select the most suitable one and apply it in delivering visual analytic tools and reports for the benefit of medical curriculum designers and innovators.
Results: A formal description of a particular curriculum of general medicine is based on 1347 learning units covering 7075 learning outcomes. Two data-analytical reports have been developed and discussed, showing how the curriculum is consistent with the MeSH thesaurus and how the MeSH thesaurus can be used to demonstrate interconnectivity of the curriculum through association analysis.
Conclusion: Although the MeSH thesaurus is designed mainly to index medical literature and support searching through bibliographic databases, we have proved its use in medical curriculum mapping as being beneficial for curriculum designers and innovators. The presented approach can be followed wherever needed to identify all the mandatory components used for transparent and comprehensive overview of medical curriculum data.

1. Introduction

Medical coding and classification systems are used in a variety of applications in medicine, public health, education and medical informatics, including their use in statistical analyses, data mining, and expert and decision support systems or knowledge engineering. Creation and maintenance of the coding systems, nomenclatures and standardized vocabularies involves the processes of transforming descriptions of diagnoses and procedures into a universal medical standardized scheme [1]. When talking about standardization, it is necessary to mention also ontologies, which formalize the meaning of terms used in practice, expected to play a major role in the automated integration of medical data with relevant information to support basic discovery and clinical research, drug formulation, and drug evaluation through clinical trials [2]. If we want information systems in health care to process data automatically, to sort them and produce correct statistics and if the principal data should be understood globally, it is necessary to use international coding systems and nomenclatures in the process of data acquisition [3]. From the perspective of medical informatics, the topic covering the broad use and global acceptance of medical coding standards still involves two main unresolved challenges: (i) the number of coding standards developed for medicine and health care is high, whereas their widespread adoption has been rather slow; (ii) the developed standards and vocabularies vary in their coverage, often being contradictory and competing. The basic principles of medical informatics and discussions on medical coding can be found in books and guides, such as [4] or [5]. One working group inside the European Federation for Medical Informatics (EFMI) is focusing on codes, classifications, terminologies and nomenclatures in their overview article [6]. They provide definitions and history of the International Classification of Primary Care (ICPC), and of the Read code and the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). The review [7] analyses the use of SNOMED CT over time from 1966 until June 2006. A more recent review [8] then reports on the increasing use of SNOMED CT in literature, showing that the adoption of the nomenclature needs additional efforts toward harmonization with other standardized terminologies.

Medical education is an active area of education research, the importance and value of which has been demonstrated by Harden [9,10] and later by Wartman [11] already in the late eighties and the early nineties of the previous century. Medical curriculum development, mapping and harmonization are the natural components of

* Corresponding author. Tel.: +420 549492854.
E-mail address: schwarz@iba.muni.cz (D. Schwarz).

http://dx.doi.org/10.1016/j.compbiomed.2015.05.006
0010-4825/

medical education, regardless of whether the underlying model is systems-based [12], outcome-based [13] or different (competency-based, socially accountable, distributed, etc.). Increasing variety in the form and resulting proficiency of different curriculum models has recently prompted the medical education community to introduce curriculum mapping guides [14] and interoperability standards [15]. The MedBiquitous Curriculum Inventory Standard [16], approved by the American National Standards Institute (ANSI), provides the technical syntax through which a wide range of different curricula can be expressed and subsequently analysed.

While there are a plethora of references, surveys and books on medical coding and classification for healthcare informatics applications, much less can be found on standardized vocabularies employed in medical education. The existing cataloguing schemas and thesauri relating, at least partly, to medical education are reviewed in [17] by the Medical Education Taxonomy Research Organization (METRO). Following the results of their research among stakeholders, METRO uses two existing thesauri: (i) the Medical Subject Headings (MeSH) thesaurus, being generic for medicine, and the British Education Thesaurus (BET), being generic for education. The thesaurus of descriptors for topics in medical education resulting from METRO's efforts is not available today, but their set of descriptors for assessment in medical education [18] using as much as possible from the MeSH and BET thesauri is still a living project with results available on-line. The need for identifying a suitable ontology or taxonomy for annotating the content of medical educational resources is emphasized in the systematic review [19], and especially from the perspective of sematic interoperability of information systems used in medical education and the potential of the semantic web (Web 3.0) to support medical curriculum mapping. The review provides a tabular overview of all existing standardized vocabularies, including their features, such as language, scope and the underlying structure, dividing them according to their context into biomedical, educational and administrative. Education resources shared through social media platforms such as YouTube have undergone exponential growth in recent years. Annotating education videos from YouTube and enriching them with other resources through Linked Data approach [20,21] is the main focus of research presented in [22]. SNOMED CT nomenclature is used to bridge YouTube tagging data and the metadata of enriching education resources. The resulting web-based architecture includes also the application interface (API) of BioPortal providing services from the National Center for Biomedical Ontology [23].

In this research, we focus on exploiting the well-known polyhierarchical MeSH thesaurus while helping the medical curriculum designers and innovators to navigate through multidimensional and extensive datasets embedded in the curriculum mapping systems. We are building on our previous experience with various nomenclatures and standards used in the development of the education web portal platform [24] and its extensions [25] adopted by all medical faculties in the Czech Republic and Slovakia under the umbrella of the MEFANET network [26]. In terms of methods and tools for curriculum mapping and reengineering, we are using our original web-based OPTIMED platform [27] and the recently adopted visual analytics tools [28–30] with the aim to develop new insights into curricular datasets and to demonstrate that the use of standardized thesaurus can facilitate curriculum reengineering processes as well as inter-institutional curriculum comparisons. In short, this paper shows an innovative way how to easily identify all essential components of a medical curriculum by adopting the MeSH biomedical nomenclature.

The remainder of this paper is as follows. In the methods section, detailed information is provided regarding: (i) existing and widely-accepted medical terminologies; (ii) our implementation of a web-based system for medical curriculum management; (iii) background and statistical methodology used to develop analytical reports over curricular datasets. In the following section, two data-analytical reports are shown, depicting the results obtained from frequency and association analyses of the particular medical curriculum while employing the MeSH thesaurus. The importance of these achievements together with their current limitations and future prospects are discussed and concluded in the final sections.

2. Materials and methods

The main contribution of this paper lies in the use of a standardized medical thesaurus in the analyses of a dataset extracted from a curriculum mapping system. We believe that our approach can be generalized, although we present results obtained from data stored in a particular curriculum mapping system, OPTIMED, and a specific thesaurus, MeSH. The selection of this specific thesaurus is the result of our scoping review carried out according to the principles laid out in the framework for scoping studies by Arksey and O'Malley [31] and recently extended and clarified elsewhere [32,33].

2.1. Overview of the existing medical terminologies

In this subsection, we refer to the summary overview relating to the up-to-date use of various specialized terminologies in information sciences and education. Our close attention is primarily focused on medically oriented fields where these taxonomies promote an efficient way of organizing and understanding data. It is clear that specialized vocabularies matter, not only in the traditional library and information sciences world and but also for many different digital information stakeholders [21]. For instance, the use of the vocabulary of a particular domain is an important initial step in creating formalized knowledge representations as an essential part of the education process. These vocabularies follow the ratchet principle: it moves from basic understanding to thorough understanding, from simple to complex education [34]. When a virtual learning environment (VLE) turns to the task of consolidated education data collection (for instance content management, curriculum mapping and planning, student engagement and administration, communication and collaboration domain), these vocabularies will present considerable challenges to standardizing medical education. A prerequisite to more comprehensive categorization of the education content is the implementation of standardized terminology directly in the VLE systems. One of the primary aims is to overcome two significant barriers to effective retrieval of machine-readable and processable information: (i) the variety of names used to express the same concept, (ii) the absence of a standard format for distributing terminologies [35]. The purpose of medical vocabularies is to embody what has been known in the past about every phase of medicine [36]. These vocabularies continue to increase and grow, not only in its technological aspects, but also from the perspective of medical education quality, which is logically reflected at the global level of health care. Below, the most widespread, suitable and commonly used standardized methodologies are introduced.

2.1.1. UMLS

Unified Medical Language System (UMLS) brings together many health and biomedical vocabularies, ontologies and standards to enable interoperability between computer systems. It was developed by the National Library of Medicine and covers the entire terminology domain by integrating more than 60 families of

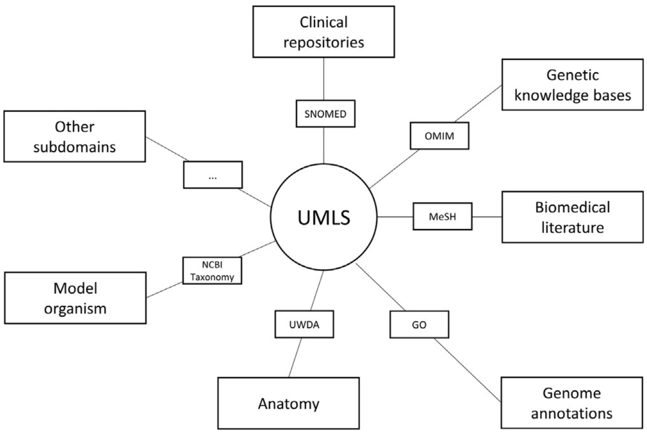


Fig. 1. The various subdomains integrated in UMLS [37].

biomedical vocabularies. The three major components of UMLS are the Metathesaurus (repository of inter-related biomedical concepts covering extensive list of terms and codes from various vocabularies), the Semantic network (high-level categorization of Metathesaurus concepts) and the Specialist Lexicon (generating the lexical variants of biomedical terms) [37–39]. For illustration, Fig. 1 shows UMLS Metathesaurus integrated sub-domains.

2.1.2. MeSH

Although geared specifically for information retrieval, MeSH can almost be seen as a general purpose vocabulary with concepts from all segments of the biomedical domain. This classification is a rich and controlled vocabulary generated through an intense indexing process performed by examiners. Terms (namely descriptors) are assigned to documents to delineate their content at 16 different branches of specificity. The 2014 MeSH vocabulary is specifically composed by more than 27 thousand descriptors, which are organized in a tree-like structure. Descriptors may be also complemented by one or more qualifiers. These terms further contextualize the meaning of the descriptors to which they are assigned in relation to the content of the considered document. The MeSH vocabulary is often used to delineate samples of documents in a number of medical areas and, as discussed, at various levels of specificity [40–42].

2.1.3. ICD-10

International Classification of Diseases (ICD) is used to classify and code mortality information worldwide. The tenth revision (ICD-10) is more complex than the previous one (ICD-9). Specifically, ICD-10 contains more codes as well as re-structured chapters and changes in rules for coding. ICD-10 as a whole is designed to be the core classification for a family of disease- and health-related classifications. Some members of the family of classifications are derived by using a fifth or even sixth character to specify more detail. In others, the categories are condensed to render the broad groups suitable for use, for instance, in primary health care or in general medical practice [43,44].

2.1.4. SNOMED-CT

SNOMED CT as a coding system for clinical problems was created by the merger, expansion and restructuring of two large-scale terminologies: SNOMED RT (Systematized Nomenclature of Medicine Reference Terminology) and CTV3 (Clinical Terms Version 3). Since it contains more than 380 thousand concepts, with a total of about 800 thousand descriptions or terms, the practical use of SNOMED will demand new types of tools to search and navigate intuitively in the term "collection". It is probably the most comprehensive clinical healthcare terminology in the world because it systematically supports the development of comprehensive high-quality clinical content for health records and provides a standardized way of representing clinical phrases captured by the clinician, enabling automatic interpretation. SNOMED CT is a clinically validated, semantically rich, controlled vocabulary that facilitates evolutionary growth in expressivity to meet the emerging requirements [45,46].

2.1.5. Emtree

Emtree is a hierarchically structured and controlled terminology for biomedicine and related life sciences, developed by Elsevier as a biomedical and pharmacological online database. It includes a whole range of terms for drugs, diseases, medical devices and essential life sciences concepts. Emtree thesaurus contains about 48 thousand preferred terms and 200 thousand synonyms in 15 main domains of drugs and diseases, organized in a multilevel hierarchy. All terms, strictly synonymous, are organized in a concept-based structure [47–49].

2.1.6. RadLex

RadLex [50] is a standardized vocabulary of radiological terms, which includes highly detailed terms for anatomy, pathology, and radiological observations. It already contains over 8000 anatomic and pathology terms, many of which are not currently available in other controlled medical terminology systems. Generally, RadLex was designed to fill the gaps in other medical terminology systems, thereby creating a single source for medical imaging terminology. Another key distinguishing feature of RadLex is that it is designed to be continuously supplemented and updated with the incorporation of new concepts, including harmonization with other popular medical vocabularies and term sets. The goal of this method is to establish a uniform, consistent terminology to improve communication of results and to better integrate clinical practice with education and scientific literature [51].

2.1.7. Summary

The integration of various specialized nomenclatures, vocabularies, and terminologies allows more precise analysis and improvement of the education data content. Furthermore, automated systems can apply the knowledge encoded with the use of mentioned taxonomies and human users can easily search and browse the available data in a simple and straightforward manner. It will have both academic and clinical implications by enhancing the retrieval as well as the indexing of information. Besides the introduced approaches, there exist many other vocabularies that are specifically focused on a particular biomedical discipline. Examples include Gene Ontology (GO), Gene expression data (eVOC), Online Mendelian Inheritance in Man (OMIM), Merged Disease vocabulary (MEDIC), London Dysmorphology Database (LDDB), Foundational Model of Anatomy (FMA), Logical Observation Identifier Names and Codes (LOINC) and the NCI thesaurus for cancer research. Contrary to these approaches, other concepts are strongly related to administrative data, i.e. when was what created by whom and who is the intended audience - examples: Dublin Core [52] as a pioneering standard for metadata about electronic resources; Learning Object Meta-data (LOM) ontology developed by IEEE association for describing the learning resources [53]; Shareable Content Object Reference Model (SCORM) representing a widely adopted collection of specifications for web-based e-learning systems [54]. On the grounds of our scoping review and composed requirements, we chose the most suitable thesaurus for the complex process of medical curriculum harmonization.

2.2. OPTIMED: web-based curriculum management system

For the purposes of enhancing the long-term process of medical curriculum mapping and harmonization, we decided to design, develop, implement and run our original web-based OPTIMED platform (http://opti.med.muni.cz/en). This system describes and categorizes all learning activities (lectures, seminars, clinical practices and self-study) in the theoretical and clinical sections of medical curriculum. We have aimed to capture the systematic transmission of medical/clinical knowledge to students during their courses in General Medicine at the Faculty of Medicine of the Masaryk University. By development we refer to the Extreme programming methodology [55], which is a popular agility-oriented framework. Based on the gathered requirements, a modular structure of the OPTIMED system was proposed. It consists of three fundamental modules, which provide easy management as well as fast and transparent browsing through the extensive domain of curriculum metadata: (i) Learning outcome register, (ii) Learning unit register, (iii) Curriculum browser. Overall, OPTIMED gives students, teachers, guarantors, curriculum designers and the faculty management a detailed look at where specific topics and learning outcomes are addressed and how education objectives are being met.

2.3. Statistical and computational methodologies

We have processed in-depth analysis on medical curriculum data supported by standardized vocabulary. We came up with a systematic approach that allows formally describing and effectively optimizing the education structure through appropriate metadata attributes, and provides direct interconnection to essential terms in the thesaurus. The mentioned analytical outputs can significantly assist in detecting potentially problematic areas and subsequently, compile comprehensive reports or instructions for onwards in-depth global inspection, which is necessary for maintaining continuity of the process of first-rate and guaranteed studies at higher education institutions. The derived aggregated and comprehensible information, such as clear tables, transparent graphs and other interactive visualizations, serve as supporting material for the long-term evaluation process under the supervision of senior guarantors and the faculty management. Moreover, the integration of proper vocabulary also provides a unified terminology for instituting curricular change and innovation.

We based all reports on a set of proven analytical and statistical techniques encompassing robust computational and visualization software. The occurrence of MeSH terms and linked categories in the description of learning units were described by absolute and relative frequencies (percentage representation). Associations among all MeSH keywords were computed using the Jaccard similarity coefficient, which represents a statistic method for comparing the similarity and diversity of data sets. It provides robust measurement of the association widely used in ecology (measuring biodiversity) or market research (market basket analysis). This coefficient is based on positive similarity (subject to co-occurrence) and not negative similarity (based on double absence) and we used it in order to avoid the double-zero problem. All data merging, manipulation, transformation and computations were conducted by IBM SPSS Statistics version 22.0.0. For interactive visualization of association maps, we used the yEd Graph Editor version 3.10.2, a freely available desktop application developed by yWorks GmbH typically used for diagramming large adjacency matrixes with automated layout.

3. Results

Since the coding systems, nomenclatures and standardized vocabularies can be used for formal content description of any object, they may be also utilized together with traditional education resources such as books, media, and programmes – to annotate courses, learning units and outcomes, exam questions, or any other form of online resource. Thus, they allow searching for thematically related objects [19]. From the perspective of the education process, implementation of controlled vocabularies significantly influences normalization of the achieved results, because students and teachers very often use different theoretical and clinical terms that in fact mean the same. For example, the terms heart attack, myocardial infarction, and MI abbreviation may represent the same meaning to academics, but, to a computer, they are different. The main purpose of the thesaurus concept integration is allowing users to search one term and to find and retrieve learning activities that use synonymous terms. It also enhances the indexing and aggregation process of data across various medical fields of study. In our particular case, we introduce the practical use of medical terminologies, namely the MeSH thesaurus, in medical curriculum harmonization. We have developed a web-based curriculum management system supporting correctly compiled and balanced structure of obligatory and optional courses across medical fields. We have also adopted the MeSH nomenclature for standardizing metadata stored in the system database, which provides a set of descriptive attributes of the whole curriculum with respect to international framework.

3.1. Case study: MeSH classification in medical curriculum mapping

The outputs of this case study represent systematic identification of content related items across curriculum boundaries and bring an effective and proven approach enabling us to better understand the domain of continual improvement of medical education using modern information and communication technologies. As stated above, the implementation of standardized vocabulary can significantly improve the quality of further analytical processing in order to understand stored data. We adopted the MeSH biomedical dictionary, where the main objective is to classify learning activities appearing in the curriculum. In the past, keywords were defined and structured in many forms and there was a growing need for their unification with respect to international framework. The main requirement for standardized dictionary integration was regular updates of the Czech version, which MeSH fulfils as the only available solution. No other language mutations are foreseen at the moment, but a possible change should not bring too many complications in terms of the proposed structure. The National Library of Medicine produces annual editions, including the development and implementation of a concept-centred vocabulary maintenance system for MeSH. This system has been extended to create a multilingual database of translations and allows continual updating of the translations, as well as facilitating tracking of the changes within MeSH from one year to another. In this way, the vocabulary is useful for the non-English user and enables potential connection between various languages (e.g. Czech and English).

3.1.1. The MeSH structure

MeSH is a controlled vocabulary thesaurus comprising over 27,000 primary terms called descriptors, which are arranged in hierarchical structure and permit searching at various levels of specificity. More specific headings are found at narrow levels of the twelve-level hierarchy. Currently, the MeSH hierarchy is divided into 16 categories (A for anatomic terms, B for organisms,

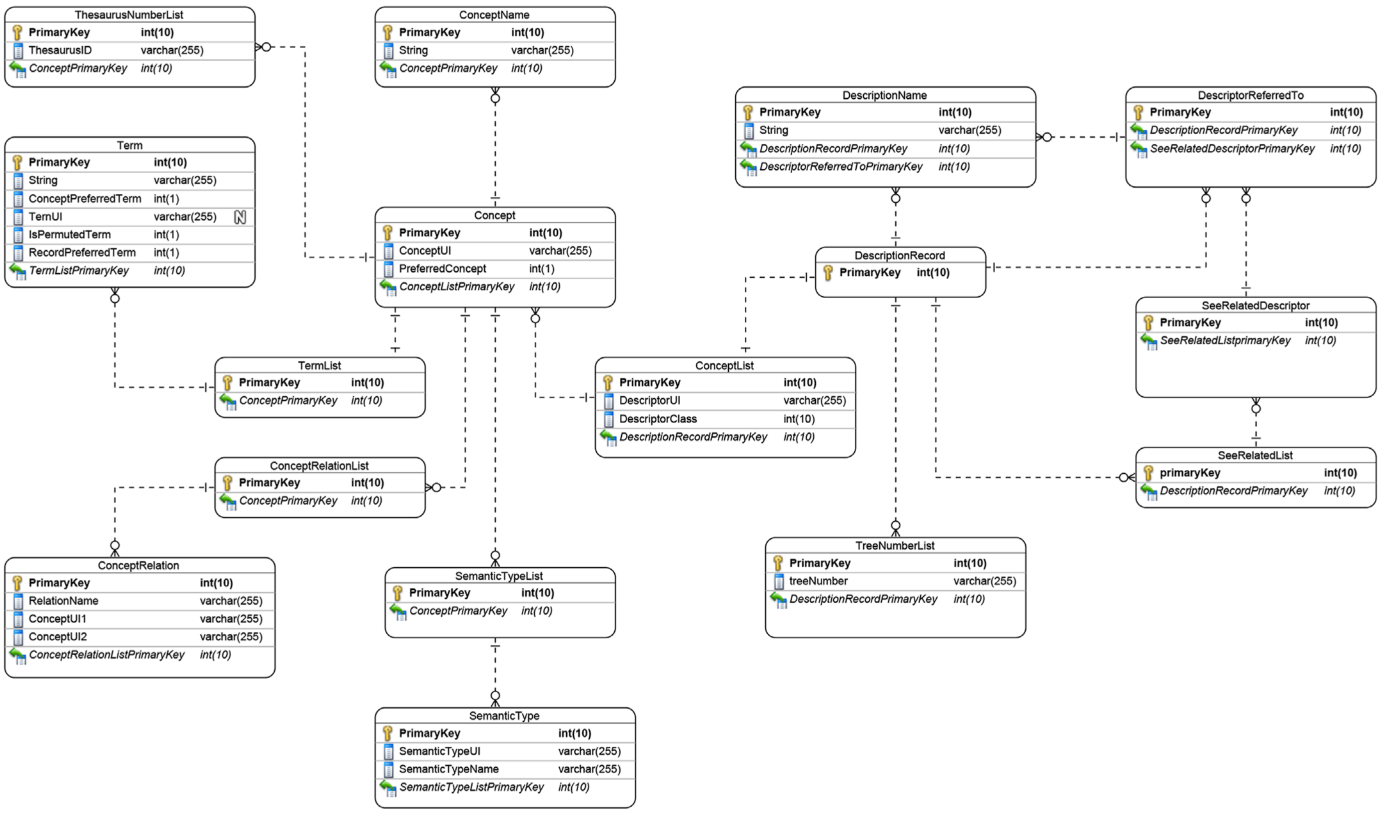


Fig. 2. Complete data model of MeSH entities in the OPTIMED database.

Table 1
Attributes describing the medical curriculum used for case study analysis. (*Source*: the OPTIMED PostgreSQL database).

Name of attribute	PostgeSQL data type	Sample value
unitName	character varying(255)	Skin anatomy
unitSection	character varying(255)	Internal medicine
unitClass	character varying(100)	Dermatovenerology
meshName	character varying(255)	Epidermis
meshPreferred	Integer	Yes
meshSemanticType	character varying(50)	Tissue
meshTreeId	character varying(100)	A10.272.497
meshTreeLevel	Integer	3
meshCategory	character varying(50)	Anatomy
meshSubcategory	character varying(50)	Tissues

C for diseases, D for drugs and chemicals, etc.). Each category is further divided into subcategories (117 in total). For instance, category B consists of the following subcategories: eukaryota, archaea, bacteria, viruses, organism, and forms. These hierarchies are extremely valuable as they allow associated data to be viewed at various levels of granularity, with data annotated to children of a branch to be aggregated at each higher level of the hierarchy [42]. Generally, the entire MeSH vocabulary is available for download in XML, ASCII or MARC formats. In order to better understand the thesaurus construction, we decided to model a complicated structure of the descriptors, concepts and terms. Using the diagramming framework Visual paradigm, we carefully identified all important entities and attributes of spreading the MeSH structure and transformed the XML form into a set of entity-relationship models, thus describing the data in systematic and more transparent manner. Fig. 2 represents a conceptual data model covering all essential MeSH parameters for curriculum description.

3.1.2. Advanced data-analytical reports

Implementing MeSH vocabulary into the OPTIMED system requires specific validation tasks based on standardized terms that we tried to resolve by curriculum data analyses. There are several interesting domains that the selected statistic and analytical methods may answer: coverage and relations between learning activities and the MeSH classification, objective evaluation of the MeSH keywords occurrence, model term integration into education, etc. Two particular reports based on MeSH classification, introduced below, assist in finding potentially problematic areas in the curriculum and provide a comprehensive overview for subsequent in-depth global medical curriculum inspection. The extracted information from the OPTIMED curriculum management system serves as supporting material for the evaluation process under the supervision of the expert committee and the faculty management.

3.1.3. Medical curriculum data set

We have already designed and implemented formal specification of the database scheme as a basic building element of the OPTIMED platform, which covers all elements pertaining to global curriculum harmonization including detailed metadata arrangement down to the level of learning units, interconnections to the learning outcomes in accordance with Bloom's taxonomy and direct interconnection to specialized MeSH nomenclature. The Bloom's taxonomy is often used when designing educational and learning processes, where educational objectives are divided into three domains: cognitive (mental skills known as knowledge), affective (growth in feelings or emotional areas known as attitude or self), and psychomotor (manual or physical skills known as skills). We proposed a special script for transforming the MeSH tree structure to linear representation of individual terms. The main reason was to simplify the links between the keywords and the particular learning units. For the purpose of further analysis, we merged the data (see Table 1) about learning units (unitName, unitSection, unitClass) with data describing MeSH keywords (meshName, meshPreferred, meshSemanticType, meshTreeId, meshTreeLevel, meshCategory, meshSubcategory). The final dataset contains information about 1347 learning units, which were described by 3224 varied MeSH terms (from one to five terms per learning unit) and classified into 133 categories and subcategories.

3.1.4. Results

Two reports relating to MeSH categories in accordance with the created learning units are described here. The primary motivation of our effort was to allow medical curriculum designers, senior guarantors and the faculty management answer questions based on aggregated data from the OPTIMED portal platform. Several analytical dashboards were built and prepared for further review and evaluation. The implications of these findings can help academics in their curriculum redesign activities, and provide a clear overview of the curriculum structure.

The first analytical output is a view of the representation of categories and subcategories according to the MeSH biomedical dictionary based on learning units, which have been created with the aim of describing in detail the education in General Medicine field of study at the Faculty of Medicine of the Masaryk University. In the OPTIMED project, the study of medicine was systematically divided into four sections (theoretical sciences, surgical sciences, diagnostics and neurosciences, internal medicine). With respect to the fact that suitable MeSH classifiers must be obligatorily assigned to each learning unit, the available links to a tree structure and to multilevel categorisation make it possible to draw summarising graphs. Fig. 3 shows the comparison of two sections of education (surgical sciences and theoretical sciences) according to the most frequent subcategories.

The main MeSH categories are very general (only 16 areas); in this case, therefore, we decided to work with subcategories (a total of 117 subcategories), which are more detailed and thus make it possible to define clearer and more understandable reports. Similar to other outputs of this analysis, a fully standardised dictionary is employed containing set phrases and specialised terminology commonly used in medicine. These features provide a simple overview to the guarantor of a given section, describing the thematic categorisation of all learning units belonging to his/her expertise; the guarantor should, therefore, be able to define potential discrepancies very quickly, on relatively general level of curriculum description, without the necessity of performing in-depth analysis. A view across sections, such as the comparison of categorizations between selected sections, can be classified as outputs primarily intended for the faculty management. Like in the first case, potential overlaps—whether more or less desirable—can be identified at first glance.

The second analysis consists in identifying links based on the assigned MeSH keywords in order to reveal the structure of individual sections. At the present time, the OPTIMED portal contains a very comprehensive description of the entire

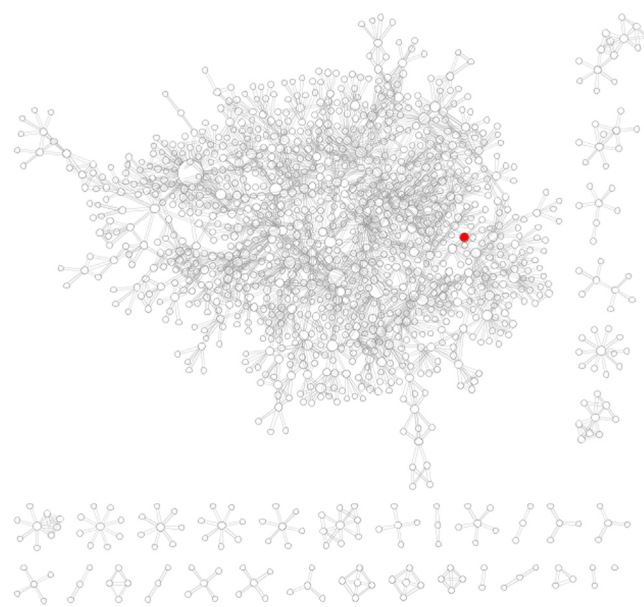


Fig. 4. Global association map of all MeSH terms occurring in one particular section of the medical curriculum. Individual clusters represent various subsets of the MeSH keywords describing the same learning units. Less often used keywords are located in smaller clumps out of the main cluster including the majority of MeSH terms. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

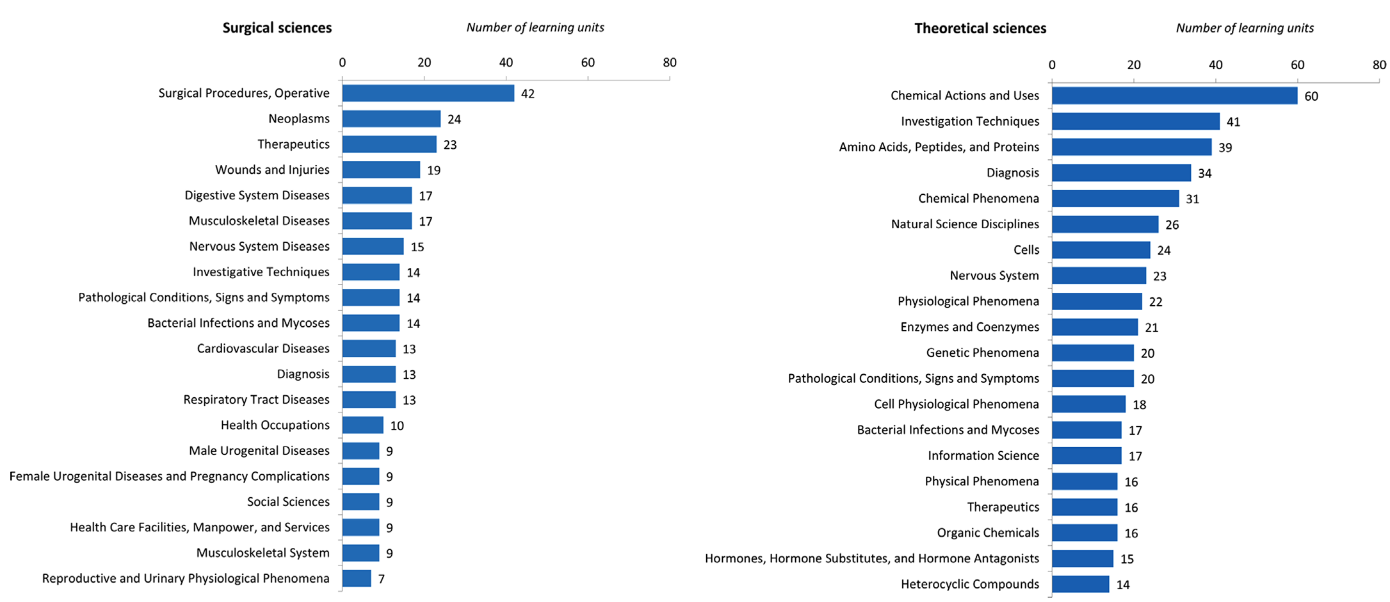


Fig. 3. The MeSH categories according to number of learning units (only 20 most frequent categories are visualized).

curriculum. It is, thus, not humanly possible to go through the entire contents or to imagine how the individual learning units, courses, or branches of medicine are interrelated. We have, therefore, again used the MeSH classification to suggest a report displaying the so-called association maps, i.e. links among MeSH keywords according to their use in individual learning units. For the sake of clarity, each of the resulting visualizations is divided according to the four above-mentioned sections. The graphs are interactive, making it possible to zoom in on a selected segment, up to the level of a specific MeSH keyword. This feature makes it very simple for an expert in a given field to critically evaluate whether the coverage of a given section makes sense, and whether all relevant areas are represented. Figs. 4 and 5 demonstrate the association map of the Theoretical Sciences section. The first view is mostly general, showing all connections. As an example, the term "inflammation" is marked in red. The second view corresponds to a several-fold zoom that, however, is still illegible, and it is not obvious how the terms are interconnected. By contrast, Fig. 6 represents the most detailed view of a given element of the graph. At maximum zoom, the user can easily reveal all terms that have been used to describe the same or related learning units. In this specific case, the term "inflammation" is directly linked to 10 other MeSH terms. Generally, these association maps visualize interconnections among the standardized MeSH classification used for describing the medical curriculum. Each node represents one keyword and its size is calculated from an absolute sum of associations (measured by the Jaccard coefficient) for the given keyword. Since an association between two terms is always reversible, also each edge is doubled. At the bottom of this visualization are located the MeSH keywords, which are used only for describing a single learning unit and are not, in our data set, associated with any other MeSH keyword.

3.1.5. Discussion and future visions

For the purposes of optimizing medical education encompassing a correctly compiled and balanced curriculum, we proposed the formal database metadata arrangement including standardized MeSH vocabulary, which was chosen in accordance with the proven scoping review. In our particular case, the database of medical curriculum contains more than 140 courses, which are described by 1347 learning units and 7075 learning outcomes, i.e.

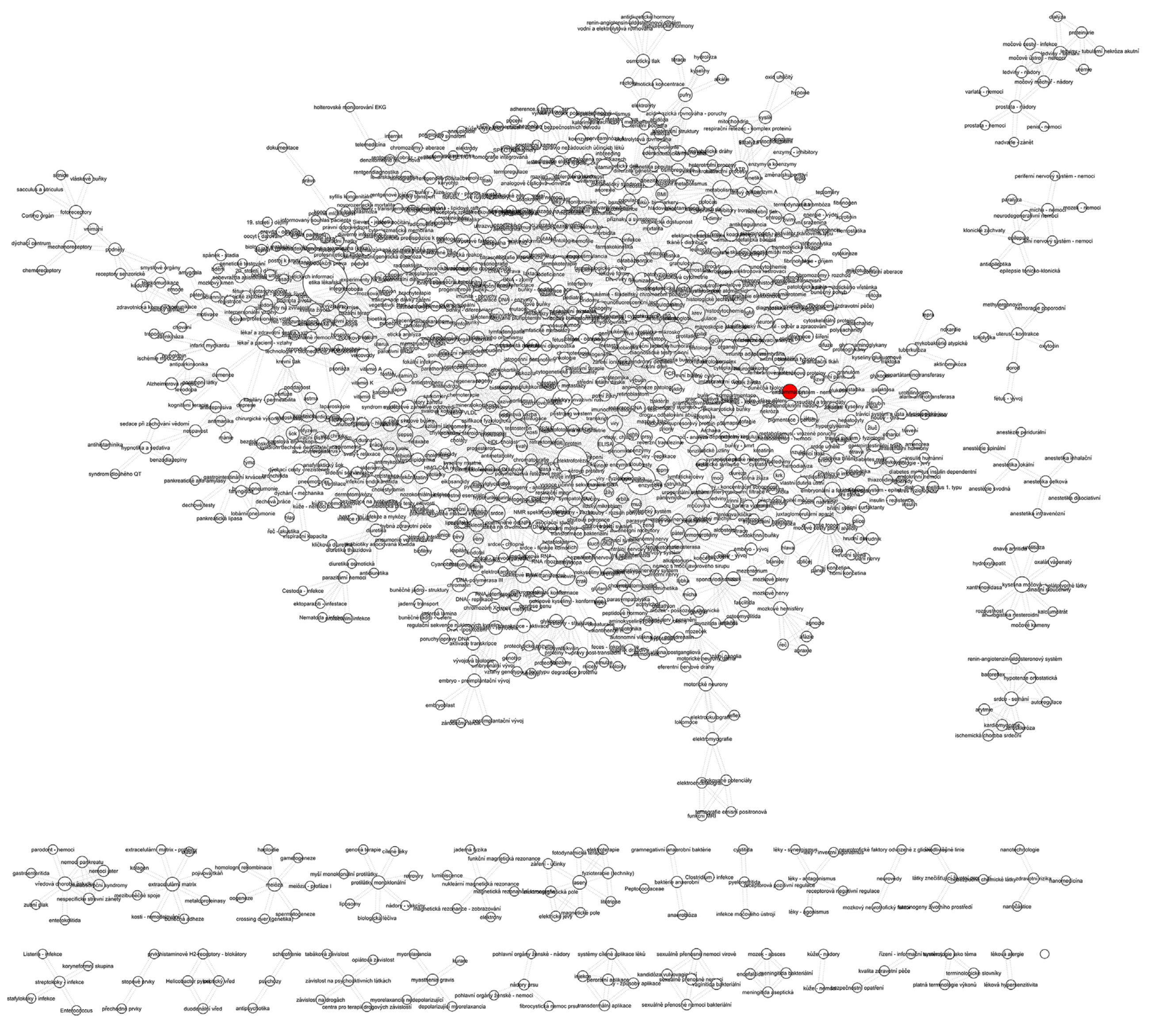


Fig. 5. A more detailed association map of selected MeSH terms occurring in one particular section of the medical curriculum. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

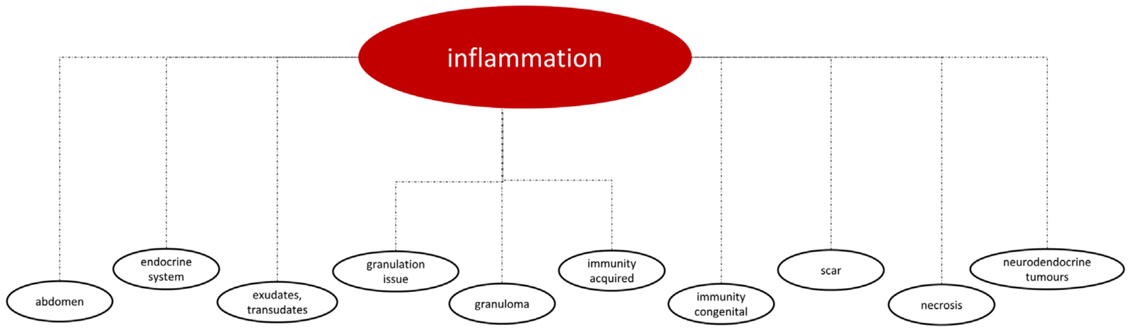


Fig. 6. The deepest level of association map representing all connections between MeSH keywords occurring in theoretical sciences from the perspective of the term “Inflammation”.

more than 2500 pages of text. From the perspective of human cognition abilities, it is not possible to carefully read and verify all of the medical curriculum content. Classification by standardized medical vocabulary can significantly influence effectiveness and shorten the total time requirements placed on curriculum inspection. Before we integrated the MeSH thesaurus, we had to take into consideration the fundamental properties and characteristics of potential candidates such as linear/branched structure (how the terms are organized and interconnected), annual updates and multilingual translation at regular intervals, scope (specific focus or general purpose), data format and licensing. MeSH fully met our visions and fulfilled the requirements. Two analytical reports show different unique approaches on how to clearly visualize a completed structure and the relations between the vast seas of data. By applying the described dashboards, we are able to determine the similarity between individual sections of medical education explored at the level of selected attributes, such as MeSH keywords or subcategories. We have also classified the set of learning units into one of several predefined classes according to Bloom's taxonomy, which represents intellectual behaviour into three various levels: cognitive, affective, and psychomotor.

In the future, the idea of Automatic Term Mapping (ATM) of medical curriculum data is planned. The ATM technique [56] was developed by the US National Center of Biotechnology Information for mapping end-user queries to the MeSH thesaurus and other search fields. The basic aim will be to improve information retrieval as structured information: searching indexes instead of only the free text. Furthermore, with proper use of structured vocabularies and medical curriculum standards, which define the core curriculum data for education programmes in health professions, we will be able to automatically compare and exchange data between various educational systems. For example, a particular comparative analysis of a medical curriculum against topics in the MeSH (or other) lexicon, where the number of articles is rapidly increasing, could help to identify areas within the already crowded curriculum that may need more (or less) emphasis, or perhaps just a mention to students as an area to watch.

4. Conclusions

In this paper, the global summary of specialized medical terminologies, nomenclatures, coding and classification systems was introduced. We identified the significant problem of selecting and integrating proper vocabulary for computer-processable description in compliance with particular requirements. Based on the scoping review, we adopted the MeSH biomedical dictionary, where the main objective was to classify learning activities appearing in the medical curriculum of General Medicine. Thus, we have integrated MeSH into our original web-based system for innovating and harmonizing the medical curriculum. Currently, this system provides search assistance through a rich set of lexical look-up facilities, which are based on MeSH's complex structure including terms with tree structured contexts. Moreover, two data-analytical reports depicting the results obtained from frequency and association analyses of the particular medical curriculum while employing the MeSH thesaurus are described. They represent the aggregated tabular and graphic outputs including interactive visualizations, which play an essential role during evaluation of the whole curriculum. The main aim is to identify information rich data relations and offer clear and transparent overview for simpler and easier understanding of the curriculum structure. We believe that our approach can be easily generalized and used by any higher education institution without any methodological or technological restrictions.

Conflict of interest

None.

Acknowledgement

The authors were supported from the grants project OPTIMED - OPTImized MEDical education: horizontal and vertical connections, innovations and efficiency in practice reg. no: CZ.1.07/2.2.00/28.0042, which is funded by the European Social Fund and the state budget of the Czech Republic. Special thanks are dedicated to the management of Faculty of Medicine at Masaryk University represented by a vice-dean for education in clinical branches prof. Jaroslav Štérba and dean of the faculty prof. Jiří Mayer.

References

[1] A. Holzinger, Biomedical Informatics: Discovering Knowledge in Big Data, Springer, 2014. ⟨http://www.google.com/books?hl=cs&lr=&id=-FMlBAAAQBAJ&oi=fnd&pg=-PR7&dq=Biomedical+Informatics:+Discovering+Knowledge+in+Big+Da&ots=POzGZuK9Bn&sig=w5mDOXcVJskGqgnzFH-G4L6g_Sc⟩ (accessed 01.22.2015).
[2] J.S. Luciano, B. Andersson, C.R. Batchelor, O. Bodenreider, T. Clark, C.K. Denney, et al., The translational medicine ontology and knowledge base: driving personalized medicine by bridging the gap between bench and bedside., J Biomed. Semant. 2 (2011) S1.
[3] M. Dvořák, Progress of information technology in healthcare, current state, outlook toward future, Mefanet J. 1 (2013) 23–28.
[4] Handbook of Medical Informatics, in: J.H. van Bemmel, M.A. Musen, J.C. Helder (Eds.), Bohn Stafleu Van Loghum ; Springer Verlag, AW Houten, Netherlands : Heidelberg, Germany, 1997.
[5] E. Coiera, Guide to Health Informatics, 2nd ed., Arnold ; Distributed in the USA by Oxford University Press, London : New York, NY, 2003.
[6] S. de Lusignan, Codes, classifications, terminologies and nomenclatures: definition, development and application in practice, Inform. Prim. Care 13 (2005) 65–70.
[7] R. Cornet, N. de Keizer, Forty years of SNOMED: a literature review, BMC Med. Inform. Decis. Mak. 8 (Suppl 1) (2008) S2. http://dx.doi.org/10.1186/1472-6947-8-S1-S2.

[8] D. Lee, N. de Keizer, F. Lau, R. Cornet, Literature review of SNOMED CT use, J. Am. Med. Inform. Assoc. JAMIA 21 (2014) e11–e19. http://dx.doi.org/10.1136/amiajnl-2013-001636.

[9] R.M. Harden, Approaches to research in medical education, Med. Educ. 20 (1986) 521–531. http://dx.doi.org/10.1111/j.1365-2923.1986.tb01394.x.

[10] R.M. Harden, Ten questions to ask when planning a course or curriculum, Med. Educ. 20 (1986) 356–365.

[11] S.A. Wartman, Research in medical education: the challenge for the next decade, Acad. Med. 69 (1994) 608–614. http://dx.doi.org/10.1097/00001888-199408000-00003.

[12] M.H. Davis, R.M. Harden, Planning and implementing an undergraduate medical curriculum: the lessons learned, Med. Teach. 25 (2003) 596–608. http://dx.doi.org/10.1080/0142159032000144383.

[13] R.M. Harden, Learning outcomes as a tool to assess progression, Med. Teach. 29 (2007) 678–682. http://dx.doi.org/10.1080/01421590701729955.

[14] R.M. Harden, AMEE Guide No. 21: curriculum mapping: a tool for transparent and authentic teaching and learning, Med. Teach. 23 (2001) 123–137.

[15] R.H. Ellaway, S. Albright, V. Smothers, T. Cameron, T. Willett, Curriculum inventory: modeling, sharing and comparing medical education programs, Med. Teach. 36 (2014) 208–215. http://dx.doi.org/10.3109/0142159X.2014.874552.

[16] R. Ellaway, V. Smothers, ANSI/MEDBIQ CI.10.1-2013 Curriculum Inventory Specifications, 2013, ⟨http://www.medbiq.org/sites/default/files/files/CurriculumInventorySpecification.pdf⟩.

[17] A. Haig, R. Ellaway, M. Dozier, D. Liu, J. McKendree, METRO-the creation of a taxonomy for medical education, Health Inf. Libr. J. 21 (2004) 211–219. http://dx.doi.org/10.1111/j.1471-1842.2004.00534.x.

[18] A. Haig, M. Dozier, D. Liu, J. Mckendree, T. Roper, C. Selai, METRO taxonomy – progress report on assessment, Med. Teach. 27 (2005) 155–157. http://dx.doi.org/10.1080/01421590500091102.

[19] W.E. Blaum, A. Jarczweski, F. Balzer, P. Stötzner, O. Ahlers, Towards Web 3.0: taxonomies and ontologies for medical education - a systematic review, GMS Z. Med. Ausbild. 30 (1) (2013) Doc13: 9–16. http://dx.doi.org/10.3205/zma000856.

[20] C. Bizer, T. Heath, T. Berners-Lee, Linked data - the story so far, Int. J. Semant. Web Inf. Syst. 5 (2009) 1–22. http://dx.doi.org/10.4018/jswis.2009081901.

[21] E. Méndez, J. Greenberg, Linked data for open vocabularies and HIVE's global framework, El Prof. Inf. 21 (2012) 236–244.

[22] S. Konstantinidis, L. Fernandez-Luque, P. Bamidis, R. Karlsen, The role of taxonomies in social media and the semantic web for health education. A study of SNOMED CT terms in YouTube health video tags, Methods Inf. Med. 52 (2013) 168–179. http://dx.doi.org/10.3414/ME12-02-0005.

[23] P.L. Whetzel, N.F. Noy, N.H. Shah, P.R. Alexander, C. Nyulas, T. Tudorache, et al., BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications, Nucleic Acids Res. 39 (2011) W541–W545.

[24] M. Komenda, D. Schwarz, J. Feberová, S. Štípek, V. Mihál, L. Dušek, Medial faculties educational network: Multidimensional quality assessment, Comput. Methods Progr. Biomed. 108 (2012) 900–909.

[25] D. Schwarz, P. Štourač, M. Komenda, H. Harazim, M. Kosinová, J. Gregor, et al., Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET, J. Med. Internet Res. 15 (2013) e135. http://dx.doi.org/10.2196/jmir.2590.

[26] D. Schwarz, L. Dušek, M. Bareš, Computer applications, Systems and Networks for Medical Education, Masaryk University, Institute of Biostatistics and Analyses, Facta Medica, Brno, 2014.

[27] M. Komenda, D. Schwarz, C. Vaitsis, N. Zary, J. Štěrba, L. Dušek, OPTIMED Platform: Curriculum Harmonisation System for Medical and Healthcare Education, in: Madrid, 2015.

[28] C. Vaitsis, G. Nilsson, N. Zary, Visual analytics in healthcare education: exploring novel ways to analyze and represent big data in undergraduate medical education, PeerJ 2 (2014) e683. http://dx.doi.org/10.7717/peerj.683.

[29] C. Vaitsis, G. Nilsson, N. Zary, Big data in medical informatics: improving education through visual analytics, Stud. Health Technol. Inf. 205 (2014) 1163–1167.

[30] M. Komenda, D. Schwarz, Visual analytics in environmental research: a survey on challenges, methods and available tools, in: J. Hřebíček, G. Schimak, M. Kubásek, A.E. Rizzoli (Eds.), Environmental Software Systems Fostering Information Sharing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 618–629, accessed 25.02.2015.

[31] H. Arksey, L. O'Malley, Scoping studies: towards a methodological framework, Int. J. Soc. Res. Methodol. 8 (2005) 19–32.

[32] D. Levac, H. Colquhoun, K.K. O'Brien, Scoping studies: advancing the methodology, Implement. Sci. 5 (2010) 69. http://dx.doi.org/10.1186/1748-5908-5-69.

[33] H.M. Daudt, C. van Mossel, S.J. Scott, Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework, BMC Med. Res. Methodol. 13 (2013) 48. http://dx.doi.org/10.1186/1471-2288-13-48.

[34] C. Adelman, The Bologna Process for US Eyes: Re-Learning Higher Education in the Age of Convergence, Institute for Higher Education Policy, Washington, DC, 2009, accessed 22.01.2015.

[35] D.A. Lindberg, B.L. Humphreys, A.T. McCray, The unified medical language system, Methods Inf. Med. 32 (1993) 281–291.

[36] J. Anderson, The computer: medical vocabulary and information, Br. Med. Bull. 24 (1968) 194–198.

[37] O. Bodenreider, The unified medical language system (UMLS): integrating biomedical terminology, Nucleic Acids Res. 32 (2004) D267–D270.

[38] Unified Medical Language System (UMLS), (n.d.), ⟨http://www.nlm.nih.gov/research/umls/⟩ (accessed 22.01.2015).

[39] Performance evaluation of unified medical language system's synonyms expansion to query PubMed, BMC Med. Inform. Decis. Mak. 12, 2012 12–17. http://dx.doi.org/10.1186/1472-6947-12-12.

[40] Fact SheetMedical Subject Headings (MeSH®), (n.d.). ⟨http://www.nlm.nih.gov/pubs/factsheets/mesh.html⟩ (accessed 12.07.2012).

[41] D. Rotolo, L. Leydesdorff, Matching MEDLINE/PubMed data with web of science (WoS): a routine in R language, J. Assoc. Inf. Sci. Technol. (2014), http://dx.doi.org/10.1002/asi.23385.

[42] A.P. Davis, T.C. Wiegers, M.C. Rosenstein, C.J. Mattingly, MEDIC: a practical disease vocabulary used at the Comparative Toxicogenomics Database, Database: The Journal of Biological Databases and Curation 2012 (2012) bar065, http://dx.doi.org/10.1093/database/bar065.

[43] W. ICD, Classification of mental and behavioural disorders, Edinb. Lond. Melb. N. Y. Tokyo Churchill Livingstone, 1991.

[44] F. Gjertsen, S. Bruzzone, M.E. Vollrath, M. Pace, Ø. Ekeberg, Comparing ICD-9 and ICD-10: the impact on intentional and unintentional injury mortality statistics in Italy and Norway, Injury 44 (2013) 132–138. http://dx.doi.org/10.1016/j.injury.2012.01.010.

[45] P. Ruch, J. Gobeill, C. Lovis, A. Geissbühler, Automatic medical encoding with SNOMED categories, BMC Med. Inform. Decis. Mak. 8 (2008) S6. http://dx.doi.org/10.1186/1472-6947-8-S1-S6.

[46] P.L. Elkin, S.H. Brown, C.S. Husser, B.A. Bauer, et al., Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists, Mayo Clin. Proc. 81 (2006) 741–748.

[47] C. Fluit, M. Sabou, F. van Harmelen, Ontology-Based Information Visualization: Toward Semantic Web Applications, in: V.G.Ds. MSc, C.C. BSc MSc (Eds.), Vis. Semantic Web, Springer London, 2006, pp. 45–58. ⟨http://link.springer.com/chapter/10.1007/1-84628-290-X_3⟩ (accessed 22.01.2015).

[48] T. Bekhuis, D. Demner-Fushman, R. Crowley, Comparative effectiveness research designs: an analysis of terms and coverage in Medical Subject Headings (MeSH) and Emtree, J. Med. Libr. Assoc. 101 (2013) 92–100. http://dx.doi.org/10.3163/1536-5050.101.2.004.

[49] M. Taboada, R. Lalin, D. Martinez, An automated approach to mapping external terminologies to the UMLS, IEEE Trans. Biomed. Eng. 56 (2009) 1598–1605. http://dx.doi.org/10.1109/TBME.2009.2015651.

[50] C.P. Langlotz, RadLex: a new method for indexing online educational materials 1, Radiographics 26 (2006) 1595–1597.

[51] M.W. Shore, D.L. Rubin, C.E. Kahn Jr, Integration of imaging signs into RadLex, J. Digit. Imaging 25 (2012) 50–55.

[52] Dublin Core Metadata Element Set, Version 1.1: Reference Description, (n.d.). ⟨http://dublincore.org/documents/1999/07/02/dces/⟩.

[53] X. Ochoa, J. Klerkx, B. Vandeputte, E. Duval, On the use of learning object metadata: the GLOBE experience, in: Proceedings of the Ubiquitous Learn., Springer, 2011, pp. 271–284. http://link.springer.com/chapter/10.1007/978-3-642-23985-4_22 (accessed 18.02.2015).

[54] A. Maratea, A. Petrosino, M. Manzo, Automatic generation of SCORM compliant metadata for portable document format files, in: Proceedings of the 13th International Conference on Computer Systems and Technologies, ACM, 2012, pp. 360–367. ⟨http://dl.acm.org/citation.cfm?id=2383328⟩ (accessed 18.02.2015).

[55] K. Beck, C. Andres, Extreme Programming Explained: Embrace Change, Addison-Wesley Professional, 2004 ⟨http://dl.acm.org/citation.cfm?id=1076267⟩, accessed 21.03.2014.

[56] N.C. for B. Information, U.S.N.L. of M. 8600R. Pike, B. MD, 20894 USA, PubMed Help, (2014). ⟨http://www.ncbi.nlm.nih.gov/books/NBK3827/⟩ (accessed 22.01.2015).

Original Paper

Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice

Martin Komenda[1,2], PhD; Vojtěch Bulhart[1,2]; Matěj Karolyi[1,2], MSc; Jiří Jarkovský[1,2], PhD; Jan Mužík[1,2], PhD; Ondřej Májek[1,2], PhD; Lenka Šnajdrová[1,2], PhD; Petra Růžičková[1,2], MSc; Jarmila Rážová[3], PhD; Roman Prymula[3], Prof Dr; Barbora Macková[4], PhD; Pavel Březovský[4], PhD; Jan Marounek[3], MBA, MD; Vladimír Černý[5,6], Prof Dr; Ladislav Dušek[1,2], Prof Dr

[1]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[2]Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic
[3]Ministry of Health of the Czech Republic, Prague, Czech Republic
[4]National Institute of Public Health, Prague, Czech Republic
[5]Department of Anesthesiology, Perioperative Medicine and Intensive Care, Masaryk Hospital, Ústí nad Labem, Czech Republic
[6]Jan Evangelista Purkyne University, Ústí nad Labem, Czech Republic

Corresponding Author:
Martin Komenda, PhD
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Kamenice 3
Brno
Czech Republic
Phone: 420 725174940
Fax: 420 549492855
Email: komenda@iba.muni.cz

Abstract

Background: The beginning of the coronavirus disease (COVID-19) epidemic dates back to December 31, 2019, when the first cases were reported in the People's Republic of China. In the Czech Republic, the first three cases of infection with the novel coronavirus were confirmed on March 1, 2020. The joint effort of state authorities and researchers gave rise to a unique team, which combines methodical knowledge of real-world processes with the know-how needed for effective processing, analysis, and online visualization of data.

Objective: Due to an urgent need for a tool that presents important reports based on valid data sources, a team of government experts and researchers focused on the design and development of a web app intended to provide a regularly updated overview of COVID-19 epidemiology in the Czech Republic to the general population.

Methods: The cross-industry standard process for data mining model was chosen for the complex solution of analytical processing and visualization of data that provides validated information on the COVID-19 epidemic across the Czech Republic. Great emphasis was put on the understanding and a correct implementation of all six steps (business understanding, data understanding, data preparation, modelling, evaluation, and deployment) needed in the process, including the infrastructure of a nationwide information system; the methodological setting of communication channels between all involved stakeholders; and data collection, processing, analysis, validation, and visualization.

Results: The web-based overview of the current spread of COVID-19 in the Czech Republic has been developed as an online platform providing a set of outputs in the form of tables, graphs, and maps intended for the general public. On March 12, 2020, the first version of the web portal, containing fourteen overviews divided into five topical sections, was released. The web portal's primary objective is to publish a well-arranged visualization and clear explanation of basic information consisting of the overall numbers of performed tests, confirmed cases of COVID-19, COVID-19-related deaths, the daily and cumulative overviews of people with a positive COVID-19 case, performed tests, location and country of infection of people with a positive COVID-19 case, hospitalizations of patients with COVID-19, and distribution of personal protective equipment.

Conclusions: The online interactive overview of the current spread of COVID-19 in the Czech Republic was launched on March 11, 2020, and has immediately become the primary communication channel employed by the health care sector to present the current situation regarding the COVID-19 epidemic. This complex reporting of the COVID-19 epidemic in the Czech Republic also shows an effective way to interconnect knowledge held by various specialists, such as regional and national methodology experts (who report positive cases of the disease on a daily basis), with knowledge held by developers of central registries, analysts, developers of web apps, and leaders in the health care sector.

(J Med Internet Res 2020;22(5):e19367) doi: 10.2196/19367

KEYWORDS

coronavirus disease; COVID-19; Czech Republic; web app; interactive reporting; epidemiological overview; CRISP-DM; public health; app; epidemiology; virus; health data; data mining; modeling

Introduction

In early 2020, the pandemic of the coronavirus disease (COVID-19) started to spread all over the world. COVID-19 is caused by a novel type of coronavirus, referred to as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This highly infectious disease is mainly manifested by fever, respiratory difficulty (cough, dyspnea), muscle pain, and fatigue. The disease can be rather serious for people who are older or have chronic illnesses and can even be fatal. The beginning of the COVID-19 epidemic dates back to December 31, 2019, when the first cases were reported in the city of Wuhan in Hubei Province in the People's Republic of China. Over the next 4 months, more than 1,120,000 people became infected across the world and almost 60,000 patients died from the disease [1]. In the Czech Republic, the first 3 cases of infection with the novel coronavirus were confirmed on March 1, 2020. A team of experts from the Institute of Health Information and Statistics of the Czech Republic (IHIS CR), together with researchers from the Institute of Biostatistics and Analyses at the Faculty of Medicine of the Masaryk University, focused on the design and development of a web app intended to provide a regularly updated overview of COVID-19 epidemiology in the Czech Republic to the general public. The joint effort of state authorities and researchers gave rise to a unique team, which combines methodical knowledge of real-world processes with the know-how needed for effective processing, analysis, and online visualization of data. The requirements on functionalities of this tool were mainly defined by the seriousness of an unexpected situation; an urgent need emerged for a tool that would make it possible to present important reports based on valid data sources only. To achieve this goal, it was necessary to ensure that individual graphs, maps, and tables would be easy to understand and could be correctly interpreted by the public and the media, and that misinterpretation of outputs would be avoided. An extensive review of tools available abroad was a valuable input for the development of a Czech tool. Worldwide, most papers published until recently [2-10] have been focused on the research of populations affected by COVID-19, on the structure of SARS-CoV-2 and its comparison with similar viruses (Middle East respiratory syndrome-related coronavirus and severe acute respiratory syndrome-related coronavirus), and on treatment and the overall mortality of COVID-19. From the technological point of view, several articles have been dedicated to the issue of data collection and sharing, together with their online presentation [1,11-15]. There are several examples of portals documenting the COVID-19 epidemiology clearly on a nationwide level, such as the Icelandic COVID-19 in Iceland – Statistics; the Korean Coronavirus Disease-19, Republic of Korea; or the Singaporean Dashboard of the COVID-19 Virus Outbreak in Singapore.

The aim of the research team was to adopt one of the time-tested methodologies for data mining, analytics, knowledge discovery, and data science projects, and to apply it in the process of mapping the current COVID-19 epidemic situation in the Czech Republic. This paper describes all essential steps from methodological as well as technical points of view. The Czech approach to the design, development, and implementation of online monitoring of the COVID-19 epidemic is based on a verified methodology for the acquisition, processing, and presentation of information. The methodology discussed in this paper made it possible to interconnect knowledge held by various specialists such as regional and national methodology experts from the National Institute of Public Health and regional public health authorities, who report positive cases of the disease on a daily basis, with knowledge held by developers of central registries, analysts, developers of web apps, and leaders in the health care sector.

Methods

Methodological Background for the COVID-19 App: An Overview of the Current Situation in the Czech Republic

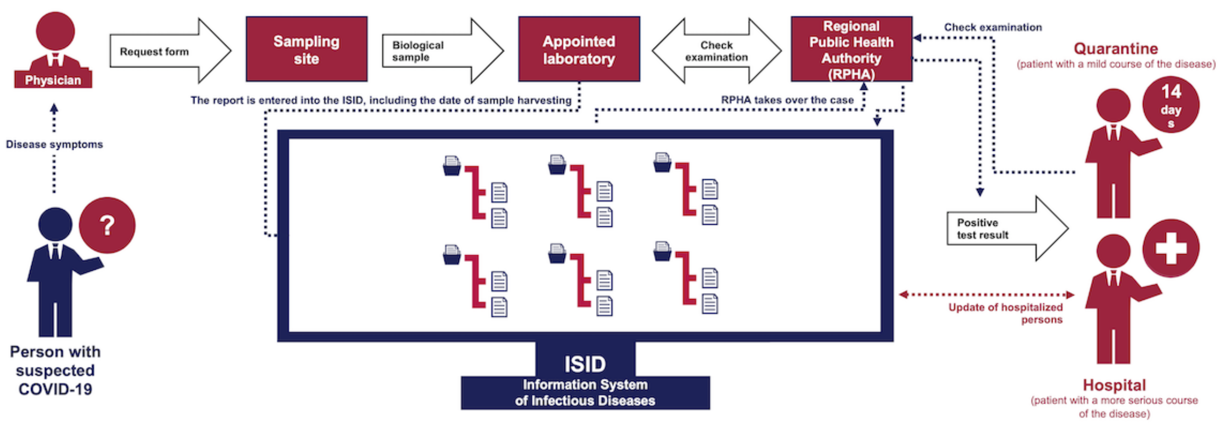
When the urgent need emerged to map the current situation regarding the COVID-19 epidemic, it was essential to consider several factors that might have a significant impact on the resulting reports. The infrastructure of a nationwide information system run by the public administration is one of the most important factors in this regard; data from basic registries as well as data from health service providers are collected and processed in this information system. The cooperation between organizations involved in the process of data reporting, collection, processing, validation, analysis, evaluation, and visualization is another key aspect. The methodological setting of communication channels, duties resulting from legal measures, clearly defined competence, the sequence of steps to be made, and the overall management of the health care sector are complex at the time of an unexpected pandemic. In addition, it is important to select appropriate procedures for knowledge mining from database structures and for an undistorted interpretation of data provided to individual target groups. These groups involve not only the general public and health care professionals but also the media, as the resulting reports are published on behalf of the Ministry of Health of the Czech Republic and can, therefore, be considered as guaranteed and entirely reliable.

Among several known methodological recommendations and standardized procedures, which can be cited as the implementation of the knowledge discovery in databases, the process Sample, Explore, Modify, Model, Assess and cross-industry standard process for data mining (CRISP-DM) are used most frequently in practice [16]. The team of authors chose the latter methodology (CRISP-DM) due to its higher versatility. The individual stages of this methodology directly correspond to the complex solution of analytical processing and visualization of data, which provides validated information on the COVID-19 epidemic across the Czech Republic. The power of CRISP-DM is demonstrated by the fact that great emphasis is put on the understanding and on a correct implementation of all six steps needed in the process (see Multimedia Appendix 1); furthermore, CRISP-DM provides the option to return back to previous steps, and importantly, it does not leave out the frequently omitted process of checking the achieved results before their publication.

Business Understanding

In the initial stage, maximum attention is paid to mapping the situation from the managerial point of view. In this case, the task is focused on online visualization of data on the current state of the COVID-19 epidemic in the Czech Republic. The Ministry of Health of the Czech Republic is responsible for the methodological setting of regular reports and processing data on newly identified cases of COVID-19 across the Czech Republic (see Figure 1), record keeping on testing locations, and overviewing the purchases and distribution of personal protective equipment (PPE). The reporting process always starts with a person with suspected COVID-19 being referred by a physician or a regional public health authority (RPHA) to a testing location, where a biological sample is taken from them, in compliance with the RPHA methodology. The sample is then analyzed by one of the appointed laboratories, which determines whether or not it is positive for SARS-CoV-2. The result is subsequently entered into the central Information System of Infectious Diseases (ISID), which is then taken over by the respective RPHA. In the next step, the same RPHA carries out a second investigation to verify the result, informs the person about the result, and, if the sample was confirmed to be positive for SARS-CoV-2, provides information about the next steps to be taken. The systematically designed architecture of the National Health Information System (NHIS) made it possible for the result of each performed test to be processed by a laboratory and verified by the respective RPHA in the central ISID, both procedures being done with only a minimum delay. The primary objective of the ISID is to obtain information on the incidence of infectious diseases to assess the epidemiological situation across the Czech Republic, to monitor the population's health status, and to control the provision of health care. In compliance with section 70, paragraph 3 of the Act No 372/2011 Coll, on Health Services and Conditions of Their Provision (Act on Health Services), the administration of NHIS has been delegated to the IHIS CR. NHIS is a fully computerized system involving components that are enshrined in the legislation. Each person in the Czech Republic can be unequivocally identified based on their birth certificate number; based on this number, essential links can be found between the NHIS and other relevant databases run by the public administration, such as the National Register of Hospitalized Patients, the Registry of Inhabitants, the National Register of Health Services Providers, the National Register of Health Care Professionals, or the Death Records Database. Data from these registries provide a comprehensive—and, most of all, up-to-date—data basis for subsequent analytical processing. On top of that, the unequivocal identification of patients, which is identical across these registries, made it possible to obtain information on the infection rate among health care workers in real time, for example.

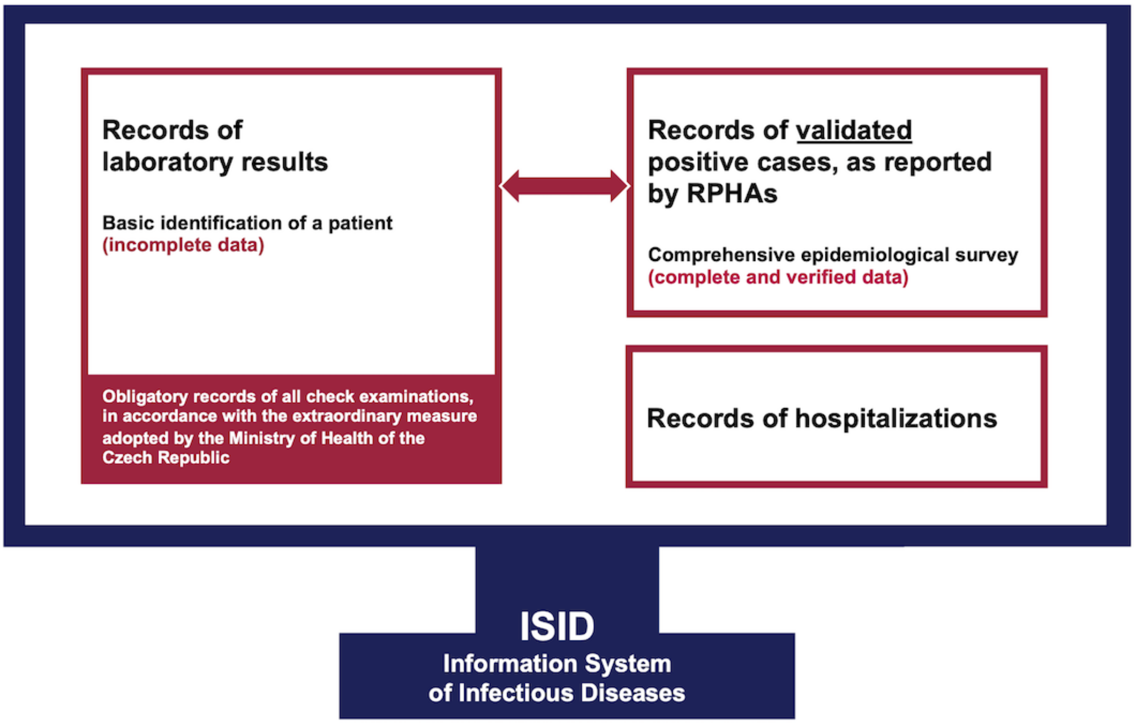
Figure 1. Simplified schema of newly identified cases of COVID-19. COVID-19: coronavirus disease.



Data Understanding

For a correct collection and processing of information on the COVID-19 epidemiology in the Czech Republic, it is essential that all reporting processes are well understood, that communication with all involved institutions runs smoothly, that the architecture of central databases is well designed from the technical point of view, and that the resulting reports are correctly interpreted. At the same time, all interactive outputs, whether in the form of graphs, maps, or tables, which are available for the general public, must meet the condition that it will never lead to a direct identification of any patient. It is therefore not possible to show detailed information on a district (or even a town) in combination with sex or age of a given person. What can be published, on the other hand, are summary data on the entire Czech Republic (or possibly on individual regions or districts), as the population of the entire country (or region or district) is large enough and the total number of positive cases is higher than 10, which cannot lead to the direct identification of a particular patient. Other examples of outputs that can be published involve daily reports on cumulative numbers and on the increase or decrease in the number of confirmed cases or division according to sex or age groups. When processing personal data in the various part of NHIS, every individual record must be processed in accordance with Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Act No 372/2011 Coll, on Health Services and Conditions of Their Provision (Act on Health Services), as subsequently amended, and the Act No 110/2019 Coll, on the Processing of Personal Data, as subsequently amended. The interactive online reporting is based on the five following data sets, which, combined together, characterize the COVID-19 epidemic in the Czech Republic. First, reports by RPHAs contain daily records of persons with confirmed COVID-19 (eg, date of report, age, sex, region, location, and country of infection). Second, reports by laboratories (LAB) contain deidentified records on persons with confirmed COVID-19, which have not yet been taken over by a RPHA in the respective region. In relation to reports by RPHAs, these are disjointed sets of records (each particular record goes to the LAB repository and consequently is either approved or directly confirmed by a respective RPHA); in other words, none of the records are present in both data sets, and undesirable duplicates are, therefore, avoided. Figure 2 shows how records provided by RPHAs and LABs are transferred into ISID. Among necessary adjustments, patient identification against the Registry of Inhabitants is performed at this stage, aiming to determine the region of a citizen who has been confirmed as a positive case by the laboratory. Third, a report on performed COVID-19 tests contains the number of all samples tested by laboratories across the Czech Republic on individual days. Fourth, reports on persons hospitalized in health care facilities contain daily summaries of currently hospitalized persons, persons in a serious condition or receiving highly intensive care (eg, mechanical ventilation, extracorporeal membrane oxygenation), and hospitalized persons who have been cured or discharged to home quarantine. Fifth, a report by the Ministry of Health of the Czech Republic contains purchases and distribution of PPE, and provides up-to-date numbers of face masks, respirators, goggles, bottles with disinfectant, face shields, and other equipment.

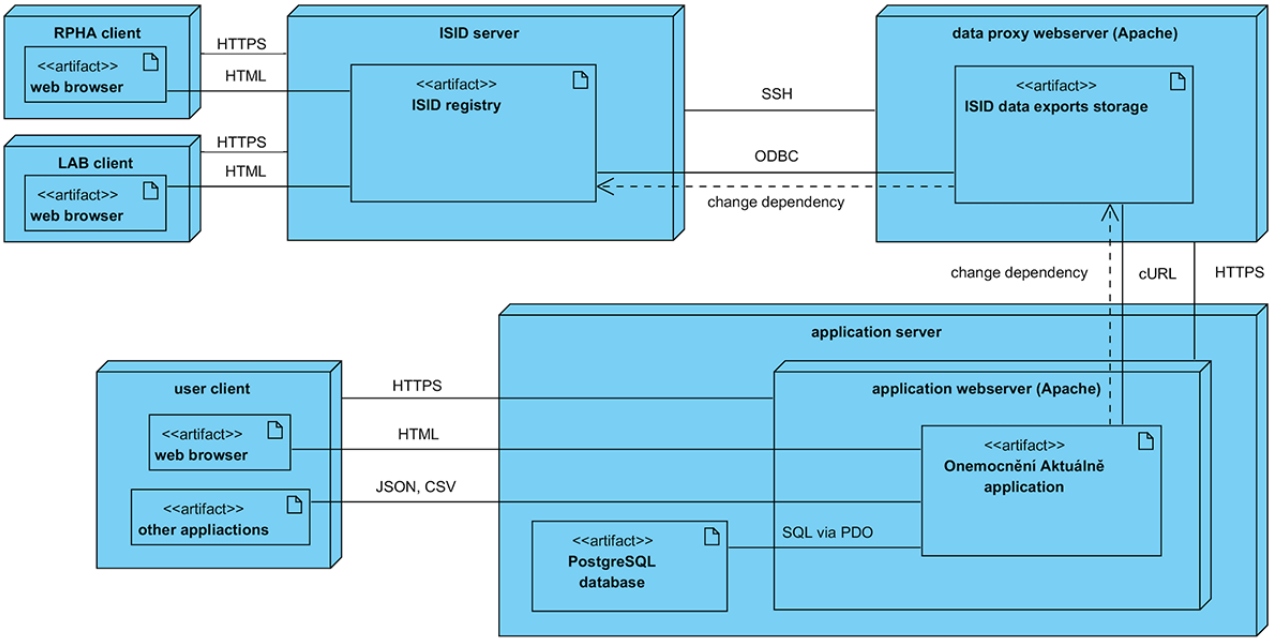
Figure 2. Diagram of data transfer within the Information System of Infectious Diseases.



Data Preparation

Several different data sources are used as input at this stage; these data sources must be validated thoroughly and then transformed into a form suitable for computer processing and subsequent visualization. In particular, the validation process involves a set of control mechanisms ensuring the completeness of individual records. This stage involves data cleaning (incomplete records are detected, corrected, or removed), construction of particular data views (selected attributes with primary identification are derived and merged), and data integration (final data sets are automatically generated). This process ensures that no invalid records are included in the stage of further processing and web visualization. Syncing tests, which are launched at regular intervals, are among the main control mechanisms. These tests provide information on whether or not data presented in the end report are consistent with data on the input. Their objective is to detect possible inconsistencies in the input records, such as a patient's incomplete or incorrectly entered birth certificate number; their permanent address; or distinction between a Czech citizen, a foreigner, or even a person who is homeless. If possible, standardized lists of values (ie, those valid on a nationwide level) are integrated, making it possible to name values in a unified way. As an example, a standardized classification of territorial units in the Czech Republic has been used (Nomenclature of Units for Territorial Statistics) based on unique codes for each of the 14 Czech regions (eg, CZ064 was used as a code for the South Moravian Region). Valid records in the ISID form the basis for the selection of descriptive attributes and their final adaptation into a format in which they are sent by export tools at regular intervals and in a secured manner to the web server. Figure 3 shows a diagram of implementation of individual components, showing the communication among three independent servers transmitting data sources that are necessary to draw the final visual outputs and to provide open data sets to the public.

Figure 3. Deployment diagram. CSV: comma-separated values; cURL: Client URL; ISID: Information System of Infectious Diseases; JSON: JavaScript Object Notation; LAB: laboratories; ODBC: Open Database Connectivity; PDO: PHP Data Objects; RPHA: regional public health authority; SQL; Structured Query Language; SSH: Secure Shell.



Modelling

In this step, detailed static instructions were prepared for each report (graph, map, table), always involving its data source, computational algorithm, description of visualization, explanation of its meaning, and information on the last update. Afterward, the entire process of data processing and dynamic drawing of the online report was transformed into a fully automatic one. All predefined data sets are transferred into an internal data storage device, which is dedicated to epidemiological analyses and reporting; individual overviews are subsequently generated in the form of interactive graphs, maps, and data tables. The resulting presentation is a set of interactive graph visualizations and tabular outputs, which can be viewed online by anyone using only a standard web browser. The portal has been developed using the Symfony 4.4 PHP framework. Graphical outputs are processed by the NVD3 JavaScript library, which is based on d3.js. Graphs are slightly adjusted to meet the report's needs, mainly in the responsive environment of the contemporary imaging technology. Tabular data are subsequently processed using the DataTables library, which internally employs jQuery (The jQuery Team). In this case, the data layer mainly consists of preprocessed and saved files in the JavaScript Object Notation (JSON) format. This approach significantly accelerates the access to data, making it possible to accelerate the app's response markedly. On each update, data are transformed straight away, source files being replaced by new ones. The number of computational operations that are carried out at the app layer is kept at a minimum. The presentation layer deals with the graphical appearance of the user interface with control elements as well as the interactivity of the visualization. All reports prepared in this way are first implemented on development servers, where functionality and data correctness are thoroughly verified. Development servers are also used to check syncing with external data sources. Individual versions are subsequently published, always within the logical structure of the web app, which is divided into thematic sections. The app itself has been designed to be responsive and is fully supported by all types of devices, including the portable ones (mobile, tablet, and desktop).

Evaluation of Results

It is extremely important to validate the results before their publication, particularly when providing information about the spread of an epidemic. Despite several optimization and transformation processes performed in the preceding stages, data must remain consistent and all reports must correspond with the original input as well as with calculations carried out to check agreement with primary data. User testing not only revealed some inconsistency in original values obtained by static calculations, but also brought new ideas on how to improve the intelligibility, technical implementation, and user interface. Experts on accessibility of web apps were called in too. The target group was also taken into consideration, as it consists of disabled users, among others, for whom navigation on the internet might be difficult. A correct structure of headings, an adequate contrast between texts and the background, the overall legibility, a properly defined alternative information for images, and the availability of complementary tabular overviews next to graphs and maps are the most important accessibility attributes that have been thoroughly applied. One must keep in mind that web apps guaranteed by the state administration are required to provide maximum accessibility for all citizens without exception.

Deployment of Results

Primary data are entered into the ISID by LAB and RPHAs. After their validation and the unification of data formats, two data exports are created, and these are periodically sent from the ISID to a location from where they can be finally used for the purposes of the web portal. The periodicity of updates is set to 2 hours, with a nighttime shutdown between 2 am and 6 am The cron tool launches the updating script on the side of the "Disease at the Moment" ("Onemocnění aktuálně" in Czech) app at predefined times, which results in the transformation of provided data into preprocessed data sets. The previously mentioned syncing tests are launched during this process as well. The entire process is launched on the production instance of the "Disease at the Moment" app, as well as on the development instance and the stage instance. The latter two instances are automated by Jenkins, which is a continuous integration tool. In case of any error in data assembly or in the update of the entire app, an email notification is sent to the development team, allowing it to react appropriately. At the time of writing this paper (April 2020), primary data are updated three times a day on the production instance (12:30 am, 8:30 am, and 5:30 pm), and secondary data are usually updated once a day.

Open Data for COVID-19

Science is built on data, namely their collection, analysis, publication, reanalysis, critique, and reuse. Barriers include inability to access data, restrictions on usage applied by publishers or data providers, and publication of data that are difficult to reuse, for example, because they are poorly annotated or "hidden" in unmodifiable tables like PDF documents [17]. For that reason, the concepts of open access and open data are strongly emphasized and play a key role in the complex web-based reporting of the COVID-19 epidemic in the Czech Republic. The Ministry of the Interior of the Czech Republic guarantees and maintains the National Catalogue of Open Data, which consists of 24 local instances, including the Catalogue of Open Data run by the Ministry of Health of the Czech Republic. The main goal is to collect metadata about data sets published as open data throughout the whole country and to show transparency and effectiveness of government services [18]. There are three major target groups of potential users who can access various data sets freely: (1) those who are interested in information describing the entire concept together with benefits of broad usage of open data sets; (2) those who want to publish and update various data sets in accordance with the given open data rules, such as diagrams, data formats, keywords, and metadata description; and (3) those who can browse and freely use available data sets for the purpose of further analysis and development. The authors of this paper decided to adopt a standardized methodology for the publication of open data, and thus facilitate the use of COVID-19 epidemiological data sets by different public sector bodies, academic institutions, and business sector companies.

Google Analytics as the Monitoring Service

Monitoring and further analysis of web usage is one of the crucial points in terms of reflecting users' behavior and requirements. Systematic tracking and web analysis significantly improve the efficiency and quality during a long-term design and development of robust web apps. Google Analytics is a third-party service that measures and generates up-to-date statistics, reports, and analyses based on website traffic and on the behavior of its visitors. It tracks the visitors' activity, collects statistical data in real time, and stores them for a later analysis. Using different types of metrics, we can easily determine the number of visitors over any period, which pages they viewed, and how long was their visit. The flow of visitors is an important factor, showing user transitions between pages and the rate of abandonment of every single page [19]. Based on Google Analytics' powerful features, advanced visualizations mapping the analytics intelligence, dashboard, mobile device tracking, referrers, and geographic tracking capabilities have been used [20].

Results

Description of the COVID-19 App: An Overview of the Current Situation in the Czech Republic

On March 11, 2020, the first version of the web portal was released [21]. It provides a set of outputs in the form of tables, graphs, and maps intended for the general public and the media. Its primary objective is to provide a well-arranged visualization and clear explanation of basic information included in the basic overview of COVID-19 epidemiology in the Czech Republic (see Multimedia Appendix 2). The overall numbers of performed tests, confirmed cases of COVID-19, persons who have recovered from the disease, and COVID-19-related deaths are displayed first to the user. The rest of the webpage is divided into several sections with coherent topics, each of them providing a synoptic overview from a selected perspective. First, basic overviews according to reports by RPHAs and positive results from laboratories show the overall number of persons who were tested positive for COVID-19, the incidence of confirmed cases of COVID-19 by region per 100,000 people, absolute numbers of people with positive COVID-19 cases by region, the overall number of performed tests (including repeated tests in the same person) for COVID-19 across the entire Czech Republic, and an overview of COVID-19-related deaths by age group and by region. Second, daily overviews according to reports by RPHAs and positive results from laboratories show a daily overview of number of persons with newly confirmed COVID-19, a daily overview of the number of performed tests (including repeated tests in the same person), the overall (cumulative) number of persons with laboratory-confirmed COVID-19 and the daily percentage change, and the daily percentage of persons with confirmed COVID-19 in the number of performed tests on a given day. Third, the number of persons with laboratory-confirmed COVID-19 according to reports by RPHAs shows an overview of cumulative numbers of persons with laboratory-confirmed COVID-19 that have been verified by RPHAs (not the number of all persons with laboratory-confirmed COVID-19), the location and country of infection of people with a positive COVID-19 case, and the number of persons with laboratory-confirmed COVID-19 by sex and age group. Fourth, an overview of hospitalizations of patients with COVID-19 shows the current number of hospitalized persons, the number of persons in a serious condition or receiving highly intensive care, and the number of hospitalized persons who have been cured or discharged to home quarantine. Fifth, an overview of the distribution of PPE shows how much PPE has been purchased and distributed by the Ministry of Health of the Czech Republic across individual regions.

Because published data are based on several independent sources, it is obvious that updates must be performed at different times. Nevertheless, fixed time intervals have been set at which updated values are presented on the web. The latest numbers summing up the incidence of COVID-19 in the Czech Republic are updated 3 times a day: at 12:30 am, at 8:30 am, and at 8:30 pm. The overviews showing total numbers from the previous day are prepared each morning at 8:30 am, and the latest numbers of recovered persons and COVID-19-related deaths are published at 8:30 am and at 5:30 pm.

Types of Visualization

In the stages of design and development of this web app, particular emphasis was placed on the character of presented data, which might be viewed by users anywhere and anytime. That was one of the reasons why responsive web design was among the main requirements, together with the overall optimization for mobile devices. The majority of line graphs and bar charts, therefore, primarily display an overview from the last 14 days, which is always adjusted even to small screens of mobile phones without the user having to manipulate the graph in any way. An additional complex view of the entire period of epidemic follow-up means that users of smaller devices may need to manipulate the graph if they want to display all values (Figures 4 and 5). A tabular overview is the last option available, displaying all values in rows and columns of a table. Other types of charts are also available on the website, such as the pie chart, stacked bar chart, or grouped bar chart, as well as map visualization and standard tables. Moreover, all published data are available to download in an open data format.

Figure 4. Daily trend in the percentage of persons with confirmed coronavirus disease in the overall number of persons tested on a given day.

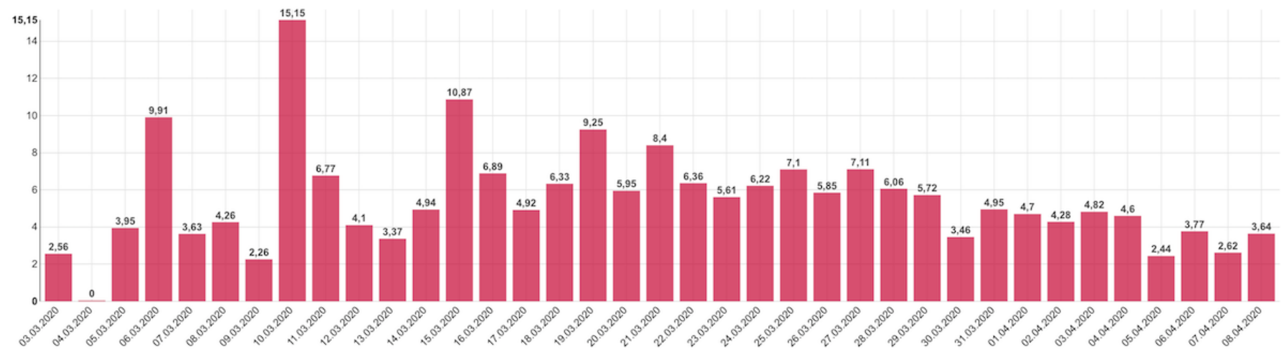
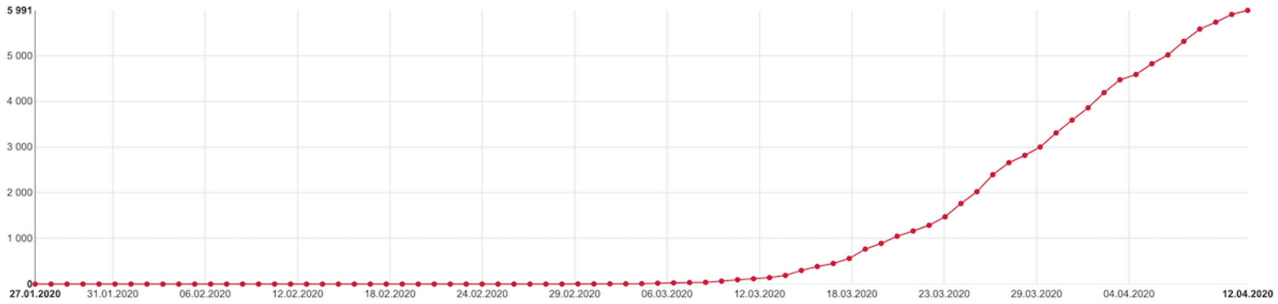


Figure 5. Overall (cumulative) number of persons with laboratory-confirmed coronavirus disease according to reports by regional public health authorities and laboratories.



Open Data Sets

Open data sets intended for further processing are an integral part of this information website. Source data in .CSV (comma-separated values) and .JSON formats are published daily at regular intervals, and can be downloaded by anyone who wants to process them either by a computer or manually. Data set schema are also included, describing the structure of records. Normalized schemas contain the list and definitions of all descriptive attributes as well as the technological representation of the data schema. All data sets published in this way are linked to records in the National Catalogue of Open Data (provided by the Ministry of Health of the Czech Republic), which is administered by the IHIS CR.

Analysis of the Number of Visits

The Google Analytics component is used to monitor and to analyze the users' behavior on the website. The objective of the development team was to monitor all basic characteristics such as the number of sessions, page views, ways of user acquisition (direct, organic, referral), type of device used (mobile, tablet, desktop), display resolution, and web browser. An overview of sessions on a given day and time makes it possible for the development team to plan timely releases of new versions of the portal and to update information. The website was officially launched on March 11, 2020, and has immediately become the primary communication channel employed by the health care sector to present the current situation regarding the COVID-19 epidemic. In the period between March 11, 2020, and April 12, 2020, the web portal registered 13,634,325 sessions and 19,662,351 page views. Returning visitors accounted for 16,286,551 (more than 80%) of users. This trend can also be confirmed by the graph showing how visitors landed on the website. A total of 7,158,020 (more than 50%) of visits were direct (ie, the website URL was typed into a browser) or were the result of an organic search (ie, users employed search engines such as Seznam.cz or Google.com). From the beginning, the development team designed the website as mobile-first. With regard to the portal contents, we assumed that its visitors would want to see the information fast and at any time. A total of 8,248,766 (more than 60%) of visitors used a mobile device or a tablet. A focus on graphs being drawn on displays with a certain resolution was no less important. The most usual width of mobile devices was 360 px, which comfortably accommodates a summary graph containing information from the last 14 days. Despite the clear preference of mobile devices, we also had to bear in mind that all graphs had to be well displayed in desktop browsers. Although Chrome was the most frequently used browser (8,112,423 of users, almost 60%), we also had to consider that some users prefer Internet Explorer (version 11.0 or lower), which accounted for more than 160,000 sessions. We used the Google Data Studio tool to create a dashboard that presents all needed information (see Multimedia Appendix 2) and thus provides a clear and easily available report. After the publication of open data sets and of a publicly available application programming interface on March 28, 2020, more than 100,000 page views were recorded over the next few days.

Discussion

The web-based app introducing an overview of the current spread of COVID-19 in the Czech Republic has been designed, developed, and implemented in accordance with the CRISP-DM methodology. All interactive graphs, maps, and tables fully respect strict rules of data management in the health care sector, where data reporting, collection, processing, validation, analysis, evaluation, and final publishing are under the supervision of the Ministry of Health of the Czech Republic. The online interactive overview of the current spread of COVID-19 in the Czech Republic [21] provides comprehensive information to the general public in a well-arranged manner. Since the launch of the first version of the website, the development team has not only systematically collected and evaluated suggestions on improvements from the general public but has also responded to the needs of leadership in the health care sector and of the media. Selected requirements (ie, those that are not contrary to legislation on personal data protection and that do not lead to a direct identification of an individual) are subsequently implemented and released in the next version. At the beginning of March 2020, the epidemiological situation in the Czech Republic was deemed rather serious, so many measures have been put in place aiming to curb the epidemic on a nationwide level and as effectively as possible. This is why an entire family of web apps focusing on COVID-19 has been under development. In addition to the existing overview of the current situation in the Czech Republic, two more online systems are planned to be launched in a short time (ie, in the next few weeks): (1) an epidemiological portal on COVID-19, which will present the descriptive demography, mathematical prediction models, overviews of incidence, prevalence and mortality, the R coefficient, and other relevant information; and (2) an online control room for intensive care, containing the latest reporting on occupancy and availability of beds in real time, including an interface for a quick entry of the currently free capacity for patients with COVID-19 positive cases versus patients with COVID-19 negative cases. All three online tools have been primarily designed with the objective to keep everyone across the Czech Republic informed and to provide objective, data-based views for further decisions made by the leadership of the health care sector and by the emergency committee dealing with the COVID-19 epidemic in the Czech Republic.

Acknowledgments

This app was directly supported by the leadership of the Ministry of Health of the Czech Republic and the Institute of Health Information and Statistics of the Czech Republic. In particular, we would like to thank all regional public health authorities, laboratories, and hospitals for their enormous commitment during the COVID-19 epidemic.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Diagram of the cross-industry standard process for data mining reference model.
[PNG File , 166 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Selected visualization of coronavirus disease epidemiology in the Czech Republic. A: incidence of confirmed cases of coronavirus disease in the Czech Republic by region per 100,000 people. B: Overall number of persons who tested positive for coronavirus disease in the Czech Republic. C: Daily overview of the number of persons with newly confirmed coronavirus disease over the last 2 weeks. D: Overall number of persons with laboratory-confirmed coronavirus disease by age group.
[PNG File , 313 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Overview of monitored statistics regarding the number of visits (period: March 12-April 12, 2020).
[PNG File , 595 KB-Multimedia Appendix 3]

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
2. Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19 - studies needed. N Engl J Med 2020 Mar 26;382(13):1194-1196. [doi: 10.1056/NEJMp2002125] [Medline: 32074416]
3. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). J Gen Intern Med 2020 May;35(5):1545-1549. [doi: 10.1007/s11606-020-05762-w] [Medline: 32133578]
4. Jin H, Liu J, Cui M, Lu L. Novel coronavirus pneumonia emergency in Zhuhai: impact and challenges. J Hosp Infect 2020 Apr;104(4):452-453 [FREE Full text] [doi: 10.1016/j.jhin.2020.02.005] [Medline: 32068012]
5. Stebbing J, Phelan A, Griffin I, Tucker C, Oechsle O, Smith D, et al. COVID-19: combining antiviral and anti-inflammatory treatments. Lancet Infect Dis 2020 Apr;20(4):400-402 [FREE Full text] [doi: 10.1016/S1473-3099(20)30132-8] [Medline: 32113509]
6. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med 2020 Mar 13;27(2):e [FREE Full text] [doi: 10.1093/jtm/taaa021] [Medline: 32052846]
7. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
8. Sun J, He W, Wang L, Lai A, Ji X, Zhai X, et al. COVID-19: epidemiology, evolution, and cross-disciplinary perspectives. Trends Mol Med 2020 May;26(5):483-495 [FREE Full text] [doi: 10.1016/j.molmed.2020.02.008] [Medline: 32359479]
9. Mavragani A. Tracking COVID-19 in Europe: infodemiology approach. JMIR Public Health Surveill 2020 Apr 20;6(2):e18941. [doi: 10.2196/18941] [Medline: 32250957]
10. Hernández-García I, Giménez-Júlvez T. Assessment of health information about COVID-19 prevention on the internet: infodemiological study. JMIR Public Health Surveill 2020 Apr 01;6(2):e18717. [doi: 10.2196/18717] [Medline: 32217507]
11. Moorthy V, Henao Restrepo AM, Preziosi M, Swaminathan S. Data sharing for novel coronavirus (COVID-19). Bull World Health Organ 2020 Mar 01;98(3):150 [FREE Full text] [doi: 10.2471/BLT.20.251561] [Medline: 32132744]
12. Heymann DL. Data sharing and outbreaks: best practice exemplified. Lancet 2020 Feb 15;395(10223):469-470. [doi: 10.1016/S0140-6736(20)30184-7] [Medline: 31986258]

13. Johns Hopkins University. 2020. OVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) URL: https://coronavirus.jhu.edu/map.html [accessed 2020-04-10]
14. Worldometer. 2020. Coronavirus update (live) URL: https://www.worldometers.info/coronavirus/ [accessed 2020-04-10]
15. Our World in Data. 2020. Total confirmed COVID-19 cases URL: https://ourworldindata.org/grapher/total-cases-covid-19 [accessed 2020-04-10]
16. Azevedo A, Santos M. Repositório P.Porto. 2008. KDD, SEMMA and CRISP-DM: a parallel overview URL: http://recipp.ipp.pt/handle/10400.22/136 [accessed 2013-07-09]
17. Molloy JC. The Open Knowledge Foundation: open data means better science. PLoS Biol 2011 Dec;9(12):e1001195. [doi: 10.1371/journal.pbio.1001195] [Medline: 22162946]
18. Nečaský M, Klímek J, Škoda P. Practical use cases for linked open data in eGovernment demonstrated on the Czech Republic. In: Kő A, Francesconi E, editors. Electronic Government and the Information Systems Perspective. Switzerland: Springer, Cham; 2017:80-96.
19. Plaza B. Google Analytics for measuring website performance. Tourism Manage 2011 Jun;32(3):477-481. [doi: 10.1016/j.tourman.2010.03.015]
20. Gaur L, Singh G, Kumar S. Google Analytics: a tool to make websites more robust. 2016 Mar Presented at: Second International Conference on Information and Communication Technology for Competitive Strategies; 2016; Udaipur, India. [doi: 10.1145/2905055.2905251]
21. Komenda M, Karolyi M, Bulhart V, Žofka J, Brauner T, Hak J, et al. Onemocnění Aktuálně. COVID-19: overview of current situation in the Czech Republic URL: https://onemocneni-aktualne.mzcr.cz/covid-19 [accessed 2020-05-09]

Abbreviations

COVID-19: coronavirus disease
CRISP-DM: cross-industry standard process for data mining
CSV: comma-separated values
IHIS CR: Institute of Health Information and Statistics of the Czech Republic
ISID: Information System of Infectious Diseases
JSON: JavaScript Object Notation
LAB: laboratories
NHIS: National Health Information System
PPE: personal protective equipment
RPHA: regional public health authority
SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

Edited by G Eysenbach; submitted 15.04.20; peer-reviewed by F Dankar, M Pradhan, TH Kwan; comments to author 07.05.20; revised version received 14.05.20; accepted 14.05.20; published 27.05.20

Please cite as:
Komenda M, Bulhart V, Karolyi M, Jarkovský J, Mužík J, Májek O, Šnajdrová L, Růžičková P, Rážová J, Prymula R, Macková B, Březovský P, Marounek J, Černý V, Dušek L
Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice
J Med Internet Res 2020;22(5):e19367
URL: http://www.jmir.org/2020/5/e19367/
doi: 10.2196/19367
PMID: 32412422

Original Paper

Development of the Czech Childhood Cancer Information System: Data Analysis and Interactive Visualization

Denisa Krejčí[1,2], MSc; Matěj Karolyi[1,2,3], MSc; Lucie Pehalová[1,2], MSc; Jakub Ščavnický[1,2], MSc; Michaela Zapletalová[1,2], MSc; Ivana Katinová[1,2], MSc; Jaroslav Štěrba[4], PhD; Jan Starý[5], PhD; Lenka Šnajdrová[1,2], PhD; Martin Komenda[1,2], PhD; Ladislav Dušek[1,2], PhD

[1]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[2]Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic
[3]Facilty of Informatics, Masaryk University, Brno, Czech Republic
[4]Department of Paediatric Oncology, Institutions shared with the Faculty Hospital Brno, Faculty of Medicine, Brno, Czech Republic
[5]Department of Paediatric Haematology and Oncology, Charles University and University Hospital Motol, Prague, Czech Republic

Corresponding Author:
Denisa Krejčí, MSc
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Kamenice 3
Brno, 62500
Czech Republic
Phone: 420 608273417
Email: Denisa.Krejci@uzis.cz

Abstract

Background: The knowledge of cancer burden in the population, its time trends, and the possibility of international comparison is an important starting point for cancer programs. A reliable interactive tool describing cancer epidemiology in children and adolescents has been nonexistent in the Czech Republic until recently.

Objective: The goal of this study is to develop a new web portal entitled the Czech Childhood Cancer Information System (CCCIS), which would provide information on childhood cancer epidemiology in the Czech Republic.

Methods: Data on childhood cancers have been obtained from the Czech National Cancer Registry. These data were validated using the clinical database of childhood cancer patients and subsequently combined with data from the National Register of Hospitalised Patients and with data from death certificates. These validated data were then used to determine the incidence and survival rates of childhood cancer patients aged 0 to 19 years who were diagnosed in the period 1994 to 2016 (N=9435). Data from death certificates were used to monitor long-term mortality trends. The technical solution is based on the robust PHP development Symfony framework, with the PostgreSQL system used to accommodate the data basis.

Results: The web portal has been available for anyone since November 2019, providing basic information for experts (ie, analyses and publications) on individual diagnostic groups of childhood cancers. It involves an interactive tool for analytical reporting, which provides information on the following basic topics in the form of graphs or tables: incidence, mortality, and overall survival. Feedback was obtained and the accuracy of outputs published on the CCCIS portal was verified using the following methods: the validation of the theoretical background and the user testing.

Conclusions: We developed software capable of processing data from multiple sources, which is freely available to all users and makes it possible to carry out automated analyses even for users without mathematical background; a simple selection of a topic to be analyzed is required from the user.

(JMIR Public Health Surveill 2021;7(6):e23990) doi: 10.2196/23990

KEYWORDS

cancer epidemiology; children; data visualization; software development

Introduction

Childhood and adolescent cancers (or *childhood cancers* for short) are classified among rare diseases as their incidence rates are orders of magnitude lower than cancer incidence rates in adults. About 400 new cases of childhood cancers are diagnosed in the Czech Republic each year. Although childhood cancers are rare in terms of absolute numbers, they are the second leading cause of death among children (after injuries).

The web portal SVOD (System for Visualisation of Oncology Data) [1] provides representative epidemiological data on cancer in the Czech Republic. However, it is not entirely convenient for childhood cancers. The main issue is a different classification system of childhood cancers, which takes into consideration differences between childhood and adult cancers; unlike the classification system of adult cancers, which is based on the site of the primary tumor, the classification system of childhood cancers is primarily based on morphology.

For this reason, the Institute of Health Information and Statistics of the Czech Republic (IHIS CR), together with the Institute of Biostatistics and Analyses at the Faculty of Medicine of the Masaryk University (IBA FM MU), the Department of Paediatric Oncology at the University Hospital Brno, and the Department of Paediatric Haematology and Oncology at the University Hospital in Motol, decided to develop a new web portal entitled the Czech Childhood Cancer Information System [2] (CCCIS), which would provide information on childhood cancer epidemiology in the Czech Republic.

The primary objective of the CCCIS portal is to provide comprehensible overviews of epidemiological data on the incidence of childhood cancers in the Czech Republic and mortality and survival data related to cancers in childhood and adolescents. The CCCIS portal also aims to provide relevant information on childhood cancer in the Czech Republic to the international community; this is the reason why the portal is available not only in the Czech language but also in English. The authors of this paper asked the following exploratory questions, which are answered in the Discussion section: how can detailed analytical views—covering incidence, mortality, and survival—be methodically and technically designed and subsequently implemented; is it possible to determine epidemiological trends of selected cancer diagnoses, based on available representative data; and what are the current survival rates in childhood cancer patients in the Czech Republic?

Methods

Data Sources

Data on childhood cancers, which are used on the portal, have been obtained from the Czech National Cancer Registry (CNCR), which is administered by IHIS CR [3]. These data were validated using the clinical database of childhood cancer patients [4] and subsequently combined with data from the National Register of Hospitalised Patients [5] and with data from death certificates [6]. These validated data were then used to determine incidence and survival rates of childhood cancer patients aged 0 to 19 years who were diagnosed in the period 1994 to 2016. Data from death certificates were used to monitor long-term mortality trends. Demographic data on the population of interest were obtained from outputs of the Czech Statistical Office [7]. International data sources were used as well: incidence data from the International Incidence of Childhood Cancer [8], mortality data from the European Cancer Information System [9], and survival data from the international comparative study CONCORD-3 [10].

Classification

Cancers were classified into 12 main groups, according to the International Classification of Childhood Cancer, 3rd Edition (ICCC-3) [11]. All diagnostic groups (I-XII) with behavior 3 (primary malignant tumors) plus diagnoses from groups III (central nervous system [CNS] and miscellaneous intracranial and intraspinal neoplasms) and Xa (intracranial and intraspinal germ cell tumors) with behavior 0/1 (benign neoplasms or those of uncertain or unknown behavior) were considered to be malignant tumors. As regards mortality data, the most common causes of childhood cancer deaths according to the International Statistical Classification of Diseases, Tenth Revision (ICD-10) [12] are shown, namely, the following list of diagnoses: all malignant neoplasms (C00-C97); malignant neoplasms of bone and articular cartilage (C40-C41); malignant neoplasms of connective and soft tissue (C47, C49); malignant neoplasms of brain, spinal cord, and other parts of the CNS (C70-C72); non-Hodgkin lymphoma (C82-C86); and leukemias (C91-C95).

Analysis of Epidemiological Data

In terms of statistical analyses, the web portal CCCIS focuses on three epidemiological indicators: incidence, mortality, and survival. Incidence is the number of new cases diagnosed in a given period in a given population. The CCCIS portal makes it possible to express the incidence in several different ways. First, absolute numbers show the overall burden of the population with a given disease. Second, crude incidence is the number of new cases arising per 1 million children in a given population. If the population of interest only comprises persons in a given age interval (eg, 10-14 years), we are talking about an age-specific incidence. Third, the age-standardized incidence is the theoretical incidence rate that a given population would have if it had a standard age structure. The portal makes it possible to calculate the European age-standardized incidence rate (ASR-E) [13] and the world age-standardized incidence rate (ASR-W) [14]. Mortality is the number of deaths from a given diagnosis (the so-called cause-specific mortality) occurring in a given period in a given population. As is the case with incidence rates, mortality on the CCCIS portal can also be expressed in absolute numbers, rates per 1 million children, age-specific rates, and age-standardized rates to European or world standard population. Overall survival rates were used to evaluate the patients' survival, corresponding to the overall monitored survival, regardless of the cause of death. The overall 1-, 2-, 3-, 4-, and 5-year survival was calculated using the life tables method with 1-year intervals, where death from any cause was the event of interest.

Design and Development

CCCIS is a web portal equipped with an online data browser, which has been developed using the modern and practice-proven Symfony PHP framework in version 3.4 [15]. Using this framework for systematic design, development, and implementation of web applications significantly accelerates individual stages and generally facilitates the applications' administration and extensibility. It is therefore possible to react relatively quickly to users' needs and requirements to adapt current functionalities or to add new functionalities. The Twig template engine has been used to create page templates, and the Doctrine Object Relational Mapper has been used to map the objects—both had been released together with the Symfony framework. The data repository has been built on an open-source object-oriented system, PostgreSQL, which is currently routinely used to organize more complex data structures [16]. The main advantages of PostgreSQL include the support of the developer community, the possibility of advanced performance optimization, and a high quality of technical documentation, making it possible to administer the entire database system and the individual databases without problems.

The import itself into the database is performed by automatic scripts, which upload the new database contents. Maintaining the uniformity of the data model and data purity and quality (thoroughly validated by analysts and developers) are essential requirements for a successful import. The dependence on third-party libraries is dealt with by Yarn (front end) and Composer (back end). A large proportion of the portal has been designed to be responsive (ie, displayed content is automatically adjusted depending on the user's device—desktop, tablet, mobile phone). However, responsiveness is not complete in several parts containing the data browser; the mobile version is not fully supported in this instance because the user interface is far too complex (graphics, filters, analysis settings). The responsive front end has been developed using the Zurb Foundation framework and the jQuery library. The webpack tool has been used to compile the final package of styles and Javascript functions. The interactive data browser requires a special functionality, which is provided by extension components of Javascript libraries; in particular, d3.js, NVD3, and Datatables have been used for interactive data visualizations. We had already applied a similar approach to the development of web applications in the past, namely, in interactive data browsers focused on several issues in Czech health care [17-20].

Data and Application Security

One of the basic requirements of this project was to secure the entire application, including the data layer. The application has been designed to resist potential third-party attacks and to respond flexibly. The communication between the client and the server takes place in a secure way via the HTTPS protocol. This encrypted transmission is nowadays used as standard, and it is common practice to secure the flow of data in this way. The application itself, built on the Symfony framework, has other security mechanisms that are built into it. Respecting standard implementation approaches ensures that exposure to basic attacks such as cross-site scripting, cross-site request forgery, or various types of injection, especially SQL injection, is avoided.

Server-side protection is provided mainly by a network firewall. Another effective way to detect possible third-party attacks is to set up well-configured resource and traffic monitoring, log errors and accesses from the external internet environment, and alert the administrator to nonstandard events.

Protection of the data itself is another necessary requirement to be met. The underlying data, which are used by the portal for rendering visualizations, are cleaned of all personal and sensitive information on patients and their hospital stays. Therefore, it is impossible to directly connect the records to a specific patient. The data are securely stored in a database that is accessible only from predefined locations (these are always part of an internal network) and to a limited set of users whose permissions are restricted to certain data operations. The data are always sent to the client side in an aggregated form, as required for the final visualization. It is important to perform the aggregation operation before sending so that the data cannot be broken down into individual rows at the client side. At the same time, with potentially small numbers of records, the result of the analysis is not sent, and the user is notified of this fact. Thus, the identification of a specific person is effectively prevented.

Validation and User Testing

Feedback was obtained and the accuracy of outputs published on the CCCIS portal was verified using the following methods, which are generally suitable to identify potential shortcomings not only in terms of contents but also in terms of design of the user interface and control elements.

The *validation of theoretical background*, which describes the basic terms, the methodology of cancer classification, and static analytical reports, was performed internally (ie, in cooperation with the analytical team and a group of senior doctors who are specialists in childhood cancers and have many years of experience with the methodology of childhood cancer classification). At the same time, all three sections of the interactive data browser were thoroughly checked; complex analytical views of incidence, mortality, and survival rates according to user settings were extensively tested.

User testing involved simple instructions to go through individual sections of the portal and to provide subjective feedback as regards the overall visual style, control elements, and user-friendliness.

The outputs from both assessments were extensively discussed by the team of authors, and selected suggestions, which fit in with the overall concept of the portal, were subsequently implemented.

Results

Basic Description of the CCCIS Portal

The web portal CCCIS [2] is a stand-alone online presentation, which has been freely available on the internet since November 2019, without the necessity of user authentication. The portal is allowed to be indexed and therefore to be found by standard search engines. Users can access the published contents via a web browser, and all communication takes place via the HTTPS protocol (ie, in a secured and encrypted manner). The CCCIS portal is divided into several sections:

- The *Introduction* section provides basic information about the portal objectives and contents. Participating institutions and the team of authors are introduced. This section also describes the source of data that have been used for statistics and for interactive data views. News related to childhood cancers in the Czech Republic are also involved.
- The *Methodology* section describes how childhood cancers are classified according to the international classification system.
- The *Statistics* section contains an overview of information and descriptive attributes on provided views of the data set. The section is divided into incidence, mortality, and survival. Static analyses are also available for download; however, this feature is only available in the Czech language.
- The *Interactive data views* section provides graphical outputs, which make it possible for users to go through available data sets in an interactive manner. All data sets are regularly updated, based on data from the CNCR and data from a clinical database. Like the Statistics section, the Interactive data views section is divided into incidence, mortality, and survival subsections. This section of the portal is described in more detail in the next section of this paper.
- The *Publications* section provides a list of articles published in research journals and a list of conference papers.

Introduction of the CCCIS Browser

The interactive browser is the principal component of the portal, containing predefined analytical tools that make it possible for the user to look into epidemiological data from different points of view, both in graphical and tabular representation. From the user's viewpoint, this is how the interactive browser is used:

- Selection of the main module (incidence, mortality, survival)
- Selection of analysis type
- Selection of the analyzed group of patients, setting the analysis outputs

Selection of the main module is the first step to begin with any analysis. The principal epidemiological analyses, the so-called modules, cover the following topics: incidence, mortality, and survival. After selecting the main module, the user needs to select the required type of analysis, which means analysis by year of diagnosis, by sex, by cancer type, by age and cancer type combined (this option is only available in the incidence and mortality modules), or by international comparison. Individual types of analyses can be selected in the upper part of the screen. After selecting the required analysis type, an analytical window is displayed, showing the results with basic settings. These settings can be further adjusted on two levels.

Options for the Analyzed Group of Patients

Options for the analyzed group of patients can be selected using the following filters:

- Diagnosis (or cancer type): selection of diagnosis by ICCC-3 (in the incidence and survival modules) or by ICD-10 (in the mortality module)
- Sex: the entire population of children, boys only, or girls only
- Age: selection of age categories 0 to 19 years, 0 to 14 years, <1 year, 1 to 4 years, 5 to 9 years, 10 to 14 years, or 15 to 19 years (in the incidence and mortality modules), or 0 to 19 years, 0 to 14 years, or 15 to 19 years (in the survival module)
- Period/years: a scrollbar can be used in the incidence and mortality modules to select individual years or a span of years (currently between 1994 and 2016); as for the survival module, only the predefined periods 1999 to 2004, 2005 to 2010, and 2011 to 2016 can be selected

These filters can be combined, and a detailed view of selected topics can be obtained in this way. Unsuitable or illogical variables in the context of the selected analysis are inactive (gray). The *reset filters* button can be used to restore the original analysis settings.

Detailed Settings of the Analysis Output

Depending on the selected analysis, the software offers a suitable computational method such as absolute numbers; annual numbers; percentages; rate per 1 million children; ASR-E; ASR-W; and 1-, 2-, 3-, 4-, or 5-year overall survival rate. The toggle switch *Group years?* in the Incidence and Mortality modules makes it possible to visualize data for individual years (the *off* position) or for years grouped together (the *on* position). In the *International comparison* analysis, this toggle switch is always in the *on* position, making it possible to compare data from the Czech Republic with data from other European countries.

The primary output of the interactive tool is a graph displayed in the center of the working window, including the description of applied filters and the data source. Apart from this graphical output, results can also be displayed in the form of a data table. Graphical outputs can be downloaded as images, whereas tabular outputs can be copied, printed, or downloaded as *.csv or *.xlsx files. Short reports describing the main epidemiological indicators have been written to provide basic overviews of selected diagnoses.

As an illustration, Figure 1 shows the Incidence module, namely, the analysis by cancer type, with the toggle switch *Group years?* being in the *on* position, *annual numbers* selected as the computational method, the time filter set to the period 2007 to 2016, and the graph sorted in descending order. We can see that leukemias (ICCC I) are the most common childhood cancers (ie, those diagnosed in patients aged 0 to 19 years), followed by CNS tumors (ICCC III), other malignant epithelial neoplasms, and malignant melanomas (ICCC XI) and lymphomas (ICCC II).

For illustrative purposes, the Survival module is described (see Figure 2), namely, the analysis by age, with *Leukemias* selected as the cancer type and 1-, 2-, 3-, 4-, and 5-year survival selected as the computational method. We can see that survival rates in the period from 2011 to 2016 differ significantly: patients aged

15 to 19 years have markedly lower 5-year overall survival rates than younger patients.

Figure 1. Czech Childhood Cancer Information System Interactive data views, incidence by cancer type. ICCC: International Classification of Childhood Cancer.

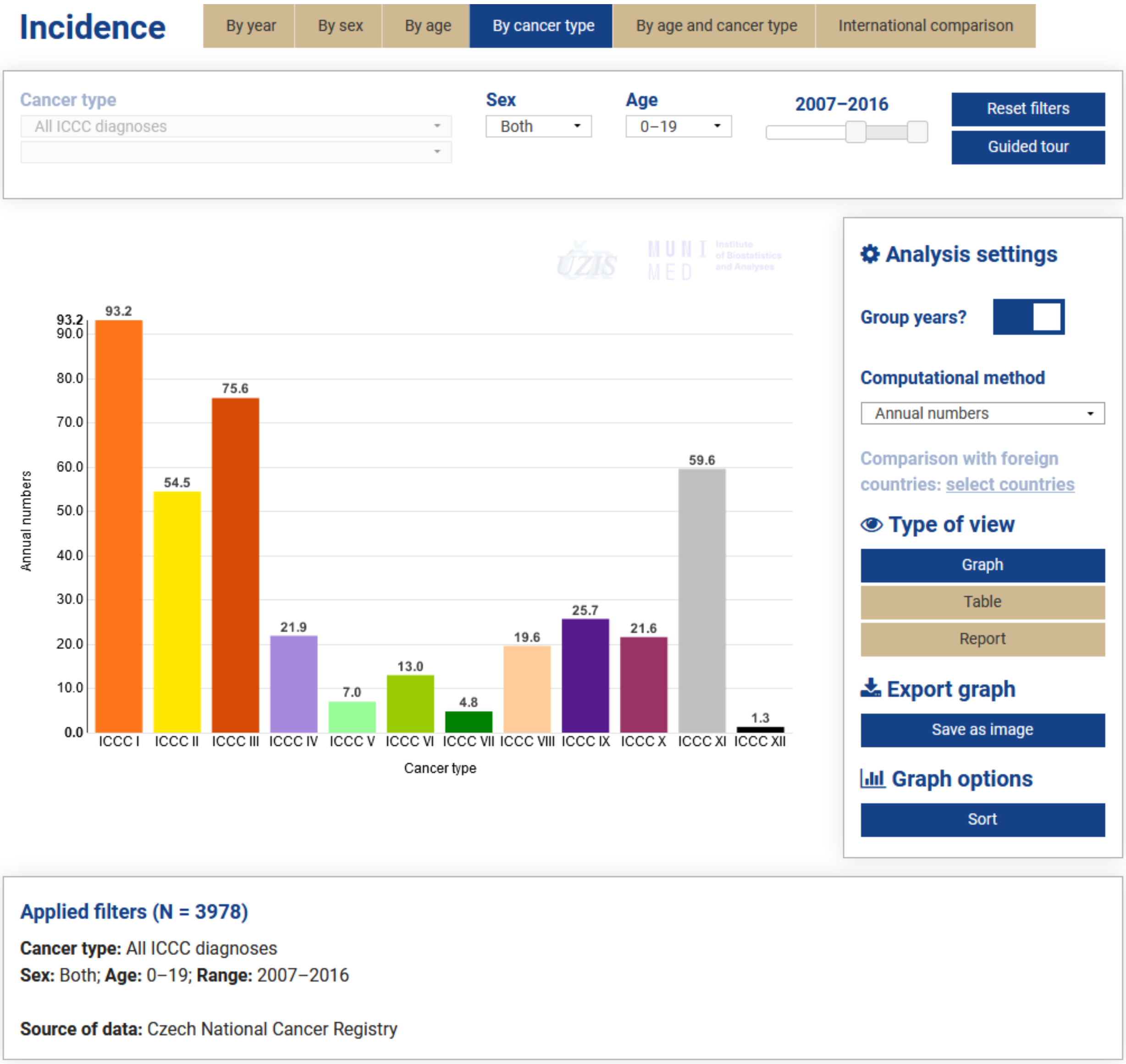
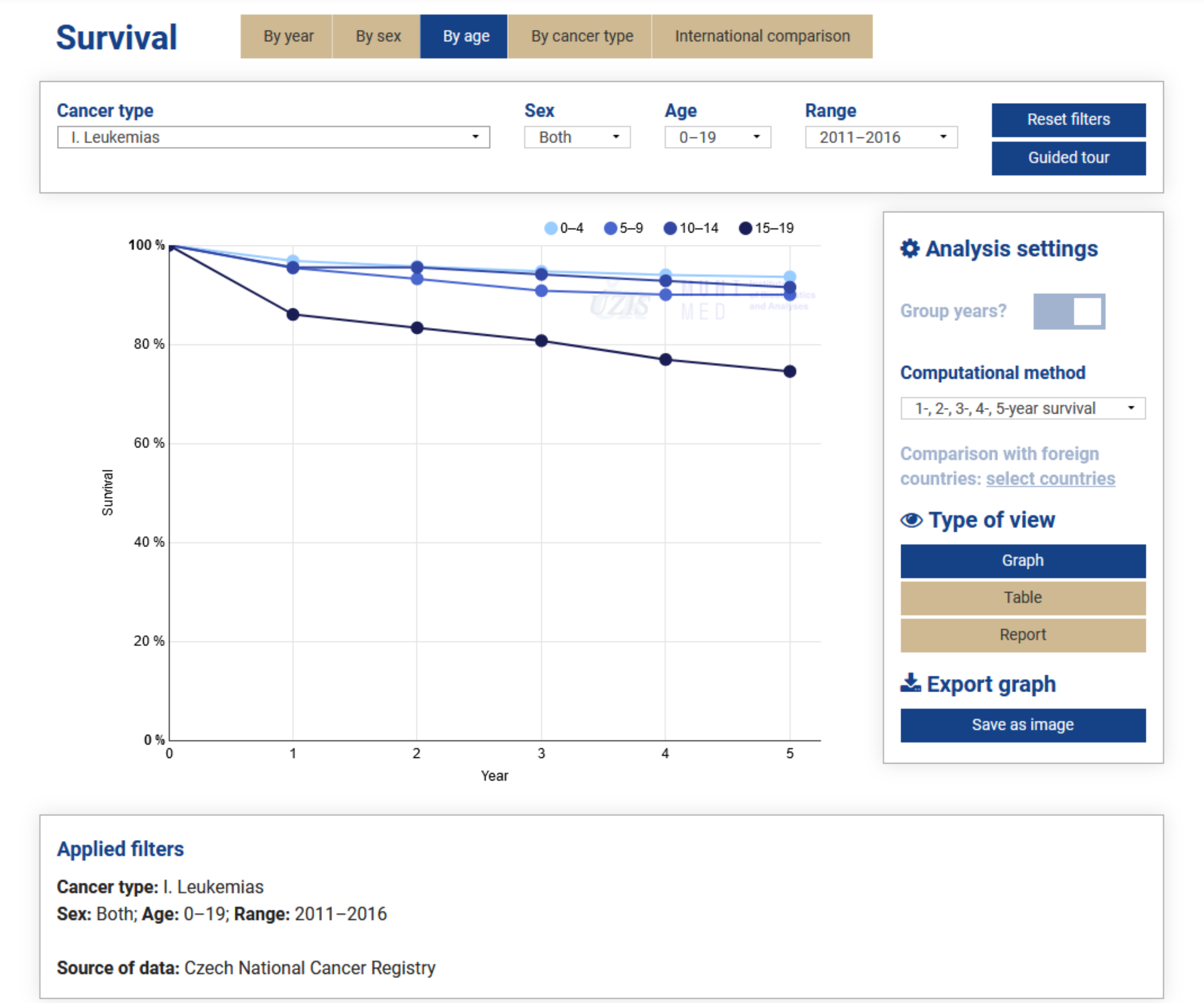


Figure 2. Czech Childhood Cancer Information System Interactive data views, survival by age.



Benefit for Users

The software is capable of processing data from multiple sources, is freely available to all users, and makes it possible to carry out automated analyses even for users without mathematical background; a simple selection of a topic to be analyzed is required from the user. This online available software tool makes it therefore possible for anyone to display long-term trends of childhood cancer incidence, mortality, and survival, and to compare data from the Czech Republic to those from abroad. All analytical outputs are displayed in both graphical and tabular outputs. The aggregation of data over time periods is another indisputable advantage of the interactive browser.

Evaluation of the CCCIS Portal

The validation of the static contents was carried out as an internal review of static texts and analytical reports. The comments were of rather formal character, aiming to unify the published information from the linguistic and visual points of view. The CCCIS portal was also assessed by experts, namely, by two senior doctors dealing with childhood cancer patients on a daily basis, who were asked to follow a given scenario using the CCCIS portal (full description of this scenario including feedback is available at [21]).

Discussion

Principal Results

The main findings of this study are summarized in the following discussion in the form of answers to three exploratory questions, which were asked in the Introduction section.

How can detailed analytical views—covering incidence, mortality, and survival—be methodically and technically designed and subsequently implemented? The basic methodical concept of the CCCIS portal lies in the division of the presented contents into three parts (theoretical background, interactive data views, analytical reports), which make up a compact set of information comprehensible to experts in the field and health care professionals. Clear and appropriate definitions of basic terms, together with the description of the International Classification of Childhood Cancer, explain relevant issues to users. The interactive browser represents a unique dynamic tool that makes is possible to set different views of available data sets, including the type of view (graph, table, report) and subsequent download. Analytical reports provide a summary overview of childhood cancer epidemiology in the Czech Republic and detailed reports for individual ICCC groups. The technical solution itself is based on the robust PHP development Symfony framework, with the PostgreSQL system used to accommodate the data basis.

Is it possible to determine epidemiological trends of selected cancer diagnoses based on available representative data? The interactive browse makes epidemiological data available through predefined analytical tools. Users can look into epidemiological indicators—incidence, mortality, and survival of selected cancer diagnosis—by selecting the main module (incidence, mortality, or survival), followed by the selection of the required type of analysis (by year of diagnosis, sex, age, cancer type, age and cancer type combined, and international comparison), and the selection of the patient group to be involved in the analysis. Absolute numbers; annual numbers; rates per 1 million children; age-standardized rates to European or world standard population; and the overall 1-, 2-, 3-, 4-, and 5-year survival can be displayed, depending on the selected analysis.

What are the current survival rates in childhood cancer patients in the Czech Republic? The overall 1-, 2-, 3-, 4-, and 5-year survival rates for 12 main ICCC groups can be obtained using the main *Survival* module of the interactive browser, after selecting the type of analysis and possibly by being more specific about the group of patients to be involved in the analysis. The highest 5-year survival in the period 2011 to 2016 was observed in retinoblastoma (ICCC V), and in other malignant epithelial neoplasms and malignant melanomas (ICCC XI). By contrast, the lowest 5-year survival rates were observed in soft tissue and other extraosseous sarcomas (ICCC IX) and in hepatic tumors (ICCC VII).

Future Visions

The CCCIS portal will be further developed within the joint workplace of the IHIS CR and IBA FM MU. Comments and suggestions provided by users themselves will play an important role in this process. Specific questions will be formulated in cooperation with the expert society, probably requiring additional analytical outputs.

Comparison With Other Works

Publicly available portals describing childhood cancer epidemiology in other European countries, namely, in Ireland [22], Switzerland [23], and the United Kingdom [24], have become our motivation. The Irish web portal provides an interactive analytical reporting for incidence; the Swiss and the UK websites are static but provide information not only on incidence but also on mortality and survival of childhood cancer patients. The new CCCIS portal combines the approaches previously mentioned and thus enables interactive analytical reporting of incidence, mortality, and survival.

Conclusions

The CCCIS portal is the result of a long-term cooperation of a state organization directly coming under the Ministry of Health of the Czech Republic and selected specialized workplaces in the academic sphere. Doctors, representatives of a department administering the National Health Information System, data analysts, systems analysts, graphic designers, and developers have worked together to create a platform that makes accessible valuable and interesting views of available data in a user-friendly form. The web portal is available for anyone at [2], providing basic information for experts (ie, analyses and publications) on individual diagnostic groups of childhood cancers. It involves an interactive tool for analytical reporting, which provides information on the following basic topics in the form of graphs or tables: incidence, mortality, and overall survival.

Acknowledgments

Special thanks to Lenka Snajdrova for her language support. The authors are also thankful to all specialists involved in data collection and administration. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

None declared.

References

1. Dusek L, Muzik J, Kubasek M, Koptikova J, Zaloudik J, Vyzula R. Epidemiology of Malignant Tumours in the Czech Republic.: Masarykova univerzita; 2005. URL: https://www.svod.cz/ [accessed 2020-03-20]
2. Krejci D, Scavnicky J, Zapletalova M, Svobodova I, Karolyi M, Muzik J, et al. Czech Childhood Cancer Information System. 2018. URL: https://ccc-is.uzis.cz [accessed 2020-03-20]
3. Czech National Cancer Registry. National Health Information System. URL: http://www.uzis.cz/registry-nzis/nor [accessed 2020-03-20]
4. Bajciova V, Dusek L, Janotova I, Kabickova E, Kepak T, Klimes D, et al. Czech National Informational and Educational Portal. URL: http://detskaonkologie.registry.cz/ [accessed 2020-03-20]
5. National Register of Hospitalised Patients. National Health Information System. URL: http://www.uzis.cz/registry-nzis/nrhosp [accessed 2020-03-20]
6. Death certificate system. National Health Information System. URL: http://www.uzis.cz/registry-nzis/list-prohlidce-zemreleho [accessed 2020-03-20]

7. Czech demographic handbook - 2016. Czech Statistical Office. URL: https://www.czso.cz/csu/czso/czech-demographic-handbook-2016 [accessed 2020-03-20]
8. Steliarova-Foucher E, Colombet M, Ries LAG, Hesseling P, Moreno F, Shin HY, et al. Results. International Incidence of Childhood Cancer 3. 2017. URL: https://iicc.iarc.fr/results/ [accessed 2020-03-20]
9. Incidence and mortality historical data. ECIS - European Cancer Information System. URL: https://ecis.jrc.ec.europa.eu/ [accessed 2020-03-20]
10. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, CONCORD Working Group. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018 Mar 17;391(10125):1023-1075 [FREE Full text] [doi: 10.1016/S0140-6736(17)33326-3] [Medline: 29395269]
11. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. Cancer 2005 Apr 01;103(7):1457-1467. [doi: 10.1002/cncr.20910] [Medline: 15712273]
12. World Health Organization. International Statistical Classification of Diseases and Related Health Problems - 10th Revision, edition 5. Geneva, Switzerland: WHO Press; 2016.
13. Eurostat. Revision of the European Standard: Population Report of Eurostat's Task Force: 2013 edition. Luxembourg: Publications Office of the European Union; 2013.
14. Doll R, Payne P, Waterhouse J. Cancer Incidence in Five Continents: A Technical Report. Verlag Berlin Heidelberg: Springer; 1966:XI, 244.
15. Symfony documentation. Symfony. URL: https://symfony.com/doc/3.4/index.html [accessed 2020-09-22]
16. Juba S, Volkov A. Learning PostgreSQL 11: A Beginner's Guide to Building High-Performance PostgreSQL Database Solutions. 3rd edition. Birmingham, UK: Packt Publishing; 2019.
17. Karolyi M, Komenda M, Janoušová R, Vita M, Schwarz D. Finding overlapping terms in medical and health care curriculum using text mining methods: rehabilitation representation - a proof of concept. MEFANET J 2017;4(2):71-77.
18. Dušek L, Mužík J, Karolyi M, Šalko M, Malúšková D, Komenda M. A pilot interactive data viewer for cancer screening. In: Hřebíček J, Denzer R, Schimak G, Pitner T, editors. Environmental Software Systems. Computer Science for Environmental Protection 12th IFIP WG 5.11 International Symposium, ISESS 2017, Zadar, Croatia, May 10-12, 2017, Proceedings. Cham: Springer; 2017:173-183.
19. Karolyi M, Krejčí J, Ščavnický J, Vyškovský R, Komenda M. Tools for development of interactive web-based maps: application in healthcare. Plzen: Vaclav Skala - UNION Agency; 2019 Presented at: WSCG 2019 - 27. International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2019; May 27-31, 2019; Plzen p. 1-8. [doi: 10.24132/csrn.2019.2902.2.1]
20. Komenda M, Bulhart V, Karolyi M, Jarkovsky J, Muzik J, Majek O, et al. Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice. J Med Internet Res 2020 May 27;22(5):e19367 [FREE Full text] [doi: 10.2196/19367] [Medline: 32412422]
21. Evaluation of the CCCIS portal. Czech Childhood Cancer Information System. URL: https://ccc-is.uzis.cz/index-en.php?pg=evaluation [accessed 2020-09-22]
22. National Cancer Registry Ireland. URL: https://www.ncri.ie/data/incidence-statistics [accessed 2020-03-20]
23. Childhood Cancer Registry. URL: https://www.childhoodcancerregistry.ch/figures-and-reports/figures-incidence-survival-mortality/ [accessed 2020-03-20]
24. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/childrens-cancers [accessed 2020-03-21]

Abbreviations

ASR-E: European age-standardized incidence rate
ASR-W: world age-standardized incidence rate
CCCIS: Czech Childhood Cancer Information System
CNCR: Czech National Cancer Registry
CNS: central nervous system
IBA FM MU: Institute of Biostatistics and Analyses at the Faculty of Medicine of the Masaryk University
ICCC-3: International Classification of Childhood Cancer, 3rd Edition
ICD-10: International Statistical Classification of Diseases, Tenth Revision
IHIS CR: Institute of Health Information and Statistics of the Czech Republic
SVOD: System for Visualisation of Oncology Data

Edited by G Eysenbach; submitted 31.08.20; peer-reviewed by BT Joyce, E Andrikopoulou; comments to author 21.09.20; revised version received 05.10.20; accepted 17.04.21; published 29.06.21

Please cite as:
Krejčí D, Karolyi M, Pehalová L, Ščavnický J, Zapletalová M, Katinová I, Štěrba J, Starý J, Šnajdrová L, Komenda M, Dušek L
Development of the Czech Childhood Cancer Information System: Data Analysis and Interactive Visualization
JMIR Public Health Surveill 2021;7(6):e23990
URL: https://publichealth.jmir.org/2021/6/e23990
doi: 10.2196/23990
PMID: 34185010

PLOS ONE

RESEARCH ARTICLE

Sharing datasets of the COVID-19 epidemic in the Czech Republic

Martin Komenda[1,2,3]*, Jiří Jarkovský[1,2], Daniel Klimeš[1,2], Petr Panoška[1,2], Ondřej Šanca[1], Jakub Gregor[1,2], Jan Mužík[1,2], Matěj Karolyi[1,2,4], Ondřej Májek[1,2], Milan Blaha[1,2], Barbora Macková[5], Jarmila Rážová[6], Věra Adámková[7,8,9], Vladimír Černý[10,11], Jan Blatný[6], Ladislav Dušek[1,2]

1 Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic, **2** Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **3** Medical Simulation Centre, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **4** Faculty of Informatics, Masaryk University, Brno, Czech Republic, **5** National Institute of Public Health, Prague, Czech Republic, **6** Ministry of Health of the Czech Republic, Prague, Czech Republic, **7** Department of Preventive Cardiology, Institute for Clinical and Experimental Medicine, Prague, Czech Republic, **8** Department of Nephrology, First Faculty of Medicine, Charles University, Prague, Czech Republic, **9** Faculty of Biomedical Engineering, Czech Technical University, Kladno, Czech Republic, **10** Department of Anaesthesiology, Perioperative Medicine and Intensive Care, Masaryk Hospital in Ústí nad Labem, Ústí nad Labem, Czech Republic, **11** Jan Evangelista Purkyně University, Ústí nad Labem, Czech Republic

* martin.komenda@uzis.cz

OPEN ACCESS

Citation: Komenda M, Jarkovský J, Klimeš D, Panoška P, Šanca O, Gregor J, et al. (2022) Sharing datasets of the COVID-19 epidemic in the Czech Republic. PLoS ONE 17(4): e0267397. https://doi.org/10.1371/journal.pone.0267397

Editor: Manuel Corpas, Universidad Internacional de La Rioja, SPAIN

Received: August 27, 2021

Accepted: April 7, 2022

Published: April 21, 2022

Data Availability Statement: All data are available from Open data catalogue: https://opendata.mzcr.cz/.

Funding: a) All sources of funding (financial or material support) related to this manuscript are covered by Ministry of Health of the Czech Republic and Institute of Health information and statistics of the Czech Republic. There are no particular grants or other organizations that supported this study. b) The funders had no role in study design, data collection and analysis, decision

Abstract

At the time of the COVID-19 pandemic, providing access to data (properly optimised regarding personal data protection) plays a crucial role in providing the general public and media with up-to-date information. Open datasets also represent one of the means for evaluation of the pandemic on a global level. The primary aim of this paper is to describe the methodological and technical framework for publishing datasets describing characteristics related to the COVID-19 epidemic in the Czech Republic (epidemiology, hospital-based care, vaccination), including the use of these datasets in practice. Practical aspects and experience with data sharing are discussed. As a reaction to the epidemic situation, a new portal *COVID-19: Current Situation in the Czech Republic* (https://onemocneni-aktualne.mzcr.cz/covid-19) was developed and launched in March 2020 to provide a fully-fledged and trustworthy source of information for the public and media. The portal also contains a section for the publication of (i) public open datasets available for download in CSV and JSON formats and (ii) authorised-access-only section where the authorised persons can (through an online generated token) safely visualise or download regional datasets with aggregated data at the level of the individual municipalities and regions. The data are also provided to the local open data catalogue (covering only open data on healthcare, provided by the Ministry of Health) and to the National Catalogue of Open Data (covering all open data sets, provided by various authorities/publishers, and harversting all data from local catalogues). The datasets have been published in various authentication regimes and widely used by general public, scientists, public authorities and decision-makers. The total number of API calls since its launch in March 2020 to 15 December 2020 exceeded 13 million. The datasets have been adopted as an official and guaranteed source for outputs of third parties, including public authorities, non-governmental organisations, scientists and online news portals. Datasets

to publish, or preparation of the manuscript. c) The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

currently published as open data meet the 3-star open data requirements, which makes them machine-readable and facilitates their further usage without restrictions. This is essential for making the data more easily understandable and usable for data consumers. In conjunction with the strategy of the MH in the field of data opening, additional datasets meeting the already implemented standards will be also released, both on COVID-19 related and unrelated topics.

Introduction

Open data, i.e. "data that can be freely accessed, used, modified, and shared by anyone for any purpose", represents a systematic approach to making selected government-produced datasets available for further manual or automatic processing in a uniform and clearly defined way [1]. At the time when the COVID-19 pandemic paralysed the entire world, providing access to data (suitably prepared from the perspective of personal data protection) plays a crucial role in providing the general public and media with up-to-date information. Open datasets represent one of the means for evaluation of the pandemic on a global level. The sources of such datasets include, for example, the WHO (World Health Organization), ECDC (European Centre for Disease Control) via Our World in Data [2], and the EU Open Data Portal [3]. Many datasets, including those published by the Johns Hopkins University [4], are available through the GitHub platform [5, 6]. Typically, websites guaranteed by governments or governmental agencies in the individual countries are used as sources of information for such international and/or global statistics. The same is true about the Czech Republic [7], presenting both visualisation and direct links to all open datasets on the COVID-19 epidemic situation on a website guaranteed by the Ministry of Health (MH). A similar approach was adopted by many other countries, including the USA [8], France [9], Germany [10], Austria [11], Spain [12], United Kingdom [13], etc. Datasets of other countries, e.g., Australia, Italy, South Africa, or Argentina, are available on GitHub [14].

From our perspective, datasets can be published as open data for several reasons: a) for the purposes of transparency (enabling verification of spending and financial management of institutions funded by the national budget, of the performance of the institutions, contracts, etc.) b) for providing complex population data (population characteristics, epidemiological statistics, medications, performance characteristics, etc.) and, last but not least, c) to offer fully aggregated data for the scientific and analytical purposes (secondarily created datasets describing groups of objects or subjects or even individual anonymised records of individuals, e.g., anonymised records of births, hospitalisations, surgeries, etc.). Some datasets are, as described below in greater detail, published in a regulated regime. Data published in this regime constitute a direct source of information about institutions, insurance payers, healthcare providers, or data intended for further processing with the aim to influence the behaviour of target subjects. In the Czech Republic, a National Catalogue of Open Data guaranteed by the MH has been developed and maintained by a joint team of the Institute of Health Information and Statistics of the Czech Republic and the Institute of Biostatistics and Analyses at the Faculty of Medicine of the Masaryk University in Brno over the last decade.

The primary aim of this paper is to describe the methodological framework including basic technical background for publishing datasets describing basic as well as advanced epidemiological characteristics related to the COVID-19 epidemic in the Czech Republic, including the use of these datasets in practice. The entire process of design, development and implementation of a robust system for tracking and reporting respiratory diseases caused by SARS-CoV-2 is based on a previously successfully used method for acquisition, processing, and presentation of information [15]. The system is based on a framework for cooperation between the Ministry of Health of the Czech Republic (MH), which collects the records from the testing points, and regional public health authorities (RPHAs). The infrastructure designed within the Czech National Health Information System (NHIS) processes these data and supports reporting of the results of testing for COVID-19 on a daily basis. Open datasets form an inherent part of the NHIS, providing information to the general public. Having this system in place has proved to be a big advantage in the current COVID-19 situation, as it was easily adaptable to suit the current requirements for information processing and presentation.

Methods

The methodological principles leading to the publication of prepared datasets presenting basic and advanced epidemiological characteristics will be described in this section.

The methods of data publishing and sharing

All regimes of the data provision from NHIS strictly require a certain degree of legislative regulation and must fully meet the criteria set for NHIS by the Czech legislation, particularly Act No. 372/2011 on health services and their provision. In other words, publishing of open data cannot be misinterpreted as publication of primary records without any regulation and standardisation; the term “open data”, therefore, does not necessarily describe primary database records (the data may be aggregated, statistically processed, etc.). The dataset design, preparation, and publishing should respect the algorithm of dataset preparation shown below, which always respects several principal rules: (i) individual natural persons must not be identifiable, (ii) individual legal persons must not be identifiable unless expressly stated by the law, (iii) secondary processing must lead to the pseudonymisation of the dataset, (iv) the purpose of the dataset publication must correspond to the NHIS purpose, and (v) the standardised process of approval and publishing must be adhered to (Fig 1). The basic requirement of a comprehensive process, which is met by this approach, is to ensure the necessary completeness, validity and overall quality of data.

Publication of a dataset consists of six consequent steps that describe key phases on the preparation and implementation. Step 1: Proposal of a dataset concept in the form of short description, which can be rised by any entity; usually a state institution, a health insurance company, a professional (medical) society, a research institution, or an academic institution. Step 2: The delivered concept is reviewed from the perspective of data availability, export feasibility, design, and personal data regulations. Step 3: After approval, a methodology for the dataset processing is proposed (data export from central registries, data pre-processing and cleaning, analytical adjustments and validation mechanisms). Step 4: The dataset is generated

Step 1 Concept design	Step 2 Concept evaluation	Step 3 Feasibility analysis	Step 4 Dataset production	Step 5 Review	Step 6 Publication
Proposed by state administration, external subjects (health insurance companies, expert societies, research institutions)	Purpose, data availability, feasibility, legal perspective	Data extraction, processing, analysis, validation	Structure, methods of production, metadata description	Personal data protection, factual content, IT solution	National Catalog of Open Data www.uzis.cz

Fig 1. A chart of dataset production and publication.

https://doi.org/10.1371/journal.pone.0267397.g001

in an open data standardised format according to the predefined scheme, including an obligatory description with metadata. Step 5: Review and validation of the dataset and its content is a mandatory procedure before publication (manual approach is always needed, as well as technical control in order to follow open data standards and best practices). Step 6: Final publication of the dataset in the National Catalogue of Open Data.

Freely available primary data. This publication regime is very rarely used for the publication of data from the NIHS internal systems. In this regime, the primary database records are published without any processing. In principle, this regime is only allowed for characteristics of "inanimate subjects" without any relations to personal data. Examples may include service providers as defined in the respective acts, machines, swimming pools, chemical substances or drugs, and their primary characteristics.

Primary data publishable after necessary processing. Such datasets are individually designed (based on proposals of groups of experts, such as clinicians or healthcare providers), reviewed, and finalised, including the descriptive parameters defining target subjects or object cohorts. Data published in this form prevent the identification of any natural or legal person. Adherence to a standardised methodology (data format unification, metadata description, publishing in central catalogue) when creating the dataset and senior expert review (advanced stakeholders in a particular healthcare domain, representatives of data management, data analysis, and development team) or approval are prerequisites for publishing in this regime. Examples of such necessary amendments include highlighting of missing values, correction of erroneous records, aggregation of records for patients' age categories, calculation of new variables or indices from the raw data, etc. Usually, these records are completely deleted (in cases of incorrect and invalid information), directly corrected in the central database or supplemented in accordance with the methodological instructions of the register.

Data requiring reference interpretation—Reference statistics. Publication of statistics characterising a group (cohort) of subjects/objects (e.g. a region, period of time, patient cohort, type of medical procedures, provider category, etc.) can be also perceived as a publication of open data and/or knowledge. The user receives summary statistics and information about methods and calculation algorithms. The "reference interpretation" in this context means "a comment or summary to data output given by an expert in the field, usually given to data and values that have more sophisticated background and require careful interpretation with respect to objective uncertainties". In the future, a special subcategory of reference departmental statistics will be formed; the list of such statistics will be approved in an executive regulation issued by the MH. These statistics must be as a rule accompanied by departmental interpretation and can be published in a way allowing the identification of an individual healthcare provider or health insurer. It is necessary to point out that such records bear a high risk of misinterpretation that could lead to erroneous decisions on the side of patients, healthcare providers, or regulatory bodies. As an example of such a parameter, we can name the hospital mortality—a figure requiring sophisticated analytical processing and interpretation with respect to case mix and other factors. Publishing such data is covered by special executive regulations.

Technical background

From the long-term perspective, the Institute of Health Information and Statistics of the Czech Republic (IHIS CR), together with the Ministry of Health, have the aim to make as much information as possible from selected registries available as open data through the deployment and maintenance of the local Open Data Catalogue of the MH [16] and its feeding with data. These datasets are subsequently included in the National Open Data Catalogue [17], which covers and indexes (on 16 December 2020) records on datasets from other 24 local catalogues, such as the local Open Data Catalogue of the Ministry of Finance, or the Vysočina region.

The technical cornerstone for the collection and storage of data about the COVID-19 epidemic on the national level is represented by the Information System of Infectious Diseases (ISID), serving for entering individual cases at the level of regional public health authorities (RPHAs). The system supports entering of individual cases with individual diagnoses. Similar to other public health registries, ISID has been developed and operated in the unified departmental environment that successfully passed the audit by the National Security Authority and the Office for Personal Data Protection of the Czech Republic.

As a reaction to the unexpected epidemic situation, a new portal *COVID-19*: *Current Situation in the Czech Republic* was developed and launched to provide a fully-fledged and trustworthy source of information for the public and media. This online tool provides an interactive overview of the current COVID-19 spreading status in the Czech Republic [7]. This portal also contains a section devoted to datasets containing basic descriptive information on each of them. In open datasets, this information always includes metadata (the name and basic description of the dataset, date and time of last update, standardised dataset scheme); the data are available for download in CSV and JSON formats. In addition, for the purposes of supporting regional and local crisis management, the portal contains an authorised-access-only section where the authorised persons (representatives of local and regional authorities, members of emergency sevices) can—through an online generated token—safely visualise or download regional datasets with aggregated data at the level of the individual municipalities. The basic condition that the data cannot contain details facilitating direct patient identification must be of course always met.

The general background of open data publishing. All methods of release and processing of open data in the NHIS are stipulated by Act No. 372/2011 Coll., detailing that the data can be published or provided i) only as a summary or in an anonymized form, ii) for purposes that are not limited by legislation, and iii) for use in commercial activities, study or scientific purposes, or for public inspection of state-funded subjects. A regime for data publishing and availability is assigned by legislation to every data source, register or component of the NHIS, using categorization as follows: (i) non-public systems, (ii) systems accessible only to legislatively designated readers/editors, (iii) sources of reference statistical data for particular (named) purposes and readers, (iv) sources of statistical data published as open datasets, (v) open sources in the "open data" regime. The NHIS as a unified nationwide information system is designated mainly for: (1) Processing data on the health status of the population, on activities of healthcare providers and on their economy, on healthcare professionals and other workers in health services, for the purpose of obtaining information about the extent and quality of provided health services, for their management and for the creation of health policy. (2) Maintaining National Health Registers and processing data kept in the registers. (3) Maintaining the National Register of Health Services Providers and the National Register of Healthcare Professionals and processing data kept in the registers. (4) Performing and processing sample surveys on health status of the population, on health determinants, on the need and consumption of health services and satisfaction with the services and on expenditure on health services. (5) The needs of science and research in the field of health. (6) Processing the data listed under numbers (1) and (4) (as mentioned above) for statistical purposes and for providing data and statistical information, including information provided to international institutions, in the extent determined by this or other legal regulations.

The process of data acquisition is always in accordance with the scheme presenting dataset production and publication (see Fig 1). The crucial step consists in data standardization and metadata description of the set according to the valid rules of data set creation and best

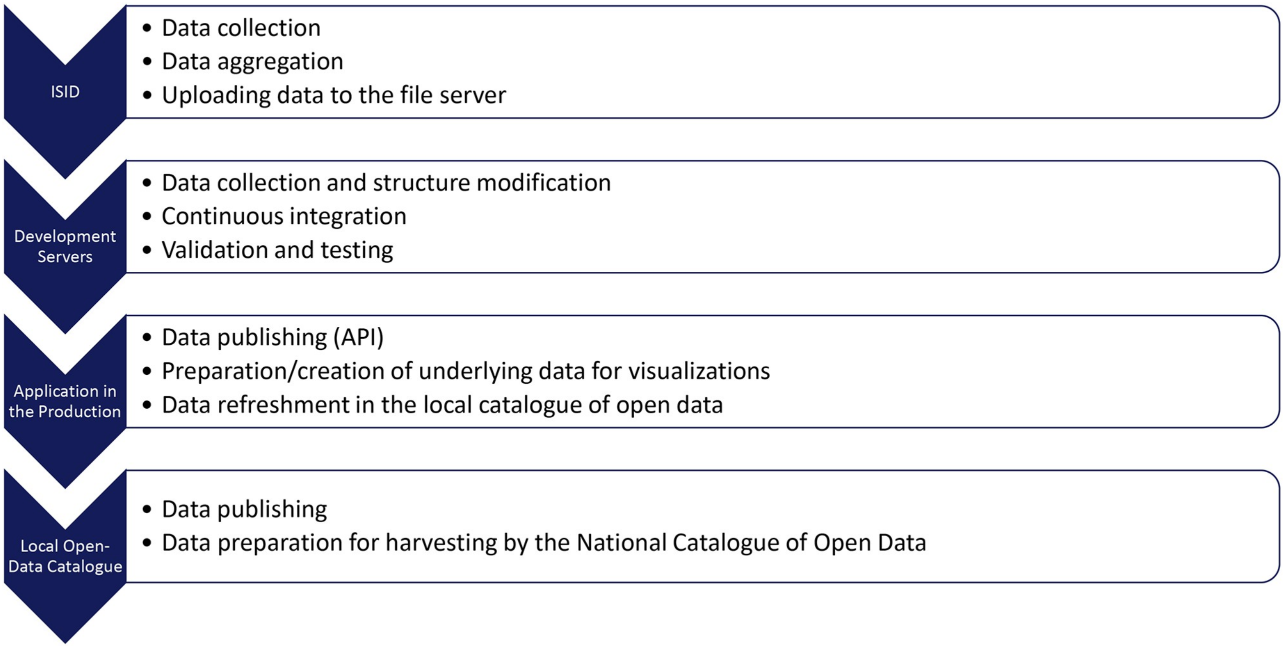


Fig 2. The process of data processing and publishing.

https://doi.org/10.1371/journal.pone.0267397.g002

practices. Once the primary data are stored in central database, they are validated both at the level of the application layer and at the level of the database, ensuring the integrity of individual records. Subsequently, the data are anonymised (i.e., patients' personal data are removed), and uploaded into the analytical database where they are supplemented with data from other registries (place of residence, death records, etc.). The data are subsequently transformed into a final dataset, which is exported into the CSV format and standardised. Finally, the dataset is stored at the internal cloud file server and a periodical transfer (for example once per day/week/month/year) to open data server is set up.

The process of COVID-19 datasets preparation, validation, and publication. The process of entering, processing, and subsequent visualisation of data describing the current epidemiological situation in relation to the new coronavirus SARS-CoV-2 requires meticulous validation. The complete workflow of how the patient suspected of being SARS-CoV-2 positive is referred by his general practitioner or regional public health authority (RPHA) for biological sampling at the testing point, how the patient record flows through the system, and how the record is further processed and stored in the Information System of Infectious Diseases (ISID), is described in detail in [15]. In terms of web-based software deployment architecture, the following tiers have been created (i) development instance, (ii) staging instance (a mirror of production instance), and (iii) production instance.

The production instance [18] of the application *COVID-19*: *An Overview of the Current Situation in the Czech Republic* for publishing freely available datasets is publicly available and provides extensive summaries shown below, in the section Results. To facilitate safe data sharing and, therefore, crisis management on both the national and regional levels (regions and municipalities with extended competence), a special interface was developed, facilitating easy access to selected authorised persons.

In the last stage of the dataset update (which takes place once to twice a day), the data are sent to the application interface of our local open data catalogue of the MH (covering only open data on healthcare, provided by the Ministry of Health). The main datasets are available in the CSV and JSON formats. From this catalogue, the data are harvested by the National Catalogue of Open Data (covering all open data sets, provided by various authorities/publishers, and harversting all data from local catalogues). Both catalogues contain links to datasets and their metadata (name, description, update frequency, keywords, and contact persons). The entire process is captured in Fig 2.

Some datasets might have been extended by additional attributes or their structure might have been changed after publishing (for example, data describing demography or region/district coding, which provide an additional point of view, regarding population distribution in this particular case). To maintain backward compatibility, the datasets are available in all versions using the application programming interface (API) [19]. Finally, after careful data design, processing, preparation, validation and standardisation, the data are published manually in the local catalogue using a graphical user interface, which is accesible only for authorised members of the open data team at the Ministry of Health.

Results

COVID-19 datasets

This section describes epidemiological characteristics that were so far published for the purposes of the online presentation and subsequent manual or automated processing on the portal *COVID-19*: *An Overview of the Current Situation in the Czech Republic*. These data are available either as open data (primary data publishable after necessary processing) or as limited access data (data requiring reference interpretation—reference statistics). The achieved results follow both methodological and well as technological perspective, which is described in the Methods section.

Public datasets. Depending on the dataset, the update is usually performed at 08.00 AM each day. It contains data validated by midnight of the previous day. Each dataset is recorded in the National Catalogue of Open Data, and is therefore fully trackable in the central catalogue system. The main benefits and added values of this catalogue include the possibility of full-text and parametric search, a direct interconnection with the National Catalogue of Open Data of the Czech Republic, and the necessary support for international standards in the publication of open data. Moreover, users can employ these open data to reconstruct almost all summaries published on the web interface *COVID-19*: *An Overview of the Current Situation in the Czech Republic*. The overview of the datasets including their structure and description is detailed in S1 Table.

- Epidemiology—datasets based on everyday reports on newly confirmed and terminated cases of the COVID-19 infection submitted to the ISID system by RPHAs and laboratories.
- Tests—daily and cumulative data on the numbers of performed antigen tests according to type, indication, and region/district.
- Vaccination—distribution and consumption of COVID-19 vaccines number of vaccinated people according to vaccination centres and profession.
- Hospital care availability mapping—daily changes in the available hospital capacity in the individual regions (the total number of available beds, machines, personnel).
- Distribution of personal protective equipment (PPE)—data about the distribution of PPE (glasses, disinfectants, masks, respirators, etc.) from the Administration of National Material Reserves to individual regions.

Restricted-access COVID-19 data. Datasets for predictive modelling—secured datasets covering the period from the beginning of the epidemic to the date one week before the current date. The one-week delay allows publication of a comprehensive, fully validated dataset; the delay is intentional because after the original report on positivity, supplementary data can be added to the original records, or the records can be amended to make them more accurate (e.g. the patient's place of residence) (S2 Table).

Datasets for regions of the Czech Republic: Administration of municipalities with extended competencies (MEC)—datasets intended for the MEC administrations includes a complex overview of the basic epidemiological parameters with a special focus on the vulnerable elderly population groups (65+, 75+) at the geographical level of the individual MECs and districts; selected data are also available at the level of individual municipalities (S3 Table).

Datasets for crisis management of the Czech Republic at the national level—a set of daily reports documenting the current condition and course of the COVID-19 epidemic in the Czech Republic to regularly supply the top management of the Czech healthcare system and of the Integrated Central Crisis Management Team of the Ministry of Health with the latest data, including the current capacities of acute hospital-based care from the Intensive Care Control Centre [20] (S3 Table).

Reports, summaries, and online interactive visualisations of available data

This section provides an overview of selected outputs of individual datasets and their real usage in practice at both regional and national levels. These include web applications primarily intended for the general and expert public, media, and news channels. Besides the MH, representatives of cities, regions, news servers, and other subjects participated in the development of these systems.

COVID-19: The overview of the current situation in the Czech Republic. An official Czech website guaranteed by the Ministry of Health provides the statistics and interactive presentation of the COVID-19 epidemic in the country, as well as time series and cumulative overviews from its beginning (see Fig 3) [7]. Data are updated on a daily basis from the Information System of Infectious Diseases. The tool has been widely used by media, national and regional authorities, and the general public [15].

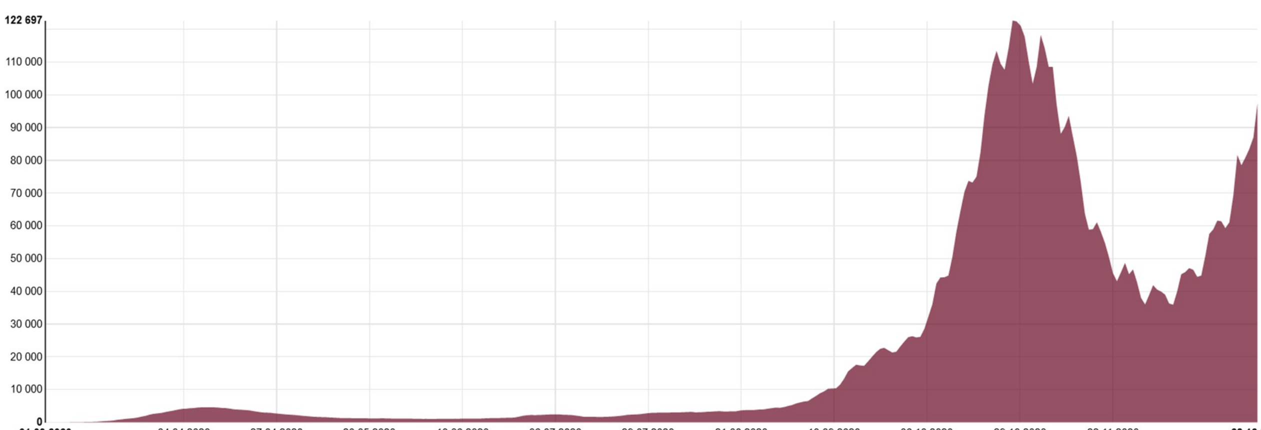


Fig 3. The trend profile of the COVID-19 active cases in the Czech Republic.

https://doi.org/10.1371/journal.pone.0267397.g003

Dashboard for Integrated Central Crisis Management Team of the Ministry of Health. The Dashboard [21] is a management tool for the Integrated Central Management Team providing a real-time analytical overview and data sharing. This tool, along with the data sources, provides complex analytical reporting for the government of the Czech Republic and the supporting workgroups and predictive models for mathematical determination of various scenarios of the epidemic. The application provides the user with the option to simply change the graph parameters and to show data based on visual filters (regions, districts, districts of Prague, age groups, etc.). Access to this application is restricted to the members of the government, top management of MH and of workgroups participating in the crisis management, representatives of the Czech Armed Forces, and other involved institutions.

Predictive models. A simple epidemiological model [22] was based on published dataset describing current epidemiological situation in the Czech Republic and developed at the Institute of Health Information and Statistics of the Czech Republic to help decision-makers understand the course of the epidemics (including the estimation of the effective reproduction number) and to facilitate short-term predictions. The model uses the classical S(E)IR approach [23] with the following compartments: S (susceptible), I (infected, set of compartments), R_{subcl} (subclinical cases), and R (removed, laboratory-confirmed COVID-19 cases). In addition, a more complex SEIR model was adapted [24] and is being used for the explanation and forecasting of the COVID-19 epidemic (see Fig 4).

The use of open data by other bodies. Nowadays, many third-party services are connected directly on selected open data, using a permanent URL address of an API secure token. All these overviews and summaries are built on the datasets introduced in the section Public datasets. The Golemio Data Platform [25] is a set of technical tools for integration, storage, visualisation, and sharing of data managed by experts on municipal data. The Golemio Data Platform, which is built on the Microsoft Power BI technology, aims to provide quality IT

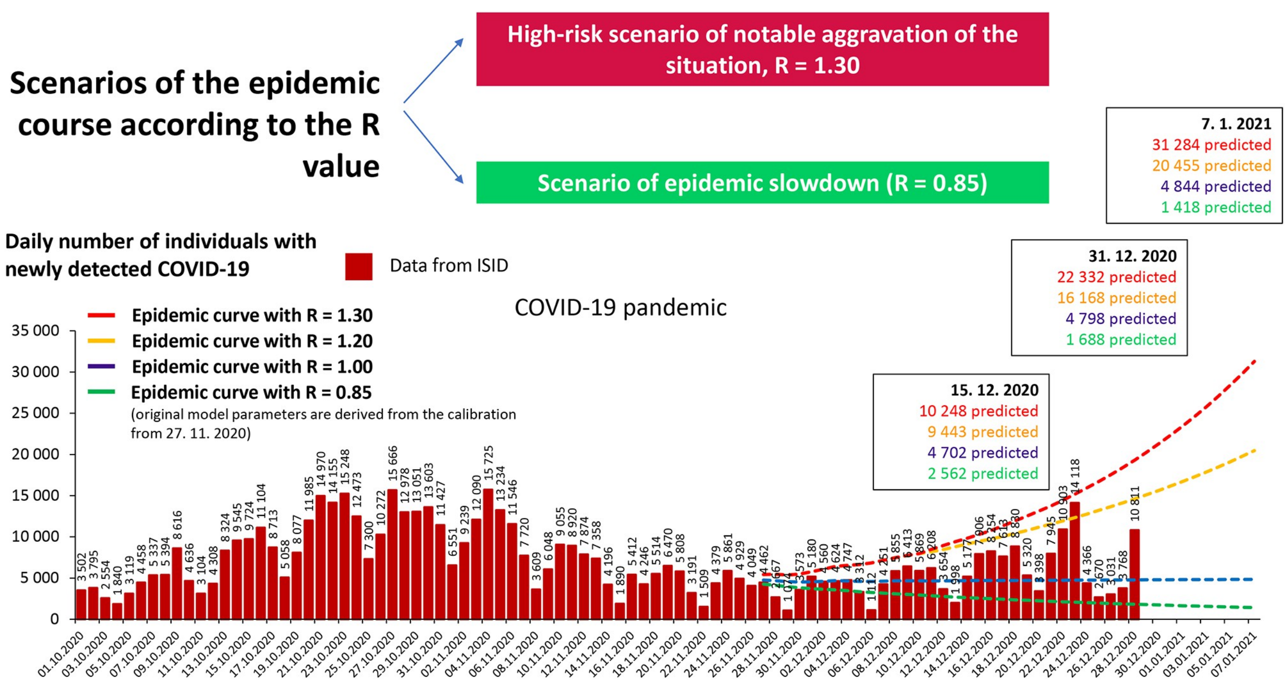


Fig 4. Situation in the entire Czech population: Possible scenarios for various values of the reproduction number R.

https://doi.org/10.1371/journal.pone.0267397.g004

services to the City and District councils in the field of Smart City data processing. Novinky.cz [26], one of the leading news servers in the Czech Republic, uses the Flourish platform for visualisation of the coronavirus spreading in the individual municipalities and Prague districts. The same technology is used e.g. by the BBC, Bosch company, or British Council. Another example of the work with open data is the authenticated map visualisation operated by the Association of Local Councils of the Czech Republic (website is accessible only for ArcGis-Members). Through map layers and the ESRI (Environmental Systems Research Institute) technology, this visualisation clearly depicts the current situation in the municipalities of the Czech Republic and the social services threatened by the COVID-19 disease. The overview allows the visualisation of nursing homes and retirement homes in relation to the current number of COVID-19 cases in the respective municipality.

Statistical overview of the access to the datasets

The datasets are available through an application interface (API) [19] on the platform *COVID-19*: *An Overview of the Current Situation in the Czech Republic*. It is one of the access points to the datasets provided in the context of the COVID-19 epidemic. The API represents an alternative access point to the datasets published in the National Catalogue of Open Data of the Czech Republic and is optimised for automated data processing. Recently implemented functions include, e.g., the introduction of improved data compression, enabling the use of the resources by other domains (Cross-Origin Resource Sharing), minification of the datasets, unification, and standardisation of the attribute names, etc.

Most datasets are provided in the CSV and JSON data formats (including the minified version). The open data schema is described in accordance with the best practices of the standardised open data publication.

As of 30 June 2021, a total of 673 registered users were permitted to use the restricted-access datasets. There were 261 representatives of national, regional and local authorities and crisis management teams, and 412 users accessing data for predictive modelling. Researchers from Czech and foreign universities were the most frequent type of users working with data modelling (34%), followed by private persons (31%), private companies (17%), public sector and institutions (10%) and media (8%). The total number of API calls from 30 September 2020 to 30 June 2021 was 861,705; the *hospitalised patients* being the most popular dataset with 323,004 calls. A total of 56 datasets have already been published on the COVID-19 epidemic, describing comprehensively the current situation in the Czech Republic (the smallest dataset "Basic overview" has 4 kB, the largest dataset "Vaccination according to profession" has 3.2 GB). Additional datasets are constantly being prepared in response to new stakeholders' and users' requirements and requests.

Discussion

The presented framework covering open data as well as secured datasets has been widely used by the general public, researchers, public authorities and decision makers. These datasets have been adopted as an official and guaranteed source for outputs provided by third parties, including public authorities, non-governmental organisations, researchers and online news portals.

Existing platforms and information systems can be employed to enhance emergency risk communication. Information and communication systems should be tailored to users' needs and involve local stakeholders to guarantee the flow of information across sectors [27]. From our perspective, there are several major points in which official and validated national datasets on COVID-19 epidemic have played a crucial role. They support communication from national, regional, and local authorities to the general public, explaining the current situation and adopted measures. Since the datasets have been used by different types of users (from individual data enthusiasts over companies and institutions to the government and mainstream media), the same information reaches different target populations through various communication channels. One of the secondary aims of the authors has therefore been setting up a community interested in processing, analysis and visualisation of health data in order to maximise the impact of these activities across stakeholders, media, and the general public. Based on our long-term experience with involved users and stakeholders in the field of health data, we have identified the following major technical and methodological limitations that complicated the proper and smooth use of published data: (i) The size of selected data files in combination with MS Excel tool that do not allow their retrieval and subsequent processing. (ii) A metadata description in a machine-readable and standardised format that several users have not been able to find and understand. (iii) Reverse changes in data related to data collection and reporting methodology caused distrust and misunderstanding of published information. (iv) The inability to correctly interpret data by representatives of selected media channels due to their lack of understanding of data collection methodology and the importance of epidemiological characteristics.

An extension of existing datasets and the publication of new ones have closely followed the development of the COVID-19 epidemic, the requirements of national and local authorities and the interest of the general public—starting with overviews of the epidemic in spring 2020, over numbers of hospitalisations and availability of intensive care during the autumn 2020 surge, to the progress in vaccination in early 2021. There has been a high demand by media and data analysts for open datasets on a particular agenda right after its start. The preparation of a properly structured and, above all, validated dataset always requires an individual assessment. It must be thoroughly assessed what information the dataset contains, including the level of detail (with respect to the EU General Data Protection Regulation). With any combination of filters over the individual attributes of the dataset, the number of records found must be at least 10, otherwise a risk might occur that a specific person would be traceable. The mere existence of source data is far from being equivalent to publishing a dataset. The complete listing of source databases should always be carefully checked, bearing in mind the subsequent automation in export, preprocessing, analytic procedures, and the deployment of final publication scripts. If the above prerequisites are met, the dataset can be fully published as open data. Some data are included in internal decision-making systems and their publication may be subject to different processes and approvals. This is often in contradiction with requirements for an "absolute" openness and building trust of professionals and the general public. In particular, this was the case of data on the availability of acute care in individual hospitals, or data on the vaccination process in different types of healthcare providers. Thus, if a publication of any dataset is planned, its design and preparation should be an integral part of the whole agenda or process of collection and publication of new data.

Datasets currently published as open data meet the 3-star open data requirements, which makes them machine-readable and facilitates their further usage without restrictions. This is essential for making the data more easily understandable and usable for data consumers. Nevertheless, most datasets contain also geographical data, typically used for territorial differentiation. Such data need additional processing or context. Hence, transformation of these datasets to more complex RDF formats is planned as these formats support/facilitate contextualisation by linking to existing resources. The current course of the epidemics makes the importance of such datasets ever more and more apparent and hence, our goal is to proceed to publish 5-star linked open data in the future versions of the API. This approach will allow us to ensure the necessary interoperability and thus remove the limitations associated with dataset mapping and interconnection. We are moving towards a state where it will be possible to easily interconnect different systems at the data layer, which is absolutely crucial for an effective communication among various web applications.

Personal data processing and ethics are always an important issue in acquisition of health data based on particular method of publishing and sharing. Each dataset must be individually proposed, designed, analysed and validated. Moreover it must comply with general regulations (such as General Data Processing Regulation in EU countries) and national legislation, which are sometimes in contradiction with requirements for data openness and detailedness. This is the reason why three different methods of publishing and sharing have been defined. The mode of publication and legally regulated data availability is clearly defined for each dataset from the National Health Information System; in this regard, the law distinguishes among non-public data, data accessible only to legally defined readers/editors, reference statistical data for identified purposes and identified applicants, and statistical data published in the form of open datasets. For example, too detailed information about a person's sex, age and place of residence can lead to his/her identification, particularly in small communities. It is therefore important to consider the proper degree of geographical (region, district, municipality) and demographic (specific age vs. age group) type of information that is published. Open data represent the publication of secondarily produced datasets and processed statistics, which must always comply with the legislation in force.

Based on literature review, documented and shared public health data can help to prevent catastrophic events. Such data can also help health researchers and, by extension, improve health status of a city, a state, a nation, or the world at large [28]. In the current COVID-19 pandemic, the need to share and collaborate openly has proved to be more important than personal careers or organisational goals [29]. The COVID-19 pandemic has made it clear that the health informatics community agrees with and strongly demands unified frameworks for sharing and exchanging digital epidemiological data and, accordingly with data protection regulations, facilitating the flow of information between health workers, stakeholders, policy makers and the public. The demand for digital data sharing has also raised some crucial points of discussion [30]. Nowadays, many country-specific open data guidelines have been created (e.g. in the United States [31] or Canada [32]). The publication of data in the Czech healthcare sector as an integral part of building e-government—and ensuring the necessary interoperability across the European Union—is based on methodological recommendations under the guarantee of the Department of the Chief Architect of eGovernment of the Ministry of the Interior of the Czech Republic. Specifically, this includes the strategic framework Czech Republic 2030, the action plan for Society 4.0 and the legal regulation of open data in the form of Act No. 106/1999 Coll., on free access to information [33]. In accordance with the strategic document EU implementation of the G8 Open Data Charter [34], essential datasets with maximum information added value for users at national and EU level are identified and subsequently published with the aim of prospective linking with the EU Open Data Portal. As part of a long-term effort to publish not only COVID-19 data, internationally valid recommendations such as the FAIR principles [35], which stress various preconditions for data sharing and emphasise the methodological and transparency requirements in reporting scientific research into the domain of data stewardship [36], are being considered.

Data sharing is one of the pillars of scientific progress, and cooperation between different countries and cultures is the fastest way to accumulate valuable knowledge and face challenges such as the current pandemic [37]. In conjunction with the strategy of the MH in the field of data opening, additional datasets meeting the already implemented standards will be also released, both on COVID-19 related and unrelated topics. The Ministry and its subsidiary organisations possess an immense amount of data in the National Health Information System, which contains long-term records in tens of registries. The plan is to gradually make data of these registries available over the next few years by automated publishing of selected open datasets that will go through the full life-cycle of the dataset according to the given methodology (see section Methods).

Supporting information

S1 Table. Content of public open datasets on COVID-19 in the Czech Republic.
(DOCX)

S2 Table. Content of datasets on COVID-19 in the Czech Republic for predictive modelling.
(DOCX)

S3 Table. Content of regional datasets on COVID-19 in the Czech Republic for local authorities and regional disease management.
(DOCX)

Author Contributions

Conceptualization: Martin Komenda, Vladimír Černý.

Data curation: Martin Komenda, Jiří Jarkovský, Daniel Klimeš, Ondřej Šanca, Jan Mužík, Ondřej Májek.

Formal analysis: Martin Komenda, Jiří Jarkovský.

Investigation: Martin Komenda.

Methodology: Martin Komenda, Jakub Gregor, Jan Mužík, Ondřej Májek, Milan Blaha, Barbora Macková, Jarmila Rážová, Věra Adámková, Vladimír Černý, Jan Blatný, Ladislav Dušek.

Project administration: Martin Komenda.

Resources: Martin Komenda, Daniel Klimeš.

Software: Martin Komenda, Petr Panoška, Matěj Karolyi.

Supervision: Martin Komenda.

Validation: Martin Komenda.

Visualization: Martin Komenda.

Writing – original draft: Martin Komenda, Jakub Gregor.

References

1. Altayar MS. Motivations for open data adoption: An institutional theory perspective. Gov Inform Q. 2018 Oct; 35(4):633–643. https://doi.org/10.1016/j.giq.2018.09.006
2. Our World in Data [Internet]. 2020 [Cited 2021 Jun 28]. Coronavirus pandemic (COVID-19)—the data. https://ourworldindata.org/coronavirus-data.
3. European Centre for Disease Prevention and Control [Internet]. 2020 [Cited 2021 Jun 28]. COVID-19 Coronavirus Data—Weekly (from 17 December 2020. https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data-weekly-from-17-december-2020.
4. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020 May; 20(5):533–534. https://doi.org/10.1016/S1473-3099(20)30120-1 PMID: 32087114

5. CSSEGISandData [Internet]. 2020 [Cited 2021 Jun 28]. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. https://github.com/CSSEGISandData/COVID-19.
6. COVID-19 Data Hub [Internet]. 2020 [Cited 2021 Jun 28]. https://github.com/covid19datahub/COVID19.
7. Komenda M, Karolyi M, Bulhart V, Žofka J, Brauner T, Hak J, et al. COVID-19: Overview of Current Situation in the Czech Republic. Diseases at the Moment. Prague: Ministry of Health of the Czech Republic. 2020 [Cited 2020-12-30]. https://onemocneni-aktualne.mzcr.cz/covid-19.
8. Centers for Disease Control and Prevention [Internet]. 2020 [Cited 2021 Jun 25]. CDC COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/.
9. Etalab [Internet]. 2020 [Cited 2021 Jun 25]. Coronavirus epidemic monitoring dashboard in France. https://www.data.gouv.fr/en/reuses/tableau-de-bord-de-suivi-de-lepidemie-de-coronavirus-en-france/.
10. Robert Koch Institute [Internet]. 2020 [Cited 2021 Jun 25]. COVID-19 (Coronavirus SARS-CoV-2). https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/nCoV.html.
11. Federal Ministry for Digitization and Business Location [Internet]. Austrian COVID-19 Open Data Information Portal [Cited 2021 Jun 25]. https://www.data.gv.at/covid-19/.
12. Open Data Initiative of the Government of Spain [Internet]. [Cited 2021 Jun 25]. https://datos.gob.es/.
13. Public Health England [Internet]. Coronavirus (COVID-19) in the UK [Cited 2021 Jun 25]. https://coronavirus.data.gov.uk.
14. Alamo T, Reina DG, Mammarella M, Abella A. Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. Electronics 2020; 9(5):827. https://doi.org/10.3390/electronics9050827
15. Komenda M, Bulhart V, Karolyi M, Jarkovský J, Mužík M, Májek O, et al. Complex reporting of coronavirus disease (COVID-19) epidemic in the Czech Republic: use of interactive web-based application in practice. J Med Internet Res. 2020 May; 22(5):e19367. https://doi.org/10.2196/19367 PMID: 32412422,
16. Ministry of Health of the Czech Republic [Internet]. 2019 [Cited 2021 May 14]. Open Data Catalogue. https://opendata.mzcr.cz/.
17. Ministry of Interior of the Czech Republic [Internet]. 2018 [Cited 2021 May 14]. National Catalogue of Open Data. https://data.gov.cz/datasets.
18. Ministry of Health of the Czech Republic [Internet]. 2020 [Cited 2021 May 14]. Diseases at the Moment: Overview of Current Situation in the Czech Republic. https://onemocneni-aktualne.mzcr.cz/.
19. Ministry of Health of the Czech Republic [Internet]. 2020 [Cited 2021 May 14]. COVID-19 in the Czech Republic: Open datasets and datasets for download. https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19.
20. Ministry of Health of the Czech Republic [Internet]. 2020 [Cited 2021 May 16]. COVID-19: Online Controlling of Intensive Care. https://dip.mzcr.cz/.
21. Ministry of Health of the Czech Republic [Internet]. 2020 [Cited 2021 May 14]. COVID-19 Dashboard. https://dashboard.uzis.cz.
22. Májek O, Ngo O, Jarkovský J, Komenda M, Rážová J, Dušek L, et al. 2020. Modelling the first wave of the COVID-19 epidemic in the Czech Republic and the role of government interventions. Available from: https://www.medrxiv.org/content/10.1101/2020.09.10.20192070v1.
23. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis. 2020 May; 20(5):553–558. https://doi.org/10.1016/S1473-3099(20)30144-4 PMID: 32171059
24. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health. 2020 May; 5(5):e261–e270. https://doi.org/10.1016/S2468-2667(20)30073-6 PMID: 32220655
25. Golemio Data Platform [Internet]. [Cited 2021 Jun 19]. https://golemio.cz/en.
26. Novinky cz [Internet]. [Cited 2021 Jun 19]. https://www.novinky.cz/
27. World Health Organization. Communicating risk in public health emergencies: a WHO guideline for emergency risk communication (ERC) policy and practice. Geneva: World Health Organization; 2017. ISBN 978-92-4-155020-8.
28. D'Agostino M, Samuel NO, Sarol MJ, de Cosio FG, Marti M, Luo T, et al. Open data and public health. Rev Panam Salud Publica. 2018; 42: e66. https://doi.org/10.26633/RPSP.2018.66 PMID: 31093094
29. Park JJH, Mogg R,Smith GE, Nakimuli-Mpungu E, Jehan F, Rayner CR, et al. How COVID-19 has fundamentally changed clinical research in global health. Lancet Glob Health. 2021;May; 9(5):e711–e720. https://doi.org/10.1016/S2214-109X(20)30542-8 PMID: 33865476

30. Dagliati A, Malovini A, Tibollo V, Bellazzi R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. Brief Bioinform. 2021 Mar 22; 22(2):812–822. https://doi.org/10.1093/bib/bbaa418 PMID: 33454728

31. Government of the USA [Internet]. [Cited 2021 Jun 19]. The home of the U.S. Government's open data. https://www.data.gov/.

32. Government of Canada [Internet]. [Cited 2021 Jun 19]. Open Data. https://open.canada.ca/en/open-data.

33. Legislativní prostředí otevřených dat [Legislative environment of open data] [Internet]. [Cited 2021 Jun 19]. https://opendata.gov.cz/legislativa:start.

34. European Commission [Internet]. [Cited 2021 Jun 24]. EU implementation of the G8 Open Data Charter. https://digital-strategy.ec.europa.eu/en/library/eu-implementation-g8-open-data-charter.

35. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15; 3:160018. https://doi.org/10.1038/sdata.2016.18 PMID: 26978244

36. Boeckhout M, Zielhuis GA, Bredenoord AL. The FAIR guiding principles for data stewardship: fair enough? Eur J Hum Genet. 2018 Jul; 26(7):931–936. https://doi.org/10.1038/s41431-018-0160-0 PMID: 29777206

37. Rios RS, Zheng KI, Zheng MH. Data sharing during COVID-19 pandemic: what to take away. Expert Rev Gastroenterol Hepatol. 2020 Dec; 14(12):1125–1130. https://doi.org/10.1080/17474124.2020.1815533 PMID: 32842793

Viewpoint

Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic

Martin Komenda[1,2,3], PhD; Vladimír Černý[4,5], MD, PhD; Petr Šnajdárek[6], MSc; Matěj Karolyi[1,2], MSc; Miloš Hejný[1,2], BSc; Petr Panoška[1,2], BSc; Jiří Jarkovský[1,2], PhD; Jakub Gregor[1,2], PhD; Vojtěch Bulhart[1,2]; Lenka Šnajdrová[1,2], PhD; Ondřej Májek[1,2], PhD; Tomáš Vymazal[7], MD, PhD; Jan Blatný[4,8], MD, PhD; Ladislav Dušek[1,2], PhD

[1]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[2]Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic
[3]Department of Simulation Medicine, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[4]Ministry of Health of the Czech Republic, Prague, Czech Republic
[5]Clinic of Anaesthesiology, Perioperative and Intensive Medicine, Masaryk Hospital in Ústí nad Labem, Ústí nad Labem, Czech Republic
[6]General Staff, Czech Armed Forces, Prague, Czech Republic
[7]Clinic of Anaesthesiology, Resuscitation and Intensive Medicine, University Hospital in Motol, Second Faculty of Medicine of Charles University, Prague, Czech Republic
[8]Department of Paediatric Haematology and Biochemistry, University Hospital Brno, Brno, Czech Republic

Corresponding Author:
Martin Komenda, PhD
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Kamenice 126/3
Brno, 62500
Czech Republic
Phone: 420 549 49 4469
Email: komenda@iba.muni.cz

Abstract

In the Czech Republic, the strategic data-based and organizational support for individual regions and for providers of acute care at the nationwide level is coordinated by the Ministry of Health. At the beginning of the COVID-19 pandemic, the country needed to very quickly implement a system for the monitoring, reporting, and overall management of hospital capacities. The aim of this viewpoint is to describe the purpose and basic functions of a web-based application named “Control Centre for Intensive Care,” which was developed and made available to meet the needs of systematic online technical support for the management of intensive inpatient care across the Czech Republic during the first wave of the pandemic in spring 2020. Two tools of key importance are described in the context of national methodology: one module for regular online updates and overall monitoring of currently free capacities of intensive care in real time, and a second module for online entering and overall record-keeping of requirements on medications for COVID-19 patients. A total of 134 intensive care providers and 927 users from hospitals across all 14 regions of the Czech Republic were registered in the central Control Centre for Intensive Care database as of March 31, 2021. This web-based application enabled continuous monitoring and decision-making during the mass surge of critical care from autumn 2020 to spring 2021. The Control Center for Intensive Care has become an indispensable part of a set of online tools that are employed on a regular basis for crisis management at the time of the COVID-19 pandemic.

(J Med Internet Res 2022;24(2):e33149) doi: 10.2196/33149

KEYWORDS

COVID-19; coronavirus; intensive care; inpatient care; online control center; prescription; open data; ICU; monitoring; strategy; development; app; function; Czech Republic; inpatient; crisis management

Introduction

The COVID-19 pandemic has placed an urgent burden on health care systems, communication among their components, and health care management. Although the majority of patients have a mild or even asymptomatic course of the disease, approximately 8% to 15% of patients require hospitalization during the infection or afterward [1]. Approximately 21% to 36% of patients hospitalized due to COVID-19 are admitted to the intensive care unit (ICU) [2-4], and the average mortality of those admitted to the ICU is approximately 30% [5,6].

In the majority of developed countries, ICU beds are often close to their full capacity even under normal circumstances, let alone during the COVID-19 pandemic. Modeling studies suggest that a pandemic of severe influenza or another similar disease would require ICU and mechanic ventilation capacity that is significantly greater than what is available and that many patients who would require a ventilator might not have access to one [7,8].

By definition, epidemics and disasters will result in many patients arriving in a continuous stream, with shortages of necessary technology, beds, oxygen support, ventilators, medications, as well as trained health care personnel. The high likelihood of such conditions emphasizes the importance of having a framework that is suitably constructed to allow users to anticipate and adapt to these inevitable complexities and challenges [8]. Beds, equipment, and medications should be monitored ideally in “real time” and electronically, with knowledge of both their absolute numbers and whereabouts immediately available to system leaders [9]. For example, ventilatory support is an absolute necessity for the survival of critically ill patients and may be the single most important therapy that dictates the outcomes. Moreover, ventilators are highly demanding for staff qualification; they are therefore likely to be the limiting factor in any hospital’s ability to accommodate a large surge of mass critical care [10]. The COVID-19 pandemic has generated a sense of urgency that will allow the adoption of innovation without the logistical barriers and path dependencies that we have become accustomed to. Web-based solutions can fill a critical gap for the allocation of health care resources [11].

The Czech National Control Centre for Acute Inpatient Care (CNCC-AIC) was established under the auspices of the Ministry of Health of the Czech Republic in May 2020 to provide strategic data-based and organizational support for individual regions and for providers of acute care at the nationwide level. Its main purpose is to monitor hospital capacities, to provide analytical reporting on these capacities, and to enable the overall management of health care facilities run by all inpatient care providers in the Czech Republic.

The operative online tools of key importance employed by the CNCC-AIC involve the Control Centre for Intensive Care (CC-IC) [12], an online database updated on a daily basis, which provides an overview of capacities of inpatient care (classified into categories of acute, long-term, and supportive care) and a dedicated “Clinic” module, which is part of the Information System of Infectious Diseases (a database of people with laboratory-confirmed COVID-19).

In accordance with the defined methodology, the CC-IC was brought into operation in April 2020 to enable the continuous monitoring as well as daily reporting of available data on the occupancy rate of inpatient beds. The aim of this viewpoint paper is to present the methodology of managing capacities in Czech hospitals during the ongoing COVID-19 pandemic. A newly developed and robust technical solution, namely the CC-IC, is also introduced. The methodological background of the design and implementation of this unique practical operative tool, which is built on the long-term experience of experts across the Czech health care system, and experience with its practical use are presented.

Structure of the CC-IC

Overview

The CC-IC is an entirely new platform that has been developed since the start of the COVID-19 pandemic. This has been achieved in mutual cooperation among the following groups of involved stakeholders and experts to assess recent progress of the epidemic and recommend measures for crisis situations caused by the COVID-19 epidemic in the Czech Republic: (1) a team of regional coordinators of intensive care; (2) the Integrated Central Management Team, which includes representatives from the Ministry of Health, Czech Armed Forces, National Institute of Public Health, National Agency for Communication and Information Technologies, and Institute of Health Information and Statistics of the Czech Republic; and (3) a team of developers from a joint workplace of the Institute of Health Information and Statistics of the Czech Republic and the Institute of Biostatistics and Analyses at the Faculty of Medicine of Masaryk University.

The CC-IC provides two tools of key importance to offer effective support for the CNCC-AIC: (1) a module for regular online updates and overall monitoring of currently free capacities in hospitals (health care technology/medical devices, beds, staff) in real time, and (2) a module for online entering and overall record-keeping of requirements on medications for COVID-19 patients. No patient personal data or medical records are stored in the CC-IC application; thus, no anonymization or pseudonymization algorithms are used.

The CC-IC also specifically monitors the reprofiled capacity (ie, beds that had previously, under normal circumstances, in a given health care facility) been intended for the provision of care of another type or another specialty.

The CC-IC is a web application that has been continuously developed and run by the Institute of Health Information and Statistics of the Czech Republic, which cooperates closely with the Ministry of Health of the Czech Republic. Since the very beginning, the development has been based on requirements and needs for the coordination of intensive care, as required by health care facilities in the Czech Republic (Figure 1).

Use of the application (available in the Czech language only) relies on several basic principles: (1) user authentication and authorization, (2) definition of user roles and rights to edit and read, (3) each new user registration is approved by the CNCC-AIC, and (4) private and free email addresses are not allowed. CC-IC users include representatives of individual health care facilities (management, coordinators of intensive care, physicians, and nurses), emergency medical service workers, pharmacy workers, and members of the Integrated Central Management Team.

The use case diagram in Figure 2 provides an overview of the entire CC-IC system, including key functions assigned to users in specific roles. The main objective of this diagram is to depict the various actors as well as interactions that are available to these actors.

Use cases provide insight into the basic structure of CC-IC functional requirements. Based on modeling of individual use cases, which are associated with the primary actors, we can see individual parts of the system so that we can decompose and divide the whole system into separate submodules. Moreover, the CC-IC provides an authenticated application programming interface (API) for secure sharing of selected data sets, which consists of access to a list of requirements on medications and an export of currently free capacities in hospitals, including their history.

Figure 1. Schematic diagram of the Control Centre for Intensive Care (CC-IC).

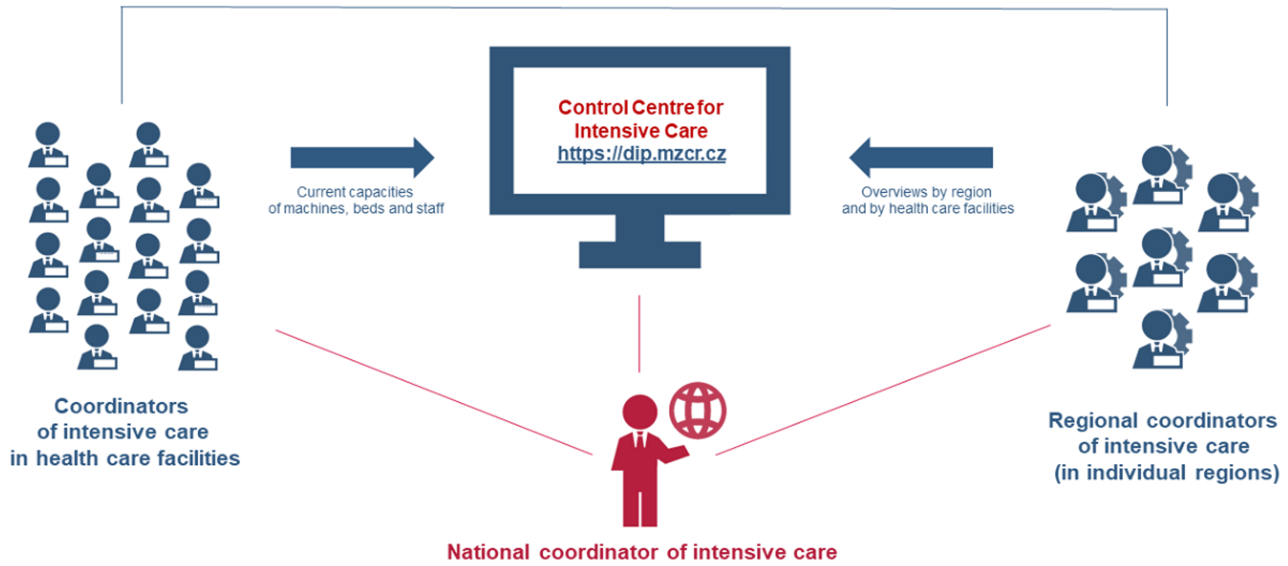
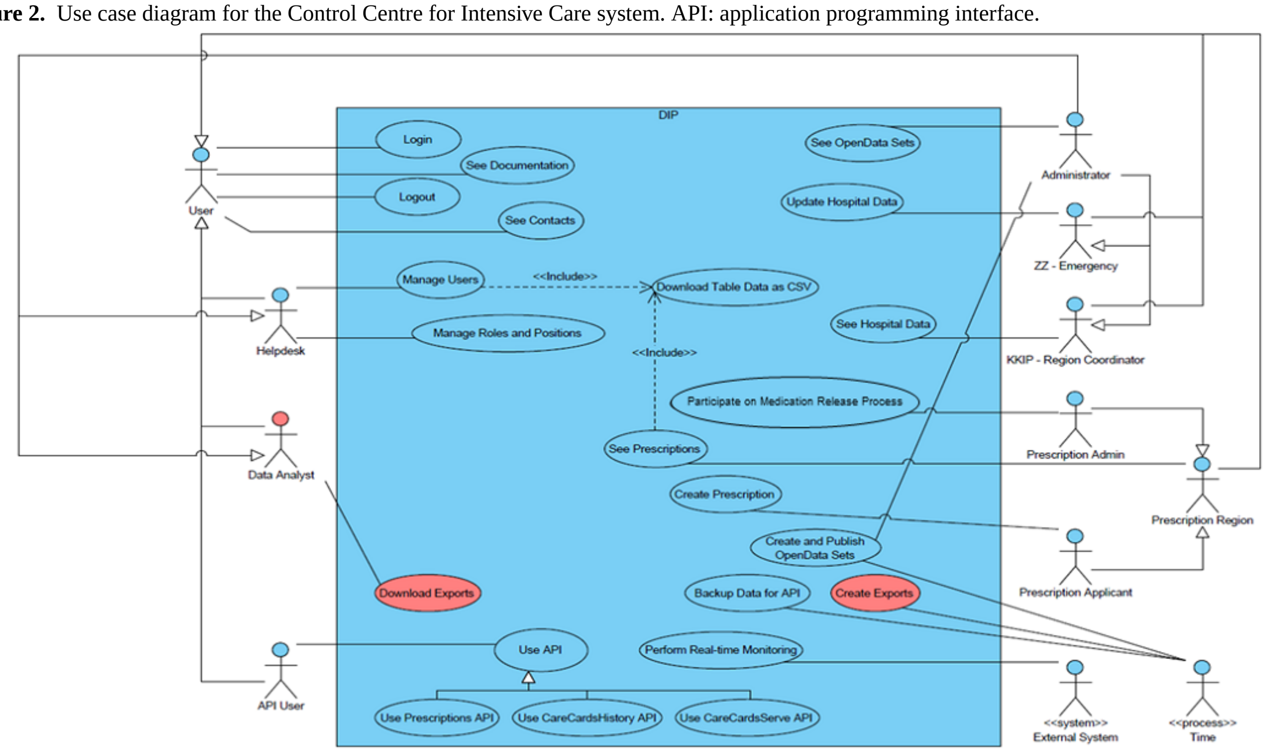


Figure 2. Use case diagram for the Control Centre for Intensive Care system. API: application programming interface.



Regular Monitoring of Health Care Providers and the ICU Network

On April 9, 2020, the Ministry of Health of the Czech Republic adopted an exceptional measure, ordering all acute care providers to use the CC-IC to immediately report any changes in capacities of intensive care related to COVID-19 patients. Since the launch of the CC-IC, the availability of free capacity of machines, beds, and staff has been monitored on a regular basis (Textbox 1). Hospitals have been requested to report their intensive care capacities on a daily basis, usually at the end of the day. The system allows for continuous updates at any time; more frequent entering of the latest data would nevertheless bring an unnecessary burden to hospital staff at the time of a significant shortage of capacities in terms of personnel, time, and equipment.

Based on data in the CC-IC, it is possible to coordinate the provision of inpatient care centrally, on a nationwide level. In each individual region, hospitalizations of patients who need admission to an intensive care facility (ICU and/or emergency department) are managed and coordinated by the so-called regional coordinator of intensive care, who is a physician specialized in anesthesiology and intensive care medicine. These physicians cooperate closely with regional emergency medical services and with representatives of regional authorities. Apart from that, each health care facility has prepared emergency beds (ie, inpatient beds that could be used to provide care for COVID-19 patients if all existing capacities are occupied).

The system reacts to immediate needs by individual regions, depending on the current occupancy rate of machines and beds, combined with the number of COVID-19 patients. If an increase in the number of new patients is detected within the CC-IC, it is necessary to prepare health care facilities for a crisis situation. This involves limiting capacities for patient admissions (ie, only admit patients who require intensive and emergency care, and suspend the provision of nonemergency care); discharging as many patients as possible to home care; plan an increased bed supply in health care facilities for an increase in COVID-19 patients; and provide a sufficient number of competent health care professionals.

The module for regular online updates and overall monitoring of the currently free capacities of intensive care in real time is predominantly used by coordinators of intensive care in individual health care facilities. A total of 134 intensive care providers from all 14 regions of the Czech Republic are registered in the central CC-IC database. As of March 31, 2021, a total of 927 users were registered who used this module to update data on intensive care capacities in predefined intervals (ie, at least once a day during critical periods of the COVID-19 pandemic), who made 86,565 sessions in total.

A user-friendly and simple interface (Figure 3) is available to users, who can therefore update occupancy rates of available capacities (in terms of machines, beds, and staff) very easily. The entire application is fully responsive and therefore entirely compatible with all types of devices, namely smartphones, tablets, laptops, and desktops. Free capacities for COVID-19–positive versus COVID-19–negative patients are distinguished only for beds. Entering a free bed into the system means that this bed is fully functional (ie, all required equipment and corresponding health care professionals are available). Monitoring of staff and equipment in and out of service has turned out to be ineffective and unsustainable on a long-term basis; this function was therefore removed from the system upon the decision of the Central Integrated Management Team. Across the entire module, each individual card has one of three background colors to provide a quick overview of free capacities: white, <30% to 100%> free capacity; yellow, <10% to 30% free capacity; and red, <0% to 10% of free capacity.

Textbox 1. Overview of monitored information on capacities of health care facilities providing inpatient care in the Czech Republic.

Health care technology/medical devices

- Extracorporeal membrane oxygenation
- Mechanical ventilation in accident and emergency (A&E) departments and in intensive care units (ICUs) for adults
- Continuous renal replacement therapy
- Intermittent renal replacement therapy
- Transport ventilators
- Anesthesia machine with mechanical ventilator

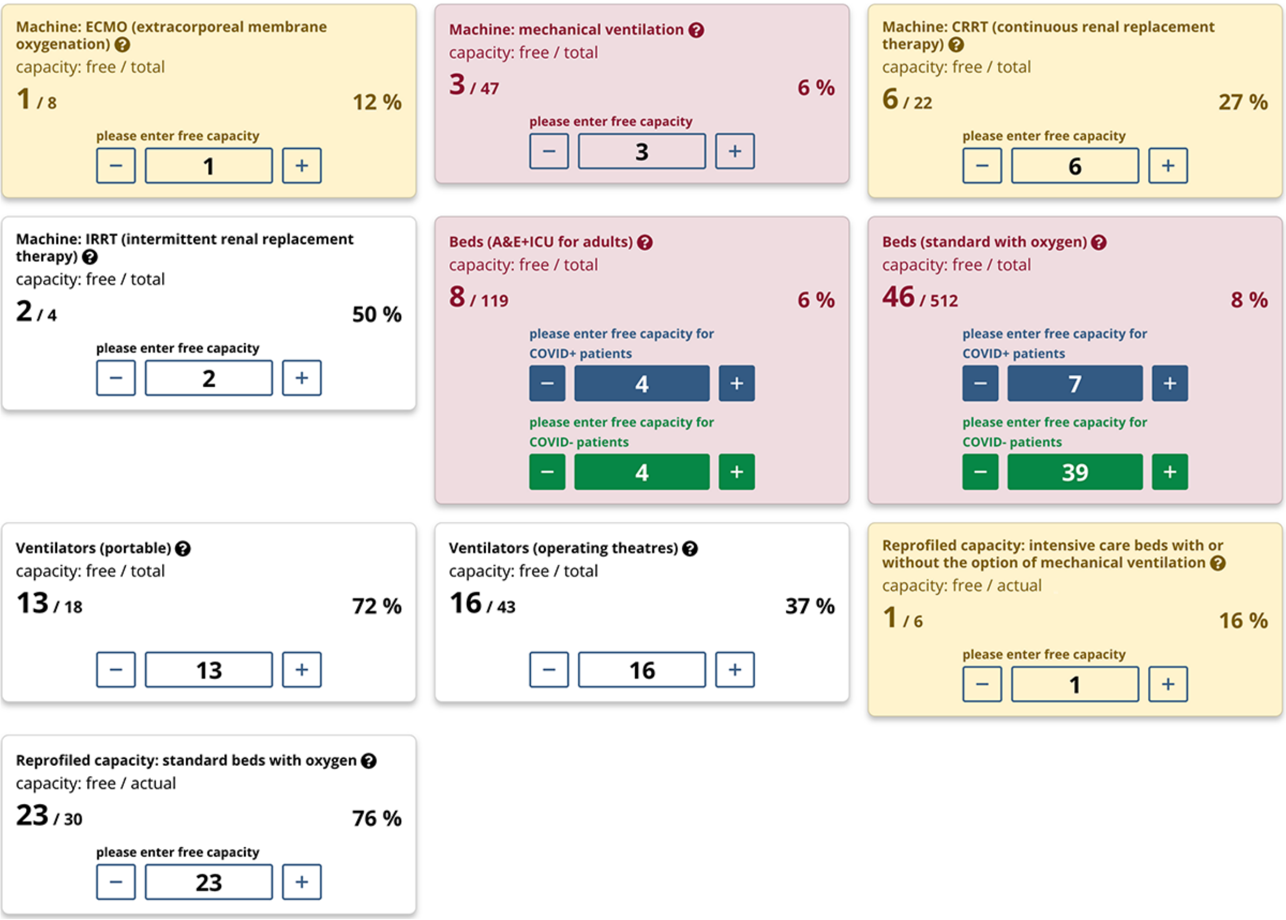
Hospital beds

- A&E and ICU beds for adults; a distinction is made between beds for COVID-19–positive and COVID-19–negative patients
- Standard beds with oxygen; a distinction is made between beds for COVID-19–positive and COVID-19–negative patients
- Reprofiled capacity: beds with or without the option of ventilatory support devices

Staff

- Physicians (A&E and ICU for adults)
- Nurses (A&E and ICU for adults)

Figure 3. Interface for updates of free capacities. Overview of intensive care capacities for the capital, Prague, as of December 30, 2020. A&E: accident and emergency; ICU: intensive care unit.



Two different reports are available to provide a quick overview of the latest data. These reports are primarily used by regional coordinators of intensive care, members of the operational team of the CNCC-AIC, members of management of individual health care facilities, and top officials of the Ministry of Health of the Czech Republic. The first report is a summary data table providing a clearly arranged view of all health care facilities, including real-time absolute and relative occupancy numbers as regard to available capacities. The second is an aggregated visualization in the form of context cards, which also show real-time absolute and relative occupancy numbers in regard to available capacities, but users can further choose among the nationwide view, views for individual regions, or those for individual health care facilities.

The example screenshot from the CC-IC application in Figure 3 also demonstrates an aggregated visualization of key parameters monitored in real time for the capital, Prague. Four cards have a yellow background because the available capacity of corresponding parameters was reported to be below 30% at the time.

Owing to the timely and strict measures at the national level, the first COVID-19 wave in the spring of 2020 did not bring significant pressure to hospital-based care, with a maximum daily prevalence of around 4500 active cases and less than 100 ICU beds being occupied with COVID-19 patients among the Czech population of over 10 million people. The situation changed in the autumn of 2020; the number of active cases peaked twice to approximately 120,000, first in early November 2020 and then in early January 2021 (Figure 4, top). The number of patients requiring intensive care at the ICU fluctuated at around 1200 in these periods. Requirements for intensive care increased to an even more critical level in the second half of February 2021, with 1574 ICU beds occupied by COVID-19 patients on March 1, 2021 (Figure 4, middle).

The increasing numbers of hospitalized patients with COVID-19 during October 2020, as steadily monitored by the CC-IC, required increased capacities of standard beds and ICU beds. The total numbers of both types of beds were increased by repurposing the existing beds and adding new beds, according to the capabilities of individual hospitals. The availability of mechanical ventilators as an essential equipment of intensive care followed the overall trends: it decreased from nearly 1200 machines on September 1, 2020, to less than 700 at the end of February 2021 (Figure 4, bottom). If a temporary depletion of ICU capacities occurred in some regions, patients were transferred to hospitals in other regions in which free capacities were available according to the CC-IC report. Over 100 patients were transferred between hospitals during autumn 2020 and winter 2021, usually from small local hospitals to large hospitals in regional capitals or in Prague. All of these decisions were taken by the national and regional intensive care coordinators in cooperation with the Central Integrated Management Team.

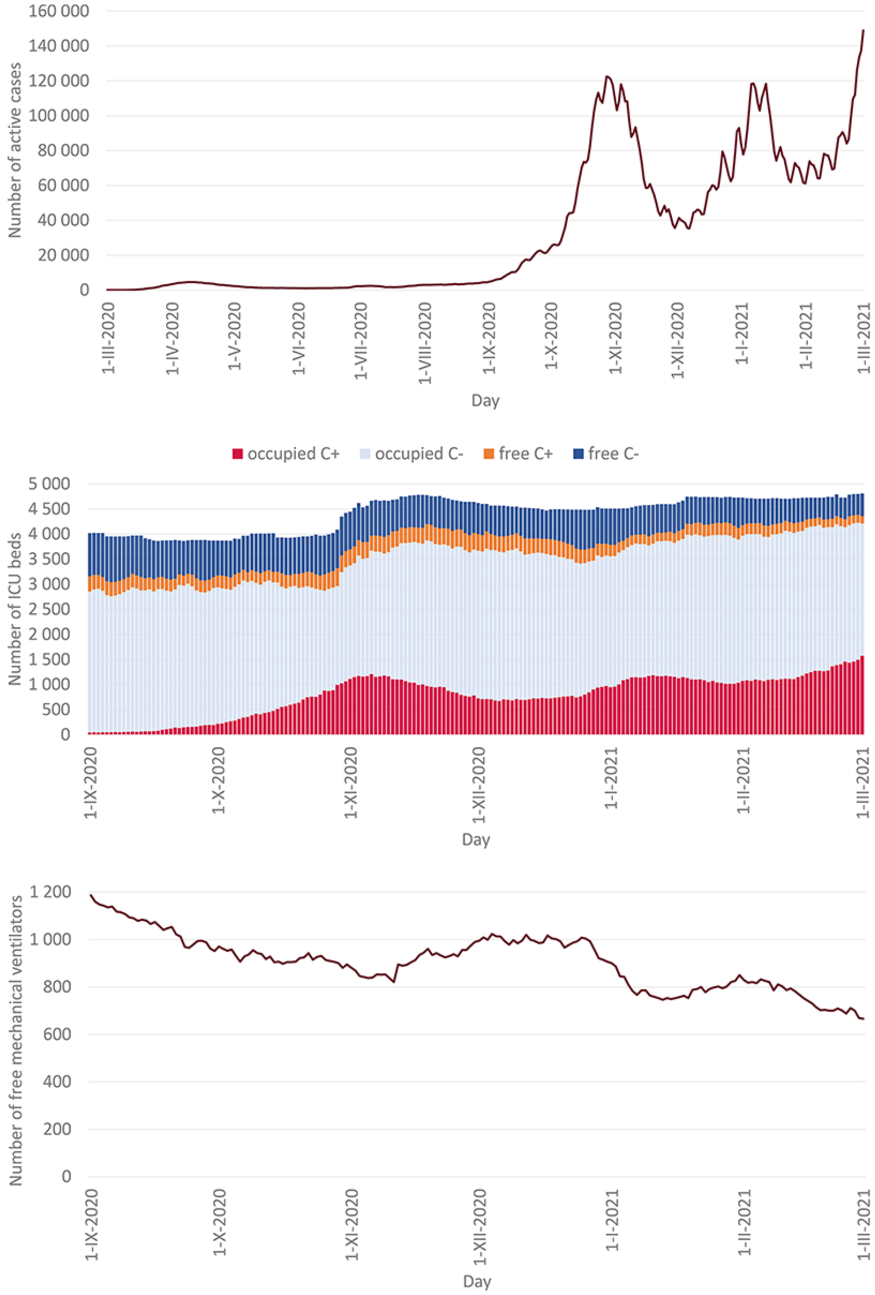


Figure 4. Time trends of COVID-19 cases and related intensive care, from top to bottom: number of active cases (prevalence), occupied and free intensive care unit (ICU) beds for COVID-19 patients and other patients, and availability of mechanical ventilators.

Prescription of COVID-19 Medications

With an increasing number of hospital admissions of COVID-19 patients who need medication to manage the disease, it was necessary to secure the systematic record-keeping and distribution of these medications to individual health care facilities. For this reason, a dedicated module was developed within the CC-IC, making it possible to request any of three therapies against COVID-19: remdesivir, convalescent plasma, and favipiravir. The tool has been intentionally designed in a rather flexible manner so that it can be easily adapted and employed after the end of the pandemic, thus providing a generally applicable ordering system for any medication.

The CC-IC enables physicians from any involved health care facility to request a certain medication for a specific patient. The module has three basic user roles: (1) applicant for medication (physician), which has been assigned to an approved group of registered users from individual health care facilities who are thus entitled to enter new requests from their respective health care facility; (2) regional coordinator of intensive care, who is entitled to enter new requests from health care facilities from across the entire region and can also see an overview of requests not only from the region but from the Czech Republic as a whole; (3) virtual indication group, with members entitled to approve requests for medications from the respective region and they can also see an overview of requests from across the Czech Republic; and (4) contact persons from pharmacies linked to specific health care facilities, who are entitled to supply the approved medications.

Mutual cooperation among representatives of the Ministry of Health of the Czech Republic, regional coordinators of intensive care, members of the virtual indication group, and representatives of pharmacies led to the development of a structured form that is employed to send a request for a specific medication. The form consists of three parts: (1) health care facility, containing information on the physician and hospital requesting the medication; (2) specification of therapy, containing information on the requested medication and the date of supply; and (3) patient description, which is an anonymized description of the patient's condition and their risk factors (without mentioning any personal data in conformity with personal data protection according to General Data Protection Regulation).

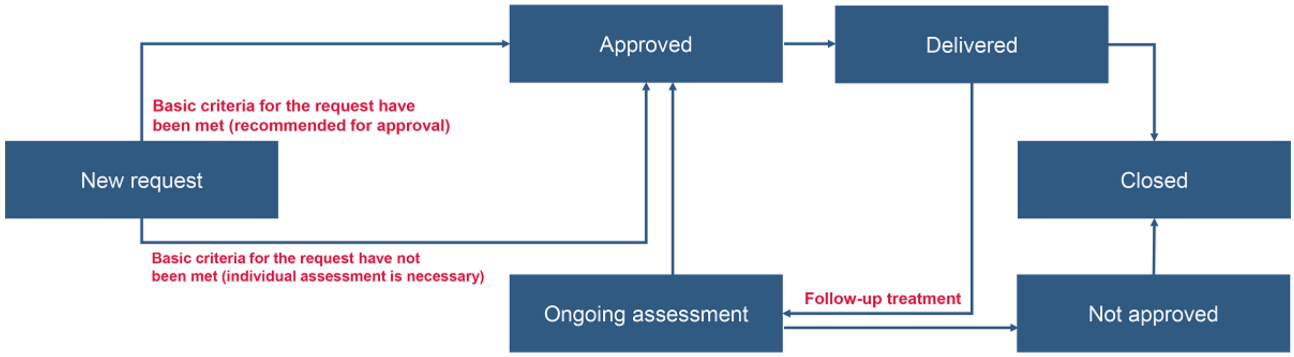
After the form is filled out and sent, validation of individual items is performed within the CC-IC, the request is stored in the central database, and a notification is sent to all users involved (physician, member of the virtual indication group in the respective region, pharmacy). The medication must be explicitly approved before being supplied by a respective pharmacy that secures supplies for a given health care facility. In the final step, the request is closed for the purpose of archiving and subsequent statement of charges for the respective health insurance company. In exceptional cases, request for follow-up treatment can occur after the administration of remdesivir; such requests can also be processed by the CC-IC. The life cycle of a request for a medication within the CC-IC is shown in Figure 5.

All health care facilities providing acute care began to report, by means of the CC-IC, the current use of these capacities on April 9, 2020, when the first version of the system was launched. The first version only contained one module, which enabled regular online updates and overall monitoring of currently free capacities of intensive care in real time. On June 10, 2020, the second module was launched for the online entering and overall record-keeping of requirements on medications for COVID-19 patients.

A major added value of the entire CC-IC system is the fact that it was designed to meet the current needs of acute medicine in response to the unexpected burden on inpatient care capacity caused by the COVID-19 pandemic. The data that the CC-IC allows to collect provide a key support for data-oriented decision-making processes within ICU occupancy crisis management. Thanks to the CNCC-AIC and a proper methodological setup, it is possible to effectively coordinate, for example, elective care, which can have a negative effect on patients' health when being moved or postponed. The following paragraphs briefly describe the functions of both of the above-mentioned modules as well as analytical reports and open data sets that have been provided as an outcome of the CC-IC database.

As of December 30, 2020, a total of 4759 requests for medication were registered in the CC-IC, out of which 4666 (98.05%) were requests for remdesivir. The remaining 93 (1.95%) requests were for convalescent plasma. On that date, a total of 496 users entitled to work with requests for medications were registered in the system; these users included physicians in individual health care facilities, regional coordinators of intensive care, pharmacy workers, and members of the virtual indication group. Between the first wave and the second wave of the COVID-19 pandemic, there was a major change in the functioning of the module for the sending and record-keeping of requests for medications. A new version was implemented in September 2020, fully supporting the entire process of requests for medications (entering, validation of entered data, sending, approval/rejection, supply, conclusion), including email notifications. The form itself for sending a request contains a list of key information necessary for subsequent approval and the supply of a specific medicine (see Multimedia Appendix 1).

Figure 5. Description of the life cycle of a request for a medication in the Control Centre for Intensive Care (CC-IC).



Practical Use of Open Data

Data on free capacities are available as records in the National Catalogue of Open Data [13,14] in a regularly updated data set that describes two waves of the COVID-19 pandemic in the Czech Republic.

The data set contains overviews of machine equipment (extracorporeal membrane oxygenation, mechanical ventilation, continuous renal replacement therapy, intermittent hemodialysis, ventilators—portable and those in operating theaters) and occupancy rate of inpatient beds (only beds in the emergency department and ICU for adults and standard beds with oxygen in the entire hospital). Data are regularly updated for beds suitable for the provision of the requested type of care (ie, necessary staff as well as corresponding material and technical equipment must also be available). The definition of the necessary number of staff for a given type of bed can be different among individual hospitals. These data have been available since September 1, 2020, when the obligation for hospitals to update free capacity was reset to a 1-day interval.

On top of that, for the purpose of the transfer of record-based (nonaggregated) data directly from the CC-IC to the Institute of Health Information and Statistics of the Czech Republic analytical team and to external authorized systems (run by the leadership of the Capital of Prague, for example), an API was developed to provide a machine-readable alternative to user access. For a secure transfer of data, each user must be authenticated and authorized, and the following conditions must be met: (1) each new user of the API must first register in the CC-IC system and must be explicitly approved for a given user role; (2) each request must contain the so-called API key (api_key), which is an integral part of the query (query string); (3) other parameters can also be part of the query string parameter, which are listed at specific endpoints; and (4) for each user, there is a defined list of allowed endpoints, IP addresses, and parameter values. If the user's request does not pass through the authorization process, the user will obtain the response "401 Unauthorized."

Data from the CC-IC, along with information from other registries (eg, the Information System of Infectious Diseases), are regularly processed for the purpose of daily reporting, which is provided to the team of the CNCC-AIC. Among dozens of analytical outputs (as of December 23, 2020), one of them monitors—on a daily basis—the situation in the Czech Republic from the points of view of currently hospitalized patients, new admissions to hospitals, new discharges from hospitals, as well as the utilization of health care facilities in individual regions, expressed by their free capacity [15].

Discussion

The Czech Republic was one of the most affected countries during the COVID-19 pandemic waves in autumn 2020 and winter 2021 in terms of incidence, hospitalizations, and deaths. The CC-IC was one of the key tools developed to manage the capacities of standard and ICU beds in the entire country, and to ensure that appropriate care was provided not only to patients with COVID-19 but also to other patients.

The Czech Republic has by far the highest density of ICU beds (43.2 beds per 100,000 population) and one of the longest average hospital stays (9.5 days) among the Organization for Economic Cooperation and Development countries [16]. There is also a high level of centralization of hospital-based care, because all large hospitals are under the control of either the Ministry of Health or regional authorities. These factors enabled a very quick response and management of intensive care, including the compulsory reporting of current capacities into the central system.

The spread of COVID-19 has usually not been uniform, and there have been different ICU demands within the affected countries and their regions. Central reporting systems for the monitoring of ICU capacities were therefore established in some countries [17]. The absence of such a platform, often combined with a fragmented health care system and small ICU facilities, may slow down the response and hold back the provision of appropriate care to critically ill patients, both with and without COVID-19 [18].

The systems for ICU capacities monitoring employed in various countries are usually operated under the auspices of national authorities (ie, government, ministries, and armed forces). They have been designed to provide near real-time data for crisis management and decision-making [19-21]. Short- and long-term estimates of ICU bed occupancy in the upcoming days and weeks using statistical models are another important output [21,22].

The CC-IC platform has been designed and developed in a tailor-made manner, responding to needs of the leadership of the Ministry of Health of the Czech Republic and crisis teams that are actively involved in the management of the COVID-19 epidemic in the Czech Republic. Based on the need to quickly and effectively resolve various crisis situations related to the capacity of ICUs, the CC-IC system was incorporated into the Czech legislation. Experience with the COVID-19 pandemic has clearly shown the requirement for long-term monitoring of ICU bed occupancy. This is the main reason why the obligation to update the occupancy on a daily basis for all health care facilities providing inpatient care in the Czech Republic is now enshrined in the Czech legislation. As part of international cooperation, the possibility of transfer of this online tool to other countries is being discussed. With respect to the general concept of the CC-IC design, the transferability of this tool to other countries should be trouble-free provided that the basic methodology for capacity monitoring in health care facilities is adhered to. It is of key importance that new requirements are collected from health care professionals themselves, even in the future. Intensivists should be part of strategic planning committees before, during, and after pandemics to coordinate ICU responses with hospital and regional efforts for triage, clinical care, and infection control [7]. Regular evaluation includes the analysis of user behavior and further optimization of the whole platform, so that it can continue to function as the primary support tool for all involved stakeholders.

Conclusions

The web platform CC-IC has been developed as a reaction to the urgent need of strategic data-based and organizational support for individual regions and providers of acute care on a nationwide level. The platform is intended for online real-time reporting of changes of free capacity of hospitals, real-time reporting of occupancy rate and availability of beds, a clearly arranged reporting of health care facilities, and the basis for quick decision-making of crisis management during the COVID-19 pandemic. Delegated representatives of health care facilities (coordinators of intensive care) are tasked with reporting the remaining free capacity of machines and beds that are available for patients hospitalized with COVID-19. Information on newly reported current capacity is immediately available to the Integrated Central Management Team, which continuously monitors the situation across the Czech Republic. Online entering and nationwide record-keeping of requests for medications intended for COVID-19 patients are processed by a standalone module. This system can be accessed by users from individual health care facilities providing acute care; the users typically involve several physicians and management representatives from the same health care facility, so that they can stand in for each other and share the latest information.

Development of the CC-IC is far from complete; the system has been further improved, partly based on requirements by members of the team of the CNCC-AIC and partly in response to suggestions by users themselves.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Key information necessary for subsequent approval of a specific medication.
[DOCX File , 15 KB-Multimedia Appendix 1]

References

1. Stawicki S, Jeanmonod R, Miller A, Paladino L, Gaieski D, Yaffee A, et al. The 2019-2020 novel coronavirus (Severe Acute Respiratory Syndrome Coronavirus 2) pandemic: a joint American College of Academic International Medicine-World Academic Council of Emergency Medicine Multidisciplinary COVID-19 Working Group Consensus Paper. J Glob Infect Dis 2020;12(2):47-93 [FREE Full text] [doi: 10.4103/jgid.jgid_86_20] [Medline: 32773996]
2. Boëlle PY, Delory T, Maynadier X, Janssen C, Piarroux R, Pichenot M, et al. Trajectories of hospitalization in COVID-19 patients: an observational study in France. J Clin Med 2020 Sep 29;9(10):3148 [FREE Full text] [doi: 10.3390/jcm9103148] [Medline: 33003375]
3. Gottlieb M, Sansom S, Frankenberger C, Ward E, Hota B. Clinical course and factors associated with hospitalization and critical illness among COVID-19 patients in Chicago, Illinois. Acad Emerg Med 2020 Oct;27(10):963-973. [doi: 10.1111/acem.14104] [Medline: 32762106]
4. Nachtigall I, Lenga P, Jóźwiak K, Thürmann P, Meier-Hellmann A, Kuhlen R, et al. Clinical course and factors associated with outcomes among 1904 patients hospitalized with COVID-19 in Germany: an observational study. Clin Microbiol Infect 2020 Dec;26(12):1663-1669 [FREE Full text] [doi: 10.1016/j.cmi.2020.08.011] [Medline: 32822883]
5. Abate SM, Ahmed Ali S, Mantfardo B, Basu B. Rate of intensive care unit admission and outcomes among patients with coronavirus: a systematic review and meta-analysis. PLoS One 2020 Jul 10;15(7):e0235653 [FREE Full text] [doi: 10.1371/journal.pone.0235653] [Medline: 32649661]
6. Serafim RB, Póvoa P, Souza-Dantas V, Kalil AC, Salluh JI. Clinical course and outcomes of critically ill patients with COVID-19 infection: a systematic review. Clin Microbiol Infect 2021 Jan;27(1):47-54 [FREE Full text] [doi: 10.1016/j.cmi.2020.10.017] [Medline: 33190794]
7. Kain T, Fowler R. Preparing intensive care for the next pandemic influenza. Crit Care 2019 Oct 30;23(1):337 [FREE Full text] [doi: 10.1186/s13054-019-2616-1] [Medline: 31665057]
8. Daugherty Biddison EL, Faden R, Gwon H, Mareiniss DP, Regenberg AC, Schoch-Spana M, et al. Too many patients…a framework to guide statewide allocation of scarce mechanical ventilation during disasters. Chest 2019 Apr;155(4):848-854. [doi: 10.1016/j.chest.2018.09.025] [Medline: 30316913]
9. Dichter J, Kanter R, Dries D, Luyckx V, Lim ML, Wilgis J, Task Force for Mass Critical Care. System-level planning, coordination, and communication: care of the critically ill and injured during pandemics and disasters: CHEST consensus statement. Chest 2014 Oct;146(4 Suppl):e87S-e102S [FREE Full text] [doi: 10.1378/chest.14-0738] [Medline: 25144713]
10. Einav S, Hick J, Hanfling D, Erstad BL, Toner ES, Branson RD, Task Force for Mass Critical Care. Surge capacity logistics: care of the critically ill and injured during pandemics and disasters: CHEST consensus statement. Chest 2014 Oct;146(4 Suppl):e17S-e43S. [doi: 10.1378/chest.14-0734] [Medline: 25144407]
11. Krausz M, Westenberg J, Vigo D, Spence R, Ramsey D. Emergency response to COVID-19 in Canada: platform development and implementation for eHealth in crisis management. JMIR Public Health Surveill 2020 May 15;6(2):e18995 [FREE Full text] [doi: 10.2196/18995] [Medline: 32401218]
12. COVID 19: online intensive care dispatching. Control Centre for Intensive Care. URL: https://dip.mzcr.cz/dip-login [accessed 2022-02-11]
13. Klímek J. DCAT-AP representation of Czech National Open Data Catalog and its impact. J Web Semant 2019 Mar;55:69-85. [doi: 10.1016/j.websem.2018.11.001]
14. National Catalogue of Open Data. URL: https://opendata.mzcr.cz/ [accessed 2022-02-11]
15. Komenda M, Bulhart V, Karolyi M, Jarkovský J, Mužík J, Májek O, et al. Complex reporting of the COVID-19 epidemic in the Czech Republic: use of an interactive web-based app in practice. J Med Internet Res 2020 May 27;22(5):e19367 [FREE Full text] [doi: 10.2196/19367] [Medline: 32412422]
16. Health at a Glance 2021: OECD Indicators. Paris: OECD Publishing; 2021.
17. Winkelmann J, Webb E, Williams G, Hernández-Quevedo C, Maier CB, Panteli D. European countries' responses in ensuring sufficient physical infrastructure and workforce capacity during the first COVID-19 wave. Health Policy 2021 Jul 09:online ahead of print. [doi: 10.1016/j.healthpol.2021.06.015] [Medline: 34311982]

18. Yamamoto T, Ozaki M, Kasugai D, Burnham G. Assessment of critical care surge capacity during the COVID-19 pandemic in Japan. Health Secur 2021;19(5):479-487. [doi: 10.1089/hs.2020.0227] [Medline: 34346775]
19. Pilcher D, Coatsworth N, Rosenow M, McClure J. A national system for monitoring intensive care unit demand and capacity: the Critical Health Resources Information System (CHRIS). Med J Aust 2021 Apr;214(7):297-298 [FREE Full text] [doi: 10.5694/mja2.50988] [Medline: 33774832]
20. Raposo G, Castellanos J, Hernandez H. Near real-time telemetry of nation-wide ICU's capacity strain in Argentina. Stud Health Technol Inform 2021 May 27;281:820-821. [doi: 10.3233/SHTI210291] [Medline: 34042787]
21. Zhao C, Tepekule B, Criscuolo NG, Wendel Garcia PD, Hilty MP, RISC-ICU Consortium Investigators In Switzerland, et al. icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. Swiss Med Wkly 2020 May 04;150:w20277. [doi: 10.4414/smw.2020.20277] [Medline: 32374886]
22. de Salles Neto LL, Martins CB, Chaves A, Roma de Oliveira Konstantyner TC, Yanasse HH, de Campos CBL, et al. Forecast UTI: application for predicting intensive care unit beds in the context of the COVID-19 pandemic. Epidemiol Serv Saude 2020;29(4):e2020391 [FREE Full text] [doi: 10.5123/S1679-49742020000400023] [Medline: 32997068]

Abbreviations

API: application programming interface
CC-IC: Control Centre for Intensive Care
CNCC-AIC: Czech National Control Centre for Acute Inpatient Care
ICU: intensive care unit

Edited by C Basch; submitted 25.08.21; peer-reviewed by P Štourač, M Henry, C Guinemer; comments to author 20.11.21; revised version received 10.12.21; accepted 20.12.21; published 16.02.22

Please cite as:
Komenda M, Černý V, Šnajdárek P, Karolyi M, Hejný M, Panoška P, Jarkovský J, Gregor J, Bulhart V, Šnajdrová L, Májek O, Vymazal T, Blatný J, Dušek L
Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic
J Med Internet Res 2022;24(2):e33149
URL: https://www.jmir.org/2022/2/e33149
doi: 10.2196/33149
PMID: 34995207