STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS 279

Navigating Healthcare Through Challenging Times

Proceedings of dHealth 2021 – Health Informatics Meets Digital Health

Editors: Dieter Hayn
Günter Schreier
Martin Baumgartner

IOS Press

Aside from the dramatic effects that the COVID-19 pandemic has had on the lives of people everywhere, it has also triggered and accelerated some important process changes in healthcare. Digital health has become ever more important, supporting test strategies and contact tracing, statistical analysis, prognostic modeling, and vaccination roll-out and documentation. Video calls have become more common, and it seems likely that all these changes will continue to influence healthcare in the longer-term.

This book presents the proceedings of dHealth 2021 – the 15th annual conference on Health Informatics Meets Digital Health – held as a virtual conference on 11 & 12 May 2021. The dHealth conference is where research and application meet as equals, and the conference series has been contributing to scientific exchange and networking since 2007. The 2021 edition is the second that has been organized virtually. Each year, this event attracts 300+ participants from academia, industry, government and healthcare organizations, and provides a platform for researchers, practitioners, decision makers and vendors to discuss innovative health informatics and dHealth solutions with the aim of improving the quality and efficiency of healthcare.
The 24 papers included here offer an insight into the research on digital health conducted during the COVID-19 crisis, and topics include the management of infectious diseases, telehealth services, standardization and interoperability in healthcare, nursing informatics, data analytics, predictive modeling and digital tools for rare-disease research.
The book provides new healthcare insights from both science and practice, and will be of interest to all those working in healthcare.

ISBN: 978-1-64368-180-1
9 781643 681801

ISBN 978-1-64368-180-1 (print)
ISBN 978-1-64368-181-8 (online)
ISSN 0926-9630 (print)
ISSN 1879-8365 (online)

NAVIGATING HEALTHCARE THROUGH CHALLENGING TIMES

Studies in Health Technology and Informatics

International health informatics is driven by developments in biomedical technologies and medical informatics research that are advancing in parallel and form one integrated world of information and communication media and result in massive amounts of health data. These components include genomics and precision medicine, machine learning, translational informatics, intelligent systems for clinicians and patients, mobile health applications, data-driven telecommunication and rehabilitative technology, sensors, intelligent home technology, EHR and patient-controlled data, and Internet of Things.

Studies in Health Technology and Informatics (HTI) series was started in 1990 in collaboration with EU programmes that preceded the Horizon 2020 to promote biomedical and health informatics research. It has developed into a highly visible global platform for the dissemination of original research in this field, containing more than 250 volumes of high-quality works from all over the world.

The international Editorial Board selects publications with relevance and quality for the field. All contributions to the volumes in the series are peer reviewed.

Volumes in the HTI series are submitted for indexing by MEDLINE/PubMed; Web of Science: Conference Proceedings Citation Index – Science (CPCI-S) and Book Citation Index – Science (BKCI-S); Google Scholar; Scopus; EMCare.

Volume 279

Recently published in this series

Vol. 278 R. Röhrig, T. Beißbarth, W. Brannath, H.-U. Prokosch, I. Schmidtmann, S. Stolpe and A. Zapf (Eds.), German Medical Data Sciences: Bringing Data to Life – Proceedings of the Joint Annual Meeting of the German Association of Medical Informatics, Biometry and Epidemiology (gmds e.V.) and the Central European Network – International Biometric Society (CEN-IBS) 2020 in Berlin, Germany

Vol. 277 A.J. Maeder, C. Higa, M.E.L. van den Berg and C. Gough (Eds.), Telehealth Innovations in Remote Healthcare Services Delivery – Global Telehealth 2020

Vol. 276 M. Merolli, C. Bain and L.K. Schaper (Eds.), Healthier Lives, Digitally Enabled – Selected Papers from the Digital Health Institute Summit 2020

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

Navigating Healthcare Through Challenging Times

Proceedings of dHealth 2021 – Health Informatics Meets Digital Health

Edited by

Dieter Hayn

AIT Austrian Institute of Technology GmbH, Graz, Austria

Günter Schreier

AIT Austrian Institute of Technology GmbH, Graz, Austria

and

Martin Baumgartner

AIT Austrian Institute of Technology GmbH, Graz, Austria

IOS Press

Amsterdam • Berlin • Washington, DC

ISBN 978-1-64368-180-1 (print)
ISBN 978-1-64368-181-8 (online)
Library of Congress Control Number: 2021937489
doi: 10.3233/SHTI279

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

For book sales in the USA and Canada:
IOS Press, Inc.
6751 Tepper Drive
Clifton, VA 20124
USA
Tel.: +1 703 830 6300
Fax: +1 703 830 2300
sales@iospress.com

PRINTED IN THE NETHERLANDS

Preface

For more than a year, the COVID-19 crisis has been affecting healthcare systems worldwide. Aside from its various dramatic effects on citizens and healthcare professionals, COVID-19 has also triggered and accelerated some important process changes. Digital health has played a role of the utmost importance in the last 12 months, supporting contact tracing, test strategies, statistical analyses, prognostic modeling, vaccination organization and documentation, etc. Teleconferences have become common as a way of communicating, not only on a business level, but also in our daily personal lives, even between grandparents and their grandchildren, and it seems likely that these changes will influence healthcare even in the longer-term.

These proceedings provide an insight into research on digital health as conducted during the COVID-19 crisis, including articles concerning the management of infectious diseases, telehealth services, standardization and interoperability in healthcare, nursing informatics, data analytics, predictive modeling, digital tools for rare disease research, and many others.

The dHealth conference series has been contributing to scientific exchange and networking since 2007, and the conference was organized in a virtual setting for a second time in 2021. Each year, this event attracts 300+ participants from academia, industry, government and healthcare organizations. With its motto, "Health Informatics meets Digital Health", the event provides a platform for researchers, practitioners, decision makers and vendors to discuss innovative health informatics and dHealth solutions with the aim of improving the quality and efficiency of healthcare. It is the USP of this event that it is where research and application meet as equals to provide new insights from both the scientific and practical points of view.

Dieter Hayn, Günter Schreier and Martin Baumgartner

Graz, Vienna, April 2021

This page intentionally left blank

Scientific Programme Committee

Univ.-Prof. Dr. Klaus-Peter Adlassnig, *Medical University of Vienna, Austria*
Prof. Dr. Josef Altmann, *UAS Upper Austria, Wels, Austria*
Univ.-Prof. Dr. Elske Ammenwerth, *UMIT, Hall in Tyrol, Austria*
Dr. Sergi Beltran, *Centre Nacional D'Anàlisi Genòmica, Bacelona, Spain*
Univ.-Prof. Dr. Andrea Berghold, *Medical University of Graz, Austria*
Prof. Dr. Oliver Bott, *UAS Hannover, Germany*
Univ.-Prof. Dr. Ruth Breu, *University of Innsbruck, Austria*
Prof. Anthony J. Brookes, *University of Leicester, United Kingdom*
Catherine Chronaki, *HL7 Foundation, Brussels, Belgium*
Ass. Prof. Kim Delbaere, *University of New South Wales, Sydney, Australia*
Univ.-Prof. DDr. Wolfgang Dorda, *Medical University of Vienna, Austria*
Prof. Dr. Stephan Dreiseitl, *UAS Hagenberg, Austria*
Univ.-Prof. Dr. Georg Duftschmid, *Medical University of Vienna, Austria*
Univ.-Prof. Dr. Martin Dugas, *University of Münster, Germany*
Prof. Chris Evelo, *Maastricht University, Netherlands*
Univ.-Prof. Dr. Walter Gall, *Medical University of Vienna, Austria*
Prof. José García Moros, *Universidad de Zaragoza, Spain*
Dr. Holger Gothe, *UMIT, Hall in Tyrol, Austria*
Prof. Dr. Martin Haag, *Heilbronn University, Germany*
Prof. Dr. Peter Haas, *UAS Dortmund, Germany*
Prof. Dr. Anke Häber, *UAS Zwickau, Germany*
Marc Hanauer, *INSERM, US14 – Orphanet, Paris, France*
Dr. Dieter Hayn, *AIT, Graz, Austria*
Dr. Mira Hercigonja-Szekeres, *UAS Zagreb, Croatia*
ao. Univ.-Prof. Dr. Alexander Hörbst, *UMIT, Hall in Tyrol, Austria*
Prof. Dr. Ursula Hübner, *University of Osnabrück, Germany*
Sergey Karas, MD, PhD, *Siberian State Medical University Tomsk, Russia*
Prof. Dr. Guido Kempter, *UAS Vorarlberg, Austria*
Prof. Dr. Peter Klutke, *University of Kempten, Germany*
Prof. Dr. Petra Knaup, *University of Heidelberg, Germany*
Dr. Tino Kulnik, *Ludwig Boltzmann Institute for Digital Health, Salzburg, Austria*
Prof. Dr. Werner Kurschl, *UAS Hagenberg, Austria*
Univ.-Prof. Dr. Richard Lenz, *University of Erlangen-Nürnberg, Germany*
Prof. Dr. Nicos Maglaveras, *Aristotle University, Thessaloniki, Greece*
Prof. Dr. Michael Marschollek, *Medical University of Hannover, Germany*
Prof. Dr. Christian Menard, *UAS Carinthia, Austria*
Prof. Dr. George Mihalas, *University of Medicine and Pharmacy, Timisoara, Romania*
Prof. Josef Niebauer, *Ludwig Boltzmann Institute for Digital Health, Salzburg, Austria*
Ass.-Prof. Dr. Ivana Ognjanovic, *Donja Gorica University, Podgorica, Montenegro*
Univ.-Prof. Dr. Karl-Peter Pfeiffer, *FH Joanneum, Graz, Austria*
Univ.-Prof. Dr. Ulrich Prokosch, *University of Erlangen, Germany*
Dr. Ana Rath, *INSERM, US14 – Orphanet, Paris, France*
Ass.-Prof. Stephen Redmond, *University College Dublin, Ireland*

Reviewers

We are deeply thankful for all reviewers for their significant contribution to the conference proceedings. Without them, the dHealth 2021 would not be possible:

Elske Ammenwerth
Martin Baumgartner
Andrea Berghold
Stephan Dreiseitl
Georg Duftschmid
Alphons Eggerth
Walter Gall
José García Moros
Holger Gothe
Martin Haag
Anke Häber
Dieter Hayn
Emmanuel Helm
Peter Klutke
Alexander Kollmann
Karl Kreiner
Stefan Tino Kulnik
Werner Kurschl
George Mihalas
Bernhard Pfeifer
Günter Schreier
Abbas Sheikhtaheri
Michael Shifrin
Martin Staemmler
Lacramioara Stoicu-Tivadar
Hanna Suominen
Sai Pavan Kumar Veeranki
Elena Villalba-Mora
Andreas Ziegl

This page intentionally left blank

Contents

This page intentionally left blank

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210081

eHealth Interventions for Dementia – Using WordPress Plugins as a Flexible Dissemination for Dementia Service Providers

Michael REICHOLD[a,1] Marina SELAU[a], Elmar GRAESSEL[c], Peter L. KOLOMINSKY-RABAS[b] and Hans-Ulrich PROKOSCH[a]
[a] *Chair of Medical Informatics, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany*
[b] *Interdisciplinary Center for Health Technology Assessment (HTA) and Public Health (IZPH), Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany*
[c] *Center for Health Services Research in Medicine, Department of Psychiatry and Psychotherapy, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany*

Abstract. The benefits of eHealth interventions for people with dementia and their informal caregivers have been demonstrated in several studies. In times of contact restrictions, digital solutions have become increasingly important, especially for people with dementia and their mostly elderly caregiving relatives, which are at increased risk of severe illness from COVID-19. As in many other health areas, there is a lack of digital interventions in the dementia landscape that are successfully implemented (i.e., put into practice), especially digital interventions that are scientifically evaluated. Evaluated and proven effective digital interventions exist, but these often do not find their way from research into practice and stay on low-level implementation readiness. Within the project digiDEM Bayern, a digital platform with digital services and interventions for people affected by dementia (people with dementia, caregivers, volunteers and interested citizens) is established. As one digital intervention for informal caregivers, the 'Angehörigenampel' (caregivers' traffic-light) was developed, which is able to assess the physical and psychological burden of caregivers. This can help to counteract the health effects of caregiving burden early on before it is too late. The development of the digital intervention as a WordPress-plugin was kept generic so that it can easily be adapted to other languages on further websites. The 'intervention as a plugin' approach demonstrates an easy and flexible way of deploying eHealth interventions to other service providers, especially from other countries. The implementation barriers for other service providers are low enough for them to be able to easily integrate the eHealth intervention on their website, enabling more caregivers to benefit from the disseminated eHealth intervention.

[1] Corresponding Author, Michael Reichold, Chair of Medical Informatics Friedrich-Alexander-University Erlangen-Nürnberg, Germany; E-mail: michael.reichold@fau.de.

1. Introduction

COVID-19 has clearly demonstrated the importance of digitalization in all areas of society. Moreover, digital deficiencies have become even more visible in the past months. Particularly for people with dementia and their informal caregivers (usually family members or close relatives who provide unpaid care), access to eHealth interventions is becoming more important than ever [1], especially in view of the fact that the risk factors in developing dementia [2] such as older age, cardiovascular disease, diabetes, chronic respiratory disease, and hypertension are the same that also cause a severe or fatal COVID-19 disease [3].

With 50 million people affected worldwide, dementia is one of the major challenges the healthcare systems are facing. In Germany, 1.6 million people suffer from dementia, about 240,000 of which are in Bavaria (a federal state in Germany) [4]. In 2019, the research project 'Digital Dementia Registry Bavaria - digiDEM Bayern' was launched to improve the situation of people with dementia and their caregiving relatives in Bavaria. digiDEM Bayern is based on the results of the preceding project 'Bavarian Dementia Registry – BayDem' [5–7]. The main objectives of digiDEM Bayern [8] are: First, establishing a digital patient register for the acquisition of long-term data of people with dementia and their caregivers. Second, providing a digital platform with digital interventions and services for people with dementia and cognitive impairments as well as for caregivers, volunteers and interested citizens.

The effectiveness of digital interventions for people with dementia and their family caregivers has been demonstrated in several studies [9–11]. Unfortunately, these interventions are often only available while the research study is funded or have other restrictions such as a foreign language. As a result, these interventions are only accessible to a limited group of people with dementia or caregivers. In 2019, Christie et al. showed that only 2 of 12 interventions initially published in scientific journals are still available to informal caregivers and freely accessible on dementia websites [12]. Among other things, these limitations lead to the fact that the availability of digital interventions in practice is assessed as relatively poor [13].

As one digital intervention, a so-called 'Angehörigenampel' (caregivers' traffic light) was developed for the digiDEM Bayern platform. With the help of the 'Angehörigenampel', informal dementia caregivers can assess their caregiver burden on health, namely on the physical and psychological (emotional) health. Particularly in the early stages of dementia, care for people with dementia is mainly provided by family members serving as informal caregivers [14]. Informal caregiving is associated with poorer physical and mental health outcomes [15], for example, a greater risk of becoming depressed or an increased mortality rate. For these reasons, the informal caregivers are often referred to as the 'hidden secondary patients' [16], who need and deserve protection and guidance [17]. Here, eHealth interventions can improve the situation for informal caregivers, which in turn can also lead to potential benefits for care recipients, i.e., people with dementia [18].

The objective of the study is to demonstrate a feasible method for the dissemination of eHealth interventions using the newly developed intervention 'Angehörigenampel' as an example. By lowering the barriers to a successful implementation, affected people should be able to benefit regardless of research-project and language boundaries.

2. Methods

The scientific basis of the newly developed eHealth intervention 'Angehörigenampel' is a questionnaire containing 10 questions to which answers can be given with 4 predefined answer options. A total score is then calculated from the answers given, which can vary between 0 and 30 points. This self-assessment provides a meaningful assessment of the burden experienced by caregiving relatives in the home care situation and allows classification into a low, moderate and high burden. Specific recommendations for further steps are assigned to these levels of severity. The questionnaire – a short version of the Burden Scale for Family Caregivers (BSFC) [19] - is available in more than 20 languages and has been used across the world for many years. The short version's reliability and validity were evaluated in a scientific study, which measured the burden experienced by 351 informal caregivers in home care [20]. The classification system according to 'traffic light colors' is the second principle that is scientifically verified. Pendergrass et al. reviewed the classification of how the degree of burden relates to the risk of physical psychosomatic complaints [21]. Based on the preliminary scientific work, a specification for the eHealth intervention with functional and non-functional requirements was created by a multi-professional team of digiDEM Bayern.

An easy way to bring digital offerings to those affected is by using a website. From a technical point of view, the basis of the digital platform in digiDEM Bayern is the Content Management System (CMS) WordPress (version 5.6). WordPress supports building a multimedia enriched platform by expanding its basic functions with additional plugins. Plugins are a powerful way to add additional features and services to WordPress using PHP-based scripts. In principle, there are almost no limits to what plugins can do in WordPress [22]. Therefore, the digital intervention 'Angehörigenampel' was developed as a WordPress plugin to be easily integrated into our digiDEM Bayern platform. For this purpose, the scripting language PHP was used in compliance with the WordPress PHP Coding Standards [23]. To allow an easy adaptation to local requirements of other dementia service providers, we kept the plugin as generic as possible. Previous to the publication of the intervention on the website, the technical implementation and functionality of the developed plugin were tested by 5 project members.

3. Results

Based on the above-mentioned scientific evidence of Gräßel and Pendergrass, a WordPress plugin was developed in PHP, which allows informal caregivers an assessment of their caregiver burden by answering 10 questions of the short version of the BSFC questionnaire. Depending on the result, the caregivers receive further recommendations to strengthen their physical and mental health.

The developed plugin is available as a zip file and can be easily installed on any WordPress website using the WordPress built-in plugin manager. To ensure an easy and flexible integration of the plugin on any WordPress website page at any place, a shortcode support has been added for the plugin. Shortcodes are commands enclosed by square brackets, which are converted by WordPress into elements of a website. This allows calling a PHP function by adding '[digiDEM_Angehoerigenampel]' in the WordPress page editor. As illustrated in Figure 1, additional regular WordPress content

such as additional details about the intervention or the scientific evidence can be added before and after the embedded plugin-area [24].

The following four functions were implemented in the plugin 'Angehörigenampel': First, having completed the assessment in the selected language, the caregiver receives a rating of their burden and a corresponding visual indication symbolized by a traffic light (low burden = green light, moderate burden = yellow light, high burden = red light) as well as short advice. To monitor and evaluate the usage of the intervention, the results of the assessment are anonymously stored in a separate table.

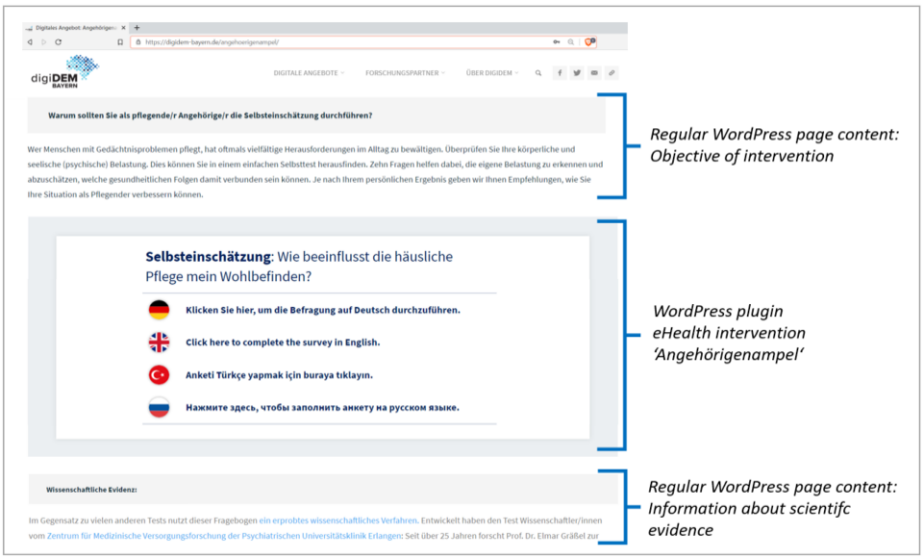


Figure 1. Integration of the plugin 'Angehörigenampel' on the digiDEM Bayern website [24]

Second, the caregiver can download a document corresponding to the identified burden for getting a more detailed assessment and recommended measures for reducing the burden. The recommendations were developed in a panel of experts together with family care advisors. They were then checked for comprehensibility by a group of informal caregivers. Caregivers can take this document with them to their physician or other consulting facilities to take further steps, if necessary. Technically, the document is a structured PDF, generated from a structured HTML and CSS layout file, where individual text areas are combined depending on the individual assessment result. The text modules for the PDF can also be changed directly in WordPress via the plugin's settings.

Third, an optional reminder function has been included after the assessment, which automatically reminds the caregiver to repeat the assessment by email after 6 months. It is recommended to perform the assessment regularly to see how the caregiver burden is developing over time. The period of 6 months was chosen because the progression of dementia caregivers' burden is relatively sluggish [25]. If the caregiver uses the reminder function and enters their email, it will be encrypted (encryption standard: AES-256-CBC) together with the date of the assessment in a separate table of the MySQL WordPress database. By using a cron-job, it is checked daily in this table if 6 months have passed since the assessment. If so, the table entry will be decrypted, and a reminder

email is sent to the caregiver using the wp_mail function provided by WordPress and the table entry is deleted.

Fourth, a feedback section has been included after the assessment to determine how the caregivers accept the eHealth intervention. Here, the caregiver can answer five questions anonymously and on a voluntary basis. Among other things, they can indicate whether the self-assessment accurately reflects their perceived burden of caregiving and whether they would recommend the assessment to other caregivers by means of the Net Promoter Score [26]. The feedback data is stored in an internal MySQL table, and a CSV file can be downloaded for evaluation purposes in the plugin settings. The feedback option of the plugin can be enabled or disabled in the WordPress plugin settings as needed.

To pay attention to the solution's flexibility during the development, the intervention was published initially in four languages (German, English, Turkish, Russian). Due to the multi-language approach, it was possible to immediately test and ensure that an adjustment of the wording or language is possible without programming knowledge for each function. As illustrated in Figure 2, an adaptation to other languages is easily possible via the plugin settings. For this purpose, only the corresponding text modules have to be translated and replaced. If one wants to provide the WordPress plugin to another dementia service provider, the provider only has to upload the plugin to their WordPress installation, activate it and embed it on their website. In the settings of the plugin, they can select whether they want to activate the feedback function. They can also specify which page the reminder email should link to and where the privacy page is, which is referred to in the reminder function.

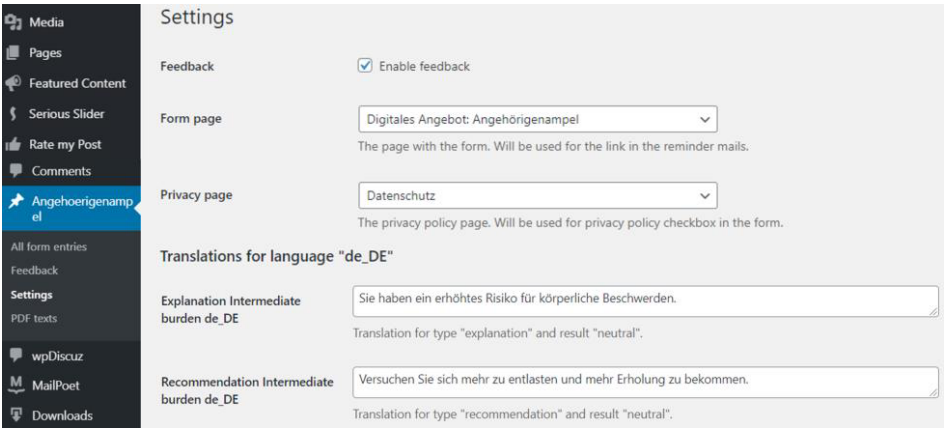


Figure 2. Settings of the WordPress plugin 'Angehörigenampel' of digiDEM Bayern

By deciding to develop the eHealth intervention as a plugin, the following dissemination advantages could be achieved. The website can easily be extended with additional content and functions without having to modify the core of the WordPress system. It also makes it much easier to share the plugin and updates, because one does not have to replace individual pieces of code, which can be very error-prone.

4. Discussion

The objective of this paper was to demonstrate a practical-based example of how an eHealth intervention, which was developed in the context of a dementia research project, can be made available to a broader target group in a simple way.

Talking about digitalization and age, stereotypes always arise about whether the digital way is still the right way to go and whether there are still digital barriers. In contrast to the young generation, older adults are said not to be ready to use the Internet not only technically but also mentally. The latest research on internet use in relation to age shows that older adults are becoming increasingly digital [27]. Corresponding initiatives to overwhelm barriers and further increase eHealth literacy and digital participation are being launched by the policy [28,29]. Particularly for informal caregivers who cannot leave home due to caregiving responsibilities, internet-based interventions provide a cost-effective and flexible way to access interventions from home. Digitalization certainly cannot replace personal contact in every service situation. However, in many cases, digitalization can overcome the limitations of non-digital services such as time, distance and human resources.

Despite evidence of the effectiveness and efficacy of eHealth interventions in numerous studies, there is still a mismatch between the availability of eHealth interventions in research on the one hand and implementation in practice in dementia settings [9]. Gitlin et al. have revealed that for less than 3% of proven interventions for informal caregivers of people with dementia, the translation from research studies into practice is successful [30]. Although studies demonstrate the effectiveness and benefits of digital interventions, the interventions often remain at a low implementation readiness level [31]. Even though there are already considerations on how to achieve sustainable implementations, these are elaborated only on theoretical examples [32]. The need for digital interventions in practice was reaffirmed in a survey from 2019 among service providers in the dementia care landscape, which found a gap between digital and non-digital dementia services and interventions [13].

Therefore, from the beginning of the development, the focus was on factors enabling a sustainable implementation of the intervention and an adaptation to regional conditions as flexible as possible. To minimize the barriers to implementation, it was ensured that the solution did not demand any major technical requirements in terms of software and hardware and that country-specific adaptations, such as the language, were possible without any programming knowledge. The example of a WordPress plugin was used to show the possibility of developing a generic plugin with a high level of implementation readiness. Country-specific adaptations such as the texts' language can be easily realized via the settings menu of WordPress.

WordPress has by far the largest market share among the top 10 content management systems worldwide, with around 63.9%. The CMS in second place only had a market share of 5.1% [33]. The use of WordPress plugins makes it possible to reuse content and solutions easily. The widespread use of WordPress is reflected, for example, in the number of page views of WordPress articles, which in November 2020 was about 25.96 billion [33]. Due to its high penetration, ease of use, low implementation costs, low system requirements, a WordPress-based repository of scientifically validated interventions-plugins in the dementia environment could contribute to the digitalization of other healthcare organizations with similar eHealth needs for their customers.

As a limitation of the solution as WordPress plugin data privacy should be considered and corresponding protective measures such as anonymization must be taken into account during development. WordPress is the most frequently used CMS in the world, which makes it a valuable target for security vulnerabilities and attacks. Although there are security components and best practices provided by WordPress [34], we store sensitive information only in encrypted or anonymized form in our WordPress environment. A further limitation can be seen in the fact that no data can yet be provided to prove dissemination. However, there are already initiated discussions with Alzheimer's associations in two countries to implement the plugin outside of digiDEM Bayern. Furthermore, there is a request from a palliative care service provider. Due to its easy adaptability, the 'Angehörigenampel' will be made available in an adapted form for informal caregivers of palliative care patients in the next step.

5. Conclusion

Digital interventions are a valuable resource to support people with dementia and their informal caregivers. A CMS such as WordPress with its plugins offers a suitable solution in which a sustainable eHealth intervention can be developed rapidly and both disseminated and implemented quickly. The fact that from the beginning of the development the focus was placed on disseminating the eHealth intervention to other providers - together with the fact that it is a customizable WordPress plugin - a very easily adjustable solution was created. This simplifies the installation of eHealth interventions for all interested parties without being restricted by any borders, whether research-project or geographical borders, because dementia does not respect such borders.

Acknowledgments

The project digiDEM Bayern is funded by the Bavarian State Government as part of the funding initiative 'BAYERN DIGITAL II' (funding code: G42d-G8300-2017/1606-83).

The present work was performed by Michael Reichold in (partial) fulfillment of the requirements for obtaining the degree "Dr. rer. biol. hum." from Faculty of Medicine of the Friedrich-Alexander-University Erlangen-Nürnberg.

Disclosure Statement

The authors have no conflict of interest to declare.

References

[1] H. Wang, T. Li, P. Barbarino, S. Gauthier, H. Brodaty, J.L. Molinuevo, H. Xie, Y. Sun, E. Yu, Y. Tang, W. Weidner, and X. Yu, Dementia care during COVID-19, *Lancet*. **395** (2020) 1190–1191. doi:10.1016/S0140-6736(20)30755-8.

[2] S. Norton, F.E. Matthews, D.E. Barnes, K. Yaffe, and C. Brayne, Potential for primary prevention of Alzheimer's disease: an analysis of population-based data, *Lancet Neurol*. **13** (2014) 788–794. doi:10.1016/S1474-4422(14)70136-X.

[3] Z. Wu, and J.M. McGoogan, Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention, *JAMA*. **323** (2020) 1239–1242. doi:10.1001/jama.2020.2648.

[4] Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit, Gesundheitsreport Bayern: 2/2019 – Update Demenzerkrankungen, (2019). https://www.lgl.bayern.de/publikationen/doc/gesundheitsreport_2_2019.pdf (accessed January 8, 2020).

[5] L. Karrer, N. Dietzel, F. Wolff, A. Kratzer, M. Hess, E. Gräßel, and P. Kolominsky-Rabas, [Use of Outpatient Care Services by People with Dementia: Results of the Bavarian Dementia Survey (BayDem)], *Gesundheitswesen*. **82** (2020) 40–49. doi:10.1055/a-1071-7851.

[6] F. Wolff, N. Dietzel, L. Karrer, A. Kratzer, M. Hess, E. Gräßel, and P. Kolominsky-Rabas, Zeitgerechte Diagnosestellung bei Menschen mit Demenz: der Bayerische Demenz Survey (BayDem), *Gesundheitswesen*. **82** (2020) 23–29. doi:10.1055/a-1031-9559.

[7] N. Dietzel, L. Karrer, F. Wolff, A. Kratzer, M. Hess, E. Gräßel, and P. Kolominsky-Rabas, [Predictors of Caregiver Burden in Dementia: Results of the Bavarian Dementia Survey (BayDem)], *Gesundheitswesen*. **82** (2020) 30–39. doi:10.1055/a-1071-7886.

[8] N. Dietzel, L. Kürten, M. Reichold, L. Köhler, C. Chmelirsch, K. Seebahn, M. Hladik, S. Meuer, A. Kirchner, K. Holm, M. Selau, M. Wendel, J. Trinkwalter, H.-U. Prokosch, E. Gräßel, P.L. Kolominsky-Rabas, and (in press), The Digital Dementia Registry Bavaria – digiDEM Bayern: study protocol for a multicentre, prospective, longitudinal register study, *BMJ Open*. (2021).

[9] H.L. Christie, S.L. Bartels, L.M.M. Boots, H.J. Tange, F.R.J. Verhey, and M.E. de Vugt, A systematic review on the implementation of eHealth interventions for informal caregivers of people with dementia, *Internet Interventions*. **13** (2018) 51–59. doi:10.1016/j.invent.2018.07.002.

[10] E. Parra-Vidales, F. Soto-Pérez, M.V. Perea-Bartolomé, M.A. Franco-Martín, and J.L. Muñoz-Sánchez, Online interventions for caregivers of people with dementia: a systematic review, *Actas Esp Psiquiatr*. **45** (2017) 116–126.

[11] J.L. Scott, S. Dawkins, M.G. Quinn, K. Sanderson, K.-E.J. Elliott, C. Stirling, B. Schüz, and A. Robinson, Caring for the carer: a systematic review of pure technology-based cognitive behavioral therapy (TB-CBT) interventions for dementia carers, *Aging Ment Health*. **20** (2016) 793–803. doi:10.1080/13607863.2015.1040724.

[12] H.L. Christie, J.L. Martin, J. Connor, H.J. Tange, F.R.J. Verhey, M.E. de Vugt, and M. Orrell, eHealth interventions to support caregivers of people with dementia may be proven effective, but are they implementation-ready?, *Internet Interventions*. **18** (2019) 100260. doi:10.1016/j.invent.2019.100260.

[13] M. Reichold, N. Dietzel, L. Karrer, E. Graessel, P.L. Kolominsky-Rabas, and H.-U. Prokosch, Stakeholder Perspectives on the Key Components of a Digital Service Platform Supporting Dementia - digiDEM Bayern, *Stud Health Technol Inform*. **271** (2020) 224–231. doi:10.3233/SHTI200100.

[14] Statistisches Bundesamt, Pflege im Rahmen der Pflegeversicherung - Deutschlandergebnisse, (2017). https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/Publikationen/Downloads-Pflege/pflege-deutschlandergebnisse-5224001179004.pdf?__blob=publicationFile (accessed January 8, 2020).

[15] J. Bom, P. Bakx, F. Schut, and E. van Doorslaer, The Impact of Informal Caregiving for Older Adults on the Health of Various Types of Caregivers: A Systematic Review, *The Gerontologist*. (2018). doi:10.1093/geront/gny137.

[16] H. Brodaty, and M. Donkin, Family caregivers of people with dementia, *Dialogues Clin Neurosci*. **11** (2009) 217–228.

[17] S.C. Reinhard, B. Given, N.H. Petlick, and A. Bemis, Supporting Family Caregivers in Providing Care, in: R.G. Hughes (Ed.), Patient Safety and Quality: An Evidence-Based Handbook for Nurses, Agency for Healthcare Research and Quality (US), Rockville (MD), 2008. http://www.ncbi.nlm.nih.gov/books/NBK2665/ (accessed December 19, 2020).

[18] M. Leng, Y. Zhao, H. Xiao, C. Li, and Z. Wang, Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Systematic Review and Meta-Analysis, *Journal of Medical Internet Research*. **22** (2020) e19468. doi:10.2196/19468.

[19] E. Graessel, R. Chiu, and R. Oliver, Development and validation of the Burden Scale for Family Caregivers (BSFC), *Comprehensive Rehabilitation and Mental Health Services*. (2003) Toronto (Canada).

[20] E. Graessel, H. Berth, T. Lichte, and H. Grau, Subjective caregiver burden: validity of the 10-item short version of the Burden Scale for Family Caregivers BSFC-s, *BMC Geriatr*. **14** (2014) 23. doi:10.1186/1471-2318-14-23.

[21] A. Pendergrass, C. Malnis, U. Graf, S. Engel, and E. Graessel, Screening for caregivers at risk: Extended validation of the short version of the Burden Scale for Family Caregivers (BSFC-s) with a valid classification system for caregivers caring for an older person at home, *BMC Health Serv Res*. **18** (2018) 229. doi:10.1186/s12913-018-3047-4.

[22] B. Williams, J. Tadlock, and J.J. Jacoby, Professional WordPress plugin development, 2020.

[23] PHP Coding Standards, *Make WordPress Core*. (2013). https://make.wordpress.org/core/handbook/best-practices/coding-standards/php/ (accessed December 16, 2020).

[24] Digitales Angebot: Angehörigenampel, *digiDEM Bayern*. (n.d.). https://digidem-bayern.de/angehoerigenampel/ (accessed December 20, 2020).

[25] M.H. Connors, K. Seeher, A. Teixeira-Pinto, M. Woodward, D. Ames, and H. Brodaty, Dementia and caregiver burden: A three-year longitudinal study, *Int J Geriatr Psychiatry*. **35** (2020) 250–258. doi:10.1002/gps.5244.

[26] D.F. Hamilton, J.V. Lane, P. Gaston, J.T. Patton, D.J. MacDonald, A.H.R.W. Simpson, and C.R. Howie, Assessing treatment outcomes using a single question: the Net Promoter Score, *The Bone & Joint Journal*. **96-B** (2014) 622–628. doi:10.1302/0301-620X.96B5.32434.

[27] O. Huxhold, and K. Otte, Internetzugang und Internetnutzung in der zweiten Lebenshälfte. (dza-aktuell: Deutscher Alterssurvey, 01/2019), Deutsches Zentrum für Altersfragen, Berlin, 2019.

[28] Bundesministerium für Familie, Senioren, Frauen und Jugend Referat Öffentlichkeitsarbeit, Achter Altersbericht - Ältere Menschen und Digitalisierung, (2020). https://www.achter-altersbericht.de/fileadmin/altersbericht/pdf/aktive_PDF_Altersbericht_DT-Drucksache.pdf.

[29] A. Efthymiou, I. Menikou, A. Stylianides, N. Geronikola, I. Ilieva, I. Kancheva, M. Tagliani, R. Seneca, B. Mintus, P. Sakka, and E. Papastavrou, eHealth Literacy Training Among Carers of Older People and People with Dementia: A Modified Delphi Survey, *Stud Health Technol Inform*. **272** (2020) 119–122. doi:10.3233/SHTI200508.

[30] L.N. Gitlin, K. Marx, I.H. Stanley, and N. Hodgson, Translating Evidence-Based Dementia Caregiving Interventions into Practice: State-of-the-Science and Next Steps, *The Gerontologist*. **55** (2015) 210–226. doi:10.1093/geront/gnu123.

[31] J.V. Lacey, and K.E. Savage, 50 % Response rates: half-empty, or half-full?, *Cancer Causes Control*. **27** (2016) 805–808. doi:10.1007/s10552-016-0748-z.

[32] H.L. Christie, L.M.M. Boots, K. Peetoom, H.J. Tange, F.R.J. Verhey, and M.E. de Vugt, Developing a Plan for the Sustainable Implementation of an Electronic Health Intervention (Partner in Balance) to Support Caregivers of People With Dementia: Case Study, *JMIR Aging*. **3** (2020) e18624. doi:10.2196/18624.

[33] Content-Management-Systeme (CMS) - Marktanteile 2020, *Statista*. (n.d.). https://de.statista.com/statistik/daten/studie/320670/umfrage/marktanteile-der-content-management-systeme-cms-weltweit/ (accessed December 15, 2020).

[34] WordPress is Secure, *WordPress.Org*. (2018). https://wordpress.org/about/security/ (accessed December 17, 2020).

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210082

Operations Management in Ambulatory Care in Switzerland

Vinzenz UHR[1], Marc T. GAUCH[1], and Kerstin DENECKE[2]
Bern University of Applied Sciences, Bern, Switzerland

Abstract. Practice efficiency is influenced by its operations management. We aim at studying implementation of operations management in Swiss medical practices and we develop a dashboard that allows controlling and managing resources. To study operations management and relevant indicators in ambulant care, we distributed questionnaires by e-mail and conducted 6 interviews. In collaboration with a group practice, we collected requirements regarding a dashboard for operations management, developed a mockup and finally a prototype. This prototype was deployed and implemented in daily routine. From the assessments we learned that practice information systems (PIS) are not sufficiently supporting production planning and control. Relevant indicators include processing time per patient or waiting time for quantifying efficiency and identify potential improvements in production. Within 5 weeks of implementation of our dashboard in a group practice, we learned that calculating indicators and support of operations management by means of a dashboard is well appreciated by practice employees. Indicators are considered extremely useful for operations management.

Keywords. Efficiency, general practice, organization and administration, outpatients

1. Introduction

Since 2000, health care costs in Switzerland are rising continuously in average by 3.5% per year. Their share of the gross domestic product (GDP) also increased by 2%. In 2018, a new peak was reached with CHF 80'242 million health care costs per year (11.2 % of the GDP) [1]. Furthermore in ambulant care, inefficiency is prevalent in nonclinical processes involved in patient care, from scheduling to test result reporting to prescription refills. In this context, inefficiency can be defined as using more inputs (or resources) than is necessary to deliver patient care or a health care service [2] and it is connected to unnecessary variation in operational and clinical processes [3].

Registered doctors in ambulant care have a dual role: as physicians they are concerned about the health of their patients. As entrepreneurs, the economic health of their practice in terms of revenue growth is essential. Improving the efficiency of a medical practice not only has economic advantages. For example, optimising processes leads to more satisfied employees, reduction of stress and burnout [4]. Satisfied employees are friendlier, more pro-active and make less mistakes. This in turn has an impact on patients, who in turn are more satisfied with the medical attendance [5].

---

[1] Contributed equally

[2] Corresponding author: Kerstin Denecke, Bern University of Applied Sciences, Biel / Bienne, Switzerland; E-mail: kerstin.denecke@bfh.ch

In principal, practice efficiency is influenced by operations management. Operations management is an area of management concerned with designing and controlling the process of production. Production as an operational function has the task of planning and controlling the healthcare service provision process [6]. In a medical practice, production consists of the provision of healthcare services. While production planning is responsible for scheduling patient appointments and the associated planning of resources (staff, equipment, rooms and materials), production control focuses on the smooth execution of processes. This includes room allocation, equipment use and the allocation of orders to an medical practice assistant (MPA). Production management is carried out dynamically by the MPA. In principal, medical practices are supported in operations management by their practice software, the practice information system (PIS). Such systems for storing and managing patient data are in use in 77% of Swiss medical practices [7].

Process improvements related to inefficiency for reducing health care costs have been considered mainly for inpatient processes [8,9]. Limited work considered operations management and the resources dedicated to primary care [4]. James et al. identified several sources for inefficiencies in medical practices [4]. They include activities related to making appointments and scheduling during pre-visit of a patient, practice layout, communication staffing during the visit and medication refills, managing examination results during post-visit.

In this work, the operations management in Swiss medical practices is investigated, specifically focussing on the activities and processes during the visit of patients. We want to find out to what extent the operations management is supported by available PIS. To assist operations management, a dashboard is developed, deployed and implemented in a medical practice. A quantitative study is conducted to determine whether an increase in efficiency can be achieved using the application.

2. Methods

2.1. Analysis of current situation of operations management in Swiss medical practices

To collect information on processes and problems in operations management in Swiss medical practices, we sent e-mails to chains of medical practices and ambulatory clinics. In these e-mails, we asked for problems in operations management that occur in daily routine. Further, we asked for a telephone interview. The interviews were based on a questionnaire that was extended when new questions popped up in an interview. Questions concerned problems in the internal coordination of personnel and structural resources, problems in the process control and important indicators for process control. To identify relevant medical practices, we searched for chains of clinical practices in Switzerland using Google (search terms: “Gruppenpraxis Schweiz”, “Gruppenpraxen Ketten schweiz”, “Grösste Arztpraxisketten Schweiz”). A list with the largest chains could be identified [10]. From this list, we contacted 12 practices located in Switzerland. Besides these chains, we contacted 9 ambulatory care clinics and polyclinics. In total, 21 practices were contacted; 13 (62%) did not replied. Three practices described their processes by e-mail and with six practices, we conducted interviews with either MPA, director or physicians. Additionally, 20 German-speaking provider of PIS were contacted by e-mail and were asked how their PIS supports the process control. The provider were identified using the FMH Swiss Medical Association Softwarekatalog [11]. 8 of 20 provider replied by e-mail.

2.2. Development and implementation of a dashboard for process control and planning

The development of the dashboard for process control and management was realized in an iterative process following principles of human-centred design (DIN EN ISO 9241-210). This means future users were involved in the whole process and contributed actively with information and feedback. In a first phase, requirements were collected in interviews with employees of a concrete group practice; existing process management and control procedures and their limitations were analyzed. Mockups were created and a scenario-based usability test was conducted with the mockup and MPAs from the group practice. We adapted the design and functionalities of the dashboard according to the feedback and implemented a prototype.

We conducted a deployment phase of 2 weeks followed by an implementation phase of 5 weeks. The dashboard was deployed in a group practice (8 physicians and 10 other healthcare staff) between 02.11.2020 and 15.11.2020. One MPA in this practice is responsible for production control. She allocates patients to rooms in time and ensures that relevant documents for the patient-doctor encounter are available. She ensures a smooth routine of processes which requires concentration. If successful, at the end of the day all patients have been treated within the consultation hours. If not successful, the leading MPA has to support the team in treating patients. During that time, she cannot complete her actual tasks which leads to overtime hours. In the deployment phase, staff was trained using a user manual. All MPA were able to use the tool after these two weeks. During five following weeks (implementation phase), the application was used in routine and we collected data on the efficiency of processes. As indicator to assess efficiency, we compared the number of overtime hours of the leading MPA before and after implementation of the tool. For this purpose, we accessed the overtime hours during the period 01.01.2020 and 18.12.2020. We removed days, where the MPA worked less than 7 hours or took time off in lieu.

3. Results

3.1. Production control and management in Swiss medical practices

As means to support operations management, we identified: agenda, careful scheduling, and consultation templates. In the following, more details on these means are provided together with challenges.

Four (n=8) medical practices consider careful scheduling and an agenda as the most important factor for successful operations management. The agenda should contain the patient appointments with doctors and the reservation of required resources. Making appointments with patients is carried out by telephone. In case the scheduling for a specific appointment is inappropriate, time for consultations might be too short or too long. Thus, the required resources might be unavailable when needed resulting in waiting times for patients or additional appointments. In five group practices, the telephone service is realized by the MPA and supported with templates of appointments. These templates are derived from process control data: For the most frequent appointments, past consultations are used to record the appointment duration, required resources, or material (e.g. duration for a normal examination 15’). All interview partners confirmed that preparing consultation templates is more trivial for a practice with specialists than a practice in family medicine. Specialists mainly treat patients with a concrete admission diagnosis

which allows for more precise planning of equipment and staff. A family doctor is confronted with more heterogeneous and diffuse complaints and the list of diagnoses can often only be narrowed down during the consultation. For more precise scheduling, one of the outpatient clinics employs nursing staff on the telephone service. The nursing staff, who have more advanced training than the MPA, can perform already some kind of triage and can allocate resources to the consultations more precisely.

Production control is responsible for carrying out the planning. In six practices, this task is distributed among all MPA, and in one group practice it is taken over by a single MPA. One problem with a group practice is the overcrowded waiting room. In stressful situations, the overview is lost, patients are forgotten and the routine is disrupted. Five medical practices consider standardisation of processes as the best means to increase efficiency. According to them, defined roles, task and priority lists for the MPAs would enable a smooth process.

Table 1. Indicators on operations management and use for process control and management.

Indicator	Description
Average processing time per patient	Average duration of a stay of a patient in the practice from entering the practice until leaving the practice
Average processing time of a patient per reason of visit	Average processing time per reason of visit (enables more precise process planning and scheduling)
Average processing time of a specific patient	Average processing time of a specific patient independent of reason per visit.
Processing time of a patient per specific employee	Average processing time per employee aims to identify bottlenecks and processing delays
Waiting time	Average time spent in the practice starting with the scheduled appointment time in the waiting room. Aims at identifying potentials for optimiziation.
Delay of patient	Average delay of a patient shows potentials for optimization.
No show without giving notice	Number and reasons for no shows to identify cluster
Percentage of time used for administrative tasks per patient	Identify optimization potentials (e.g. by hiring an employee that can help with administrative tasks to relieve the healthcare staff)
Accounted time versus actual time used by physician for one patient	Used to identify wrong or not accounted consultations and to adapt documentation of services in time

One outpatient clinic distributes patients to decentralised waiting areas, as opposed to the classic waiting room. Depending on the required action, a patient takes a seat in a different area. There are different waiting areas, e.g. a waiting area for consultations with the doctors, one for consultations with the nursing staff and in front of each consulting room there is space for one person to wait for access to the corresponding consulting room.

Empty times (unproductive time of the doctors) occur when patients are late or in case of unannounced absences (no show). For two practices, this has become the biggest problem in the area of operations management. Sending reminders in the form of e-mails and text messages reduces but does not completely prevent of no shows.

One group practice is planning to implement an alternative approach to operations management from January 2021 based on lean principles [12]: the patient no longer moves through the different rooms within a practice, but is placed in an empty room and the healthcare professionals and required resources comes to him.

A practice's processes depend on the functionalities of its practice software. All PIS surveyed offer an agenda module for production planning. The appointments can be stored with a status (e.g. 'in the waiting room' or 'cancelled'), which can be freely configured in some cases depending on the software. This allows to allocate rooms and locate the patient. Five PIS additionally support the process control with an open issue tracker system. This system can be used to digitally transmit orders from the doctor to the MPA. In one of the PIS, priorities can be assigned to the orders by the physician. This replaces manual prioritisation by an MPA.

The interviews with the medical practices revealed that they would be interested in the indicators listed in table 1. However, none of the PIS calculates indicators useful for operations management from production data.

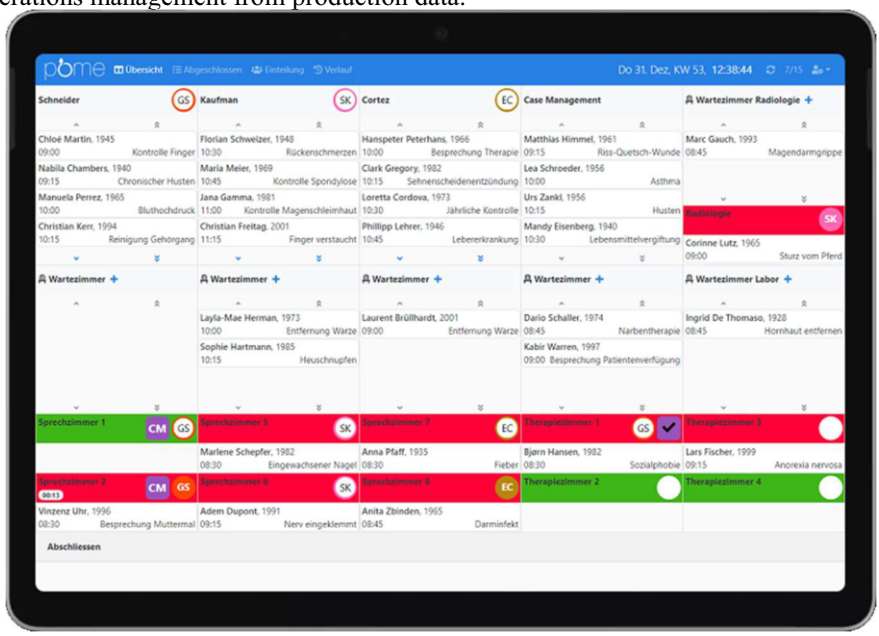


Figure 1. Dashboard for process control and management. For each physician, one column is assigned. Each column comprises three sections: agenda, waiting room, consulting room. Green colour indicates an empty consulting room; red colour an occupied consulting room.

3.2. Dashboard for process control and management

Our application is web-based and supports production control in a group practice. In particular, the functionalities are as follows (Fig. 1): The dashboard provides an overview on the production status of resources (availability of rooms, medical devices), patients (scheduled, awaiting treatment, under treatment), physicians and MPA. Patient names show up as scheduled and are assigned to the physician as scheduled. Location of the patient can be changed by drag and drop, i.e. the patient name can be moved to a waiting room once the patient shows up in the practice, moved forward to the consulting room or can even be checked out when the patient leaves the practice. Beyond, the location of physicians can be changed. The process control, i.e. changing the location of patient or physician names has to be realized by an MPA. A user can trigger manually a synchronisation. Production and process data can be extracted to calculate indicators.

Vue.js was used to create the Javascript single-page application. The library VueDraggable was exploited to realize user control by touch and mouse controller. The system is running under Windows. It is supposed to be used on a tablet. An ICS interface has been implemented to connect the dashboard to the PIS. An ICS interface allows to import calendar data into corresponding programs such as Microsoft Outlook. The concrete interface has been implemented for the PIS provided by the software company Vitodata. Appointment data is send from the PIS to the ICS-Server of Vitodata. Data from this server can be accessed via Internet. Appointment data can be downloaded from the server as text file. The dashboard can be adapted easily to the infrastructure (rooms, medical devices, personnel) of a practice. When the dashboard application is started, the ICS data is collected from the server and shown in the dashboard. After that initialization, updates of appointments are made every 15 minutes.

3.3. Evaluation of Efficiency

Overtime hours of leading MPA are shown in Figure 2. 145 data points before deployment of the tool and 18 data points in the deployment phase have been considered. In an implementation phase of five weeks, we could not yet proof that efficiency can be significantly increased by using the dashboard. Reasons for this might be due to the chosen indicator (overtime hours of the leading MPA). This indicator can be impacted also by other factors such as unusual increase of patients due to the current pandemic situation. Another indicator to be considered would be processing time per patient. We did not used this indicator in our study since we were missing the baseline data which would have been too complicated to collect in a reasonable amount of time and during the current COVID-19 pandemic situation might lead to non-representative data. Other research demonstrated that patient flow analysis is well suited for assessing efficiency of patient visits in primary care [13]. We can recognize some peaks in the overtime hours; unfortunately, no information on reasons of overtime hours are recorded in the time recording system.

4. Discussion

In this paper, we introduced a dashboard that supports production control in a medical practice. The application offers several benefits: at any time information on the production status of patients, physicians, MPA and resources is accessible and they can be easily localised. Indicators for operations management and control such as processing time per patient, time spent in the waiting room, or average delay of a patient can be calculated based on the production data that are stored by the application.

The dashboard can be easily adapted to other configurations of medical practices. Names of physicians or MPA and their working days can be added as well as additional waiting rooms and consulting rooms. In its current implementation, a user has to actively change the status of physicians and patients. The workflow could be facilitated when for example the localisation of the physician (in which consulting room he/she is at the moment) changes automatically. This would require some passive tracking sensors in the various locations.

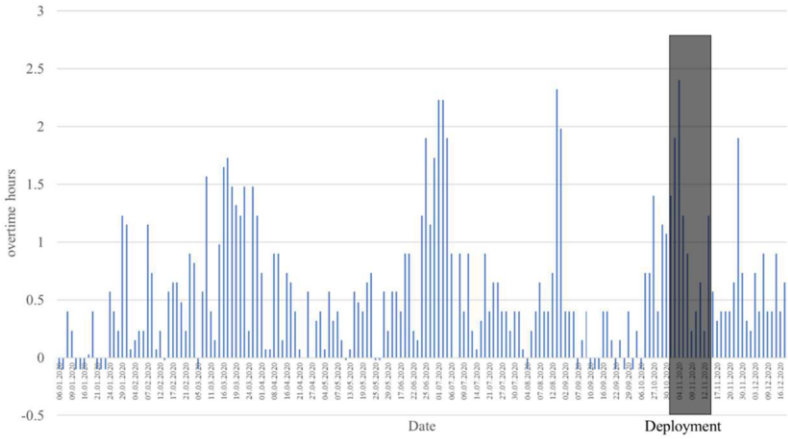


Figure 2. Overtime hours of leading MPA

Our application was developed as locally running tablet application. A server-based application might have been another, if not even better option since this would allow several users to interact with the dashboard. Our decision regarding system architecture has been made based on the interviews in the group practice where we found out that the application will only be used by one MPA. Throughout the implementation phase, it turned out that more users need access to the tool.

Our approach directly transfers knowledge from other assessments of factors that impact physicians' well-being and for increasing practice efficiency. These studies found out that interventions such as pre-visit planning, or optimization of patient flow through clinics, may reduce burnout and improve engagement [14]. Our work focused on improving patient flow during a visit in a medical practice by means of a dashboard. Use of dashboards in medical settings has already been suggested for displaying information for emergency medical services [15] or for supporting decision making. Pestana et al. [16] developed dashboards to improve health care management in hospitals. To the best of our knowledge, no approach similar to our dashboard has been reported so far for ambulatory care.

5. Conclusion

In this work, we assessed the current state and needs of operations management in Swiss medical practices. We recognized that there is a need for support in operations management that existing PIS do not sufficiently provide, mainly provision of indicators for process control. Extensions of PIS by a dashboard as we suggested in this work would be desirable. At least calculating and controlling indicators such as waiting time, processing time of patients etc. should be enabled for improving operations management. These indicators are essential for operations management.

Conflict of interest

VU is working as ICT Business Analyst at Swisscom Health AG in the context of the practice information system curaMED.

References

[1] Bundesamt für Statistik, Kosten und Finanzierung des Gesundheitswesens seit 1960, https://www.bfs.admin.ch/bfs/de/home/aktuell/neue-veroeffentlichungen.assetdetail.14840802.html, last access: 15.01.2021
[2] James B, Bayley KB Cost of poor quality or waste in integrated delivery system settings. AHRQ Publication No. 290-00-0018. Rockville, Md.: Agency for Healthcare Research and Quality; 2006
[3] Litvak E, Buerhaus PI, Davidoff F, Long MC, McManus ML, Berwick DM. Managing unnecessary variability in patient demand to reduce nursing stress and improve patient safety. The Joint Commission Journal on Quality and Patient Safety. 2005 Jun 1;31(6):330-8.
[4] James KA, Ross SE, Vance B, Radcliffe T, Harrison MI, West D. Inefficiency in primary care: common causes and potential solutions. Family Practice Management. 2015 Apr;22(2):18-22.
[5] Nowak T. Optimierungsmöglichkeiten einer Arztpraxis: Organisation, Mitarbeiterführung und Marketing, http://www.wiso-net.de/document/DIPL,ADIP%5F_9783836615242197, last access: 15.01.2021
[6] Thommen J-P. Managementorientierte Betriebswirtschaftslehre. 8. edition, Versus Verl, Zürich, 2008. pp. 275
[7] Golder L, Jans C, Melanie I, Kress J. Swiss eHealth Barometer 2020, gfs.bern, Bern, 2020
[8] McCulloch P, Kreckler S, New S, Sheena Y, Handa A, Catchpole K. Effect of a "Lean" intervention to improve safety processes and outcomes on a surgical emergency unit. Bmj. 2010 Nov 2;341.
[9] Spagnol GS, Min LL, Newbold D. Lean principles in Healthcare: an overview of challenges and improvements. IFAC Proceedings Volumes. 2013 Sep 1;46(24):229-34.
[10] Seiler R. Das rentable Geschäft mit Schweizer Arztpraxen, https://www.handelszeitung.ch/panorama/long-read/das-rentable-geschaft-mit-schweizer-arztpraxen?page=2#page, last access: 15.01.2021
[11] FMH Services, Software 2020 für die Praxisadministration, https://www.fmhservices.ch/softwarekatalog, last access: 15.01.2021
[12] Endsley S, Magill MK, Godfrey MM., Creating a lean practice, Fam Pract Manag. 2006:13(4), 34–8.
[13] Potisek NM, Malone RM, Shilliday BB et al. Use of patient flow analysis to improve patient visit efficiency by decreasing wait time in a primary care-based disease management programs for anticoagulation and chronic pain: a quality improvement study. BMC Health Serv Res 7, 8 (2007).
[14] Tawfik DS, Profit J, Webber S, Shanafelt TD. Organizational factors affecting physician well-being. Current treatment options in pediatrics. 2019 Mar 15;5(1):11-25.
[15] Christen O, Mösching Y, Müller P, Denecke K, Nüssli S. Dashboard Visualization of Information for Emergency Medical Services , Stud Health Technol Inform. 2020 Nov 23;275:27-31.
[16] Pestana M, Pereira R, Moro S. Improving Health Care Management in Hospitals Through a Productivity Dashboard. Journal of Medical Systems. 2020 Apr;44(4):1-9.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210083

Developing Intelligent Interviewers to Collect the Medical History: Lessons Learned and Guidelines

Floriana GASHI[a,1], Selina F. REGLI[a,1] , Richard MAY[b] , Philipp TSCHOPP[c] and Kerstin DENECKE[a,2]
[a] *Bern University of Applied Sciences, Bern, Switzerland*
[b] *Hochschule Harz, Wernigerode, Germany*
[c] *Soccin22 - Zentrum für Hausarztmedizin, Basel, Switzerland*

Abstract. Background: Physicians spend a lot of time in routine tasks, i.e. repetitive and time consuming tasks that are essential for the diagnostic and treatment process. One of these tasks is to collect information on the patient's medical history. Objectives: We aim at developing a prototype for an intelligent interviewer that collects the medical history of a patient before the patient-doctor encounter. From this and our previous experiences in developing similar systems, we derive recommendations for developing intelligent interviewers for concrete medical domains and tasks. Methods: The intelligent interviewer was implemented as chatbot using IBM Watson assistant in close cooperation with a family doctor. Results: AnCha is a rule-based chatbot realized as decision tree with 75 nodes. It asks a maximum of 44 questions on the medical history, current complaints and collects additional information on the patient, social details, and prevention. Conclusion: When developing an intelligent digital interviewer it is essential to define its concrete purpose, specify information to be collected, design the user interface, consider data security and conduct a practice-oriented evaluation.

Keywords. Medical history, dialog-based system, chatbot, conversational user interface, intelligent system

1. Introduction

Collecting information from a patient is an essential part in the diagnostic process. It normally comprises the present condition, personal history, family history, and a social anamnesis. The accuracy and completeness of this information significantly affects the quality of the diagnosis [1]. Normally, all relevant information is collected during the patient-doctor encounter. However, patients not always reveal all concerns and problems in these encounters due to language barriers [2], fear of stigmatism, nervousness, or other reasons [3,4]. Beyond, time in the encounter is limited. In order to improve the situation of incomplete and error-prone information on the medical history paper-based or electronic questionnaires were used [4] enabling "offline" reporting of complaints and medical history. Furaijat et al. implemented a digital communication assistance tool that

[1] Contributed equally

[2] Corresponding author: Kerstin Denecke, Bern University of Applied Sciences, Biel / Bienne, Switzerland; E-mail: kerstin.denecke@bfh.ch

lists complaints users can select and were questions with predefined answers are provided [2]. Another approach are chatbots, in this context often used in so-called conversational user interfaces that simulate a dialog and collect information from the user in this dialog. The first chatbot was developed in the 1960s. The computer program, named ELIZA, was designed to mimic a conversation with a psychotherapist [5]. Chatbots have also proven to be successful in today's COVID-19 pandemic to inform about symptoms, to encourage behavior change, or as monitoring tools [6].

There are chatbot systems available that make a presumptive diagnosis based on the information entered by the user. Ada (https://ada.com/) and the Healee (https://www.healee.com) are apps that collect symptoms from patients, derive a set of possible diseases from them, and show them to the user. A physician can be contacted if needed. However, we are pursuing another approach: We aim at supporting physician and patient in the diagnostic process by means of an intelligent interviewer. In previous work, we have already developed a concept for self-anamnesis and implemented a prototype for collecting the music biography of a person [7-9]. With the experiences gained from this work, we developed an intelligent interviewer for another medical specialty, which is family medicine. In this paper, we aim to contribute the following: 1) We present the developed prototype called AnCha, including its concept, technical specialties, as well as benefits and limitations. 2) Additionally, we aggregate our experiences from the development of AnCha and our previous developments in order to extract recommendations for designing and developing intelligent interviewer for the medical domain. Thus, we aim to facilitate the development of similar systems in the future.

2. Methods

We developed our concept and prototype in three steps: requirement analysis, concept development with mockup generation, and implementation of the prototype. We identified requirements regarding the chatbot AnCha based on two interviews with a general practitioner (GP). In this way, we determined required functionalities and we specified the dialog flow, i.e. which questions the chatbot has to ask a patient to collect information on the medical history and complaints. In addition, we assessed relevant literature to collect requirements and review functionalities of existing apps. Based on the requirements, a mockup was created using Axure RP 9. By means of the mockup and in further meetings with the GP as well as his feedback, the conversation flow of the chatbot was improved. The goal was to find the most effective way to gather all the important information of a patient from the GP'S perspective. The chatbot was implemented using the IBM Watson Assistant. Among other things, IBM Watson Assistant enables creating, training and implementing chatbots. Using integrated machine learning models, user questions posted in a chat can be understood and answered accurately.

In order to come up with recommendations for developing intelligent interviewers for the medical domain, we additionally considered our previous work on the chatbot Ana, who collects the music biography of a person [7-9]. Together with experiences gained in this work, we derive our recommendations for future developments of intelligent interviewers for the medical domain.

3. Results

3.1. Requirements and use case

The main objective of AnCha is to ask a patient questions on his or her medical history and to provide a physician access to the responses in a digital manner as preparation for the patient-doctor encounter. The chatbot is supposed to be rule-based, i.e. a set of specified questions has to be asked in a given order. This specification has the advantage that the patient is guided through the conversation without digressing and the doctor receives the information which is important to him. Several response types are required: predefined answers for selecting one or more options and free text. To avoid misunderstandings, the majority of questions will not require a free text answer. The collected information is not supposed to be a complete medical history. The aim instead is to have an initial set of information in order not to overlook relevant aspects and to focus the communication on the encounter. The information collected by the chatbot will help the physician to create a mind map of the patient's current problems. The patient should be given the opportunity to express all current worries and complaints. In addition, the chatbot should be accessible by patients through the website of the doctor's practice and data security has to be ensured.

The chatbot is supposed to be integrated into the workflow as follows: When a patient comes to a doctor's office with a complaint, he usually spends time in the waiting room before the actual medical appointment. During this waiting time, the patient accesses AnCha. In communication with the chatbot, the patient provides information on current complaints and if needed the medical history using a tablet provided or the patient's own smartphone. As soon as the patient has transmitted the data, it is available to the physician. Shortly before the actual appointment with the patient, the physician accesses the received information. This allows him to prepare for the consultation and to ask specific questions to the patient in the encounter. Another option is to ask the patient for interacting with the chatbot when she makes the appointment.

3.2. Conversation with AnCha

AnCha is running on a website, i.e. it is accessible by patients without installation. The main functionality of AnCha is the collection of current symptoms and complaints as well as aspects of the medical history. The current prototype of AnCha contains a total of 44 questions covering seven topics (see Table 1). At the beginning of the chatbot conversation, the patient is welcomed and receives initial instructions about the chatbot. Formalities for the interaction with the chatbot are clarified, i.e. the patient is asked for her name, date of birth and whether this is the first visit to this practice. In case it is the first visit, additional personal and social information is requested, general questions on medication, alcohol consumption etc. are asked as well as questions on the current status of prevention (topics 2 to 5 in Table 1). The questions about prevention are adjusted according to the entered gender and age. For female patients, questions about gynaecological check-ups are asked, which are omitted for male patients. Patients over 50 years of age are asked if they already possess a health care directive or if they regularly attend a colonoscopy. This means, the number of questions asked to a concrete patient depends on the basic demographic data. Minimum and maximum number of queries per topic are shown in Table 1.

In order to establish an emotional bond of trust to the patient, the chatbot has a name (AnCha, acronym of Anamnesis Chatbot) and is visualized by an avatar. Questions are formulated in a way that they could originate rather from a friend than from a doctor. Specific medical terminology that the patient might not understand is avoided. AnCha tries to motivate the patient to complete the entire dialog. For example, when the user claims, she is not yet willing to continue, AnCha asks a question out of topic or fun facts are posted as distraction. The chatbot not only asks specific questions, but encourages to provide more details. The aim of this is, that the patient feels well accepted and perceived with her concerns and complaints. Once all questions have been asked, the collected information is aggregated, grouped along topics (questions and corresponding answers of the patient) and transformed into a more readable format for the physician. The aggregated data is sent via secure E-Mail to the physician. The conversation with AnCha takes a maximum of ten minutes.

Table 1. Subjects covered by the chatbot as well as minimum and maximum number of questions

	Subject	Questions	Minimal number of questions	Maximum number of questions
(1)	Welcome	- Name - Date of birth - First visit	3	3
(2)	Personal information	- Gender - Age - Size and weight	0	4
(3)	Social information	- Occupation - Marital status and children - Leisure activities	0	4
(4)	General questions	- Alcohol and tobacco consumption - Recent hospitalization - Medication - Pre-existing conditions	0	11
(5)	Prevention	- Vaccinations - Medical check-up - Health care directive - Other doctors	0	10
(6)	Specific questions about the medical consultation	- Reason for visiting the doctor - Questions about the symptoms	7	10
(7)	Other concerns		0	2
		Total	**10**	**44**

3.3. Prototype implementation

Figure 1 shows the system architecture of AnCha. The chatbot is implemented with IBM Watson Assistant, a service of IBM Cloud. The content is arranged in a decision tree which consists of 75 nodes. Currently, the patient accesses AnCha via a separate website and not through the GP practice's website. In detail, the chatbot is integrated into a HTML file using a script. At the end of the questionnaire, the data is transmitted using Node.js. A new action is created in IBM Cloud which uses the module Nodemailer and allows to send e-mails via Node.js. To be able to send e-mails, the mail server, the sender e-mail address, and password have to be stored.

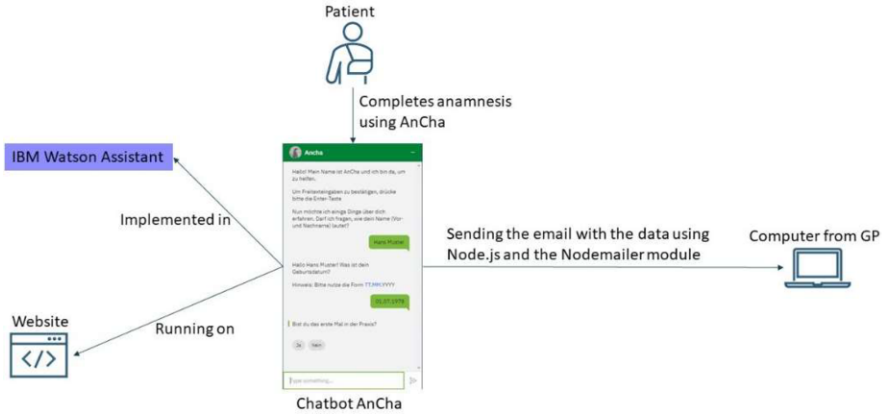


Figure 1. System architecture of AnCha

4. Limitations and benefits of AnCha

In this paper, we introduced the intelligent interviewer AnCha that collects information on the medical history from a patient. In contrast to apps like Ada or Heala, our application AnCha is not supposed to autonomously make any decisions in terms of deriving diagnoses or making a triage. Its aim is to support collection of data on the medical history of a patient and current complaints. The application offers several benefits: The patient can take time to think about the questions and to answer while staying in the waiting room. This approach has the potential to reduce feelings of nervousness which otherwise might lead to wrong or missing information. Further, the patient gets the opportunity to explain everything of relevance to him related to the current complaint. In the patient-doctor encounter, the physician normally has limited time and not everything can be discussed. The information collected by AnCha helps to clarify the patient's complaints in advance, so that the doctor already has an initial overview. This way, he can address these concerns. With the help of AnCha, the patient no longer forgets important information and the doctor is prepared in advance. AnCha thus fulfills the goal of helping the doctor, not replacing him. AnCha is supposed to be an offer to patients. It still has to be assessed, 1) how many patients possess the necessary technical competency for interacting with AnCha, 2) how many patients are willing to use the chatbot in advance, 3) whether 10 minutes of interaction are sufficient to generate a bond of trust to an intelligent assistant and to encourage a patient delivering the requested information. These research questions will be addressed in a future study.

The chatbot could also help in dealing with language barriers: The questions can be provided in multiple languages with appropriate translation. In addition, it is possible to provide explanations or to include images to facilitate communication with non-native speaking patients.

Since AnCha is a rule-based chatbot, the flow of questions is clearly defined. This ensures that all relevant data is collected. However, the communication flow is fixed and might be perceived as unusual by patients. Such perception can be reinforced by the fact that the actual content of the user input is neither analyzed nor interpreted by AnCha. This issue has the following reasons: First, analyzing free text using artificial intelligence

such as IBM Watson Assistant can be error-prone. Second, for our specific use case it is not relevant that the chatbot interprets the information. The primary purpose is to collect information that is then interpreted by the physician.

Since this work was realized during the COVID-19 pandemic, we could not test the application within a family medicine practice. We tested AnCha only with 4 people in the circle of acquaintances of the authors (age: 15, 19, 57 und 60). They had to go through the interview process with AnCha. Their feedback was positive; only the process of sending the e-mail was too complicated since the chatbot asked several times whether the mail should be sent. We are currently planning a study in daily routine with real patients of a medical practice to study patient and physician satisfaction with the adapted process. This evaluation will help to find out whether AnCha would be used by patients and whether it really supports the physician. Based on the results, we will improve our prototype. Depending on the user feedback, we will also extend the motivation capacities of AnCha.

In its current implementation, the data is not encrypted which has to be realized before the system is used in daily routine. At least secure e-mail such as HIN-Mail should be used (https://www.hin.ch/services/hin-mail/). We still have to assess how to securely transfer the chatbot protocol to the physician and whether an integration with the practice information system (or electronic medical records (EMR)) is helpful. Still an open issue is the data transfer and the connection with such system. The chatbot protocol could be stored as note in the EMR which would require a corresponding interface or the sent e-mail might be read by the physician in his mail browser and then stored in the document section of the patient's EMR. AnCha can also be applied in other areas of medicine. Since it is a rule-based chatbot based on a decision tree that does not have many branches, it can be easily extended. The existing questions can be modified, adapted for other medical specialties or additional questions could be added.

5. Recommendations for developing intelligent interviewers

Based on our experiences in designing and building intelligent interviewers, e.g. AnCha as described in this paper or Ana [7-9], we are able to derive several lessons learned. In order to successfully develop intelligent interviewers, in particular based on chatbot technology, it is essential to collect and document all requirements in terms of a comprehensive requirements engineering process [10]. The whole requirements engineering process should at least include the following steps that can be seen as essential guidelines for a successful development process:

First, the concrete purpose of the intelligent interviewer has to be defined. The precise definition depends on the specific use case which includes potential users, e.g. physicians or patients, the environment, e.g. a waiting room, and the medical specialty, e.g. music therapy in the context of mental health [7-9]. If the planned system is only expected to collect information, e.g. personal information, it is sufficient to build a rule-based system. The main advantages of these systems are that their conversations are reliable, easily reproducible and comparable. There are several frameworks or scripting languages available that are appropriate to build such systems, e.g. AIML, SIML, or RiveScript [11]. In our previous work, we developed chatbots based on these technologies and gained several in-depth insights into the technical advantages and disadvantages of these technologies [7-9,11,12,13].

Second, the information to be collected by the intelligent interviewer has to be precisely specified according to the use case. We recommend to use questionnaires or to conduct personal interviews, e.g. with experts. Based on the collected data, a questionnaire can be generated and transferred into one of the already mentioned rule-based chatbot languages. We also recommend to consider the motivation of the later user. In this context, the wording of the intelligent agent's responses should vary in order to create an immersion effect, e.g. in terms of gamification [14]. Moreover, we suggest to use short questions and as few questions as possible. Adding motivating words, funny comments and enabling a reasonable conversation flow, e.g. by creating a logic order without irrelevant questions, also motivates the user to answer all questions satisfactorily. It also makes the intelligent interviewer appear more natural and human.

Third, the user interface has to be designed according to the needs of the future users [15]. This design process includes the general interface design of the whole application and in particular the selection of the input type. The latter can be based on free text responses or pre-defined answers for selection, e.g. based on buttons or sliders. We recommend to provide at least a few free text responses to give the user the opportunity to deliver specific details in his or her own words. However, too many free text responses can be demotivating and confusing for the user. Regarding the app design, we recommend to use round shapes with smooth animations, e.g. for the speech bubbles, and comforting colors, e.g. green [16]. The general appearance of the application should also be based on already known messengers, e.g. WhatsApp or Telegram, in order to enable easy handling and to base upon experiences of users. By adding additional input or output technologies, e.g. speech recognition, the agent's immersion effect can also be increased [17]. However, the accuracy of such services should be considered before their implementation. Combining an intelligent interviewer with immersive technologies, e.g. virtual reality and voice user interfaces, can also be useful in some use cases, e.g. for patient education [13].

Fourth, data security has to be considered. Although often ignored in prototypes, data security, especially for personal patient data, has to be ensured. Therefore, encrypting the collected data, e.g. by using algorithms such as AES-256, can provide an appropriate security standard. In addition, implementing secure mail services, e.g. by using certificates or public and private keys, are essential to realize a secure information exchange of the data requested from the intelligent interviewer [18]. We do not recommend using third party services that can access (personal) data collected by the interviewer (e.g. Facebook messenger) or realize logins using Google or Facebook accounts. For instance, this includes the choice of chatbot technology, pre-built user interfaces, as well as input and output services. Fifth, by conducting a practice-oriented evaluation, e.g. based on a quantitative or qualitative questionnaire and participants according to the target group, potential issues or space of improvements can be identified. In this context, the evaluation should focus on both content-related tasks as well as general usability [8]. Accordingly, further requirements that were previously hidden can be identified and addressed. Based on the evaluation results, the gap between the theoretical concept of the intelligent interviewer in a specific use case and the actual practice can be minimized or even closed.

6. Conclusions

Providing intelligent digital interviewer in healthcare gains in interest. We suggest to consider 5 essential aspects that will help in faster development of reliable and secure intelligent interviewer for the medical domain: define concrete purpose, specify information to be collected, design user interface and address future user's needs, consider data security and conduct a practice-oriented evaluation. Regarding our chatbot AnCha, we will assess in future whether use of the artificial intelligence capabilities of IBM Watson can be used. This requires additional analysis of possible user questions and decisions on how the chatbot should react to them. AnCha is not a competitor to the patient-doctor interaction and thus the patient-doctor relationship should not be negatively impacted.

References

[1] P.J. Moonen, L. Mercelina, W. Boer, T. Fret, *Diagnostic error in the emergency department: follow up of patients with minor trauma in the outpatient clinic*. Scand J Trauma Resusc Emerg Med 2017; 25(01):13

[2] G. Furaijat, E. Kleinert, A. Simmenroth. et al. *Implementing a digital communication assistance tool to collect the medical history of refugee patients: DICTUM Friedland - an action-oriented mixed methods study protocol*. BMC Health Serv Res 2019(19):103

[3] T.M. Palermo, D. Valenzuela, P.P. Stork, *A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability*. Pain 2004;107(03):213–219

[4] D. Zakim, *Development and significance of automated history taking software for clinical medicine, clinical research and basic medical science*. J Intern Med 2016;280(03):287–299

[5] J. Weizenbaum, *Eliza - a computer program for the study of natural language communication between man and machine.* Commun ACM 1966;9(01):36–45

[6] A.S. Miner, L. Laranjo, A.B. Kocaballi, *Chatbots in the fight against the COVID-19 pandemic*. npj Digit. Med. 3, 65 (2020). https://doi.org/10.1038/s41746-020-0280-0

[7] K. Denecke, *A Mobile System for Music Anamnesis and Receptive Music Therapy in the Personal Home*, Stud Health Technol Inform. 2017;245:54-58

[8] K. Denecke, A. Pöpel, S. Lutz Hochreutener, R. May, *Self anamnesis with a Conversational User Interface: Concept and Usability Study*. Methods Inf Med. 2018 Nov;57(5-06):243-252

[9] K. Denecke, R. May, A. Pöpel, S. Lutz Hochreutener, *Can a chatbot increase the motivation to provide personal health information?* pHealth 2020, Studies in Health Technology and Informatics, Volume 273: pHealth 2020, 10.3233/SHTI200619, 85-90

[10] A. Van Lamsweerde, *Goal-oriented requirements engineering: A guided tour*. Proceedings of the IEEE International Conference on Requirements Engineering (RE), 249–261 (2001)

[11] K. Denecke, M. Tschanz, T. Dorner, R. May, *Intelligent Conversational Agents in Healthcare: Hype or Hope?* Studies in health technology and informatics. 2019;259:77-84

[12] K. Denecke, R. May, Y. Deng, *Towards Emotion-Sensitive Conversational User Interfaces in Healthcare Applications.* Studies in health technology and informatics. 2019;264:1164-1168

[13] R. May, K. Denecke, *Extending Patient Education with CLAIRE: An Interactive Virtual Reality and Voice User Interface Application*. Lecture Notes in Computer Science, Volume 12315: Springer, Cham 2020. 10.1007/978-3-030-57717-9_49.

[14] S. Deterding, D. Dixon, R. Khaled, L. Nacke, *From game design elements to gamefulness: defining "gamification"*. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, 9–15 (2011)

[15] J. Nielsen, *10 usability heuristics for user interface design.* Nielsen Norman Group 1, 1 (1995)

[16] A. Akers, J. Barton, R. Cossey, P. Gainsford, M. Griffin, D. Micklewright, *Visual Color Perception in Green Exercise: Positive Effects on Mood and Perceived Exertion*. Environment Science Technology, 2012;46(1):8661–8666

[17] R. Nakatsu, N. Tosa, *Active immersion: the goal of communications with interactive agents.* IEEE International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies 2000;1(1):85–89

[18] A. Dan, S. Gupta, S. Rakshit, S. Banerjee, Toward an AI Chatbot-Driven Advanced Digital Locker. *Advances in Intelligent Systems and Computing*, Springer Singapore 2019: 811:37–46

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210084

Topic Discovery on Farsi, English, French, and Arabic Tweets Related to COVID-19 Using Text Mining Techniques

Hamoon JAFARIAN[a,1], Mahin MOHAMMADI[b] , Alireza JAVAHERI[c], Makram SUKARIEH[d], Mohsen YOOSEFI NEJAD[e], Abbas SHEIKHTAHERI[f], Mehdi HOSSEINZADEH[g], Elaheh MOMENI-ORTNER[h] and Reza RAWASSIZADEH[d,2]

[a] *Independent Researcher, Vaughan, Canada*
[b] *School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran*
[c] *Independent Researcher, Rouen, France*
[d] *Department of Computer Science, Metropolitan College, Boston University, MA, USA*
[e] *Department of Computer Engineering and Information Technology, Payame Noor University, Tehran, Iran*
[f] *Health Management and Economics Research Center, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran*
[g] *Mental Health Research Center, Psychosocial Health Research Institute, Iran University of Medical Sciences, Tehran, Iran*
[h] *University of Vienna, Vienna, Austria*

Abstract. Background: Social networks are a good source for monitoring public health during the outbreak of COVID-19, these networks play an important role in identifying useful information. Objectives: This study aims to draw a comparison of the public's reaction in Twitter among the countries of West Asia (a.k.a Middle East) and North Africa in order to make an understanding of their response regarding the same global threat. Methods: 766,630 tweets in four languages (Arabic, English French, and Farsi) tweeted in March 2020, were investigated. Results: The results indicate that the only common theme among all languages is "government responsibilities (political)" which indicates the importance of this subject for all nations. Conclusion: Although nations react similarly in some aspects, they respond differently in others and therefore, policy localization is a vital step in confronting problems such as COVID-19 pandemic.

Keywords. COVID-19, Natural Language Processing, Social Networking, Epidemics

---

[1] Contributed equally
[2] Corresponding Author: Reza Rawassizadehi, Department of Computer Science, Metropolitan College, Boston University, MA, US, E-Mail: rezar@bu.edu

1. Introduction

The global spread of the COVID-19 pandemic, an infectious disease caused by the pathogen severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has already unleashed an unprecedented impact on public health, economy, and human society worldwide[1]. Social media platforms (such as Twitter, Facebook, Reddit, Tumblr, Pinterest and Instagram) have seen unprecedented growth in the era of big data. For example, Twitter, one of the most popular social network websites, which has been growing at a very fast pace. It has 284 million monthly active users, and 500 million tweets are sent per day [2]. Twitter has been used as an early warning notifier, emergency communication channel, public perception monitor, and proxy public health surveillance data source in a variety of disaster and disease outbreaks from hurricanes[3]. Millions of people are talking about the coronavirus on social media, particularly on Twitter, where there are massive conversations around a variety of topics related to COVID-19 [1].

On the other hand, public contribution is the key to bringing under control this pandemic. Researchers are making every effort to anticipate the pandemic's trajectory [4, 5]. Our work is a step forward toward better understanding of public opinions and concerns about this pandemic. In particular, this paper aims to answer the following two questions: RQ1) Which aspects of the COVID-19 pandemic has attracted the most public attention in West Asia and North Africa? and RQ2) What are the differences and similarities of the public response among the countries of West Asia and North Africa?

To address these questions, we investigated Twitter (https://twitter.com/) as a public opinion platform. Although generalization from tweets might lead to a certain degree of inaccuracy, Twitter is a repository of billions of attitude and expressions, and serves as a practical source for topic modelling. Investigating a public opinion about a wide spectrum of topics in Twitter has been made in plenty of studies [6-13]. The advantages of social media mining have been pointed out in a number of papers [14, 15] and it has been demonstrated that the data extracted from the social media platforms is comparable with that provided otherwise (e.g. questionnaires, etc.). Furthermore, tweets are largely clear from the errors inherent in traditional means of information gathering (e.g. polls and questionnaires) [16-18], where participants' opinions are dependent upon the context of the questions, their format, wording, and ordering.

Several studies have been carried out aiming to detect the topics related to COVID-19 [3, 19-26]. However, a mere of inadequate researches have focused on several languages. Nor adequate research has been conducted to address people's reaction to COVID-19 issue particularly in West Asia and North Africa. The MENA (Middle East and North Africa) region consist of Algeria, Bahrain, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Tunisia, United Arab Emirates and Yemen. This region enjoys a broad diversity of languages. Nevertheless, most significant spoken language of the region is Arabic, spoken in Algeria, Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Tunisia, United Arab Emirates and Yemen. Persian or Farsi, is another dominant language of this region and is spoken mainly in Iran (and Afghanstan and Tajikstan but they are outside of MENA). Although English and French are not among the official languages of this region, they are widely used throughout these countries for education, diplomacy, and business.

This paper investigates tweets of four languages dominant in West Asia and North Africa: Arabic, English, French, and Farsi. These are among the most spoken languages in this region. Results of our study can help policymakers better recognize the efficiency

of their public policies, which is the key to increase the public awareness, and to encourage people to respect as much as possible the restrictive measures and, in turn, shape a better relationship with the public. Furthermore, since there is not a global and unique policy to mitigate the risks of the COVID-19 pandemic, this study provides practical information for the governments to tailor their contamination measures to the local context. In addition, the findings of this research provide helpful information for the sociologists who try to measure the social effects of this pandemic behind the scene.

2. Methods

To categorize the tweets into themes, the process presented in Figure 1 is performed for each language. To collect COVID-19 related tweets, we used Phirehose, an open source, PHP implementation of Twitter Streaming Application Programming Interface (API). Table 1 shows details about the dataset and used filters. Tweets were collected during March 2020 as it was the first huge wave of global spread of the COVID-19.

We used Python 3.7.6 to preprocess the tweets. Preprocessing steps for the four languages are quite the same including: converting letters to lower case, removing URL, mentions, stop-words and emoji, correcting repeated characters, and tokenizing and replacing negations with NOT. Arabic and Farsi language required another step that is normalization. The normalization module is used to unify those words that may be written in different alternative forms.

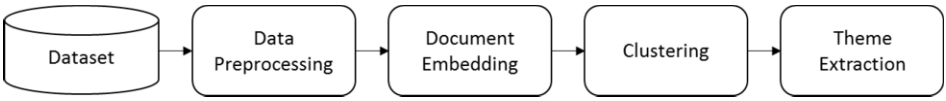


Figure 1. Process pipeline of our method

To find the most proper document embedding and clustering method for topic modeling on short texts, Curiskis et al. [27] investigated different combinations on three datasets on Twitter and Reddit and concluded that Doc2vec embedding along with k-means clustering delivered the best performance and therefore, this combination was employed in the present study. Gensim and Sklearn packages were used for Doc2vec and k-means clustering, respectively.

Table 1. Extracted dataset details

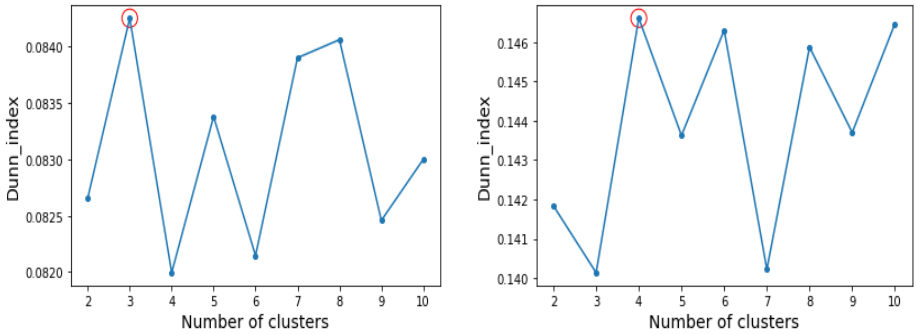
Language	Filtered hashtags	Num. of tweets	RT to Tweet Ratio	Time period
Arabic	#كورونا #كرونا #كورونا_فيروس #كرونا_فيروس #الكورونا	197,794	62.5%	2020-03-03 to 2020-03-12
Persian	#كويد19 #كرونا #ويروس_كرونا #كرونا_ويروس	199,705	57.8%	2020-03-03 to 2020-03-06
English	#corona #coronavirus #covid19 #covid2019 #covid-19	176,370	54.5%	2020-03-03 to 2020-03-09
French		192,761	65.9%	2020-03-03 to 2020-03-18

In developing the Doc2vec model, number of epochs is a case sensitive parameter that considerably affects the performance of topic modeling where it is desired for each cluster to represent the most similar texts, while for clusters to be as distinctive and different as possible. In order to find the fittest number of epochs, cosine similarity criterion [28] was employed as follows. To evaluate the tweets similarity in each cluster, 10 most frequent words in each cluster were converted to vectors using Word2vec and cosine similarity was calculated for each word-pair. The average value is the criterion for each cluster. The final value is the average of the calculated cosine similarity of all clusters. To evaluate the difference among clusters, first, the average vector of 10 most frequent words in each cluster was calculated. Next, pairwise cosine similarity of these vectors were calculated and finally, one minus the average would show how different the topics are. By applying this method, we evaluated performance for different epochs and obtain the optimum number for our Doc2vec model which in our case lead to 10 epochs. Table 2 shows the effect of number of epochs in Doc2vec model on the performance of topic modeling. Based on Table 2, 10 epochs for Doc2vec model was chosen.

Table 2. Effects of number of epochs on model performance

	Epochs	Num. topics	Cosine similarity of topics	1-cosine similarity of topics
Farsi	5	8	0.326	0.325
	10	**8**	**0.3**	**0.41**
	15	8	0.262	0,397
English	5	8	0.32	0.40
	10	**8**	**0.29**	**0.48**
	15	8	0.26	0.52
French	5	8	0.248	0.424
	10	**8**	**0.264**	**0.411**
	15	8	0.27	0.40
Arabic	5	8	0.398	0.428
	10	**8**	**0.422**	**0.472**
	15	8	0.398	0.484

To find the optimum number of clusters, Dunn index criteria [29] was utilized. For Farsi, English, French and Arabic tweets, the optimum number of clusters were obtained as 4, 5, 3 and 4, respectively. Figure 2 shows the how Dunn index is used to find the optimum number of clusters for French and Arabic tweets.



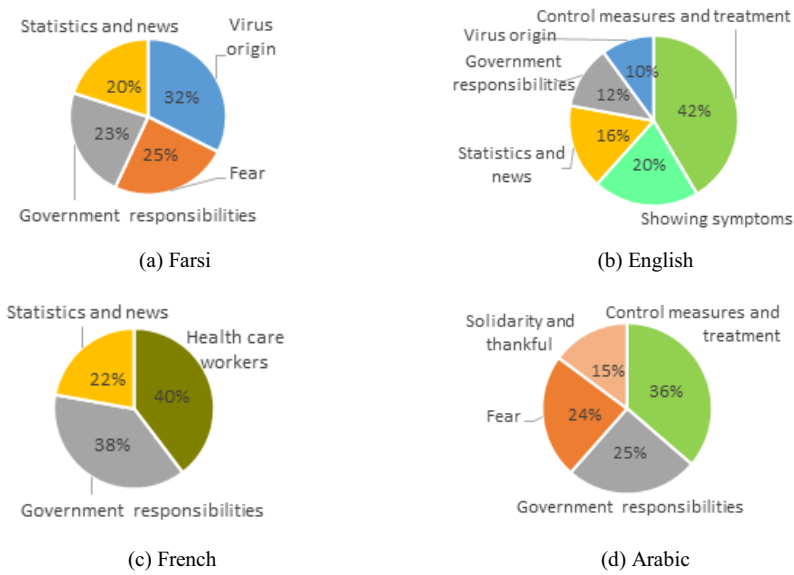
(a): French (b): Arabic

Figure 2. Dunn index versus the number of clusters in French and Arabic

After clustering tweets in each language, in order to analyze the content, each cluster should be labeled with a theme manually. To implement this task, 10 most frequent words of each cluster in addition to a random collection of 100 tweets were extracted and presented to three different native speakers of that language along with a list of suggested themes (annotators were encouraged to change the themes or add new theme as pleased and the list was just some examples to clarify the task). Candidates used the information to choose at least one theme for each cluster. In an attempt to measure the inter-rater agreement between the three raters, Fleiss Kappa testing was employed [30].

3. Results

The themes for each language and their distribution are shown in Figure 3. Farsi tweets were categorized into four themes with the Fleiss Kappa of 0.826 where "virus origin" has the largest share. English tweets with seven themes had the most versatile distribution. The Fleiss Kappa was 0.73 and most English tweets were on "control measures and treatment" theme. French and Arabic tweets were divided into three and four themes with Fleiss Kappa of 0.916 and 0.874, respectively. While the majority of French tweets were regarding "government responsibilities", Arabic tweets were mostly about "control measures and treatment", similar to English tweets.



(a) Farsi (b) English

(c) French (d) Arabic

Figure 3. Farsi, English, French and Arabic tweets analysis results pie chart

To better understand the content of the clusters, as a sample, Table 3 shows the 10 most frequent words and two samples of tweets for the "government responsibilities" theme in English and Farsi languages.

Table 3 Ten most frequent words and tweet samples for "Governmental responsibilities" theme

Language	Ten most frequent words		Tweet samples
English	outbreak	U.S.	Safety first, authorities implant strict screening measures to conduct medical checkups on passengers #dubai India suspends visas from Iran, Italy, South Korea, and Japan
	China	global	
	due	world	
	Italy	country	
	Iran	spread	
Farsi	(Iran)ایران	(US)آمریکا	ناتوانی دولت در مواجهه با ویروس صدای خبرنگار الجزیره را هم درآورد (Inability of government in confronting the virus made the al jazeera journalist to object) روحانی در جلسه هیئت دولت: قول می دهم در کوتاه ترین مدت از بحران عبور می کنیم (President Rouhani in government meeting: I promise we will pass this crisis in shortest time)
	(people) مردم	فرانسه_انگلیس (France-England)	
	(outbreak)شیوع	جمهوری_اسلامی (Islamic Republic)	
	(China)چین	(help)کمک	
	(country)کشور	مقابل (confronting)	

4. Discussion

Social media act as an appropriate source of information in dangerous situations [31]. Although at this time, the actions and reactions of the people and officials of countries are not possible in real space, the social atmosphere these days is in social media such as Twitter, and demands and actions are easily exchanged. In this study, COVID-19 related topics and discussions in English, French, Persian, and Arabic among Twitter users, inside West Asia and North Africa during the first wave of COVID-19, were analyzed and studied.

Subject of control measures and treatment is a common theme among English-speaking, and Arabic-speaking users. As COVID-19 spreads to other countries and governments try to mitigate its impact by implementing counter measures, people have also used social media platforms to express their opinion about the measures themselves, the leaders implementing them, and the ways their lives are changing [32].

The origin of the virus has been a common theme among English-speaking and Persian-speaking users. For example, in Persian language, when it was announced that the first case of the coronavirus in Iran was discovered in Qom province, the mentality of society gradually emerged that Qom province was the source of the virus, and it was also announced that the virus was transmitted by Chinese immigrants or traders. Then it has gone to other cities from Qom province. We got this in one of the Persian language clusters.

One common theme among English, French, Persian and Arabic users is the government political responsibilities, relatively authoritarian public health measures (such as physical distancing or temporary economic shutdowns) depend on societal compliance. People follow these policies when they have a good and reliable relationship with politicians and also these officials have a political economy that allows their people to stay home without suffering from hunger [33].

A common theme among Persian and Arabic users has been "fear". According to US Center for Disease Control and Prevention (CDC), this pandemic results in fear and anxiety about the new disease and also this fear may result from public health actions, and social distancing, because these actions may make people feel isolated and lonely. Therefore, governments should inform the public about the necessity of these actions and support people emotionally and provide necessary services in this regard[34].

The common theme among English-speaking, French-speaking and Persian-speaking users on Twitter is the subject of statistics and news. It shows that these communities care about the news and death statistics and the prevalence of this disease in their countries. Social media have played pivotal roles in the dissemination of information during the COVID-19 pandemic including both the rapid sharing of scientific research as well as various hoaxes and misinformation.[35].

The results of this study can be a help to improve treatment measures, macro decisions, social support, and a better understanding of people's behavior and reactions during an epidemic. For future studies, it is recommended that based on the geographical locations and time, users' opinions for this pandemic be collected and processed.

Acknowledgements
This research is supported by a grant from the Institute for Health System Innovation & Policy (IHSIP) of Boston University.

References

[1] Z. Fang, R. Costas, Tracking the Twitter attention around the research efforts on the COVID-19 pandemic, arXiv preprint arXiv:2006.05783, (2020)
[2] X. Dai, M. Bikdash, B. Meyer, From social media to public health surveillance: Word embedding based clustering method for twitter classification, in: SoutheastCon 2017, IEEE, 2017, pp. 1-7
[3] C. Ordun, S. Purushotham, E. Raff, Exploratory analysis of covid-19 tweets using topic modeling, umap, and digraphs, arXiv preprint arXiv:2005.03082, (2020)
[4] K.K. Sahu, A.K. Mishra, A. Lal, Trajectory of the COVID-19 pandemic: chasing a moving target, Annals of translational medicine, 8(11) (2020)
[5] H. Loeffler-Wirth, M. Schmidt, H. Binder, Covid-19 Transmission Trajectories—Monitoring the Pandemic in the Worldwide Context, Viruses, 12(7) (2020) 777
[6] A. Javaheri, N. Moghadamnejad, H. Keshavarz, E. Javaheri, C. Dobbins, E. Momeni-Ortner, R. Rawassizadeh, Public vs media opinion on robots and their evolution over recent years, CCF Transactions on Pervasive Computing and Interaction, (2020) 1-17
[7] C. Buntain, J. Golbeck, B. Liu, G. LaFree, Evaluating public response to the Boston Marathon bombing and other acts of terrorism through Twitter, in: Tenth International AAAI Conference on Web and Social Media, 2016,
[8] T. Mitra, S. Counts, J.W. Pennebaker, Understanding anti-vaccination attitudes in social media, in: Tenth International AAAI Conference on Web and Social Media, 2016,
[9] K. Starbird, Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter, in: Icwsm, 2017, pp. 230-239
[10] M. Bagdouri, Journalists and Twitter: A Multidimensional Quantitative Description of Usage Patterns, in: ICWSM, 2016, pp. 22-31
[11] P. Missier, C. McClean, J. Carlton, D. Cedrim, L. Silva, A. Garcia, A. Plastino, A. Romanovsky, Recruiting from the network: Discovering twitter users who can help combat zika epidemics, in: International Conference on Web Engineering, Springer, 2017, pp. 437-445
[12] S. Faralli, G. Stilo, P. Velardi, What women like: A gendered analysis of twitter users' interests based on a twixonomy, in: Ninth international aaai conference on web and social media, 2015,
[13] N. Alsaedi, P. Burnap, O.F. Rana, Automatic summarization of real world events using twitter, in, AAAI, 2016,
[14] A. Hassan, V. Qazvinian, D. Radev, What's with the attitude? identifying sentences with attitude in online discussions, in: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, 2010, pp. 1245-1255
[15] B. O'Connor, R. Balasubramanyan, B.R. Routledge, N.A. Smith, From tweets to polls: Linking text sentiment to public opinion time series, Tepper School of Business, (2010) 559
[16] H. Schuman, S. Presser, Questions and answers in attitude surveys: Experiments on question form, wording, and context, Sage, 1996.
[17] R.P. Abelson, Conviction, American Psychologist, 43(4) (1988) 267
[18] J.F. Dovidio, R.H. Fazio, New technologies for the direct and indirect assessment of attitudes, (1992)

[19] Q. Liu, Z. Zheng, J. Zheng, Q. Chen, G. Liu, S. Chen, B. Chu, H. Zhu, B. Akinwunmi, J. Huang, Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach, Journal of Medical Internet Research, 22(4) (2020) e19118
[20] J. Xue, J. Chen, C. Chen, C. Zheng, S. Li, T. Zhu, Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter, PloS one, 15(9) (2020) e0239441
[21] T. Alshaabi, J.R. Minot, M.V. Arnold, J.L. Adams, D.R. Dewhurst, A.J. Reagan, R. Muhamad, C.M. Danforth, P.S. Dodds, How the world's collective attention is being paid to a pandemic: COVID-19 related 1-gram time series for 24 languages on Twitter, arXiv preprint arXiv:2003.12614, (2020)
[22] D.R. Dewhurst, T. Alshaabi, M.V. Arnold, J.R. Minot, C.M. Danforth, P.S. Dodds, Divergent modes of online collective attention to the COVID-19 pandemic are associated with future caseload variance, arXiv preprint arXiv:2004.03516, (2020)
[23] M. Thelwall, S. Thelwall, Retweeting for COVID-19: Consensus building, information sharing, dissent, and lockdown life, arXiv preprint arXiv:2004.02793, (2020)
[24] R. Kouzy, J. Abi Jaoude, A. Kraitem, M.B. El Alam, B. Karam, E. Adib, J. Zarka, C. Traboulsi, E.W. Akl, K. Baddour, Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter, Cureus, 12(3) (2020)
[25] L. Singh, S. Bansal, L. Bode, C. Budak, G. Chi, K. Kawintiranon, C. Padden, R. Vanarsdall, E. Vraga, Y. Wang, A first look at COVID-19 information and misinformation sharing on Twitter, arXiv preprint arXiv:2003.13907, (2020)
[26] T.C. Hamamsy, R. Bonneau, Twitter activity about treatments during the COVID-19 pandemic: case studies of remdesivir, hydroxychloroquine, and convalescent plasma, medRxiv, (2020)
[27] S.A. Curiskis, B. Drake, T.R. Osborn, P.J. Kennedy, An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit, Information Processing & Management, 57(2) (2020) 102034
[28] B. Wang, A. Wang, F. Chen, Y. Wang, C.-C.J. Kuo, Evaluating word embedding models: Methods and experimental results, APSIPA transactions on signal and information processing, 8 (2019)
[29] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, (1973)
[30] J.L. Fleiss, B. Levin, M.C. Paik, Statistical methods for rates and proportions, john wiley & sons, 2013.
[31] H.W. Park, S. Park, M. Chong, Conversations and medical news frames on twitter: Infodemiological study on covid-19 in south korea, Journal of Medical Internet Research, 22(5) (2020) e18897
[32] C.E. Lopez, M. Vasu, C. Gallemore, Understanding the perception of COVID-19 policies by mining a multilanguage Twitter dataset, arXiv preprint arXiv:2003.10359, (2020)
[33] S.L. Greer, E.J. King, E.M. da Fonseca, A. Peralta-Santos, The comparative politics of COVID-19: The need to understand government responses, Global public health, 15(9) (2020) 1413-1416
[34] Pandemics can be stressful, in, 2020
[35] M.R. Jimenez‐Sotomayor, C. Gomez‐Moreno, E. Soto‐Perez‐de‐Celis, Coronavirus, Ageism, and Twitter: An Evaluation of Tweets about Older Adults and COVID‐19, Journal of the American Geriatrics Society, (2020)

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210085

First Accreditation of a Medical Informatics Program by the European Federation of Medical Informatics: Experiences

Elske AMMENWERTH[a,1]
[a] *UMIT – Private University for Health Scienes, Medical Informatics and Technology, Hall in Tirol, Austria*

Abstract. Background: European universities offer a variety of programs in Medical Informatics. The Europen Federation of Medical Informatics (EFMI) offers accreditation of these programs. Objectives: To describe the process of EFMI accreditation of a new Austrian master's program and reflect on accreditation benefits. Methods: Reflection on feedback and experiences at UMIT TIROL Results: Accreditation needs quite some preparation but offers essential opportunities for self-reflection and feedback by international experts. Conclusion: Besides national accreditation, medical informatics programs can benefit from the accreditation through international organizations.

Keywords. Medical informatics, accreditation, quality management

1. Introduction

Modern health care is not thinkable without medical informatics, and consequently, job prospects for graduates in medical informatics are excellent. The number of universities that offer master's programs in medical informatics is increasing, with > 1.000 programs in biomedical and health informatics and related specializations in Europe alone (http://efmi-ac2.bmhi-edu.org).

Accreditation determines whether a program meets specific minimum standards. National accreditation may not always include a useful international perspective on medical informatics education [1]. International accreditation may provide this international perspective and increase the attractiveness of programs for students.

To ensure transparency and quality of medical informatics programs in Europe, the European Federation of Medical Informatics (EFMI) launched an accreditation program of medical informatics programs in 2020 [2]. While the accreditation fees are comparably low, some effort still needs to be invested in such a formal accreditation process. European universities now have to consider whether this EFMI accreditation is worthwhile for them.

This contribution aims to present the process and experiences of the first EFMI accreditation of a European medical informatics master's program and discuss benefits for European universities.

[1] Corresponding Author: Elske Ammenwerth, Institute of Medical Informatics, UMIT TIROL, Hall in Tirol, Austria, elske.ammenwerth@umit.at

2. Methods

The first EFMI accreditation was conducted in summer 2020 for the new master's program Medical Informatics (http://www.umit.at/medical-informatics) at UMIT TIROL. The accreditation process comprised the following steps: Submission of the application to the EFMI board; appointment of an EFMI accreditation panel consisting of three members; preparation of a self-assessment report; panel visits in the educational institution (due to the pandemic situation, this was done online); preparation of an on-site panel's report: and decision on accreditation. Accreditation criteria comprised needs and relevance of the program, intended learning outcomes, academic environment, organization and implementation, and internal quality assurance and development. Overall, the accreditation process lasted less than four months. The self-assessment report comprised nearly 80 pages and was complemented by 22 Annexes. During the on-site visit, ten staff members from UMIT TIROL were interviewed. In the end, full EFMI accreditation was granted for the maximum possible duration of three years.

3. Results

The self-assessment report's preparation was more time-intensive than initially expected. Several aspects such as the institution, the new program's need, the intended learning outcomes, the teaching environment, organization, and quality assurance had to be described in detail and in English. However, the self-assessment report helped us reflect on our strengths, which helped us prepare national accreditation. The on-site discussions elaborated on several essential aspects of the program, and the experts' feedback helped us optimize curriculum and organization. For example, the feedback showed that the curriculum did not sufficiently show how it implemented some intended learning outcomes. It was helpful that our curriculum was in line with international education recommendations such as the IMIA recommendation [3].

4. Discussion

We benefited from the feedback of international experts on the content and organization of our curriculum. EFMI has a high reputation in medical informatics, and we expect that EFMI accreditation will help us attract the best national students. As UMIT TIROL was not involved in developing the EFMI accreditation procedure or criteria, this was also for EFMI the first “real life test” of their accreditation process.

References

[1] A. Hasman, and J. Mantas, IMIA accreditation of health informatics programs, Healthc. Inform. Res. 19 (2013) 154–161. doi:10.4258/hir.2013.19.3.154.

[2] J. Mantas, The Accreditation Procedure Ensures Quality of Education in Biomedical and Health Informatics., Stud. Health Technol. Inform. 272 (2020) 484–486. doi:10.3233/SHTI200601.

[3] J. Mantas, E. Ammenwerth, G. Demiris, A. Hasman, R. Haux, W. Hersh, E. Hovenga, K.C. Lun, H. Marin, F. Martin-Sanchez, and G. Wright, Recommendations of the international medical informatics association (IMIA) on education in biomedical and health informatics, *Methods Inf. Med.* **49** (2010).

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210086

A Customized Smart Home and Interior Design Concept Co-Designed with and for People with Autism Spectrum Disorder

Sandra Lisa LATTACHER[a,b,1], Lukas WOHOFSKY[a,b], Philip SCHARF[a,b] and Daniela KRAINER[a,b]
[a] *Research Unit Active & Assisted Living at Carinthia University of Applied Sciences, Klagenfurt, Austria*
[b] *Institute for Applied Research on Ageing, Villach, Austria*

Abstract. Autism spectrum disorder (ASD) diagnoses increased over the last decades, as reviews show comparing prevalence rates reported from different studies. Due to different effects of the disorder, personal support is required and provided by formal and/or informal caregivers in various activities of daily living. With the help of a customized smart home and interior design concept the aim is to enable people with ASD to live a more independent and self-reliant life. Following a participatory research approach, the end users are involved in the context of use and requirements definition, concept development, and later also in the implementation, and evaluation process. The solution shall assist end users in performing activities of daily living. The outcome of the work at hand is a set of modular functionalities (sensors, actuators, interior design solutions) to be integrated in a living environment specifically designed for people with ASD.

Keywords. Autism Spectrum Disorder, Smart Home, User-Centered Design, Stakeholder Analysis, Participatory Research, Assistive Technologies.

1. Introduction

Autism spectrum disorder (ASD) is a developmental disorder with a wide variation in the type and severity of symptoms. Referring to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), there are two main symptom complexes: difficulties with communication and social interactions, and restricted repetitive behavior, activities, or interests. Due to different effects of the disorder, individual support is required and provided by formal and/or informal caregivers in various activities of daily living (ADL) [1]. Although many assistive technologies for people with ASD can be found in literature [2], the field of home automation in combination with an interior design concept for people with ASD is, to the authors' knowledge, non-existent yet. Considering the whole path from initial user requirements to system design specifications, a concept of a customized smart home and interior design for assisted living is developed. This holistic approach is tailored to the needs of autistic people, aimed at increasing independency and self-reliance scalable to multifaceted life realities and diverse home environments.

[1] Corresponding Author: Sandra Lisa Lattacher, Research Unit Active & Assisted Living at Carinthia University of Applied Sciences, Klagenfurt, Austria, s.lattacher@fh-kaernten.at

2. Methods

Following a participatory research approach, the end users of the system (people with ASD, formal/informal caregivers) and associated stakeholders (architects, health professionals) were involved in the context of use and requirements definition, and the concept development based on DIN EN ISO 9241-210 [3]. Qualitative methods included individual remote focus groups targeting at various aspects of the users' life like daily routines, challenges and scenarios where smart gadgets and furnishings could support independent living. In addition, the experience of architects with the design of residential spaces and with following autism friendly design guidelines were obtained. Quantitative methods encompassed questionnaires, which were targeting at special sensitivity of individuals to thermal, acoustical, visual, and indoor air conditions.

3. Results

As a collective output, the most critical environmental factors that could trigger stress and tension were identified and corresponding technical, architectural, or combined solutions have been conceptualized, followed by a stakeholder validation of importance and usefulness and evaluation regarding feasibility and priority. The outcome is a set of modular functionalities, which covers solutions like crowd warnings, transition spaces or room condition regulation realized through sensors, actuators, and interior design components. [4]

4. Discussion

Through the continuous involvement, engagement, and participation of users we expect a sustainable, economic, and societal solution specifically designed for people with ASD and caregivers to support autonomous living. We facilitate the flow back of benefits to the end users that are in the center of our research. As a next step, the sensor technology and architectural components will be implemented, integrated, and tested in a demo apartment in a lab setting.

The research leading to these results is co-financed by the European Regional Development Fund under the framework of INTERREG IT-AT (SENSHOME project).

References

[1] I. Kamp-Becker and S. Bölte, *Autismus*, 2. Auflage. München Basel: Ernst Reinhardt, 2014.

[2] M. H. Black *et al.*, 'The use of wearable technology to measure and support abilities, disabilities and functional skills in autistic youth: a scoping review', *Scand. J. Child Adolesc. Psychiatry Psychol.*, vol. 8, pp. 48–69, 2020, doi: 10.21307/sjcapp-2020-006.

[3] ISO 9241-210, 'Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems'. 2019.

[4] S. L. Lattacher, 'Participatory Research in the Development of a Smart Home Environment for People on the Autism Spectrum - Barriers, Methods and Opportunities during a Pandemic', presented at the 2nd International Workshop on Indoor Comfort, online, Dec. 09, 2020, Accessed: Mar. 11, 2021. [Online]. Available: https://forschung.fh-kaernten.at/aal/en/senshome/workshop-indoor-comfort/.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210087

Semiautomatic Recruitment of Trial Patients Using ELGA Data: Conceptual Design and Implementation of an IT Tool

Raik MÜLLER [a,1] and Georg DUFTSCHMID [a]
[a] *Section for Medical Information Management, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria*

Abstract. Reuse of EHR data can substantially improve the recruitment process of clinical trials. As shown earlier, Shared EHR systems are particularly attractive data sources. The goal of this work was to conceptually design and implement a user-friendly tool for semiautomatic trial recruitment using ELGA data. The tool applies a web-based client (Vue and Electron frameworks) – server (Django-Python and Java server, SQLite database) architecture. Trial eligibility criteria are expressed as XPaths. Access to ELGA documents is simulated using the eHealth Connector library and the IHE XDS Open eHealth Integration Platform framework. Usability was optimized in expert interviews with investigators of two active trials. First feedback based on synthesized ELGA test data indicates suitability for clinical end users. Further insights are expected from applying the tool to real ELGA data.

Keywords. Electronic Health Records, Clinical Trial Recruitment, Standards

1. Introduction

Randomized controlled trials provide a powerful research design for the evaluation of healthcare interventions and are widely recognized as the gold standard of clinical research [1]. Clinical trials enable to test the effectiveness of medications, medical devices, and health methods. For each trial, participants with a specific profile have to be recruited [2]. The profile is defined by eligibility criteria that may for example refer to health problems, gender, age, or treatment.

Unfortunately, many trials suffer from delays or are not completed at all [3–8]. One reason are the inefficient and time-demanding recruitment methods such as advertising through flyers or brochures. Another disadvantage of traditional methods are the high costs that may arise per patient [8].

Reuse of EHR data can improve trial recruitment [9], [10]. As we have shown, Shared EHR systems represent attractive data sources for this purpose [11], [12]. The Austrian national Shared EHR system ELGA [13] in particular holds structured data for more than 60% of eligibility criteria [11] that are commonly used in clinical trials according to an analysis of the EHR4CR project [14].

[1] Corresponding Author: Raik Müller, Section for Medical Information Management, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria, E-Mail: mail@raik-mueller.com

We are currently working on a project that aims to analyze to what extent trial recruitment at the Medical University of Vienna (MedUni) can benefit from reusing ELGA data. In this course we will examine whether the promising theoretical results of Augustinov and Duftschmid [11] can be confirmed in practice. The goal of the master thesis, which is part of this project and is presented in this paper, is to conceptually design and implement the underlying recruitment tool. Hereby, our primary goal was to achieve high usability by clinical end users.

2. Methods

2.1. Conception phase

Conception of our tool was based on the IEEE software development process [15]. In the analysis phase we aimed to determine the context and requirements of the planned tool [16]. It was carried out through analysis of literature and existing recruitment systems as well as a series of expert interviews with clinical end users. An initial set of candidate requirements derived from literature research and existing systems was represented within mockups of the future user interface, which served as starting points for feedback in the interviews and a following revision of requirements. Revised requirements and correspondingly adapted mockups were then validated in a second round of expert interviews. Particular efforts were made to achieve an intuitive visualization of the results of trial eligibility checks that should allow for easy comprehension by clinical end users.

2.2. System architecture

To prepare for an easy applicability, we chose a web-based client-server system architecture. This should avoid dependencies or restrictions on the end users' hardware solutions. To alleviate reuse of our work by other researchers, widely used open source technologies were selected that should simplify potential future enhancements [17]. The source code including installation description is published on GitHub [18].

The client provides the user interface for selecting the trial and the patients to be checked for eligibility. The analysis of the patients' documents is done by the server and the results are finally displayed by the client again.

For the implementation of the client the progressive JavaScript web framework Vue [19] was used, which allows the creation of single-page web applications. In addition, the Electron Framework was employed, which enables a cross-platform desktop application [20].

In the back-end structure, a Django-Python [21] server acts as the center for all task areas. It holds the REST interface implementation for communication with the client and processes the ELGA documents.

For communication with the IHE XDS environment, a Java server based on the open-source eHealth Connector (EHC) library [22] is used. EHC is based on international implementation guidelines and standards that promote a harmonized exchange of data and documents in the healthcare sector and can be integrated into a back-end. This also makes it possible to validate CDA documents for conformance to the ELGA templates. The Open eHealth Integration Platform framework is used to simulate an IHE XDS environment [23].

An SQLite database [24] is employed for storing the trial-specific metadata including the inclusion and exclusion criteria. The criteria are represented as XPaths that refer to those components of ELGA documents that hold the required data. Hereby, we relied on a mapping of data elements commonly referenced in trial eligibility criteria to ELGA documents [11]. A tool for an interactive specification of trial criteria referring to ELGA document components that can be exported as XPaths was presented in [25]. The XPaths originate from the template IDs of the HL7 V3 templates that define the structure of the respective ELGA document component. This allows accessing any ELGA document types. Currently, ELGA lab reports and medication documents seem particularly promising as most of their components are highly structured.

2.3. Data model of trial criteria

For the representation of trial criteria, a suitable data structure was developed (compare **Figure 1**). It distinguishes between inclusion and exclusion criteria and allows a criterion to be checked by multiple alternative conditions. As an example, the criterion "person has diabetes" may alternatively be checked via conditions "diabetes was diagnosed", "diabetes-specific medication was dispensed", and "diabetes-specific lab parameter exceeds a certain threshold".

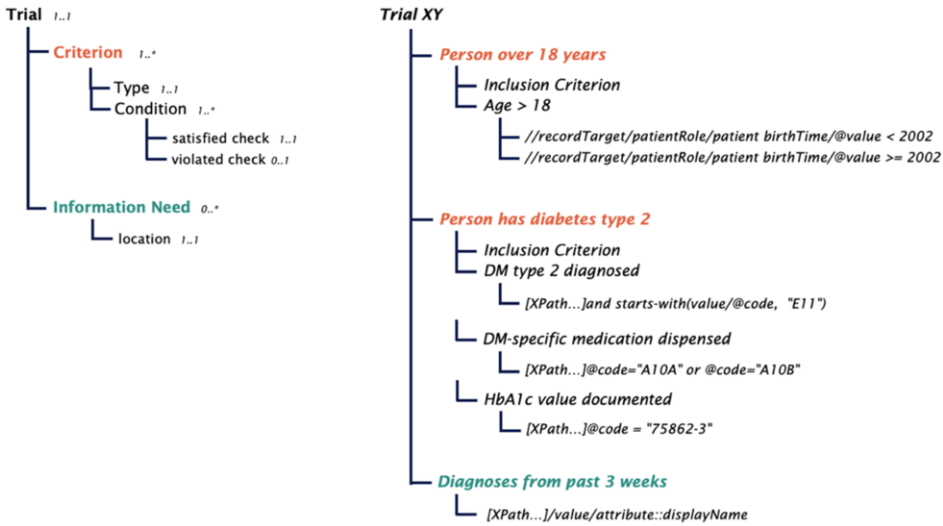


Figure 1 – Data model (left) and example instance (right) of trial criteria

In some cases, a condition may be clearly assessed to be violated based on existing data. As an example, condition "age > 18 years" can obviously be found to be satisfied or violated based on the documented birthdate of a person. Therefore, our data structure also covers checking for a conditions' violation. The latter is, however, optional as in most cases only the satisfaction of a condition will be assessable. As an example, condition "diabetes was diagnosed" cannot be rebutted just because no diabetes-specific diagnosis was recorded [26].

Further, it may sometimes be beneficial to provide a "fuzzy" alternative to strict condition checking. As an example, if condition "clinically-relevant disease in the last two weeks" is checked in a strict manner, a relevant disease recorded 15 days ago would

be ignored, even though it would probably be of interest for the trial investigator. Further, an explicit specification of a condition may sometimes be hard. As an example, a comprehensive formulation of "clinically-relevant disease" would require the listing of numerous ICD codes and would still bear the risk of missing one. Therefore, we included optional "information needs" in our data structure, in addition to criteria. In the present example, this would allow to query all recorded diseases of a person from e.g. the past three weeks and let the trial investigator decide whether they are clinically and temporally relevant.

2.4. Testing methodology

We plan to test our tool with two currently running trials of the MedUni. The trials have already been set up in our database. From those eligibility criteria that can be checked via ELGA data, we have translated about 90% to XPaths yet. We aim to analyze the practical usefulness of our recruitment tool by "re-evaluating" patients, who were already recruited for one of the two trials with conventional methods. These patients will serve as gold standard for eligible individuals. In order to measure the sensitivity of our recruitment tool, we will apply it on the ELGA documents of these patients and examine to what degree it indicates eligibility. As the trials have disjunctive study populations according to their inclusion/exclusion criteria, patients recruited for trial 1 can serve as gold standard for ineligible individuals for trial 2 and vice versa. This will allow us to analyze the specificity of our tool.

Only patients who provide written informed consent will be included in our analysis. Their ELGA documents will be downloaded in the Vienna General Hospital (AKH) information system and pseudonymized before being used in our project. The planned procedure was confirmed by the MedUni data protection officer to comply with the relevant legal regulations and received a positive vote by the MedUni ethics committee.

Currently we are working on the organizational steps of implementing an interface between the AKH hospital information system and a MedUni research platform, where the ELGA documents of trial patients will be stored for our project. As this rather bureaucratic procedure will not be completed within the limited timeframe of the master thesis presented here, we decided to focus in the thesis on the optimization of our recruitment tool's usability using test data.

Originating from the public ELGA test documents we synthesized fictive but valid documents holding data that are relevant for the eligibility criteria of our two trials. We composed our test documents in a way that should exhaust all visualization variants offered by our tool's user interface.

3. Results

As the result of the context analysis, a clear picture of the planned application environment of our recruitment tool within the MedUni clinics was achieved. In particular, the typical current recruitment procedure of trial patients was analyzed to prepare a suitable future integration of our tool.

3.1. Requirements to be covered

As the result of the requirements analysis, the following system functions were found to be essential:

- *Set up trial:* A new trial is set up in the system's database including the required trial metadata and the XPaths for the eligibility criteria and the information needs. This step is done by a technician in cooperation with the corresponding trial investigator.
- *View trial:* The clinical end user can look up all metadata of a selected trial. Further, results of selected patients from earlier eligibility checks for the trial are displayed.
- *Select patients to be checked for eligibility*: The clinical end user selects the patients, who should be checked for trial eligibility. The patients' ELGA documents can either be requested and downloaded from an IHE XDS environment or they can be accessed at a local folder (where they might have been exported from a local EHR system).
- *View summary of check results*: An intuitive overview of the eligibility check's results for all selected patients is presented to the clinical end user.
- *Inspect results in detail:* The results can be examined in detail for each patient. Hereby, a patient's results can be expanded to see which criteria are satisfied / violated or for which criteria no corresponding data are available in the patient's ELGA documents. Each criterion can further be expanded to show the results of the individual underlying conditions. For each satisfied / violated condition, the corresponding ELGA source data can be viewed directly within the embedding ELGA document.
- *Mark potential trial participants:* After going through the results of the eligibility checks, the clinical end user can mark particularly promising patients. They are saved in a shortlist that may be used later for establishing contact with the patients.

3.2. User interface design

The intuitive visualization of the results of an eligibility check proved to be a particular challenge. In particular, a balance had to be found between providing a coarse overview of the check's results for the complete selected patient cohort at a glance, and at the same time allowing a stepwise visual drilldown into each single patient to explore the respective ELGA data constellations that led to the satisfaction / violation of the individual eligibility criteria and the underlying conditions.

Figure 2 shows a screenshot of the results of checking the eligibility criteria of a diabetes-specific trial for a cohort of five fictive patients. In the header section the key trial metadata are displayed together with the trial's total number of eligibility criteria and the number of criteria for which ELGA covers the required data.

Beneath the header, the outcomes of the eligibility checks are separately displayed for inclusion and exclusion criteria. At the coarsest level, only the numbers of satisfied/violated/undecidable (due to missing data) criteria are shown for each patient. Patients are sorted according to the number of satisfied inclusion criteria. If one or more exclusion criteria are satisfied, the corresponding patients are moved to the bottom of the cohort and marked in red background color.

In order to comprehend the results of a particular patient, the corresponding row may be expanded to show the results of the individual criteria. Here, a traffic light color scheme is used to visualize whether a patient "has passed" a criterion (green dot), "has failed" on a criterion (red dot), or whether the criterion is undecidable due to missing data (gray dot). In this sense of "passing" a criterion, a satisfied *inclusion* criterion is shown as green dot, whereas a satisfied *exclusion* criterion is shown as red dot. If contradictory data are found for a patient (e.g., two blood glucose measurements within the period of interest, one above the criterion's threshold and one below), the criterion is displayed as a yellow dot.

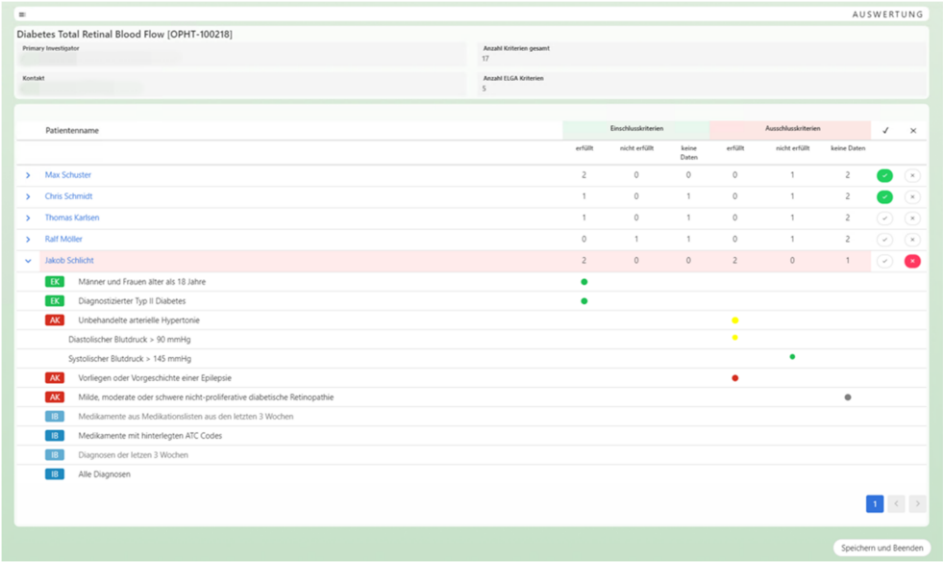


Figure 2 – Results screen of an eligibility check

If data was found for a criterion, the corresponding row may be further expanded to display the underlying conditions. The same color coding is used here to visualize the results. A condition with results can be clicked to open a pop-up window (Figure 3) that holds all source data found for the current patient that are relevant for the condition. A click on a particular value opens the embedding ELGA document and scrolls to the value within the document.

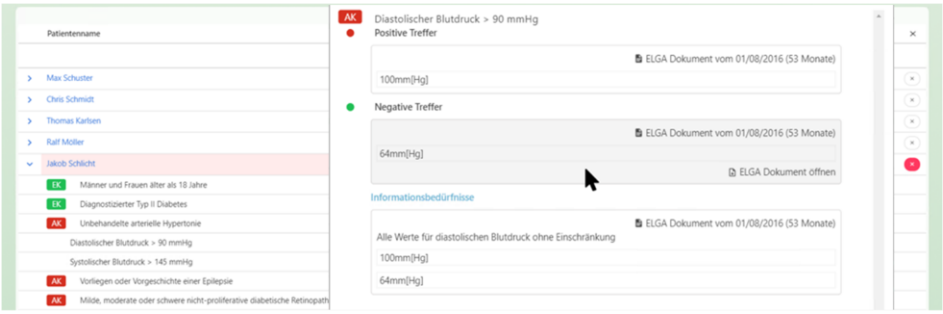


Figure 3 – For each condition the relevant ELGA source data can be retrieved

The checkbox in the right-most column can be used to manually mark those patients who were found to be most promising for a potential recruitment. The complete results of the executed eligibility check from selected patients can be saved with the trial for later processing.

4. Discussion

The present work follows a long tradition of implementing supportive tools for trial management at the Medical University of Vienna [27]. It extends these and other comparable activities for EHR-based trial recruitment [9], [28] by utilizing data of a national Shared EHR system.

We apply a rule-based approach for representing eligibility criteria and identifying suitable patients for a trial. While this is a rather straight-forward approach and knowledge bases exist that alleviate implementing these rules [29], achieving high sensitivity and specificity rates may require time-intensive fine-tuning. An alternative could be the identification of eligible patients by means of machine-learning approaches [30].

Our work is subject to several limitations. A system-immanent restriction is the fact that data available within ELGA will typically only allow a subset of a trial's eligibility criteria to be checked. However, even such kind of pre-filtering would allow the trial investigators to focus on the most promising candidates and thus in many cases entail a significant reduction of efforts in the recruitment process. We currently do not support natural language processing to locate data relevant for a trial's eligibility criteria within unstructured sections of ELGA documents.

The final goal of our project is to examine to what extent we can support patient recruitment for clinical trials at the MedUni by means of ELGA data. The master thesis presented here delivers an essential building block in this endeavor by developing a conceptual design and implementation of the underlying IT tool. The primary focus hereby was to achieve a high level of usability for clinical end users. Using the synthesized test documents, our tool was presented to the primary investigators of the two trials to gather feedback on its usability. Their preliminary feedback seems to indicate that we are on the right track in this regard.

References

[1] E. Hariton and J. J. Locascio, 'Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials', *BJOG: Int J Obstet Gy*, vol. 125, no. 13, pp. 1716–1716, Dec. 2018, doi: 10.1111/1471-0528.15199.

[2] C. A. Umscheid, D. J. Margolis, and C. E. Grossman, 'Key Concepts of Clinical Trials: A Narrative Review', *Postgrad Med*, vol. 123, no. 5, pp. 194–204, Sep. 2011, doi: 10.3810/pgm.2011.09.2475.

[3] G. Kolata, 'A Cancer Conundrum: Too Many Drug Trials, Too Few Patients (Published 2017)', *The New York Times*, Aug. 12, 2017.

[4] C. Halter, 'Liver Disease Researchers Struggle to Find Patients for Clinical Trials', *Hep*, Jan. 04, 2018. https://www.hepmag.com/article/liver-disease-researchers-struggle-find-patients-clinical-trials (accessed Oct. 14, 2020).

[5] K. Lopienski, 'Why do Recruitment Efforts Fail to Enroll Enough Participants?', *Forte*, Nov. 07, 2017. https://forteresearch.com/news/recruitment-efforts-fail-enroll-enough-patients/ (accessed Oct. 14, 2020).

[6] S. Feller, 'One in four cancer trials fails to enroll enough participants', *UPI*, Dec. 30, 2015. https://www.upi.com/Health_News/2015/12/30/One-in-four-cancer-trials-fails-to-enroll-enough-participants/2611451485504/ (accessed Feb. 21, 2020).

[7] A. M. McDonald *et al.*, 'What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies', *Trials*, vol. 7, no. 1, p. 9, Dec. 2006, doi: 10.1186/1745-6215-7-9.

[8] J. D. McAnulty, 'Bringing Patient Recruitment into Our Digital World', *Ther Innov Regul Sci*, vol. 43, no. 4, pp. 501–508, Jul. 2009, doi: 10.1177/009286150904300416.

[9] Y. S. Lai and J. D. Afseth, 'A review of the impact of utilising electronic medical records for clinical research recruitment', *Clinical Trials*, vol. 16, no. 2, pp. 194–203, Apr. 2019, doi: 10.1177/1740774519829709.

[10] S. Hussain, D. Ouagne, E. Sadou, T. Dart, and M.-C. Jaulent, 'EHR4CR: A semantic web based interoperability approach for reusing electronic healthcare records in protocol feasibility studies', *CEUR Workshop Proceedings*, p. 6, Jan. 2012.

[11] G. Augustinov and G. Duftschmid, 'Can the Austrian Nation-Wide EHR System Support the Recruitment of Trial Patients?', *Studies in Health Technology and Informatics*, pp. 87–90, 2019, doi: 10.3233/978-1-61499-961-4-87.

[12] B. Kankova and G. Duftschmid, 'Reusing Data of an EU-Wide EHR System for Clinical Trials: A Capabilities Analysis', *Stud Health Technol Inform*, vol. 271, pp. 17–22, Jun. 2020, doi: 10.3233/SHTI200069.

[13] S. Herbek *et al.*, 'The Electronic Health Record in Austria: a strong network between health care and patients', *Eur Surg*, vol. 44, no. 3, pp. 155–163, Jun. 2012, doi: 10.1007/s10353-012-0092-9.

[14] J. Doods, F. Botteri, M. Dugas, and F. Fritz, 'A European inventory of common electronic health record data elements for clinical trial feasibility', *Trials*, vol. 15, no. 1, p. 18, 2014, doi: 10.1186/1745-6215-15-18.

[15] A. Abran and J. W. Moore, *Guide to the Software Engineering Body of Knowledge*. IEEE Computer Society, 2004.

[16] U. Hammerschall and G. H. Beneken, *Software Requirements*. Pearson, 2013.

[17] 'Stack Overflow Developer Survey 2020', *Stack Overflow*. https://insights.stackoverflow.com/survey/2020/?utm_source=social-share&utm_medium=social&utm_campaign=dev-survey-2020 (accessed Jan. 22, 2021).

[18] R. Müller, *RecruitmentTool*. 2021.

[19] 'Vue.js'. https://vuejs.org/ (accessed Jan. 21, 2021).

[20] 'Electron'. https://www.electronjs.org/ (accessed Nov. 03, 2020).

[21] Django Software Foundation, 'The Web framework for perfectionists with deadlines | Django'. https://www.djangoproject.com/ (accessed Jan. 21, 2021).

[22] medshare GmbH, 'eHealth Connector Konzept zur Implementierung und API Spezifikation', Nov. 2013. [Online]. Available: https://www.yumpu.com/de/document/view/51786745/konzept-zur-implementierung-und-api-spezifikation-medshare/26.

[23] Open eHealth Foundation, 'IPF Open eHealth Integration Platform', *IPF Open eHealth Integration Platform*. https://oehf.github.io/ipf-docs/ (accessed Nov. 03, 2020).

[24] 'SQLite Home Page'. https://www.sqlite.org/index.html (accessed Jan. 21, 2021).

[25] S. Ott, C. Rinner, and G. Duftschmid, 'Expressing Patient Selection Criteria Based on HL7 V3 Templates Within the Open- Source Tool ART-DECOR', *Stud Health Technol Inform*, pp. 226–233, 2019, doi: 10.3233/978-1-61499-971-3-226.

[26] O. Dameron, P. Besana, O. Zekri, A. Bourdé, A. Burgun, and M. Cuggia, 'OWL model of clinical trial eligibility criteria compatible with partially-known information', *J Biomed Sem*, vol. 4, no. 1, p. 17, 2013, doi: 10.1186/2041-1480-4-17.

[27] G. Duftschmid, W. Gall, E. Eigenbauer, and W. Dorda, 'Management of data from clinical trials using the ArchiMed system', *Med Inform Internet Med*, vol. 27, no. 2, pp. 85–98, Jun. 2002, doi: 10.1080/1463923021000014158.

[28] M. Cuggia, P. Besana, and D. Glasspool, 'Comparing semi-automatic systems for recruitment of patients to clinical trials', *International Journal of Medical Informatics*, vol. 80, no. 6, pp. 371–388, Jun. 2011.

[29] 'PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability - PubMed'. https://pubmed.ncbi.nlm.nih.gov/27026615/ (accessed Jan. 22, 2021).

[30] J. M. Banda, Y. Halpern, D. Sontag, and N. H. Shah, 'Electronic phenotyping with APHRODITE and the Observational Health Sciences and Informatics (OHDSI) data network', *AMIA Jt Summits Transl Sci Proc*, vol. 2017, pp. 48–57, 2017.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210088

A Design Method of Tele-Rehabilitation Platforms for Post-Stroke Patients Based on Consumer Technology

Alberto ARRIGONI[a], Michele CUMETTI[a], Andrea GRECO[a], Luca SOLIVERI[a] and Andrea VITALI[a,1]
[a]*University of Bergamo, Bergamo, Italy*

Abstract. *Background*: Telerehabilitation represents a new cutting-edge method in the treatment of patients suffering from motor and cognitive disorders caused by stroke. Even if there exist dedicated devices able to track patients' movements to evaluate the performed rehabilitation exercises, they require specific settings necessary for a correct and simple use at the patient's home. If we consider the recent pandemic situation and the lockdown condition, which made difficult the access to these products, post stroke patients may be not able to perform home rehabilitation. *Objectives*: the goal of this work is the design of a specific method to develop a tele-rehabilitation platform for post-stroke patients using consumer technologies without involving ad-hoc devices. *Method*: Open-source tools have been investigated for speeding up the development starting with the medical knowledge. *Results:* a group of four healthcare technologies engineering students with no specific skills about computer science has developed a platform in four months using the design method. *Conclusion*: the presented method allowed the development of a clinical knowledge-based web platform for post-stroke patients totally based on consumer technology.

Keywords. Telerehabilitation, Stroke rehabilitation, Virtual Reality Exposure Therapy, Rehabilitation Research

1. Introduction

A systematic analysis for the Global Burden Disease Study showed that in 2016 there were 116.4 million DALYs due to stroke and 80.1 million prevalent cases of stroke globally including 5.5 million deaths [11]. Upper extremity function and cognitive impairments are common outcomes following stroke events, causing affected people a worsening of their lifestyle and a loss in work opportunities. Despite post-stroke rehabilitation therapy being generally provided in clinical facilities, many patients do not draw the maximal benefit from it. The major problems that come across during the traditional rehabilitation path are related to access issues, costs and scarce adherence.

Tele-rehabilitation-based programs may throw down the cited barriers. This form of telemedicine allows delivering remote rehabilitation services directly to the patient's home. Technology can overcome access obstacles to quality care, including the burden to travel for in-patient therapy, along with the consequent costs, scheduling demand issues and shortage of trained therapy providers. Moreover, tele-rehabilitation systems

[1] Corresponding Author: Andrea Vitali, University of Bergamo, Italy, E-Mail: andrea.vitali1@unibg.it

have proven to enhance post-stroke rehabilitation's outcomes by promoting adherence to rehabilitation therapies, boosting the brain plasticity [6]. Besides, a remotely provided rehabilitation service may positively affect the patient selection in terms of clinics, enlarging the choice opportunities by making available care providers which are located far from home. Another technological solution, that fits well along with tele-rehabilitation and may increase patient engagement during therapies, is the deployment of serious games. These are games with non-recreational intent but designed and developed pursuing a specific goal, typically linked to educational or training purposes under the scope of many contexts. These include the rehabilitation background. The key concept behind this kind of game is to provide a valuable, fun and engaging rehabilitation path. Despite all the listed advantages, virtual-aided rehabilitation cannot be considered the first therapy choice yet due to the still scarce availability of low-cost technologies and the low confidence of the target population, largely represented by elders, when it comes to manage technological solutions. Although the Human-Computer Interaction (HCI) research advances in interactive computer systems, usability barriers still obstruct access to this kind of solutions, discouraging its diffusion among users [15].

The coexistence of these pros and cons makes it interesting to design a methodological approach for the development of tele-rehabilitation platforms for taking advantage of the benefits related to the remote modality while avoiding the drawbacks. This point means achieving a system able to provide patients with efficient and stimulating treatments, whose progress can be monitored remotely by clinicians. The goal is the investigation of methods and tools to develop a platform patient and medical centered by exploiting low-cost and user-friendly technologies.

2. Scientific Background

While tele-rehabilitation proved to be a relevant resource in the clinical field, several research projects started in this regard. As shown in the study of Dodakian et al. [4], a platform for remote rehabilitation can provide patients a wide range of features besides the rehabilitative service, specifically sections dedicated to clinical education and doctor-patient interaction.

Regarding the design of serious games, to increase engagement can be useful the compliance with some key principles already examined by companies of the videogame market. These concepts transposed to the rehabilitation field represent a valuable tool to make the treatment successful in avoiding scarce adherence. The mentioned principles, indicated by Barret et al. [1] and Burke et al. [3], are the following: (i) a serious game must be meaningful to obtain clinical evidence on the patient condition from each exercise; (ii) the rehabilitation exercises need to be customizable in order to match the ability of each patient and thus, the level of difficulty must be personalized according to the rehabilitation progress; (iii) the game must be able to show the improvements with quantifiable feedback to make patients realize the value of their efforts.

To achieve medical relevance, the games must examine parameters useful to understand the rehabilitation progress during the time permitting to adjust the program as required by the specific case. Examples of these factors can be time, score, gameplay speed, the eventual number of obstacles avoided successfully in a platform game and the trajectories followed by the limb during the exercise. The challenging aspect must be correctly balanced to determine the success of the therapy: an exercise with a biased level of complexity can lead to discontent both due to an excessive simplicity or difficulty.

Boredom or frustration are harmful consequences to the observance of the training, which is not something trivial. As pointed out by Goršič et al. [8], people with chronic cognitive or arm impairment should exercise intensely to regain their abilities, but frequently a lack of motivation leads to poor rehabilitation outcome. This issue does not concern only the traditional rehabilitation approach as a recent study showed that, even if a therapist prescribes a technology-supported exercise, only about 30% of unsupervised patients will comply with the rehabilitation regimen [7]. An additional promising way to increase motivation may be to include a social component within the rehabilitation games, allowing patients to play together with other people. This implementation could have a positive impact in the acceptance and enjoyment of the therapy, leading to long term achievements. Tele-rehabilitation solutions may vary greatly from employing commercial gaming consoles to rehabilitation-specific products up to expensive Virtual reality customized programs. The survey study of Hung et al. [10] underscores a positive trend of game-based rehabilitation systems using consoles such as tablets, Kinects, and Wii remotes. The same trend was detected by Laver et al. [13] in their 2017 update of a Cochrane Review about the virtual reality in stroke rehabilitation, where the number of studies using commercially available gaming consoles increased from 6 to 22 since 2015.

On some occasions, also the webcam device embedded in each PCs has been exploited, and this is the case of Burke et al. [2], who developed a series of games based on color tracking algorithms working on webcam footages. The spread of smartphones within everyone's reach is also something to consider for preparing solutions that take advantage of mobile apps. Lee Hughes and colleagues [9] evaluated which features are worth to be included in a mobile product designed as a platform for tele-rehabilitation. According to them, this app should include a section dedicated to the patient and one to the doctor. The latter must be designed to include the privilege of modifying the treatment plan available on the patient's interface, as well as the possibility to check the training performance and the medical information which need to be kept into account, like medications and test results. A section for doctor-patient communication is also functional for efficient therapy, such as a calendar that displays the periodic appointments with the care provider team.

Considering the characteristics of the context and the scientific background, the aim of this work has been the definition of a specific method for developing an innovative multi-featured platform for cognitive and motor telerehabilitation intended for post-stroke patients.

3. Method and tools

Two target features have driven the requirement analysis to make the method pioneering in exploring alternative approaches than the current available solutions, also providing new advantages [12], i.e.:

- the use of advanced tools for facilitating and speeding up the development,
- the exclusive deployment of consumer technologies which are usually present in people's homes, without involving any investment or hardship related to the supply of the necessary equipment (e.g., the use of a Microsoft Kinect device for human motion detection). This intention arose from a nowadays striking need if considering the recent pandemic situation and the lockdown condition

which has highlighted sustainability limits by health systems in managing all the care processes without more agile solutions. This intention led the research to the employment of an artificial neural network model for achieving the motion tracking required by the rehabilitation aim, which is an area that needs further continuous explorations, despite being a recognized valuable tool for general tracking in computer vision [14].

Specifically, the realized method considers two mobile applications, addressing the doctor and the patient, one game for cognitive rehabilitation and one for the upper limb motor rehabilitation based on a hand-tracking interaction style. All the modules communicate through internet connection with a server database. By starting from the target features previously described, the first step has concerned the analysis and selection of technological tools required for respecting all the features required for the realization of the technological platform.

The chosen game editor to develop serious games has been the free Unity 3D because of: (i) no need of high skills in computer graphic development; (ii) multiplatform; (iii) useful add-ons for interfacing hand-tracking devices.

To obtain a deep learning model for the hand-tracking system, the selected tool has been the Python programming language with PyCharm as the development environment.

Android Studio has been chosen for building the mobile applications. The communication among all the applications has been implemented with MySQL DBMS and LoopBack [16]. The former allows the design of the database; the latter provides automatic REST API generation for the server.

According to the selected technological tools, a basic software architecture has been designed and depicted in Figure 1. The whole web platform is based on the standard HTTP protocol for the data exchanging.

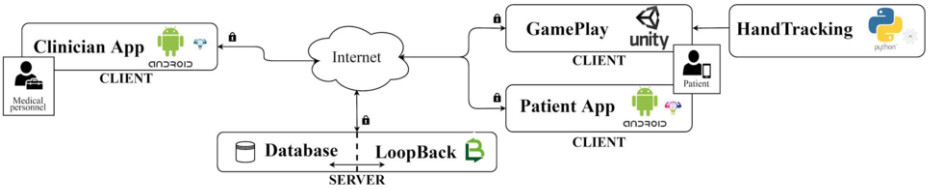


Figure 1. Basic software architecture of a tele-rehabilitation platform for post-stroke patients.

4. Discussion

By starting from the proposed method, a platform has been realized to investigate the potentialities of the proposed approach.

4.1. Development of serious games

The developed serious games are characterized by different functionalities such as different game modalities, distinct levels of difficulties, time and error-based scores, besides a simple interface featured by sounds and instructions. Two different types of games have been designed based on functional and motivational parameters:

- Cognitive rehabilitation: the game is focused on stimulating the patient's memory, inducing him to remember the exact sequence of some cooking recipes (Figure 2a). Exploiting the mouse interface interaction, is asked to the patient to remember the correct sequence of ingredients to make a specific dish, projecting the patient in a familiar home-looking virtual environment.
- Motor-skills rehabilitation: stimulating the patient to recover arms' functionalities. The webcam's footage processing algorithm automatically recognizes the hand of the patient and tracks its coordinates to lead the game's character in avoiding obstacles within a funny and stimulating virtual environment designed to enhance the adherence to the therapy (Figure 2b).

(a) (b)

Figure 2. Unity 3D serious games**:** Game's environment for cognitive rehabilitation (a); Gameplay and tracking system which determines the movement of the player for motor-rehabilitation (b).

4.2. Hand tracking by means of artificial intelligence

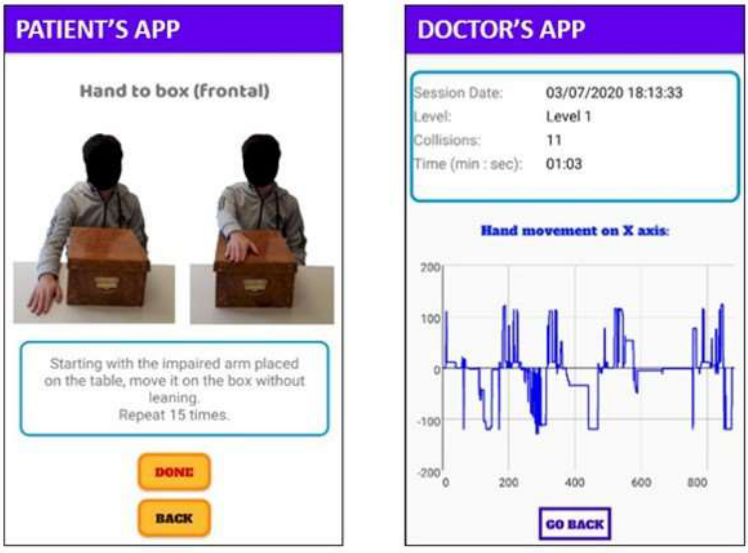
In order to avoid the use of specific hand-tracking devices, the motor rehabilitation game's design concerns the employment of an AI computer vision algorithm. Specifically, it consists of an open-source convolutional neural network based on the SSD MobileNet architecture and, according to the author, trained in cloud on a GPU machine using the Tensorflow API and employing the Egohands Dataset [5], reaching a mean Average Precision (mAP) value of 0.9686 and a total loss of 2.575 [18]. The Python scripts were appropriately modified to fit the needs of the intended application, both in terms of parameters optimization, with the number of target hands set to 1, and the assessment of a balanced certainty threshold, as well as new features implementation. Precisely, the algorithm's outcome was configured to provide, in addition to the bounding boxes drawing, also the visual information of each box centre and the tracking record of its coordinates in order to post them to the game application in real-time. About tracking this point, an interpolation was implemented for shifting the horizontal range of movement from zero up to the frame width value, to a reduced symmetrical interval that allows the movement of the game's character on both the left and right directions with respect to the centerline of the frame.

The obtained motion coordinates were sent to the rehabilitation game through web sockets based on UDP (i.e., the standard User Datagram Protocol). The optimized and adapted algorithm was then converted in a Windows executable application to make it easily accessible without installing any resources or possessing any skills. The next step was to integrate the hand tracking data to control a character moving continuously forward on a path with obstacles. The idea pursued was to convert the horizontal

coordinates information in a lateral thrust force operating on the player, which behaves like a real rigid body. This feature provided a further measure to raise the challenge of the game in terms of entertainment and competition by involving the practice of a motor-cognitive skill helpful for rehabilitation. A Unity3D script has been developed for using the UDP Socket communication to receive the position data to be exploited by the character's movement scripts in generating the motion force during the gameplay. The tracked movement is automatically saved on the MySQL database using the REST API designed with Loopback.

4.3. Development of the mobile applications

The mobile application for patients contains different features to carry out the telerehabilitation in an engaging environment: background music, button's sound and an immediate user interface. The app contains a Memory cards game developed for cognitive rehabilitation where the player, turning two cards at a time, must find the pairs with the same front image. A section with motor-rehabilitation training exercises (Figure 3a) for the improvement of daily life tasks is also available. The patient's app also offers a multi-choice quiz with a stroke-educational goal. Finally, a personal agenda shows the dates of medical examinations and the personalized rehabilitation program.



(a) (b)

Figure 3. Mobile applications' features: Training motor-rehabilitation based on the Wolf Motor Function Test (a); rehabilitation results section on the doctor's application (b).

The doctors' mobile app allows the patients' data evaluation. The physician accesses a selection list of activities that can be associated to each specific patient. The application offers a section containing the progress of the motor and cognitive rehabilitation path and the graphical representation of the coordinates, showing the movement of the patient's hand, detected by the hand tracking algorithm (Figure 3b). The doctor visualizes the patient's data and his/her improvements useful to modify the rehabilitation program. Both the mobile apps allow patients and doctors to contact each other by e-mails. All data and information are saved or modified on the MySQL database using the REST API designed with Loopback.

5. Discussion

The presented method has been designed to develop a platform for post stroke tele-rehabilitation by four students attending a master's degree course focused on medical technologies and their management. The limit of time imposed by the project deadline, as well as the complexity of developing such a product starting from a limited coding experience, were successfully addressed by the designed method and by using the proposed development tools. An initially planned approach was based on the use of a Leap Motion device [17] for realizing the motion tracking required by the motor-rehabilitation game but considering the complications that arose with the COVID-19 pandemic outbreak and the resulting impossibility to get the needed equipment during the lockdown, an innovative way has been pursued based on the use of artificial intelligence as suggested by the design method presented. The hand-tracking interaction style has been developed by using a convolutional neural network model working on a video stream. The whole platform for the rehabilitation of post-stroke patients has been developed in approximately 4 months.

The simple hand tracking algorithm implemented has limits compared to other technologies that are very precise and allow to obtain several data, however, there already exist open-source models much more sophisticated than the adopted one which enables to obtain further and useful information on the movement performance, such as pose estimators. The whole platform is based only on consumer technology: the patient can use her/his laptop to play the serious games by means of the traditional webcam to perform the hand-tracking. Therefore, no installation of specific devices is needed at the patient's home.

Specific improvements have been planned to add the cybersecurity requirement (i.e., HTTPS protocol) for data assurance and start a clinical assessment of the tele-rehabilitation platform by involving post-stroke patients under the supervision of medical experts. At present, the project has not yet been able to collect the opinion of any care provider institutions nor find their involvement due to the particular situation, which has shifted the focus to the emergency management.

6. Conclusions

The implemented method constitutes an innovative proposal for the design of a multi-functional system aimed at the telerehabilitation of post-stroke patients. The proposed approach shows how the choice of specific open-source development tools can facilitate the rapid and successful realization of these products. Regarding the utilization of serious games in motor rehabilitation, the proposed method explores a promising new strategy to track upper limbs movement given the advantage of not requiring any additional equipment other than laptops, smartphones and tablets with an embedded webcam. Furthermore, the method considers the continuous advancement of artificial intelligence models whose real-time deployment is becoming more and more accessible thanks to the improvement of the IT computing components. Moreover, the promising value of the suggested approach is clear considering the current spreading of cloud-based gaming systems which would enable in the future the easy access to powerful hardware and therefore the implementation of even more sophisticated tracking algorithms.

References

[1] N. Barrett, I. Swain, C. Gatzidis, and C. Mecheraoui. The use and effect of video game design theory in the creation of game-based systems for upper limb stroke rehabilitation. *Journal of Rehabilitation and Assistive Technologies Engineering*, 3: 2055668316643644, 2016.

[2] J. W. Burke, M. McNeill, D. Charles, P. Morrow, J. Crosbie, and S. McDonough. Serious games for upper limb rehabilitation following stroke. In *2009 Conference in Games and Virtual Worlds for Serious Applications*, pages 103–110. IEEE, 2009.

[3] J. W. Burke, M. McNeill, D. K. Charles, P. J. Morrow, J. H. Crosbie, and S. M. McDonough. Optimising engagement for stroke rehabilitation using serious games. *The Visual Computer*, 25(12):1085–1099, 2009.

[4] L. Dodakian, A. L. McKenzie, V. Le, J. See, K. Pearson-Fuhrhop, E. Burke Quinlan, R. J. Zhou, R. Augsberger, X. A. Tran, N. Friedman, et al. A home-based telerehabilitation program for patients with stroke. *Neurorehabilitation and neural repair*, 31(10-11):923–933, 2017.

[5] Egohands Dataset. URL http://vision.soic.indiana.edu/projects/egohands

[6] L. V. Gauthier, C. Kane, A. Borstad, N. Strahl, G. Uswatte, E. Taub, D. Morris, A. Hall, M. Arakelian, and V. Mark. Video game rehabilitation for outpatient stroke (vigorous): protocol for a multi-center comparative effectiveness trial of in-home gamified constraint-induced movement therapy for rehabilitation of chronic upper extremity hemiparesis. *BMC neurology*, 17(1):1–18, 2017.

[7] M. Goršič, I. Cikajlo, N. Goljar, and D. Novak. A multisession evaluation of an adaptive competitive arm rehabilitation game. *Journal of neuroengineering and rehabilitation*, 14(1):1–15, 2017.

[8] M. Goršič, I. Cikajlo, and D. Novak. Competitive and cooperative arm rehabilitation games played by a patient and unimpaired person: effects on motivation and exercise intensity. *Journal of neuroengineering and rehabilitation*, 14(1):1–18, 2017.

[9] C. M. L. Hughes, A. Padilla, A. Hintze, T. M. Raymundo, M. Sera, S. Weidner, J. Ontiveros, T. Peng, A. Encarcion, Z. A. Cruz, et al. Developing an mhealth app for post-stroke upper limb rehabilitation: Feedback from us and ethiopian rehabilitation clinicians. *Health informatics journal*, 26(2):1104–1117, 2020.

[10] Y.-X. Hung, P.-C. Huang, K.-T. Chen, and W.-C. Chu. What do stroke patients look for in game-based rehabilitation: a survey study. *Medicine*, 95(11), 2016.

[11] C. O. Johnson, M. Nguyen, G. A. Roth, E. Nichols, T. Alam, D. Abate, F. Abd-Allah, A. Abdelalim, H. N. Abraha, N. M. Abu-Rmeileh, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Neurology*, 18(5):439–458, 2019.

[12] V. Kumar. 101 design methods: A structured approach for driving innovation in your organization. *John Wiley & Sons*, 2012.

[13] K. E. Laver, B. Lange, S. George, J. E. Deutsch, G. Saposnik, and M. Crotty. Virtual reality for stroke rehabilitation. *Cochrane database of systematic reviews*, (11), 2017.

[14] P. Li, D. Wang, L. Wang, and H. Lu. Deep visual tracking: Review and experimental comparison. *Pattern Recognition*, 76: 323–338, 2018.

[15] H. Rex Hartson. Human–computer interaction: Interdisciplinary roots and trends. *Journal of Systems and Software*, 43(2): 103-118, 1998.

[16] StrongLoop. Loopback 4. URL https://loopback.io/.

[17] Ultraleap. Leap motion device. URL https://www.ultraleap.com/.

[18] D. Victor. Handtrack: A library for prototyping real-time hand tracking interfaces using convolutional neural networks. *GitHub repository*, 2017.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210089

Unsupervised Learning for Hydrogen Breath Tests

Michael NETZER[a,1], Friedrich HANSER[a], Maximilian LEDOCHOWSKI[b] and Daniel BAUMGARTEN[a]
[a] *Institute of Electrical and Biomedical Engineering, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria*
[b] *Akademie für Ernährungsmedizin, Innsbruck, Austria*

Abstract. Hydrogen breath tests are a well-established method to help diagnose functional intestinal disorders such as carbohydrate malabsorption or small intestinal bacterial overgrowth. In this work we apply unsupervised machine learning techniques to analyze hydrogen breath test datasets. We propose a method that uses 26 internal cluster validation measures to determine a suitable number of clusters. In an induced external validation step we use a predefined categorization proposed by a medical expert. The results indicate that the majority of the considered internal validation indexes was not able to produce a reasonable clustering. Considering a predefined categorization performed by a medical expert, a novel shape-based method obtained the highest external validation measure in terms of adjusted rand index. The predefined clusterings constitute the basis of a supervised machine learning step that is part of our ongoing research.

Keywords. Unsupervised Learning, Clustering, Time Series, Hydrogen Breath Tests, Carbohydrate Malabsorption

1. Introduction

The disaccharide lactose and the monosaccharide fructose play an important role in our diet. Lactose is an integral part of most dairy products whereas fructose can be directly linked to highly processed foods containing sugars in large quantities. Physiologically, lactose is broken down by the enzyme lactase into its monosaccharide components glucose and galactose which are then absorbed by the brush border of the epithelial cells lining the small intestine. About 65% of humans, however, downregulate the production of lactase after weaning resulting in lactose malabsorption [1]. In contrast, fructose is directly transported across intestinal epithelial cells by the Glucose Transporter 5 (GLUT5). This transport process is mediated via facilitated diffusion and its transport capacity is limited. On entering intestinal epithelial cells, fructose is converted to glucose to a high degree [2]. Excess fructose enters the liver via the portal vein system, possibly inducing nonalcoholic fatty liver disease. Not surprisingly, 34% of humans develop fructose malabsorption but the underlying detailed mechanisms are still elusive [3]. In general, carbohydrate malabsorptions lead to intestinal fluid retention, causing diarrhea

[1] Corresponding Author: Michael Netzer, Institute of Electrical and Biomedical Engineering, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, E-Mail: michael.netzer@umit.at

and abdominal distention [4]. Carbohydrate malabsorption (CM) is often accompanied by Small Intestinal Bacterial Overgrowth (SIBO), which is defined as an increase in the number and/or alteration in the type of bacteria in the small intestine [5]. Hydrogen breath tests are a well-established diagnostic tool to identify pathophysiology of functional gastrointestinal disorders such as CM and SIBO [6].

In this work, we analyze hydrogen breath tests resulting from lactose and fructose tolerance tests using unsupervised machine learning approaches with the aim of identifying new characteristic patterns in H_2 time series data. In general, unsupervised methods search patterns without requiring a label. Clustering methods are one popular representative of this category. Objects within one cluster have, in general, a higher similarity compared to objects of another cluster. The number of clusters can be either specified by the scientist or determined by evaluating different cluster metrics. Recently, Rubio-Escudero et al. [7] applied data mining techniques to lactose-based hydrogen breath test data including clustering based on *k*-means using Euclidean distance. The authors used three cluster indexes (i.e., Silhoutte, Davis-Bouldin, and Dunn Index) to obtain a suitable number of clusters. In our work, we compare different clustering method using a total of 26 clustering indexes for internal validation. In a next step, we perform external validation using a predefined categorization scheme. To our best knowledge, the combined application of unsupervised learning methods using internal and external validation measures to results of lactose and fructose hydrogen breath tests has not been described previously.

2. Methods

2.1. Dataset

We included 4302 breath gas samples in our study, where 1879 samples result from a lactose tolerance test (LTT) and 2423 samples from a fructose tolerance test (FTT). Each trajectory (i.e., time series) includes five H_2 measurements taken in intervals of 30 minutes. Tab. 1 summarizes the experimental details. For a detailed description of the implementation of the hydrogen breath tests the authors refer to [8]. The persons included in this study gave written informed consent. The consecutive statistical analysis was performed using anonymized data.

Table 1. Description of the considered datasets including test substance, dosage, readings and class distribution of the subgroups.

Dataset	Test substance	Dosage in 250 ml water	Readings (min)
LTT	Lactose	50 g	0-30-60-90-120
FTT	Fructose	25 g	0-30-60-90-120

2.2. Unsupervised Learning Approach

We use partitional, hierarchical and a relatively new category of shape-based clustering methods to identify characteristic patterns in the H_2 time series data resulting from LTT and FTT.

2.2.1. Partitional Methods

Partitional clustering methods separate objects into a set of disjoint clusters [11]. The most popular method of this category is the *k*-means algorithm, which is employed in this work. Each object is assigned to its nearest cluster center according to the Euclidean distance between the two [12]. The parameter *k* represents the number of clusters and is a user defined parameter. A more robust version uses centric objects in the cluster as representatives.

2.2.2. Hierarchical Methods

Hierarchical methods identify clusters by merging (agglomerative) or splitting (divisive) objects into clusters [13]. Finally, a dendrogram indicates the resulting cluster hierarchies. The similarity between objects is quantified by using distance measures such as Euclidean distance. Methods to measure the distance between clusters (each consisting of multiple objects) include single link (minimum distance) or complete link (maximum distance between objects of each cluster). We use the more advanced Wards method that is based on a classical sum-of-squares criterion [14].

2.2.3. Clustering According to Shape

Transferring classical clustering approaches to longitudinal data, the majority of methods consider similarities at local time points. In contrast, Genolini et al. propose a method to group trajectories based on similarities in shapes [10]. In particular, the authors use a generalized distance of Fréchet and curve alignment to quantify similarity of trajectories. In our work we use the R implementation of this method (package *kmlShape*).

2.3. Cluster Evaluation Approach

2.3.1. Internal Validation Measures

Internal validation measures evaluate the quality of clusters based on measures such intracluster compactness and isolation, geometric or statistical properties, number of data objects and dissimilarity or similarity measurements [15]. Thus, these measures do not need external information. Popular representatives include the Silhouette [16] index that uses comparison of cluster tightness and separation. Davies-Bouldin [17] and Dunn index [18] use intercluster (i.e., distances between clusters) and intracluster (i.e., distances within a cluster) metrics to evaluate a clustering. Internal validation measures can be used to determine a suitable number of clusters. Charrad et al. propose a method that uses different indexes and determines the final number of clusters by majority voting [15]. In particular, the method includes the following steps:

1. Calculate $k \in K$ clusters where $k = 2, 3, ..., k_{max}$ using a particular clustering method $m \in M$
2. Calculate $v \in V$ internal validation measures for each k
3. Generate a recommendation of relevant number of clusters $r \in R$ for each $v \in V$. Depending on the methodology, select r with the smallest or highest index value.
4. Summarize all relevant number of clusters by using majority voting

As an example, let the relevant number of clusters suggested by Silhouette and Dunn index (both use a number of clusters where index reaches a maximum) be 3 and the relevant number of clusters suggested by Davis-Bouldin (relevant number of clusters is where index reaches a minimum) be 4. In this case, the final number of clusters obtained by majority voting would be 3. In our analysis, we use *k*-means and hierarchical clustering as clustering methods (M). The maximum number of clusters (k_{max}) was set to 8. A total of 26 clustering indexes were used (see [15] for further information).

2.3.2. External Validation Measures

External validation measures require external information. The Rand index (RI) and the adjusted Rand index (ARI) are two common measures to quantify how close clusters are to predefined reference partitions [18]. Our reference partition was defined by a medical expert. Formally, for a 2 x 2 contingency table (see Table 2), the adjusted rand index (ARI) is computed by Equation 1 [19].

Table 2. A 2 x 2 contingency table comparing partitions U and V.

Partition	**V**	
U	**Pair in same group**	**Pair in different group**
Pair in same group	a	b
Pair in different groups	c	d

$$ARI = \frac{\binom{n}{2(a+d)} - [(a+b)(a+c) + (c+d)(b+d)]}{\binom{n}{2^2} - [(a+b)(a+c) + (c+d)(b+d)]} \quad (1)$$

3. Results

3.1. Internal Cluster Validation

Fig. 1 depicts the frequencies among all 26 considered indices using lactose and fructose tolerance test results and using *k*-means and hierarchical as clustering methods. A number of 2 relevant clusters was proposed using the lactose dataset and the k-means clustering method (a majority of 7 indexes suggested this value). Using hierarchical clustering, the relevant number of clusters was 2 and three (each suggested by 7 indexes). The number of relevant clusters in the fructose dataset was 2, as well, suggested by a majority of 10 indexes using k-means and 8 using hierarchical clustering. Fig. 2 visualizes the corresponding mean trajectories for the datasets LTT and FTT. We observed no substantial difference between k-means and hierarchical clustering setting the number of clusters $k = 2$. The mean trajectory of the first LTT cluster describes a flat pattern increasing at 120 minutes. The mean trajectory of the second cluster shows a strong H_2 increase with a maximum at 120 minutes. The mean trajectory of the first FTT cluster shows a small increase with a maximum ppm value at 90 min. The mean trajectory of the second cluster shows again a strong increase with a maximum value at 90 minutes.

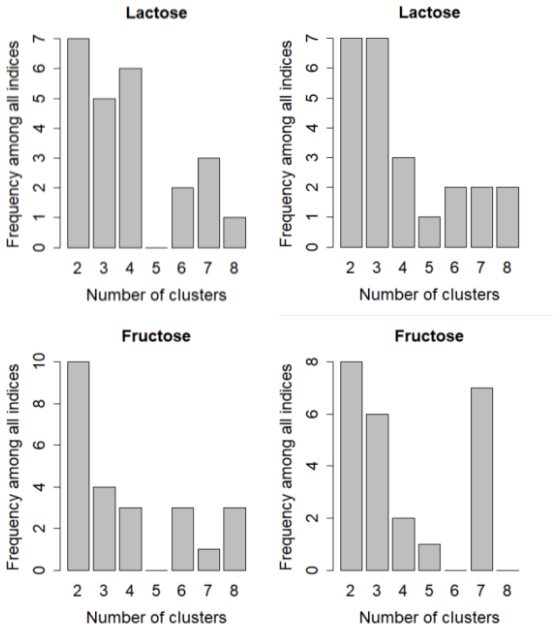


Figure 1. Frequencies among all indices using lactose tolerance test (top row) and fructose tolerance test (bottom row) using *k*-means (left column) and hierarchical (right column) as clustering methods.

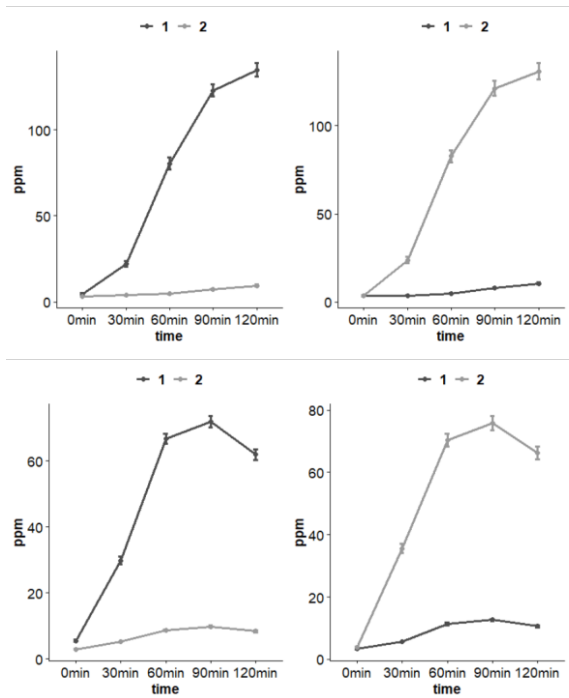


Figure 2. Mean trajectories for the resulting suggested number of clusters (n=2) using lactose tolerance test (top row) and fructose tolerance test (bottom row) using *k*-means (left column) and hierarchical (right column) as clustering methods.

Table 3. Proportions of groups NEG, POS and SIBOS for the two considered datasets (LTT, FTT).

Dataset	NEG	POS	SIBOS
	Subgroup ratios (%)		
TT	75	24	1
FTT	59	27	14

3.2. External Cluster Validation

In contrast to obtaining the number of clusters by using internal validation measures, we used a categorization scheme proposed by a medical expert. The number of manually defined clusters was set to $k = 3$ representing the following groups

1. Negative lactate or fructose test result (LNEG, FNEG)
2. Positive lactate or fructose test result (LPOS, FPOS)
3. Small intestinal bacterial overgrowth syndrome (SIBOS)

The proportion of the groups as categorized by the medical expert is given in Tab. 3. The corresponding mean trajectories for dataset LTT and FTT are depicted in Fig. 3.

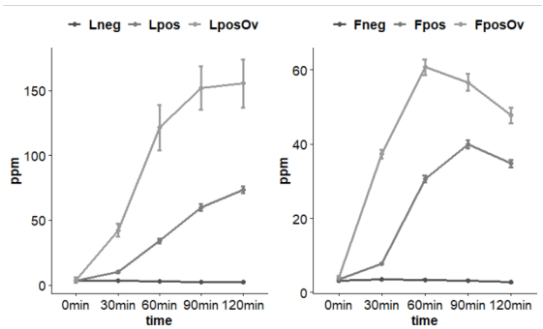


Figure 3. Mean trajectories for the manually determined number of clusters (n=3) using lactose tolerance test (left) and fructose tolerance test (right).

The resulting ARI values using k-means, hierarchical clustering and clustering according to shape are depicted in Tab. 4. The highest ARI values were obtained using clustering according to shape, which was developed to efficiently cluster longitudinal data.

Table 4. ARI values for the considered datasets using k-means, hierarchical clustering and clustering according to shape.

		ARI	
Dataset	k-means	Hierarchical	Shape based
LTT	0.54	0.17	0.68
FTT	0.53	0.21	0.67

4. Discussion

In this work we analyzed H_2 trajectories of breath gas samples resulting from lactose and fructose tolerance tests. In contrast to [7], we analyzed clusters in trajectories of lactose and fructose hydrogen breath tests using internal and external validation measures. First, we searched for number of clusters using 26 internal validation indexes. Based on the majority voting as proposed by [15], we determined a total of 2 dominant clusters. However, biologically we would expect three clusters in each dataset (LTT and FTT) representing groups NEG, POS and SIBOS. The SIBOS group was comparable small (see Tab. 4) which might explain that this group is not identified as a separate cluster by most indexes. However, the Ball index suggested a number of three clusters for all experiments (see Tab. 5). The Ball index [21] uses measurements based on sums of squares as measure of dispersion [22]. After setting the number of clusters to $k = 3$ we externally evaluated our clustering methods using ARI. In addition to k-means and hierarchical clustering, we applied a new method using shape information [10]. The shape-based method reached the highest similarity between the reference partition (ARI of 0.68 for LTT and 0.67 for FTT, respectively).

We conclude that only a small amount of internal validation indexes such as Ball index were able to obtain a reasonable number of clusters in our data. It is further important to externally validate clusters by using medical expert knowledge. Clustering methods considering shape information turned out to be the best method for clustering longitudinal data resulting from lactose and fructose tolerance tests. In our ongoing work, we will apply supervised machine learning techniques to automatically predict the result of the tolerance test (i.e., positive, negative or SIBOS).

Table 5. Listing of all clustering indexes that suggested a total of three clusters for LTT and FTT dataset and the considered clustering methods *k*-means and hierarchical clustering.

Dataset	Clustering method	Indexes suggesting k=3
LTT	*k*-Means	TrCovW, Duda, PseudoT2, Ratkowsky, Ball
	Hierarchical	Hartigan, Scott, TrCovW, TraceW, Ratkowsky, Ball, Frey
FTT	*k*-Means	Scott, TraceW, Friedman, Ball
	Hierarchical	Scott, TrCovW, TraceW, Ratkowsky, Ball, Frey

References

[1] Y. Itan, B.L. Jones, C.J. Ingram, D.M. Swallow and M.G. Thomas, A worldwide correlation of lactase persistence phenotype and genotypes, *BMC evolutionary biology* **10**(1) (2010), 1–11.

[2] C. Jang,S. Hui, W. Lu, A.J. Cowan, R.J. Morscher, G. Lee, W. Liu, G. Tesz, M. Birnbaum and J.D. Rabinowitz, The Small Intestine Converts Dietary Fructose into Glucose and Organic Acids, *Cell Metabolism* **27**(2) (2018), 351-361.

[3] R.P. Ferraris, J.-y. Choe and C.R. Patel, Intestinal absorption of fructose, *Annual review of nutrition* **38** (2018), 41–67.

[4] A. Amann, G. Poupart, S. Telser, M. Ledochowski, A. Schmid and S. Mechtcheriakov, Applications of breath gas analysis in medicine, *International Journal of Mass Spectrometry* **239**(2–3) (2004), 227–233.

[5] J. Bures, J. Cyrany, D. Kohoutova, M. Förstl, S. Rejchrt, J. Kvetina, V. Vorisek and M. Kopacova, Small intestinal bacterial overgrowth syndrome, *World journal of gastroenterology: WJG* **16**(24) (2010), 2978.

[6] S.V. Rana and A. Malik, Hydrogen breath tests in gastrointestinal diseases, *Indian Journal of Clinical Biochemistry* **29**(4) (2014), 398–405.

[7] C. Rubio-Escudero, J. Valverde-Fernández, I. Nepomuceno-Chamorro, B. Pontes-Balanza, Y. Hernández-Mendoza and A. Rodríguez-Herrera, Data Mining Techniques Applied to Hydrogen Lactose Breath Test., *PloS one* **12** (2017).

[8] A. Eisenmann, A. Amann, M. Said, B. Datta and M. Ledochowski, Implementation and interpretation of hydrogen breath tests, *Journal of breath research* **2** (2008).

[9] M.G. Omran, A.P. Engelbrecht and A. Salman, An overview of clustering methods, *Intelligent Data Analysis* **11**(6) (2007), 583–605.

[10] C. Genolini, R. Ecochard, M. Benghezal, T. Driss, S. Andrieu and F. Subtil, kmlShape: An efficient method to cluster longitudinal data (time-series) according to their shapes, *Plos One* **11**(6) (2016), 0150738.

[11] X. Jin and J. Han, *Partitional Clustering*, in: Encyclopedia of Machine Learning, C. Sammut and G.I. Webb, eds, Springer US, Boston, MA, 2010, pp. 766–766.

[12] L. Rokach and O. Maimon, *Clustering methods*, in: Data mining and knowledge discovery handbook, Springer, 2005, pp. 321–352.

[13] B. Clarke, E. Fokue, and H.H. Zhang, *Principles and theory for data mining and machine learning*, Springer Science & Business Media, 2009.

[14] F. Murtagh and P. Legendre, Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?, *Journal of classification* **31**(3) (2014), 274–295.

[15] M. Charrad, N. Ghazzali, V. Boiteau and A. Niknafs, NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set, *J Stat Softw* **61** (2014).

[16] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *Journal of computational and applied mathematics* **20** (1987), 53–65.

[17] D.L. Davies and D.W. Bouldin, A cluster separation measure, *IEEE transactions on pattern analysis and machine intelligence* (1979), 224–227.

[18] J.C. Dunn, Well-separated clusters and optimal fuzzy partitions, *Journal of cybernetics* **4**(1) (1974), 95–104.

[19] A. D'Ambrosio, S. Amodio, C. Iorio, G. Pandolfo and R. Siciliano, Adjusted Concordance Index: an Extension of the Adjusted Rand Index to Fuzzy Partitions, *Journal of Classification* (2020), 1–17.

[20] J.M. Santos and M. Embrechts, *On the use of the adjusted rand index as a metric for evaluating supervised classification*, in: International conference on artificial neural networks, Springer, 2009, pp. 175–184.

[21] G. Ball and D. Hall, *A novel method of data analysis and pattern classification,* SRI International: Menlo Park, CA, USA (1965).

[22] F.J. Abarca-Alvarez, F.S. Campos-Sánchez and R. Mora-Esteban, Survey Assessment for Decision Support Using Self-Organizing Maps Profile Characterization with an Odds and Cluster Heat Map: Application to Children's Perception of Urban School Environments, *Entropy* **21**(9) (2019), 916.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210090

Reporting of Studies Conducted on Austrian Claims Data

Andrea SUNJIC-ALIC[a,1], Karin ZEBENHOLZER[b] and Walter GALL[a]
[a] *Center for Medical Statistics, Informatics and Intelligent Systems, Department of Medical Information Management, Medical University of Vienna, Austria*
[b] *Department of Neurology, Medical University of Vienna, Austria*

Abstract. In healthcare studies, the analysis of claims data is gaining an increasingly important role. Observational studies should be reported in a manner that promotes internal and external validity assessment, with the exact and standardized description of items. Several international guidelines and checklists for reporting on secondary data are available. The aim of this work was to analyse the applicability of reporting guidelines especially for claims data. The STROSA-2 guidelines recommendations were evaluated by means of a report on a study on triptan medications in Austria. Six items were identified which could be expanded to support complete and transparent report on Austrian claims data. Therefore, we would suggest to add some details in the STROSA-2 guidelines concerning study design, legal foundations, data protection, data flow, descriptive results and risk of bias. The guidelines for reporting on Austrian claims data were successfully compiled with additional items. New guidelines should be further processed and tested with strong recommendations to focus on data limitations and legal aspects.

Keywords. Diagnoses-related groups, Austrian claims data, reporting guidelines, STROSA-2 checklist

1. Introduction

Analysing claims data can help to identify gaps in care, to reduce resource consumption and to evaluate effectiveness and safety of medications and medical procedures. Routinely collected health data (e.g. administrative data or patient records) are frequently used for observational studies and health technology assessment [1], [2]. But the limitations and quality of studies using routinely collected health data are also discussed [3]. Especially claims data have several strong limitations, like incompleteness in terms of clinical diagnoses or inaccuracy [4]. They contain information from different parties in form of bills for the care they provided, where data linkage from different sources can lead to data privacy issues [5]. Despite limitations, the information contained on a bill is still valuable. For example, such data as vital signs or clinical notes are omitted, but records of procedures and diagnoses are kept [6].

Observational studies should be reported in a manner that promotes internal and external validity assessment. The exact and standardized description of studies is very

[1] Corresponding Author: Andrea Sunjic-Alic, Center for Medical Statistics, Informatics and Intelligent Systems, Department of Medical Information Management, Medical University of Vienna, Austria, E-Mail: andrea.sunjic.alic@gmail.com

important. There are several international guidelines for reporting on secondary data available.

The STROBE Statement [7], [8] is a reporting guideline of observational studies in epidemiology, comprising a checklist of 22 items. It was developed in 2007 by the STROBE Initiative, an international collaboration of researchers, epidemiologists, statisticians, and journal editors.

RECORD [9], [10] is an international initiative that developed a STROBE-based reporting guideline for studies conducted using routinely collected health data in 2015.

RECORD-PE [11], [12] is a reporting guideline developed in 2018, based on RECORD for reporting of non-interventional pharmacoepidemiological studies.

Considerations not covered by STROBE and RECORD were described and tested by experts from the German Society for Social Medicine and Prevention, the German Society for Epidemiology and the German Network for Health Services Research. This resulted in a new checklist, STROSA-2 (revised) [13] compiled in 2016 for the reporting of secondary data analysis in Germany.

The aim of this work was to analyse and compare the existing guidelines and adapt them in order to address reporting items, which are specific for studies using the Austrian claims data.

STROSA-2 criteria were used for a preliminary report on a triptan medications study. This study was conducted on Austrian claims data for the year 2007 from the GAP-DRG database. The Austrian GAP-DRG (General Approach for Patient-oriented outpatient-based Diagnosis-Related Groups) database [14] has been used for multiple studies [15], [16], [17] in different clinical areas.

2. Methods

2.1. Reporting Guidelines

We analysed the STROBE, RECORD and STROSA criteria for missing details. A short overview of items from each checklist is presented in Table 1.

2.2. Study on triptan medications

We tested the STROSA-2 recommendations by means of a report on a study on triptan medications. The study was conducted on Austrian claims data from the GAP-DRG database. It explored the health status of the prescription patterns and linkages between triptan use, overuse, and cardiovascular diseases of triptan users over the age of 50.

2.3. GAP–DRG

The Austrian healthcare system includes a compulsory healthcare insurance and covers three major areas: inpatient care, outpatient care and the drug supply. Inpatient care is carried out mainly in hospitals. Outpatient care includes all the treatments by practitioners, specialists and outpatient clinics. The supply of prescription drugs takes place in pharmacies.

Table 1. Comparison of STROBE, STROSA 1, STROSA 2 and RECORD reporting guidelines (adapted from [13]); N = no specific recommendations, E = suggested extensions.

STROBE	STROSA-1	STROSA-2	RECORD
Title, Abstract	**Title, Abstract**	**Title, Abstract, Keywords**	**Title, Abstract, Keywords**
Title, Abstract	Title, Abstract	Title, Abstract	Title, Abstract
N	N	Keywords	N
Introduction	**Introduction**	**Introduction**	**Introduction**
Background/rationale	Background/rationale	Background/rationale	Background/rationale
Objectives	Objectives	Objectives	Objectives
Methods	**Methods**	**Methods**	**Methods**
Study design	Study design	Study design	EStudy design
N	Framework	Data origin	N
N	Legal foundation	Legal foundation	EN
N	N	Data protection	EN
N	Data flow	Data flow	EN
N	Study curriculum	Selection criteria	N
N	Unit of analysis	Unit of analysis	N
Participants	Participants	N	Participants
N	Internal validations	N	N
Variables	Variables	Variables	Variables
Data sources/ measurement	Classification Systems	N	Data sources/ measurement
Bias	Bias	N	Bias
Study size	Study size	Study population size	Study size
Quantitative variables	Quantitative variables	N	Quantitative variables
Statistical methods	Statistical methods	Statistical methods	Statistical methods
N	N	N	Data access and processing
N	N	N	Data linkage
Results	**Results**	**Results**	**Results**
Participants	Selection of study population	Selection of study population	Participants
Descriptive data	Description of participants	Descriptive results	EDescriptive data
Outcome data	Statistical measures	N	Outcome data
Main results	Main results	Main results	Main results
Other analyses	Other analyses	Other results	Other analyses
Discussion	**Discussion**	**Discussion**	**Discussion**
Key results	Key results	Key results	Key results
Limitations	Limitations	Internal validity and risk of bias	ELimitations
N	Strengthen	Strengthen and weaknesses	N
Interpretation	Interpretation	Interpretation	Interpretation
Generalisability	Generalisability	Generalisability	Generalisability
N	**N**	**Conclusion**	**Conclusion**
N	N	Conclusion	Conclusion
N	**N**	**Conflicts of interest**	**Further Information**
Funding	Funding	Funding	Funding
N	Role of data owners	Role of data owners	N
N	N	Other conflicts of interest	N
N	N	N	Accessibility of protocol, raw data, and programming code

The research database GAP-DRG of the Main Association of Austrian Social Security Institutions (HBV) contains routinely collected data from various data sources from 2006 to 2011. The data available in GAP-DRG include hospital stays, prescribed

and dispensed medications, the medical field of the physicians, social security providers, sick leaves (including their duration) and others.

The diagnosis-related model (DRG) is the regulatory framework for standardized grouping and scoring of inpatient hospital stays [18]. The assignment of a DRG depends on a patient's age and sex, their main and secondary diagnoses, procedures, comorbidities, complications and discharge status [19].

The billing data of Austrian health insurances (2006–2011) were transmitted to the HVB, integrating data from the Folgekosten-Datenbank (FoKo database), the Minimal Basic Data Set (MBDS) of the Federal Ministry of Health (BMG), billing data of public hospitals and the Private Hospital Financing funds.

All data collected from the individual health insurances and BMG were merged into one database. Double entries were removed and regional characteristics of both, the insurers and the insured parties were aggregated. The review of the data quality focused mainly on the completeness and consistency of the data.

2.4. Legal Aspects

Data protection requirements are governed by the European Data Protection Regulation (GDPR) [20]. Article 9 thereof provides an exemption for processing of personal data for scientific purposes [21]. The Austrian data protection law (DSG) [22] is compliant with European data protection regulations.

The data from the GAP-DRG are pseudonymised and stored on a multi-layered encrypted server. They can be restored to their original state, allowing for individuals to be re-identified. Direct access to the GAP-DRG servers is exclusively granted to the so-called custodians, which are specifically trained in data protection. Only data necessary for a specific study are transferred to the research server.

Access to the research server has to be granted by the HVB. Authorized persons (e.g. statisticians) can access it via a VPN (virtual private network) connection.

3. Results

The STROBE and RECORD criteria meet most of the requirements for reporting of secondary data analysis. Although RECORD was initiated as a reporting standard for routine data analysis, following the introduction of the GDPR in the EU, neither RECORD nor STROBE were fully suitable for reporting of claims data.

We identified six items not covered by STROSA-2, which focus on reporting of Austrian claims data, concerning study design, legal foundations, data protection, data flow, descriptive results and risk of bias.

The report should clearly state the original purpose of the data collection, underlining that the study is based on secondary data. We recommend to capture this under item 5 (‘Study design’).

The research question should be formulated precisely and included in the study design. This is important for the analysis, study design, data extraction, time framework and costs. A written protocol that determines the study characteristics is essential for the analysis of the secondary data. All studies conducted on Austrian claims data must be approved by the Ethics committee. We recommend to expand item 7 (‘Legal foundations’) to provide the information on the Ethics committee decision, project name, protocol, enrolees, institution and sponsors.

The analysis of billing data must comply with Austrian and European data protection law. It must not be possible to identify an individual based on their healthcare data used in the research. Safety measures, such as involving data custodians in the data selection process, anonymization or pseudonymisation of personal references, and data encryption, must be applied. Our recommendation is to expand item 8 ('Data protection') to indicate how the data from the claims database are pseudonymised and encrypted.

However, even though the data are merged into a single database, many data preparation steps are necessary for each new project. Data extraction from the database requires both expertise in database management and replication, and deep knowledge of the Austrian healthcare system. We recommend to expand the item 9 ('Data flow') to describe if and how data sources in the database were linked at a macro and micro level.

The biggest disadvantage of the data from the claims database is their incompleteness in terms of patients' outpatient diagnoses, clinical information and socio-demographic characteristics. The BMG provides the MBDS data without the personal reference (anonymized data) [14]. Since MBDS contains only information on stays and services provided but no patient reference, patients' paths cannot be completely retraced throughout the healthcare system. Thus, we recommend to include the characteristics of the study population in item 16 ('Descriptive results') and describe selection criteria based on data quality and linkage.

The STROSA-2 reporting guidelines were developed following the GPS (Good Practice in Secondary Data Analysis) recommendations [23]. Internal validity and risk of bias are aspects covered in the 'Discussion' section of the STROSA-2 guidelines. Any potential selection bias or confounders should be considered and documented. It is a fact that Austrian claims data do not contain the same information value as clinical data. The database covers most of the general population, but for example does not include the information on diagnoses for outpatient treatments. Therefore, some illnesses can only be inferred if the patient was prescribed and dispensed certain medications for the treatment of those diseases. Prescribed medications dispensed to patients who are exempt from payment due to low income or severe chronic diseases are also not included in the database. ICD codes documented for billing purposes often differ from diagnoses used in everyday clinical practice. Our recommendation is to expand item 20 ('Internal validity and risk of bias') to discuss the limitations of the study such as the missing data, and methods and quality of the data linkage.

Items resulting in the report on the study being incomplete, are marked with 'E' in Table 1. Additional recommendations with explanations are presented in Table 2.

Table 2. Recommendations for reporting on Austrian Claims Data

5. Study design (METHODS)
STROSA-2 recommendations
Demonstrate that the study is based on secondary data, what primary purpose it served, and whether you have selected a cross-sectional, cohort, case-control, or other study design within your secondary data analysis
Additional recommendations
Demonstrate that the study is based on secondary data and state the original purpose of the data collection, such as if data were administrative or claims data. Study design should include the methods of the study populations selection and elaboration on used classifications or developed algorithms. *Example*: *One of the data sources integrated in the GAP-DRG database is MBDS which consists of administrative data (e.g. personal data), patient's medical records (e.g. diagnosis), and LKF data (claims).*
7. Legal foundation (METHODS)
STROSA-2 recommendations

Describe the legal basis for data disposal and analysis.
Additional recommendations
Include the information on: Vote (Votum) from the Ethics committee, project name and protocol, enrolees, institution and sponsors. *Example: All studies conducted on Austrian health care records must be approved by the Ethics committee.*
8. Data protection (METHODS)
STROSA-2 recommendations
Indicate how personal and/or institutional data protection was ensured
Additional recommendations
Indicate how the health data from the research database are pseudonymised and stored. *Example: Pseudonymised data can be restored to their original state, with the addition of information by a data custodian, allowing for individuals to be re-identified. Removing of personal references guarantees the anonymity of the study population, but the possibility of retracing individuals has to be disabled as well. Therefore, for study on triptans, the exact date of birth and full postal codes are removed, keeping only the birth year and the district.*
Indicate if and how data from different sources in the database were linked. *Example: Linkage of the claims data from healthcare providers with the data reported through the Regional Health Funds to the Federal Ministry of Health. Data from insurance carriers include social security numbers but personal data are pseudonymised due to privacy rules and regulations. On the other hand, the social security number is completely removed (anonymized) from the records by the Regional Health Funds.*
9. Data flow (METHODS)
STROSA-2 recommendations
Represent the data flow and indicate who carried out the data provision and analysis. If necessary, describe how data from different data sources was linked
Additional recommendations
Indicate if and how the data sources in the claims database were linked at a macro level. *Example: Healthcare providers report to insurance carriers, where hospitals additionally report through the Regional Health Funds to the BMG. The data from insurance carriers include pseudonymised personal data. The data reported to BMG are anonymized and it is not possible to determine if two registered hospitalizations belong to the same patient. A record linkage was developed by the DEXHELPP team (Decision Support for Health Policy and Planning: Methods, Models and Technologies based on Existing Health Care Data) to find a unique person identifier for each event recorded in MDBS and hospital reports to the insurance carriers.*
Indicate if and which data sources were linked at a micro level, such as data not included in the database, but necessary for the study analysis. Describe the data architecture and all layers of the ETL process. Describe additional materials imported into the data staging layer. *Example: NUTS region codes can be used to link patients district of residence to the ÖSG Versorgungscode (care-supply region codes). The linkage can be performed on the research server.*
16. Descriptive results (RESULTS)
STROSA-2 recommendations
Describe the characteristics of the study population as well as exposures and possible confounders. Take into account whether there is a case or personal reference or the level of aggregation of the data
Additional recommendations
Describe in detail the characteristics of the study population and include selection criteria based on the data quality and linkage. *Example: In study on triptan medication, a cardiovascular disease was presumed if participants were dispensed certain medications. The majority of those participants didn't have the diagnosis recorded in the claims database. Characteristics of the study population can be described in the form of a diagram.*
20. Internal validity and risk of bias (DISCUSSION)
STROSA-2 recommendations
Discuss the risk of bias (selection bias, information bias, confounding, etc.) and the measures you have taken to determine its presence and extent
Additional recommendations
Discuss the limitations of the study such as missing data. If data linkage was performed, methods and quality of the outcome should be specified. *Example: The GAP-DRG claims database does not contain data on outpatient diagnoses or over-the-counter medications (OTC).*

4. Discussion

Austrian claims data can be reused in studies to serve as a foundation for better understanding and decision support in healthcare. For conducting a secondary data analyses we suggest to follow the GPS recommendations [23]. The aim of this work was to analyse the applicability of reporting guidelines specifically to claims data. Several international guidelines and checklists for reporting on secondary data are available. The STROBE and RECORD criteria meet most of the requirements for reporting of secondary data analysis, but none were fully suitable for reporting of Austrian claims data. The STROSA-2 guidelines recommendations were evaluated by means of a report on a study on triptan medications in Austria. We found more details could be provided in six items, in particular study design, legal foundations, data protection , data flow, descriptive results and risk of bias. Other items allowed complete and transparent reporting.

The adaptation of STROSA-2 reporting guidelines can provide reviewers and readers with reliable information, promote the quality of research design and help to assess the validity of the results. The report should include the information on the original purpose of the data collection, vote from the Ethics committee and data protection methods. The characteristics of the study population and selection criteria should be provided. The data architecture, all layers of the ETL process, and how the data sources in the claims database were linked both at a macro and micro level should be described. One of the most important aspects of the report is to discuss the limitations of the study in question, such as missing data (e.g., OTC medications) or inaccurate data (e.g., discrepancies caused by delays in billing dispensed drugs). If data linkage was performed, methods and quality of the outcome should be specified.

Our suggested guidelines extensions must be further analysed and discussed, especially the international applicability. E.g., the main and secondary diagnoses are the most important usable data from claims databases. However, in different countries they may be recorded for different purposes and could as well differ in meaning. It is therefore crucial to describe their interpretation precisely in order to secure the comparability of studies.

New recommendations expand the STROSA-2 criteria to reporting on observational studies conducted on Austrian claims data. The present work is an evaluation of the STROSA-2 guidelines based on a clinical study. It provides suggestions for a refinement on limitations and legal aspects. However, these guidelines could be further developed to support complete and transparent reporting of studies conducted on claims data.

References

[1] A. Makady, A. van Veelen, P. Jonsson, O. Moseley, A. D'Andon, A. de Boer, H. Hillege, O. Klungel, and W. Goettsch, Using Real-World Data in Health Technology Assessment (HTA) Practice: A Comparative Study of Five HTA Agencies, *PharmacoEconomics*. **36** (2018). doi:10.1007/s40273-017-0596-z.

[2] B. Blanch, N.A. Buckley, L. Mellish, A.H. Dawson, P.S. Haber, and S.-A. Pearson, Harmonizing post-market surveillance of prescription drug misuse: a systematic review of observational studies using routinely collected data (2000-2013), *Drug Saf.* **38** (2015) 553–564. doi:10.1007/s40264-015-0294-8.

[3] L.G. Hemkens, E.I. Benchimol, S.M. Langan, M. Briel, B. Kasenda, J.-M. Januel, E. Herrett, and E. von Elm, The reporting of studies using routinely collected health data was often insufficient, *J Clin Epidemiol*. **79** (2016) 104–111. doi:10.1016/j.jclinepi.2016.06.005.

[4] H. Gothe, P. Ihle, D. Matusiewicz, and E. Swart, Routinedaten im Gesundheitswesen, Handbuch Sekundärdatenanalyse: Grundlagen, Methoden, und Perspektiven, 2., vollständig überarbeitete und erweiterte Auflage., Verlag Hans Huber, Hogrefe AG, Bern, 2014.
[5] S. Mueller, F. Gottschalk, A. Groth, W. Meeraus, M. Driessen, T. Kohlmann, and T. Wilke, Primary data, claims data, and linked data in observational research: the case of COPD in Germany, *Respir Res*. **19** (2018) 161. doi:10.1186/s12931-018-0865-1.
[6] Claims Data - The Good, The Bad and The Ugly, *Health IT Answers*. (2014). https://www.healthitanswers.net/claims-data-the-good-the-bad-and-the-ugly/ (accessed January 10, 2021).
[7] STROBE Group, STROBE Statement, (2019). https://www.strobe-statement.org/index.php?id=strobe-home (accessed January 10, 2021).
[8] E. von Elm, D.G. Altman, M. Egger, S.J. Pocock, P.C. Gøtzsche, J.P. Vandenbroucke, and STROBE Initiative, The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies, *Int J Surg*. **12** (2014) 1495–1499. doi:10.1016/j.ijsu.2014.07.013.
[9] RECORD Group, RECORD Reporting Guidelines, (2020). https://www.record-statement.org/ (accessed January 10, 2021).
[10] E.I. Benchimol, L. Smeeth, A. Guttmann, K. Harron, D. Moher, I. Petersen, H.T. Sørensen, E. von Elm, S.M. Langan, and RECORD Working Committee, The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement, *PLoS Med*. **12** (2015) e1001885. doi:10.1371/journal.pmed.1001885.
[11] S. Langan, S. Schmidt, K.J. Wing, V. Ehrenstein, S. Nicholls, K. Filion, O. Klungel, I. Petersen, H. Sørensen, W. Dixon, A. Guttmann, K. Harron, L. Hemkens, D. Moher, S. Schneeweiss, L. Smeeth, M. Sturkenboom, E. von Elm, S. Wang, and E. Benchimol, The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE), *British Medical Journal*. (2018). doi:10.1136/bmj.k3532.
[12] RECORD-PE Reporting Checklist, (n.d.). http://www.record-statement.org/checklist-pe.php (accessed January 12, 2021).
[13] E. Swart, E.M. Bitzer, H. Gothe, M. Harling, F. Hoffmann, D. Horenkamp-Sonntag, B. Maier, S. March, T. Petzold, R. Röhrig, A. Rommel, T. Schink, C. Wagner, S. Wobbe, and J. Schmitt, STandardisierte BerichtsROutine für Sekundärdaten Analysen (STROSA) – ein konsentierter Berichtsstandard für Deutschland, Version 2, *Gesundheitswesen*. **78** (2016) e145–e160. doi:10.1055/s-0042-108647.
[14] E. Florian, GAP-DRG, *GAP-DRG Dokumentation*. (2013). http://gapdrg.endel.at/dokuwiki/doku.php/gapdrg:start (accessed January 10, 2021).
[15] A. Geroldinger, S.K. Sauter, G. Heinze, G. Endel, W. Dorda, and G. Duftschmid, Mortality and continuity of care - Definitions matter! A cohort study in diabetics, *PLoS One*. **13** (2018) e0191386. doi:10.1371/journal.pone.0191386.
[16] H. Katschnig, C. Straßmayr, F. Endel, M. Berger, G. Zauner, J. Kalseth, R. Sfetcu, K. Wahlbeck, F. Tedeschi, L. Šprah, and CEPHOS-LINK study group, Using national electronic health care registries for comparing the risk of psychiatric re-hospitalisation in six European countries: Opportunities and limitations, *Health Policy*. **123** (2019) 1028–1035. doi:10.1016/j.healthpol.2019.07.006.
[17] C. Rinner, S.K. Sauter, L.M. Neuhofer, D. Edlinger, W. Grossmann, M. Wolzt, G. Endel, and W. Gall, Estimation of severe drug-drug interaction warnings by medical specialist groups for Austrian nationwide eMedication, *Appl Clin Inform*. **5** (2014) 603–611. doi:10.4338/ACI-2014-04-RA-0030.
[18] E. Hagenbichler, The Austrian DRG system, Federal Ministry of Health, Sector I/B, 2010. https://broschuerenservice.sozialministerium.at/Home/Download?publicationId=576 (accessed January 10, 2021).
[19] Diagnosis Related Group (DRG), *HMSA Provider Resource Center*. (2018). https://hmsa.com/portal/provider/zav_pel.fh.DIA.650.htm (accessed January 10, 2021).
[20] Protecting Patients' Medical Records under the GDPR, (2019). https://idpc.org.mt/en/articles/Pages/synapse-article.aspx (accessed January 10, 2021).
[21] General Data Protection Regulation (GDPR) Compliance Guidelines, *GDPR.Eu*. (n.d.). https://gdpr.eu/ (accessed January 10, 2021).
[22] Datenschutzrecht in Österreich - Datenschutzbehörde, (2020). https://www.dsb.gv.at/gesetze-in-osterreich (accessed January 10, 2021).
[23] P. Ihle, E. Swart, T. Lampert, and S. Klug, GPS – Good Practice in Secondary Data Analysis: Revision after Fundamental Reworking, *Deutsche Gesellschaft Für Epidemiologie*. (n.d.). https://www.dgepi.de/assets/Leitlinien-und-Empfehlungen/Practice-in-Secondary-Data-Analysis.pdf (accessed January 10, 2021).

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210091

Health Data Privacy in the COVID-19 Pandemic Context: Discourses on HIPAA

Javad POOL[a,1], Saeed AKHLAGHPOUR[a] and Farhad FATEHI[a,b]
[a] *The University of Queensland, Brisbane, Australia*
[b] *Monash University, Melbourne, Australia*

Abstract. Background: Considering the impacts of the COVID-19 pandemic on health service delivery, the US Office for Civil Rights (OCR) updated the policies on health data processing, and Health Insurance Portability and Accountability Act (HIPAA). Objectives: In this study, we investigated discourses on HIPAA in relation to COVID-19. Methods: Through a search of media sources in the Factiva database, relevant texts were identified. We applied a text mining approach to identify concepts and themes in these texts. Results: Our analysis revealed six central themes, namely, Health, HIPAA, Privacy, Security, Patients, and Need, as well as their associated concepts. Among these, Health was the most frequently discussed theme. It comprised concepts such as public, care, emergency, providers, telehealth, entity, use, discretion, OCR, Health and Human Services (HHS), enforcement, business, and services. Conclusion: Our discourse analysis of media outlets highlights the role of health data privacy law in the response to global public health emergencies and demonstrates how discourse analysis and computational methods can inform health data protection policymaking in the digital health era.

Keywords. Health Information Systems, Digital Technology, Privacy, Telemedicine, Health Insurance Portability and Accountability Act, Public Health Informatics

1. Introduction

The COVID-19 pandemic, as a "Public Health Emergency of International Concern (PHEIC)" [1], demonstrated the limitations of the current business models for effective service delivery in the time of crisis [2]. As a rapid response to the pandemic, healthcare entities and business associates are seeking to transform their business models by using digital health solutions such as telehealth [3]. However, throughout the process of digital health transformation, healthcare providers and third parties still must be mindful to comply with the data protection laws and regulations during the pandemic. Given the impacts of this ongoing and evolving phenomenon, current laws and regulations on data privacy require adaptation to this public health context. In response to the COVID-19 crisis, the US Office for Civil Rights (OCR) has issued notifications of enforcement discretion, announcements, guidance, and resources for several concerns related to Health Insurance Portability and Accountability Act (HIPAA) compliance. These announcements include the use of telehealth, web-based

[1] Corresponding Author: Javad Pool, The University of Queensland, Brisbane, Australia, E-Mail: j.pool@uq.net.au

scheduling, Health Information Exchanges (HIE), and contact to COVID-19 patients for public health purposes [4]. Telehealth was among the first and key digital health areas, for which the OCR issued notifications regarding amended policies for data processing, privacy, and security under HIPAA. These notifications, in turn, triggered discourses on health information system use and compliance in the academic and media outlets [e.g., 5, 6].

This research is a response to the World Health Organization (WHO) calls for studies in the pandemic context and seeks to identify the central HIPAA-related topics in the news media discourse. Following the PHEIC, WHO has proposed a roadmap for "powering research to control the epidemic" [1]. This research, as a path towards contribution to the management of 'pandemic information systems', is aligned with the WHO roadmap and the calls for research. Existing research has explored the legal implications of HIPAA privacy for public health [7, 8]. However, in the context of pandemics, discourses of privacy regulations (e.g. HIPAA) have not been explicitly investigated. Despite the arguably critical impacts of COVID-19 as a public health emergency on health regulation compliance and privacy practices, research in this field is scant. The purpose of the present research is to identify the discourses on HIPAA in the time of a pandemic, and to contribute an understanding of relevant privacy issues and health data protection practices.

In this study, we analyzed news media discourses about HIPAA and the COVID-19 public health crisis. More specifically, through a discourse analysis empowered by a text analytics approach, we attempted to represent concepts and themes related to the privacy laws, users, health data, and health IT artifacts as indicated in the media. Given the identified concept and themes, this research informs future investigations and policies to better understand the complexity of data protection in digital health implementation, use, and transformation.

2. Methods

To discover discourses on HIPAA and COVID-19, we used the Factiva database. As the database covering region-specific and international news, Factiva has been used for discourse analysis in different disciplines from information systems to public health policy [9, 10]. Given the context and aim of our research, a discursive approach to news media offers valuable insights into the academic literature [11]. Media discourses are driven by timeliness (e.g., capturing the latest practices in privacy management in the time of COVID-19) and practicality (e.g., descriptions of the significant consequences of the pandemic and HIPAA). The search was conducted on 17 January 2021, using this search string: ((HIPAA OR "Health Insurance Portability and Accountability Act" OR "Office for Civil Rights")) AND (COVID or Coronavirus or SARS-CoV-2)). The search was limited to 'news headline' and the English language. This search strategy has resulted in the identification of 55 records. After removing duplicates and screening the news in terms of relevancy to the research aim, 42 news articles were included for text analytics and discourse analysis. We used Leximancer (www.leximancer.com) for identifying discourse concepts and themes. Leximancer is text mining software for analyzing and visually representing the concepts in collections of textual documents. It extracts concepts from words which have similar meanings and/or appear together, even when the documents use different styles and formats [12]. Before analyzing news texts by Leximancer, we applied a data cleaning step to achieve

a faithful concept map. This step involved removing irrelevant texts such as authors, and publishers' names, URLs, and advertisements about service providers. Finally, a document contains 23,507 words arranged in 42 articles was imported to Leximancer for the analysis. Our interpretations of discourses on HIPAA and COVID-19 are based on the identified concepts and themes from text analysis and reflection of illustrative quotes in the news articles. Our text analysis process is illustrated in Figure 1.

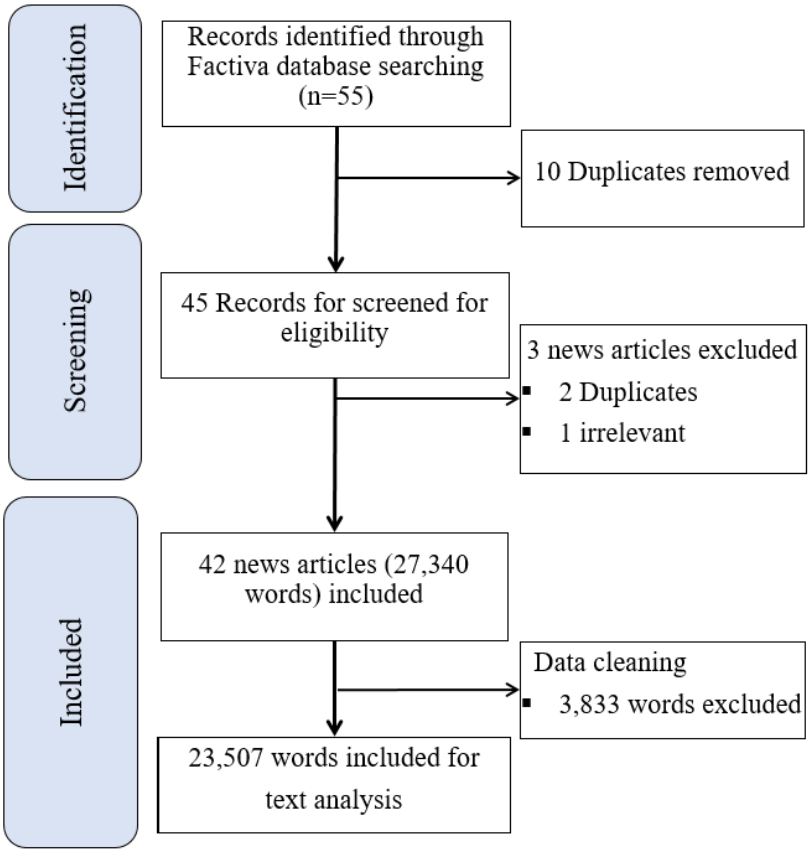


Figure 1. Text analysis process.

3. Results

3.1. Descriptive analysis

Through the descriptive analysis, we provided an overview of the findings. Textual analysis on HIPAA discourses in the context of COVID-19 shows a concept map with six clusters (themes) as shown in Figure 2. The six themes, labeled by Leximancer based on the most prominent concept in the cluster of words, are Health, HIPAA, Privacy, Patients, Security, and Need. These themes and their hits with relevant concepts are also represented in Table 1.

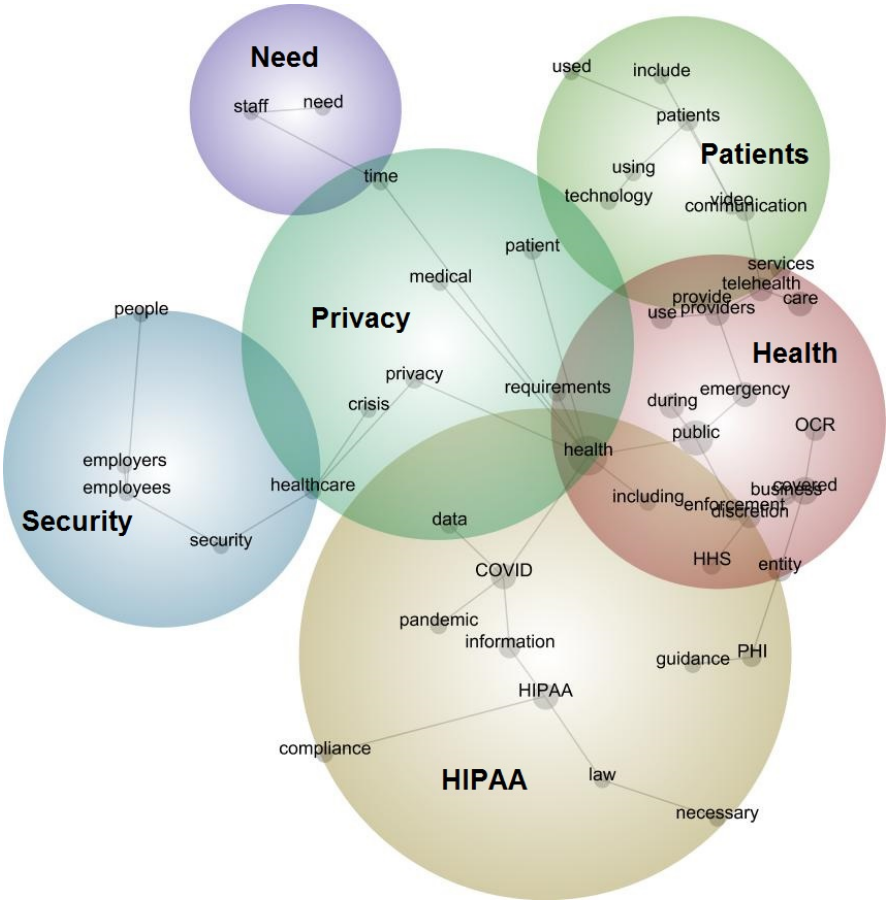


Figure 2. Concept map and network of HIPAA and COVID-19 discourses.

Table 1. Dominant themes, hits, and concepts

Themes	**Hits***	**Concepts** (number of concepts in the cluster)
Health	441	Health, public, covered, care, emergency, during, providers, telehealth, entity, use, discretion, enforcement, OCR, services, including, business, HHS, provide (n=18)
HIPAA	398	HIPAA, COVID, pandemic, information, PHI, data, compliance, guidance, law, necessary (n=10)
Privacy	214	Privacy, patient, requirements, healthcare, medical, time, crisis (n=7)
Patients	163	Patients, video, communication, technology, include, used, using (n=7)
Security	101	Security, employees, employers, people (n=4)
Need	44	Need, staff (n=2)

Note: * number of hits for dominant concepts in the map.

3.2. Thematical explanations

In this subsection, we briefly explain discourses around the main six themes and related concepts.

3.2.1. Health

Health as the main theme among six is consists of 18 concepts. The discourses around this theme are related to the department of Health and Human Services (HHS), Office for Civil Rights (OCR), and notification of enforcement discretion for using telehealth by healthcare providers during public health emergencies. The OCR is responsible for enforcing certain privacy regulations [13]. During the COVID-19 public health emergency, the notices for using telehealth platforms (e.g., Skype, Zoom, GoToMeeting) align with the good faith provision provoke discussions which represent in news. While before the pandemic, there were data protection concerns about the telehealth platforms, healthcare providers see this notification as an opportunity to improve public access to health services. These are reflecting both concerns, e.g., "*Before, it* [using Skype] *may not have been technically compliant from a security perspective* …" [14], and opportunities, e.g., "*The department now allows doctors to provide health services through audio or video on sites like Skype and FaceTime so they can see a greater number of patients and reach out to those sheltering in their homes*" [15].

3.2.2. HIPAA

The second theme that emerged in our analysis is HIPAA. Conversations around this theme are surrounded by concepts related to Protected Health Information (PHI), law, compliance, guidance, and necessary in the COVID-19 pandemic context. For instance, one discourse is upon how healthcare providers may disclose PHI about a COVID-19 patient to first responders (e.g., paramedics, medical transport personnel) and public health authorities in compliance with the HIPAA [16]. In line with this HIPAA law and PHI discourses, another data protection related concept that emerged from the text is 'minimum necessary'. This means healthcare providers "must make a reasonable effort to disclose only the 'minimum necessary' PHI to accomplish the purpose" in the context of COVID-19 [17]. Contributing to this discourse, a healthcare attorney has emphasized practicing minimum necessary in all states of the US: "*Hospitals, regardless of what state they're based in, should only share the minimum necessary information with coroners and medical examiners to ensure compliance with HIPAA*"[18].

3.2.3. Privacy

At the center of the concept map (see Figure 2), Privacy, as the third theme, shapes discourse on requirements, healthcare, patient, medical concepts in the time of crisis. While the OCR temporarily easing some regulatory requirements during the public health emergency, the discourse analysis shows perspectives on protecting health data privacy. The followings are examples from two attorneys' viewpoints concerning protecting data privacy during the pandemic: a) "*Although OCR is easing requirements during this public health emergency, healthcare providers must nevertheless take steps to ensure the privacy of their patients' medical information*" [19] and b) "Remember

that even with waivers and relaxed *requirements*, OCR still expects HIPAA compliance" [14].

3.2.4. Patients

The fourth theme (Patients) includes concepts regarding the use of new models of service delivery to patients such as remote communication technologies and platforms, especially those that have an affordance for video consultation. As in pandemic context under HIPAA, providers have more opportunities to change their business model, understanding how effectively to use digital technology in a compliant manner is critical. Regarding this discourse, a health information specialist said:

"*In cases in which the provider uses an allowed video messaging application like Facebook Messenger, it is best for staff to use a company account rather a personal account in order to prevent sharing a provider's personal information. [...] all virtual care provided should be documented like any other clinical visit*" [15].

Also, in using such third-party platforms, informed consent should be considered. For instance, one report states: "*It is important that healthcare providers notify patients that the third-party communication platforms that they are using to communicate may introduce privacy risks*" [19].

Other concerns are rising from the use of certain public video call platforms which are not authorized by the OCR for telehealth service delivery such as TikTok and Facebook live [20].

3.2.5. Security

The security theme in the analysis is people-oriented concepts in an organizational setting such as employees, employers. Like the previous theme (Patient), in this theme, we see discourses related to the use of technology. However, the main concern, here, is on the security aspects and human errors in an organizational context such as the use of telehealth and remote working. For example, experts have expressed concerns about cyber awareness and error in using technology by healthcare employees during the pandemic and potential data breaches:

"*Healthcare administrative staff working remotely sharing a picture of their new home office set-up on social media with protected information visible on their computer monitor or within paperwork on their desk*", said a director of healthcare consulting as an example of human error threatening to privacy and security [14]. On the discussions about remote working and cyber awareness, an expert recommended, "*not opening any emails with COVID-19 in the title unless a user knows exactly who sent the message*" [15].

Another security discourse elaborated in this theme is related to preparations for the post-pandemic. For example, a regulatory compliance expert suggests "*Covered entities should think about the transition back to meeting all HIPAA requirements when the public emergency is removed, particularly with respect to telehealth security requirements*" [14].

3.2.6. Need

The final theme in our analysis comprises two connected concepts of need and staff in which need is the dominant concept. As evident in Figure 2 and Table 1, a few discourses are related to this theme. From our analysis, we identified two types of

needs discussed concerning the staff: understanding the need for informed use of technology, and HIPAA compliance mobile apps which are minimizing the risk of COVID-19 exposure for staff. The followings are illustrative examples of the concepts of informed use, e.g., "*People may be surfing the web for any information about pandemic, so staff members need to be quite vigilant about inadvertently clicking on phishing-related links*" [15], and HIPAA-compliant mobile apps, e.g., "*We realized that customers where putting their staff at risk by using iPads and tablets in COVID-19 patient rooms with standard, unsecure video solutions*". "*The new features enable secure remote calling into the room without the need for a staff member to enter the room to answer or launch the call*" [21].

4. Discussion

Our research is inspired by two calls for research. The WHO calls for "*powering research to control the epidemic*". Likewise, the Information Systems discipline scholars call for studying epidemics "*based on the position on the social-technical continuum*" [22].

Drawing on discourses on HIPAA and Covid-19 pandemic, we discussed six themes and important concepts such as telehealth, minimum necessary PHI, and compliance issues in the public health emergency context. Our study extends and contributes to the existing discourse on data privacy [23]. We also offer fresh insights regarding the regulations of IT artifact and data processing in the pandemic context [24] As evident in our analysis of texts in media, an important practical implication is the centrality of the notion of privacy in the pandemic context. We highlighted the importance of including concepts such as 'public health' and 'telehealth' in policymaking and regulating data processing in the digital health era.

The global Covid-19 pandemic has accelerated the widespread adoption and use of health technologies such as telehealth [2, 25]. Inevitably, technology and data privacy regulations such as HIPAA have to adapt and adjust. Our study demonstrates that discourse analysis and computational methods such as concept mapping with Leximancer can be practical tools in identifying and analyzing important issues as reflected in public media. This approach can potentially be used by regulators for sensing and responding to public demands in an efficient and effective manner.

Gaining insights from the lens of media discourses, this research further sheds light on the complexity of data protection in the public health context. We also expect our results inform the significance of health data privacy protection, regulations, and policy-making. Finally, the concept mapping revealed in the discourses provides an indication and road map of future opportunities in data privacy research in the current pandemics and beyond.

Inherently, our research has limitations. Our identification of the HIPAA related themes is dependent on the sample data from an international news database. Understanding news media discourse is insightful for both academics and policy makers. Nevertheless, our findings are inherently biased toward journalistic perspectives and genres of text. To provide complementary perspectives, future research may analyze other sources such as peer-reviewed journals in *IEEE Xplore* and *PubMed* databases that cater to academic and specialist audiences. Future research can also employ other methods of text analysis, such as sentiment analysis, to enrich our collective understanding of the discourse formed around HIPAA and Covid-19.

References

[1] World Health Organization (WHO). COVID-19 Public Health Emergency of International Concern (PHEIC) Global research and innovation forum: Towards a research roadmap 2020 [27/01/2021]. Available from: https://www.who.int/publications/m/item/covid-19-public-health-emergency-of-international-concern-(pheic)-global-research-and-innovation-forum.
[2] Keesara S, Jonas A, Schulman K. Covid-19 and health care's digital revolution. New England Journal of Medicine. 2020;382(23):1-3.
[3] Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. Journal of the American Medical Informatics Association. 2020;27(6):957-62.
[4] The HHS Office for Civil Rights (OCR). HIPAA and COVID-19 2021 [27/01/2021]. Available from: https://www.hhs.gov/hipaa/for-professionals/special-topics/hipaa-covid19/index.html.
[5] Shachar C, Engel J, Elwyn G. Implications for Telehealth in a Postpandemic Future: Regulatory and Privacy Issues. JAMA. 2020;323(23):2375-6.
[6] Tucciarello C. OCR's Relaxed Enforcement of HIPAA During COVID-19 Paves The Way For Increase in Telehealth Services: National Law Review; 2020 [27/01/2021]. Available from: https://www.natlawreview.com/article/ocr-s-relaxed-enforcement-hipaa-during-covid-19-paves-way-increase-telehealth.
[7] Barraza L, Collmer V, Meza N, Penunuri K. The legal implications of HIPAA privacy and public health reporting for correctional facilities. Journal of Correctional Health Care. 2015;21(3):213-21.
[8] Goldstein MM, Pewen WF. The HIPAA Omnibus Rule: implications for public health policy and practice. Public Health Reports. 2013;128(6):554-8.
[9] Herzhoff J, editor Convergence and mobility–Just another fad or fashion? A systems-theoretical analysis. European Conference on Information Systems (ECIS); 2010; Pretoria, South Africa: AIS Electronic Library.
[10] Foley K, McNaughton D, Ward P. Monitoring the 'diabetes epidemic': A framing analysis of United Kingdom print news 1993-2013. PloS one. 2020;15(1):e0225794.
[11] Bednarek M, Caple H. Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. Discourse & Society. 2014;25(2):135-58.
[12] Leximancer. We aim for rapid reproducible analysis of complex text data, neither reductionist nor biased 2021 [cited 29/01/2021]. Available from: https://info.leximancer.com/science.
[13] Department of Health and Human Services (HHS). Notification of Enforcement Discretion for Telehealth Remote Communications During the COVID-19 Nationwide Public Health Emergency 2020 [cited 28/01/2021]. Available from: https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/notification-enforcement-discretion-telehealth/index.html.
[14] Schreiber SW. COVID-19 Changes HIPAA Compliance, But Caution Necessary- Healthcare Risk Management. Mondaq Business Briefing. 2020.
[15] Worth T. Working Remotely During the COVID -19 Pandemic. Renal & Urology News. 2020.
[16] Smith GRL. HHS Office For Civil Rights Issues Guidance To Help Ensure First Responders And Others Receive Protected Health Information About Individuals Exposed To COVID -19. Mondaq Business Briefing. 2020.
[17] Marotta KA. Friendly Reminder: HIPAA Still Applies During The Coronavirus Outbreak. Mondaq Business Briefing. 2020.
[18] Cohen JK. Complex web of HIPAA, state rules may impede data gathering on COVID-19 deaths. Modern Healthcare. 2020.
[19] Hoar S. OCR Announces HIPAA Telehealth Security Waiver In Response To COVID -19 Pandemic. Mondaq Business Briefing. 2020.
[20] Feldman JW. COVID-19: OCR HIPAA Enforcement Discretion For Telehealth. Mondaq Business Briefing. 2020.
[21] Harman K. Caregility Announces New Features in HIPAA Compliant iConsult Mobile App to help in fight against COVID-19. PR Newswire. 2020.
[22] Sarker S, Chatterjee S, Xiao X, Elbanna A. The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance. Mis Quarterly. 2019;43(3):695-720.
[23] Dulipovici A, Baskerville R. Conflicts between privacy and property: The discourse in personal and organizational knowledge. The Journal of Strategic Information Systems. 2007;16(2):187-213.
[24] Hassandoust F, Akhlaghpour S, Johnston AC. Individuals' privacy concerns and adoption of contact tracing mobile applications in a pandemic: A situational privacy calculus perspective. Journal of the American Medical Informatics Association. 2020:1-9.
[25] Thomas EE, Haydon HM, Mehrotra A, Caffery LJ, Snoswell CL, Banbury A, et al. Building on the momentum: Sustaining telehealth beyond COVID-19. Journal of telemedicine and telecare. 2020:1-18.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210092

Mobile Motion Tracking for Disease Prevention and Rehabilitation Using Apple ARKit

Lara Marie REIMER[a,1], Severin WEIGEL[a], Florian EHRENSTORFER[a], Malintha ADIKARI[a], Wolfgang BIRKLE and Stephan JONAS[a]
[a] *Technical University of Munich, Munich, Germany*

Abstract. Background: Physical activity helps improve the overall quality of life. The correct execution of physical activity is crucial both in sports as well as disease prevention and rehabilitation. Little to no automated commodity solutions for automated analysis and feedback exist. Objectives: Validation of the Apple ARKit framework as a solution for automatic body tracking in daily physical exercises using the smartphones' built-in camera. Methods: We deliver insights into ARKit's body tracking accuracy through a lab experiment against the VICON system as Gold Standard. We provide further insights through case studies using apps built on ARKit. Results: ARKit exposes significant limitations in tracking the full range of motion in joints but accurately tracks the movement itself. Case studies show that applying it to measure the quantity of execution of exercises is possible. Conclusion: ARKit is a light-weight commodity solution for quantitative assessment of physical activity. Its limitations and possibilities in qualitative assessment need to be investigated further.

Keywords. mHealth, Mobile Applications, Fitness Trackers, Augmented Reality

1. Introduction

Regular physical exercise is known to be beneficial for the overall quality of life. It improves functional capacity and reduces long-term risks for diseases like Diabetes mellitus Type 2 and Alzheimer's while improving overall health, health-related aspects, and mood [1, 2, 3]. Physical exercise can also positively influence hemic and oncological diseases' treatment and rehabilitation by preventing muscular atrophy and improving the patients' mood [4, 5, 6]. However, the correct execution of physical exercise is essential. Wrong exercise execution can lead to biochemical stress, injuries, and osteoarthritis in the respective joints [7]. To avoid wrong movements, regular supervision of the exercising person by experienced personnel is crucial. Various systems have been developed and evaluated to allow a more profound analysis of human motion and detect problematic movements in exercises. These systems include optical, magnetic, inertial, or mechanic sensors to detect and measure different kinds of metrics about human body motion [8, 9]. Modern motion capture systems track indoor and outdoor activities with different accuracy based on the underlying technologies [10]. However, none of those

[1] Corresponding Author: Lara Marie Reimer, Technical University of Munich, Germany, E-Mail: reimer@tum.de

systems currently allows daily usage as they require custom hardware for tracking and often are connected to high costs.

Many people track their daily lives with mobile devices like smartphones, -watches, and sleep or fitness trackers [11]. These devices are equipped with more advanced sensors such as camera systems, gyroscopes, accelerometers, and optical sensors. Also, the devices themselves are getting more powerful and allow resource-intensive applications, e.g., for on-device machine learning. These advances include powerful software applications to perform automatic detection of objects, among them the human body. Software frameworks for mobile devices like Apple ARKit[2] or Vision[3] or Google's Tensorflow Pose Estimation[4] automatically detect the joints' positions through image recognition techniques. These positions can be used for further analysis and enable various new applications, e.g., in mobile games and healthcare.

Using motion detection applications on mobile devices could add to the user's health, allowing quantitative analysis of physical activity. The application area is broad and includes prevention of injuries and sickness, progress, and aggravation tracking, e.g., in rehabilitation, during treatment, or as a motivation to be more active and exercise in general [11]. Mobile devices might be able to deliver a low-cost, commodity alternative to established solutions. Boulos et al. state that while many mobile applications for tracking physical activity have been developed, most of them rely on few sensor data. Used data is often limited to GPS and heart rate. Especially in the area of resistance exercises, personalized coaching applications are missing [12]. However, systems based on wearable sensors can recognize and count exercises [13], even without prior, exercise-specific training [14]. Current research shows that approaches using Apple ARKit can track the lower extremities' motion [15]. A more detailed evaluation of the applicability of ARKit, including the upper extremity, is missing. The aim of this work is to evaluate the suitability and applicability of on-device motion tracking using Apple ARKit and give an outlook to their application in three different use cases: goalkeeper training, physiotherapy and resistance training, and golf.

In all three use cases, complex motion needs to be tracked. If executed incorrectly, these exercises can overstress several joints and promote injuries. The usage of mobile motion capture to assess motion and detect incorrect movements could balance training, prevent injuries, and provide progress supervision. Existing approaches in goalkeeper training build on custom sensors to analyze motion [16]. In our approach, we rely on ARKit as an alternative approach using a commodity device (Case Study 1). Resistance training plays a substantial role in improving and maintaining physical strength and fitness, either during the prevention of diseases and injuries or physiotherapy. Research proves that active physical exercise is among the crucial factors in physiotherapy of common conditions, e.g., in chronic low back pain [17]. We design an algorithm to capture and assess body-weight exercises and provide individual feedback in a mobile coaching application in resistance training (Case Study 2). Golf consists of complex motion sequences (intra- and intermuscular) that require a high amount of coordination, flexibility, and strength, which needs to be practiced continuously to generate high performance. Even for professional golf trainers, an individual assessment of the golf swing and related exercises and tracking progress is challenging. Several solutions have been developed to support golf training. Existing solutions use different sensors attached

[2] https://developer.apple.com/augmented-reality/arkit/
[3] https://developer.apple.com/documentation/vision
[4] https://www.tensorflow.org/lite/models/pose_estimation/overview

to the body and the golf club to measure motion [18, 19] or focus on camera-based techniques, like the Coach's Eye application[5], which allows calculating joint angles in a single, selected video frame manually. We propose a more light-weight, automated approach by combining an optical sensor and a single Inertial Measurement Unit (IMU) sensor for qualitatively assessing motion in golf-related exercises (Case Study 3).

2. Methods

Within the scope of this paper, we aim at investigating two research questions:

- **RQ 1**: Which accuracy does Apple ARKit provide in contrast to the VICON system?
- **RQ 2**: What are potential use cases for an ARKit-based system?

To answer **RQ 1**, we performed a lab experiment in which we compared motion data generated by Apple ARKit against the VICON system's motion data. To answer **RQ 2**, we are conducting several case studies in different health and exercise science areas, each of them consisting of a mobile prototype application running the ARKit framework.

2.1. Suitability of ARKit for Motion Recognition and Tracking

A total of 12 subjects participated in the experiment, 5 females and 7 males, all of good health without physical impairments. The participants' height ranged from 1.56m to 1.96m and their weight from 52.2kg to 97.5kg. In the lab experiment, subjects had to perform 9 different exercises focusing on both the upper and lower extremities, including running on a treadmill at 3 different speeds, passing and catching a ball, jumping jacks, cuttings, and squats. All joint angles are calculated using Euler angles in 3 dimensions: x referring to flexion/extension, y referring to inversion/eversion, and z referring to the rotation. Both systems measured shoulder, elbow, neck, knee, and ankle angles. We calculated minimum, maximum, mean, standard deviation (SD), and range of motion (ROM) and compared the ARKit values against the VICON angles for all exercises.

For the study setup, we used a 10-camera VICON setup with the full-body Plug-in Gait model provided by VICON Nexus[6]. For the recording with ARKit, we placed an iPad Pro 11" 2020 with LiDAR sensor on a tripod in front of the subject with a distance of 3m.

2.2. Case Study 1: Recognition of Exercises in Goalkeeper Training

In our first case study, we wanted to identify patterns specific for individual exercises, which is the baseline for further analysis of the motion. For this, we recorded 15 different exercises with ARKit and an additional camera, which served as the basis for labeling the data afterward. The exercises consisted of 6 dive variations, 5 catch variations, 2 throw variations, and 2 kick variations. 4 goalkeepers were recorded. Their football trainer was supervising the recording sessions to guarantee the correct execution of the exercises. Through 10 session recordings in total, we created a dataset of 1050 single exercise executions mapped on the 15 exercises. We used machine learning techniques to train a classifier predicting the matching. We reduced the 6 different dive classes to 2

[5] https://www.coachseye.com
[6] https://docs.vicon.com/display/Nexus26/Full+body+modeling+with+Plug- in+Gait

classes due to similarities in the execution as a preparation. Based on research by Ronao et al. [20], we applied an approach based on Convolutional Neural Networks (CNN).

2.3. Case Study 2: Recognition of Body-Weight Exercises in Physiotherapy and Resistance Training using a Pose-based Approach

In case study 2, we developed an algorithm to recognize and assess body-weight exercises using ARKit. The algorithm allows app users to record new exercises, store them in a database, detect correct exercise executions of stored exercises, detect incorrect repetitions, and provide individual feedback to the user on improving the motion. Due to the COVID-19 pandemic, an evaluation of the prototype application remains open.

2.4. Case Study 3: Classification of Golf-Related Exercises

As part of case study 3, we developed a system consisting of an ARKit-based mobile application prototype and an IMU sensor to measure motion, rotation, and acceleration in an exercise related to the golf swing. The camera is placed in front of the player. The IMU sensor is attached to the hip. The system automatically assesses the exercise execution compared to a reference recording. The automatic assessment is performed by a Dynamic Time Warping (DTW) algorithm, which compares the exercise data against the reference data and evaluates the distance between the compared time series. Through the DTW approach, we classify the executions and provide a grading. We recorded around 300 exercise executions, and a golf trainer assessed them to generate training data. A prototype was implemented, but the automatic assessment could not be qualitatively validated due to the ongoing COVID-19 pandemic. Through an online questionnaire, we gained insights into how golf trainers assessed the prototype's usefulness.

3. Results

3.1. Suitability of ARKit for Motion Recognition and Tracking

The comparison of the Euler angles calculated by VICON against the Euler angles calculated based on the ARKit data exposed several deviations between the two systems across all tracked joints (see Table 1 for one joint). Across all joints and dimensions, the ARKit measurements revealed smaller SD and ROM values than the VICON measurements. Additionally, ARKit and VICON use different reference points to calculate the Euler angles, which led to mirrored values in the analysis (Figure 1b).

Table 1. Results of the VICON and ARKit measurements in the squat exercise. The values represent the knee joint angles. All values are calculated in degrees.

System	Angle	Dimension	Min	Max	Mean	SD	ROM
VICON	Knee	x	0.57	119.61	46.35	42.45	119.04
	Right	y	-3.60	16.52	5.86	6.73	20.12
		z	-22.86	36.30	1.53	20.97	59.16
	Knee	x	0.00	117.98	46.58	42.22	117.97
	Left	y	-3.50	13.77	4.46	5.76	17.27
		z	-17.14	41.08	7.69	20.29	58.22
ARKit	Knee	x	23.15	54.37	29.44	6.39	31.21
	Right	y	-9.07	1.28	-2.81	2.13	10.34
		z	-2.54	18.12	4.66	4.45	20.66
	Knee	x	22.77	55.88	28.79	7.11	33.11
	Left	y	-2.08	5.79	1.66	1.48	7.87
		z	-19.85	3.00	-5.20	4.31	22.85

Even though the results exhibited a considerably smaller range of motion, the motion itself was detected in every repetition of the exercises (Figure 1). The ROM was consistently smaller throughout all repetitions in all exercises, supported by the smaller

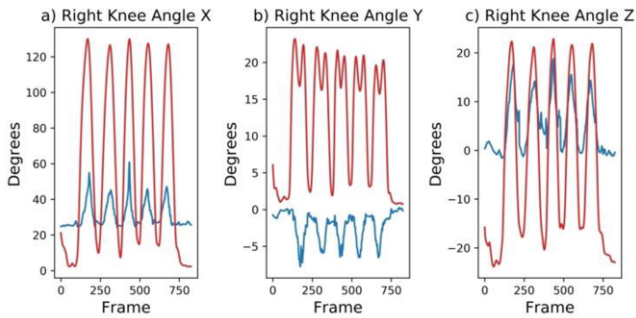


Figure 1. Knee joint angles during squat by ARKit (blue) and VICON (red).

SD values. We calculated the Pearson correlation-coefficients for shoulder, elbow, knee, and ankle angle for all participants for the squat exercise for further analysis. We calculated the mean and SD values across all participants, dimensions, and left and right sides based on the coefficients. We used the correlation coefficients' absolute values for the correct detection of the motion, as due to the different reference points of VICON and ARKit, the values of the y dimension of ARKit were mirrored for the joints on the right side of the body. The mean correlation-coefficients and SD for the shoulder, elbow, knee, and ankle angles were 0.406 ± 0.236, 0.085 ± 0.117, 0.705 ± 0.144, and 0.654 ± 0.184, respectively.

Table 2. Pearson correlation-coefficients of the VICON and ARKit measurements for all 12 participants in the Squat exercise.

	Shoulder		Elbow		Knee		Ankle	
	L	R	L	R	L	R	L	R
Min	0.017	0.011	0.000	0.000	0.297	0.410	0.222	0.207
Max	0.801	0.825	0.658	0.447	0.901	0.898	0.909	0.911
Mean	0.367	0.439	0.084	0.085	0.674	0.736	0.614	0.693
SD	0.227	0.241	0.128	0.104	0.147	0.134	0.173	0.186

3.2. Case Study 1: Recognition of Exercises in Goalkeeper Training

Figure 2 visualizes the output of the 3-dimensional positional data of a session with recordings of 5 exercises for the right hand and right foot. The different patterns are identifiable, as well as the repetitions of each exercise.

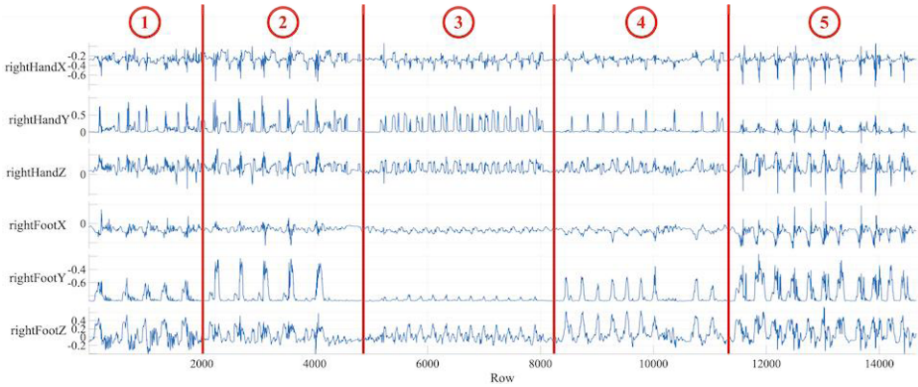


Figure 2. Dataset of a session including 5 exercises: Dive High Right (1), Dive High Left (2), Catch Hand (3), Catch Body (4), and Catch Ground (5).

Despite the comparably small dataset, the classifier achieved a validation accuracy of around 75% using the CNN in the test set (Figure 3). The dives and the two versions of the jump catch achieved high rates of correct predictions with around 70-80%. For comparison, we tested the same approach with Long Short-Term Memory Networks (LSTM) and Deep Neural Networks (DNN). Due to the small dataset, both alternatives

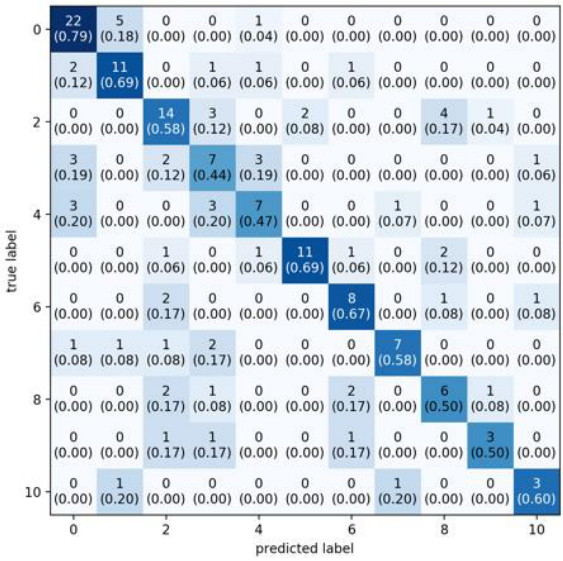


Figure 3. Confusion matrix with precision values of the CNN prediction on 11 classes: (0) Dive Right (1) Dive Left (2) Catch Hand (3) Catch Body (4) Catch Ground (5) Jump Catch (6) Jump Catch Run (7) Throw Low (8) Throw High (9) Side Kick (10) Side Kick Ground.

showed overfitting tendencies and did not reach the accuracy of the CNN with an accuracy of 55% and 65%, respectively, for all classes.

The first case study shows that the data generated by ARKit is accurate enough to distinguish between goalkeeper exercises. Even with a small dataset, a machine learning classifier can classify around 75% of the exercises correctly.

3.3. Case Study 2: Recognition of Body-Weight Exercises in Physiotherapy and Resistance Training using a Pose-based Approach

To build a mobile, ARKit-based application enabling tracking various dynamic body-weight exercises, we designed a workflow for app users to add new exercises to the app. The prerequisite for registering a new exercise is that it has to consist of clearly different poses. An exercise needs to expose motion in at least one trackable joint in three poses as observable in a squat with full extension of both knees, flexion of at least 90deg in both knees, and a full extension of both knees again (Figure 4). The system captures the three poses, including the measured angles. For each pose, a set of essential joints can be defined, e.g., knee, ankle, and hip joint in the squat. This setup should be done by qualified personnel, e.g., a physiotherapist, to ensure the exercise's correctness.

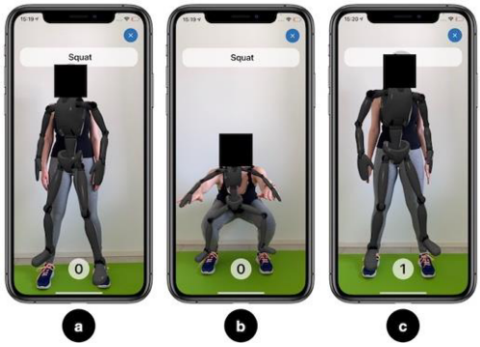


Figure 4. Exercise repetition tracking using the pose-based approach. ARKit's recognition is shown by the robot overlay. ARKit tracks the poses in a, b, and c, and updates the repetition count after pose c is matched.

For exercise recognition, we convert the 3-dimensional positional joint data provided by ARKit into angles. Our proposed algorithm observes every motion detected by ARKit and matches the motion's progression against the curve expected by the exercise. Once all poses have been reached in the correct order, the algorithm considers the repetition as completed and increases the count, as shown in Figure 4c. The algorithm is flexible enough to recognize a variety of body-weight exercises without training a specific machine learning model through this approach.

Considering the initial lab experiment results, approaches on qualitatively assessing exercise repetitions need to be investigated. Using the proposed algorithm allows the creation and tracking of dynamic body-weight exercises based on joint angle calculations.

3.4. Case Study 3: Classification of Golf-Related Exercises

We evaluated the Golf Coach app's usability in an online questionnaire. In total, 22 golf trainers. The trainers' age ranged between 19 to 66 years, with a mean of 38.5 years and a median of 32 years. Their experience as a golf trainer ranged from 1 year to 40 years, with a mean of 13.14 years and a median of 7.5 years. 95.5% stated that they would use

an app to track their trainees' progress, and 70% had used mobile coaching apps before. After a guided tour through the app, the participants were asked whether they would use it. 59% stated that they would use it, 9% would use it from time to time. 9% would prefer to test it first. 5% said that they would not use it. 18% chose not to answer. 55% stated that progress tracking would be the most important feature to them.

4. Discussion & Future Work

The lab experiment and the case studies provide evidence that mobile applications based on ARKit can track joint motion. Table 1 shows that the recognition is not as accurate as the reference values detected by VICON. ARKit does not seem to capture a full extension of the knee, as the minimum value provided is 22.77deg, compared to 0.00deg seen by VICON. Flexion is recognized, but to a much smaller extent than by the VICON system, with 119.61deg maximum in VICON and 54.37deg maximum in ARKit in the same joint. ARKit exposes smaller mean, SD, and ROM values. Similar observations apply to the inversion/eversion and rotation dimensions. The Pearson correlation-coefficients analysis shows that the ARKit and VICON measurements are strongly correlated, especially in the lower extremities. Even smaller changes in inversion/eversion are tracked in the squat motion's turning points, as shown in Figure 1b. Interestingly, the upper extremities' joint angle motion exhibit a considerably low correlation.

The ROM detected by ARKit is comparatively but reliably smaller than the ROM of VICON. This poses the question of whether an algorithm can be developed, which can approximate the ARKit values to the correct values provided by the VICON system. Enabling this would allow performing not only a quantitative motion analysis but also a qualitative analysis. A qualitative analysis would allow additional possibilities in motion analysis, mainly regarding feedback to prevent wrong motion.

Motion tracking using ARKit, especially regarding the lower extremity, seems to be a promising, light-weight approach. Even though the lab experiment shows that a quantitative assessment of exercise executions, e.g., for repetition counting, is feasible, the experiment was performed with a relatively limited number of participants and different exercises. A more extensive experiment is needed to gain further insights, which includes a more diverse set of participants and exercises. Through this, the limitations of the ARKit-based tracking need to be further investigated.

In case study 1, we showed that we can use ARKit data as reference values to recognize patterns in recordings. It remains open to which extent we can use reference data provided by ARKit to allow qualitative analysis of the motion, as proposed in case studies 2 and 3. The case studies served to explore the possibilities of ARKit-based motion tracking in health-related fields. Even though the first results show a high interest of potential app users and the lab experiment and data analysis expose such approaches' potential, the case studies need to be tested and validated in more extensive studies.

5. Conclusion

In this paper, we were able to show that even though ARKit exposes major inaccuracies in tracking the ROM, it is reliable in tracking the motion itself. Therefore, the ARKit framework can be used to assess physical exercise, recognize exercises, and count repetitions. ARKit enables various use cases for mobile applications, especially in the

prevention and rehabilitation of diseases and injuries in humans. In this paper, we presented 3 different case studies. In case study 1, we classified exercises in goalkeeper training, which shows that we can identify patterns specific to exercises in the data provided by ARKit. In case study 2, we presented an algorithm that enables the creation and automatic tracking of dynamic body-weight exercises. In case study 3, we combined ARKit with an IMU sensor to enable a qualitative assessment of an exercise related to the golf swing.

ARKit seems to be a promising, light-weight alternative to well-established motion tracking systems. Its limitations and possibilities need to be further investigated.

References

[1] W. B. Kannel, P. Sorlie, Some Health Benefits of Physical Activity: The Framingham Study. *Arch Intern Med*, **139**(8) (1979), 857–861.

[2] A.P.A. Penedo and F. Ja, Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr Opin in Psychiatr*, **18** (2005), 189 – 193.

[3] D.E.R. Warburton, C.W. Nicol, S.S.D. Bredin, Health benefits of physical activity: the evidence. *CMAJ*, **174**(6) (2006), 801-809.

[4] J. Midtgaard, M. T. Baadsgaard, T. Møller, B. Rasmussen, M. Quist, C. Andersen, M. Rørth, and L. Adamsen. Self-reported physical activity behaviour; exercise motivation and information among Danish adult cancer patients undergoing chemotherapy. *Eur J of Oncol Nurs*, **13**(2) (2009), 116 – 121.

[5] L. E. Stephenson, D. G. Bebb, R. A. Reimer, S. N. Culos-Reed, Physical activity and diet behaviour in colorectal cancer patients receiving chemotherapy: associations with quality of life. *BMC Gastroenterol*, **9**(1) (2009).

[6] C. Visovsky. Muscle Strength, Body Composition, and Physical Activity in Women Receiving Chemotherapy for Breast Cancer. *Integr Cancer Ther*, **5**(3) (2006), 183–191.

[7] R.H. Sandmeier, Osteoarthritis and Exercise: Does Increased Activity Wear Out Joints?. *Perm J*, **4**(4) (2000), 26-28.

[8] M. Yahya, J.A. Shah, K.A. Kadir, Z.M. Yusof, S. Khan, A. Warsi, ""Motion capture sensing techniques used in human upper limb motion: a review"", *Sensor Rev*, **39**(4) (2019), 504-511.

[9] M. Field, Z. Pan, D. Stirling, F. Naghdy, Human motion capture sensors and analysis in robotics, *Ind Robot*, **38**(2) (2011), 163-171.

[10] E. van der Kruk, M.M. Reijne, Accuracy of human motion capture systems for sport applications; state-of-the-art review, *Eur J Sport Sci*, **18**(6) (2018), 806-819.

[11] O.D. Lara, M.A. Labrador. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Commun Surv Tut*, **15**(3) (2013), 1192-1209.

[12] M. N. Kamel Boulos, S.P. Yang. Mobile physical activity planning and tracking: a brief overview of current options and desiderata for future solutions. *mHealth*, **7**(13).

[13] S. Ishii, A. Yokokubo, M. Luimula, G. Lopez. ExerSense: Physical Exercise Recognition and Counting Algorithm from Wearables Robust to Positioning. *Sensors,* **21**(1), 91.

[14] S. Zelman, M. Dow, T. Tabashum, T. Xiao, M.V. Albert, Accelerometer-Based Automated Counting of Ten Exercises without Exercise-Specific Training or Tuning. *J Healthc Eng*, *2020*, 8869134.

[15] S. Basiratzadeh, E. D. Lemaire, N. Baddour, Augmented Reality Approach for Marker-based Posture Measurement on Smartphones, *Annu Int Conf IEEE Eng Med Biol Soc*, 2020, 4612-4615.

[16] J. Haladjian, D. Schlabbers, S. Taheri, M. Tharr, and B. Bruegge, Sensor-based Detection and Classification of Soccer Goalkeeper Training Exercises. *ACM Trans. Internet Things*, **1**(2) (2020).

[17] M. van Middelkoop, S.M. Rubinstein, T. Kuijpers, A.P. Verhagen, R. Ostelo, B.W. Koes, M.W. van Tulder, A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non‐specific low back pain. *Eur Spine J*, **20** (2011), 19–39.

[18] H. Ghasemzadeh, V. Loseu, R. Jafari. Wearable coach for sport training: A quantitative model to evaluate wrist-rotation in golf. *J Ambient Intell Smart Environ*, **1**(2) (2009), 173–184.

[19] U. Jensen., M. Schmidt, M. Hennig, et al., An IMU-based mobile system for golf putt analysis. *Sports Eng*, **18** (2015), 123–133.

[20] C.A. Ronao, S.B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, *Expert Syst Appl*, **59** (2016), 235-244.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210093

Pathway-Supporting Health Information Systems: A Review

Tim SCHEPLITZ[a,1]
[a] *Technische Universität Dresden, Dresden, Germany*

Abstract. Care pathways and supporting health information systems (HIS) have been permeate the discipline of Health Information Systems Research (HISR) over years. Traditional objectives of workflow assistance are increasingly extended by interdisciplinary goals from technology, medicine, management and public health research. A systematic literature review is dedicated to this integrating character. It examines the interdisciplinary mesh of objectives associated with care pathways and pathway-supporting HIS in the HISR literature. From 47 identified articles, 6 thematic themes were derived. Their consolidation supports in particular design and development processes as it describes the solution space of future pathway-supporting HIS addressing requirements stated by multiple stakeholders.

Keywords. Care pathways, health information system, pathway systems, review

1. Introduction

In medical practice and research care pathways describe complex sequences of interventions of defined patient groups in defined time periods to support the organization, coordination and decision making of care processes [1]. Different terms are used to emphasize intra-organizational ("clinical pathways"), inter-institutional ("integrated pathways") or patient-centered ("patient pathways") orientations [2]–[4]. Also Health Information Systems Research (HISR) investigates care pathways especially the conception, modeling, realization, and impact of pathway-supporting Health Information Systems (HIS) [5]–[7].

In this context, the research and development work seek to bring original motivations from the process perspective in line with the requirements of innovative disciplines. For example, it is being discussed how patient integration can be intensified along care pathways [7], how data mining methods can describe care pathways retrospectively [8], or how data analytics approaches can contribute to the individualization of care plans [9]. Experiences from practice-oriented digital health projects underline the observation that the objectives of care pathways and pathway-supporting HIS go beyond the traditional process support.

This paper follows up on these observations. If research and development teams want to master the multidisciplinary field of requirements for pathways-supporting HIS, a consolidation of previous work can support them especially in conceptual design and development process. Such a compilation can describe the solution space for pathways-

[1] Corresponding Author: Tim Scheplitz, Technische Universität Dresden, Germany, tim.scheplitz@tu-dresden.de

supporting HIS, derive design implications, and point to disciplines and professions to be involved. This review paper therefore addresses the question of how the literature discusses scope and challenges of pathway-supporting HIS and which contributions can be used to face the mentioned mission of multidisciplinary requirements engineering. The review thus extends the knowledge from previous reviews on characterization of patient pathways [2] and on support opportunities of clinical pathways by Health Information Technologies [10].

2. Methods

A systematic literature review [11]–[13] for the mentioned research question was conducted in the last quarter of 2020. The following high-quality databases of ISR or HISR were chosen: AIS Senior Basket, Proceedings of the AIS Conferences, recommended eHealth Journals of the AIS SIG Health. The past decade was chosen as the publication period. After abstract and full-text screening, 47 articles were defined as final set for analysis. Only articles that prominently named care pathways as a research context were included. Review articles were not excluded from the analysis set in order to extract implications of their contributions to research and practice. Further details of the search process are given in Figure.1.

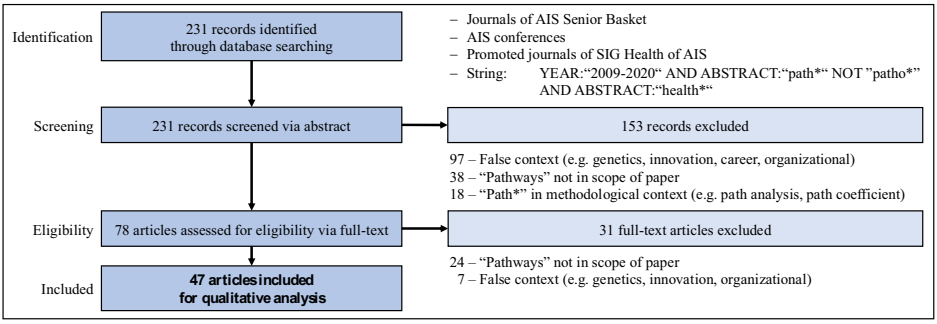


Figure 1. Review process according to PRISMA guidelines [13]

Following the process model of a summarizing content analysis with inductive category formation [14], the analysis material was reduced to those passages in which central goals or contributions related to care pathways or pathway-supporting HIS are described. Paraphrases were constructed and structured interpretatively. Through this structuring step, six themes were identified and validated in group session with three digital health experienced researchers.

3. Results

The 47 identified articles are distributed relatively evenly over the past years. With regard to the distribution across journals or conferences, it can be noted that the majority of the articles found are published in HISR journals recommended by the AIS Health SIG (n=35). Here, 13 articles could be found in the *International Journal of Medical Informatics* and 11 papers in the journal *BMC Medical Informatics and Decision Making*.

A total of 11 papers were identified in the proceedings of the AIS conferences, and only one article in the AIS Senior Basket.

Table 1 presents exemplary the results of paraphrasing step of the six articles published in 2020. It primarily shows how paraphrases have been formulated and offers secondly an insight about newest research on or with pathway-supporting HIS. A complete list of the paraphrasing results of all identified articles will be provided in additional material to this paper.

Table 1. Paraphrases examples of article analysis (only articles published in 2020)

Article [reference]: Summary of contribution's intersection to pathway-supporting HIS
Askari et al. 2020 [15]: Study for professional assessment of the effectiveness and efficiency of clinical pathways and path-supporting HIS in medical care
Cho et al. 2020 [16]: Development of a data-driven method for deriving clinical pathways from electronic health records
Gaveikaite et al. 2020 [17]: Case study along a COPD pathway for potential analysis of telemedicine
Kempa-Liehr et al. 2020 [18]: Modeling tool for clinical pathways and prediction models based on data mining methods and machine learning models for individual improvement of recovery time
Trajano et al. 2020 [19]: Process oriented modeling language for clinical pathways
Ye et al. 2020 [20]: Feasibility study on the use of deep learning prediction models for decision support and derivation of optimal, individual hypertension pathways based on electronic health records

Six themes have been inductively identified out of the paraphrasing results of all 47 articles. Some articles refer to up to two different themes due to their complexity (e.g. [18]). Table 2 shows all themes and sub-topics to give a referenced overview of pathway-related HISR of the last decade. The following additions highlight selected findings or implications for future design and implementation activities of pathway-supporting HIS.

I. Design, development and implementation of pathway-supporting HIS: This theme includes generic approaches as well as descriptions of pathway-supporting HIS and application systems for specific care scenarios. Such work has been established for years. The traditional workflow support of clinical processes is increasingly complemented by articles that present solutions for more complex, cross-institutional and patient-integrating care scenarios. Future work of this theme should concisely characterize the care scenario and parameterize targeted improvements to care-related goals or further outcomes in order to increase reusability.

II. Evaluation and assessment of pathway-supporting HIS: Theme II papers investigate the effects of pathway-supporting HIS. Increasingly complex care scenarios require future evaluative papers to consider all involved stakeholders more comprehensive (e.g. multidisciplinary care teams, patients, expanded health care market) and to discuss competing or synergistic effects and measurable outcomes.

III. Modeling and modeling languages of care pathways: Research with a focus on pathway modeling, modeling languages and tools discuss less the design and realization of application systems but address the underlying pathway models and their genesis. The established knowledge base provides already valuable guidance. However, the changing scope to integrated care scenarios with by multiple professional stakeholders and increasing technization may require additional work in this theme to offer appropriate techniques for precise, comprehensive and consistent pathway models.

IV. Data-driven pathway models and integration of data-based prediction models: Recent advances in data science, machine learning, and related disciplines drive articles in this young theme (earliest publication of this review from 2017). It includes articles that discuss the data-based derivation of care pathways from existing data, e.g.

of EHR. Such retrospective pathway analyses provide the opportunity to compare originally defined care plans with de facto care pathways and to investigate deviations. In consequence, those knowledge might be used to optimize individual care pathways or general pathway templates prospectively or to improve medical guidelines (e.g. for care quality or efficiency). On the other hand, this theme also includes discussions of how pathway-supporting HIS can provide the data basis for data-driven medical decision support systems or management-oriented predictive models (linkage to Theme V). Both sub-themes - pathway-supporting HIS as a source and as a sink of data-driven processing of health information - like to merge and are currently of increasing interest.

V. Conceptual integration of the management perspective: Traditional workflow assistance fosters goals directly related to care processes such as accelerated process flow or lower error rates. Articles of this fifth theme investigate how pathway-supporting HIS can support additional short-, medium- and long-term tasks of healthcare management on micro, meso or macro level (e.g. quality management, resource management, health program management). Future work shall explore how pathway-supporting HIS need to be designed to satisfy this information demand and to offer management decision support. It therefore will probably benefit of popular research efforts of Theme IV.

VI. Care pathways as a means for HISR: Pathway-supporting HIS can be a starting point and data source for diverse research questions. Future design and implementation activities should consider the access to and scientific usability of the processed data and integrate validation and anonymization mechanisms in particular.

Table 2. Analytical results –themes on pathway-supporting HIS

Theme (Number of papers that led to theme): Sub-topics [references]
I. Design, development and implementation of pathway-supporting HIS (n=13): - Development and implementation of clinical pathway for specific healthcare scenarios [21]–[28] - Generic conceptualizations and design recommendations [23], [25], [29]–[34]
II. Evaluation and assessment of pathway-supporting HIS (n=16): - Studies on effectiveness, efficiency and user experience [10], [15], [28], [35]–[43] - Usage analyses for coordination and communication [35], [37], [43]–[46] - Maturity model for care pathways and its implementation in HIS [47]
III. Modeling and modeling languages of care pathways (n=8): - Development of process-oriented modeling languages for care pathways [19] - Conceptualization and modeling approaches [30], [48], [49] - Modeling tools for care pathways [18] - Examples of care pathway modeling (process and final pathways) [22], [24], [40]
IV. Data-driven pathway models and integration of data-based prediction models (n=7): - Modeling methods or tools for care pathways from electronic health records [16], [18], [20], [37], [51] - Development of data-based prediction models (Data & Process Mining, Machine Learning, Deep Learning) for medical decision support [18], [20], [37], [52] - Data-based analysis and decision models from data of pathway-supporting HIS for healthcare and hospital management [50], [51]
V. Conceptual integration of the management perspective (n=5): - Intersection analysis of Information Systems, Operational Research and Industrial Engineering to solve problems related to care pathways [53] - Method conception for the embedding of quality management in care pathways [54] - Path-based data analysis for tactical and strategic hospital management [55] - Conceptualization and modeling approaches for aligning evidence-based Clinical Practice Guidelines and Clinical Pathways [48], [49]
VI. Care pathways as a means for HISR (n=7): - Analysis of the Status Quo of the Digital Transformation [56]–[58] - Analysis of technology support across care pathways [10], [17] - Analysis of key areas via patient flow pathway mapping [59] - Studies on personalization of HIS services [60]

4. Discussion

The review results can be used for different tasks by research and development consortia of pathway-supporting HIS and offers an aid for their objectives. Institutions with a need to catch up in IT-based workflow support can draw on the contributions of *Themes I* and *II* to design and implement their own pathway-supporting HIS. The traditional care process support may seem almost obsolete compared to current HISR contributions. However, these themes remain particularly relevant to healthcare actors with low digital maturity. Contributions from *Theme III* should also be considered in this outlined case to enable the creation of the necessary pathway models as a success-critical resource. Theme III can also support consortia with already established pathway-supporting HIS and contribute to improvements of the used pathway models.

From the perspective of HISR and healthcare organizations of high digital maturity, *Themes IV* and *V* are highlighted in particular. Technological advances in data science and related disciplines indicate enhanced potential for both personalized care and learning systems of a macro level for management and medicine. Future work may generate data-based individual improvements in care pathways (e.g., based on patient constitution or resource availability) or address care economic issues at the local (e.g., clinical process improvements) or global level (e.g., regionally adequate care programs). Future contributions will help pathway-supporting HIS to benefit from the technological progress while providing the required valid and consistent data for these mechanisms.

The contribution of this paper is affected by a certain number of *limitations*. For example, the analysis sample from the selected publication organs offers a solid basis for interpretation, but is still restricted. The integration of additional development-focused journals or conferences could provide more differentiation in the design of novel pathway-supporting HIS. Further, the creation of the six research themes is also subject to the limited objectivity of the author, although this influence was counteracted by group sessions with experienced digital health researchers.

How may *future research* stimulate the progress? With the assumption that concrete care scenarios will continue to be main drivers for technological innovation with and for pathway-support HIS, the relationship between care context and system design could be explored in more detail. Such investigations could determine those care scenarios that motivate the development of pathway-supporting HIS in particular or those which benefit only from dedicated functionalities. Therefore, an appropriate description model is required to define characteristics as well as types of care scenarios , which in turn could be linked to concrete care goals on micro or macro level. Those studies should focus on: care setting (outpatient, inpatient, rehabilitation); medical discipline (e.g., oncology, emergency medicine); indications (e.g., COPD, depression); patient's role (autonomous vs. paternalistic); degree of multi-professionality or number of institutions involved.

5. Conclusion

This review identified 47 articles from highly ranked literature on care pathways or pathway-supporting HIS. Through their interpretation, six themes were derived that represent the range of interdisciplinary goals with and for pathway-supporting HIS. Their consolidation supports future design and development processes by describing the solution space for pathway-supporting HIS. In particular, the design of new systems can benefit from this summary and interdisciplinary requirements management of multiple

stakeholders can be supported. Among those themes, articles with a scope on data-driven pathway models, the integration of data-based healthcare prediction models as well as the enhancement to pathway-supporting HIS for management and operations currently represent exciting fields of activity for HISR and seeks to future work.

References

[1] EPA, "Care Pathways," Eurpean Pathway Assosciation: Care Pathways, 2019. http://e-p-a.org/care-pathways/.
[2] P. Richter and H. Schlieter, "Understanding Patient Pathways in the Context of Integrated Health Care Services-Implications from a Scoping Review," presented at the Wirtschaftsinformatik 2019, Siegen, Feb. 2019.
[3] L. Kinsman, T. Rotter, E. James, P. Snow, and J. Willis, "What is a clinical pathway? Development of a definition to inform the debate," BMC Med, vol. 8, pp. 31–33, 2010.
[4] K. Vanhaecht, M. Panella, R. van Zelm, and W. Sermeus, "An overview on the history and concept of care pathways as complex interventions," International Journal of Care Pathways, vol. 14, no. 3, pp. 117–123, Sep. 2010, doi: 10.1258/jicp.2010.010019.
[5] H. Raphael, T. Lux, and V. Martin, "State-of-the-art prozessorientierter Krankenhausinformationssysteme.," in Wirtschaftsinformatik (2), 2009, pp. 689–698.
[6] M. Burwitz, H. Schlieter, and W. Esswein, "Modeling Clinical Pathways - Design and Application of a Domain-Specific Modeling Language," Wirtschaftsinformatik Proceedings 2013, Jan. 2013, [Online]. Available: http://aisel.aisnet.org/wi2013/83.
[7] M. Benedict et al., "Patientenintegration durch Pfadsysteme," presented at the Wirtschaftsinformatik (WI 2019), Siegen, 2019.
[8] E. S. Prokofyeva, R. D. Zaytsev, and S. V. Maltseva, "Application of Modern Data Analysis Methods to Cluster the Clinical Pathways in Urban Medical Facilities," in 2019 IEEE 21st Conference on Business Informatics (CBI), Moscow, Russia, Jul. 2019, pp. 75–83, doi: 10.1109/CBI.2019.00016.
[9] H. Schlieter, M. Benedict, K. Gand, and M. Burwitz, "Towards Adaptive Pathways: Reference Architecture for Personalized Dynamic Pathways," Jul. 2017, pp. 359–368, doi: 10.1109/CBI.2017.55.
[10] M. T. Neame, J. Chacko, A. E. Surace, I. P. Sinha, and D. B. Hawcutt, "A systematic review of the effects of implementing clinical pathways supported by health information technologies," Journal of the American Medical Informatics Association, vol. 26, no. 4, pp. 356–363, Apr. 2019, doi: 10.1093/jamia/ocy176.
[11] G. Schryen et al., "Literature Reviews in IS Research: What Can Be Learnt from the Past and Other Fields?," Communications of the Association for Information Systems, vol. 41, no. 1, Dec. 2017, doi: 10.17705/1CAIS.04130.
[12] G. Paré, M.-C. Trudel, M. Jaana, and S. Kitsiou, "Synthesizing information systems knowledge: A typology of literature reviews," Information & Management, vol. 52, no. 2, pp. 183–199, Mar. 2015, doi: 10.1016/j.im.2014.08.008.
[13] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and PRISMA Group, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," Ann Intern Med, vol. 151, no. 4, p. 264, Aug. 2009, doi: 10.7326/0003-4819-151-4-200908180-00135.
[14] P. Mayring, Qualitative Inhaltsanalyse: Grundlagen und Techniken, 12., Überarbeitete Auflage. Weinheim Basel: Beltz, 2015.
[15] M. Askari, J. L. Y. Y. Tam, M. F. Aarnoutse, and M. Meulendijk, "Perceived effectiveness of clinical pathway software: A before-after study in the Netherlands," International Journal of Medical Informatics, vol. 135, p. 104052, 2020, doi: https://doi.org/10.1016/j.ijmedinf.2019.104052.
[16] M. Cho et al., "Developing data-driven clinical pathways using electronic health records: The cases of total laparoscopic hysterectomy and rotator cuff tears," International Journal of Medical Informatics, vol. 133, p. 104015, Jan. 2020, doi: 10.1016/j.ijmedinf.2019.104015.
[17] V. Gaveikaite et al., "Challenges and opportunities for telehealth in the management of chronic obstructive pulmonary disease: a qualitative case study in Greece," BMC Med Inform Decis Mak, vol. 20, no. 1, p. 216, Dec. 2020, doi: 10.1186/s12911-020-01221-y.
[18] A. W. Kempa-Liehr et al., "Healthcare pathway discovery and probabilistic machine learning," International Journal of Medical Informatics, vol. 137, p. 104087, 2020, doi: https://doi.org/10.1016/j.ijmedinf.2020.104087.

[19] I. A. Trajano, J. B. F. Filho, F. R. de C. Souza, I. Litchfield, and P. Weber, "MedPath: A process-based modeling language for designing care pathways," International Journal of Medical Informatics, p. 104328, 2020, doi: https://doi.org/10.1016/j.ijmedinf.2020.104328.
[20] X. Ye, Q. T. Zeng, J. C. Facelli, D. I. Brixner, M. Conway, and B. E. Bray, "Predicting Optimal Hypertension Treatment Pathways Using Recurrent Neural Networks," International Journal of Medical Informatics, vol. 139, p. 104122, 2020, doi: https://doi.org/10.1016/j.ijmedinf.2020.104122.
[21] M. Donald et al., "Development and implementation of an online clinical pathway for adult chronic kidney disease in primary care: a mixed methods study.," BMC Medical Informatics & Decision Making, vol. 16, pp. 1–11, 2016.
[22] R. S. Evans et al., "Automated identification and predictive tools to help identify high-risk heart failure patients: pilot evaluation," Journal of the American Medical Informatics Association, vol. 23, no. 5, pp. 872–878, Sep. 2016, doi: 10.1093/jamia/ocv197.
[23] V. Gkatzidou et al., "User interface design for mobile-based sexual health interventions for young people: Design recommendations from a qualitative study on an online Chlamydia clinical care pathway.," BMC Medical Informatics & Decision Making, vol. 15, no. 1, pp. 1–13, 2015.
[24] S. Wagner et al., "Analysis and classification of oncology activities on the way to workflow based single source documentation in clinical information systems.," BMC Medical Informatics & Decision Making, vol. 15, pp. 1–13, 2015.
[25] M.-M. Bouamrane and F. S. Mair, "Integrated Preoperative Care Pathway - A Study of a Regional Electronic Implementation.," BMC Medical Informatics & Decision Making, vol. 14, no. 1, pp. 1–32, 2014.
[26] J. E. Hurwitz, J. A. Lee, K. K. Lopiano, S. A. McKinley, J. Keesling, and J. A. Tyndall, "A flexible simulation platform to quantify and manage emergency department crowding.," BMC Medical Informatics & Decision Making, vol. 14, no. 1, pp. 1–20, 2014.
[27] A. T.-H. Hao et al., "Nursing process decision support system for urology ward," International Journal of Medical Informatics, vol. 82, no. 7, pp. 604–612, 2013, doi: https://doi.org/10.1016/j.ijmedinf.2013.02.006.
[28] A. M. Ryhänen, S. Rankinen, K. Tulus, H. Korvenranta, and H. Leino-Kilpi, "Internet based patient pathway as an educational tool for breast cancer patients," International Journal of Medical Informatics, vol. 81, no. 4, pp. 270–278, 2012, doi: https://doi.org/10.1016/j.ijmedinf.2012.01.010.
[29] J. Gibbs et al., "The eClinical Care Pathway Framework: a novel structure for creation of online complex clinical care pathways and its application in the management of sexually transmitted infections.," BMC Medical Informatics & Decision Making, vol. 16, pp. 1–9, 2016.
[30] W. Li, K. Liu, H. Yang, and C. Yu, "Integrated clinical pathway management for medical quality improvement – based on a semiotically inspired systems architecture," European Journal of Information Systems, vol. 23, no. 4, pp. 400–417, Jul. 2014, doi: 10.1057/ejis.2013.9.
[31] T. Paulussen, A. Heinzl, and C. Becker, "Multi-Agent Based Information Systems for Patient Coordination in Hospitals," ICIS 2013 Proceedings, 2013, [Online]. Available: https://aisel.aisnet.org/icis2013/proceedings/HealthcareIS/2/.
[32] K. Gand and H. Schlieter, "Personalisation and Dynamisation of Care pathways-Foundations and Conceptual Considerations.," in Proceedings of ECIS 2016, 2016, p. Research-in.
[33] S. Wakamiya and K. Yamauchi, "What are the standard functions of electronic clinical pathways?," International Journal of Medical Informatics, vol. 78, no. 8, pp. 543–550, 2009, doi: https://doi.org/10.1016/j.ijmedinf.2009.03.003.
[34] P. Gooch and A. Roudsari, "Computerization of workflows, guidelines, and care pathways: a review of implementation challenges for process-oriented health information systems," Journal of the American Medical Informatics Association, vol. 18, no. 6, pp. 738–748, Nov. 2011, doi: 10.1136/amiajnl-2010-000033.
[35] M. Andellini et al., "Experimental application of Business Process Management technology to manage clinical pathways: a pediatric kidney transplantation follow up case.," BMC Medical Informatics & Decision Making, vol. 17, pp. 1–9, 2017.
[36] A. Appari, M. E. Johnson, and D. L. Anthony, "Health IT and inappropriate utilization of outpatient imaging: A cross-sectional study of U.S. hospitals," International Journal of Medical Informatics, vol. 109, pp. 87–95, 2018, doi: https://doi.org/10.1016/j.ijmedinf.2017.10.020.
[37] L. Shivers, S. S. Feldman, and L. W. Hayes, "Development of a computerized paediatric intensive care unit septic shock pathway: improving user experience," Health Systems, pp. 1–7, May 2019, doi: 10.1080/20476965.2019.1620638.
[38] J. M. Toy, A. Drechsler, and R. C. Waters, "Clinical pathways for primary care: current use, interest and perceived usability," Journal of the American Medical Informatics Association, vol. 25, no. 7, pp. 901–906, Jul. 2018, doi: 10.1093/jamia/ocy010.

[39] J. Schuld, T. Schäfer, S. Nickel, P. Jacob, M. K. Schilling, and S. Richter, "Impact of IT-supported clinical pathways on medical staff satisfaction. A prospective longitudinal cohort study," International Journal of Medical Informatics, vol. 80, no. 3, pp. 151–156, 2011, doi: https://doi.org/10.1016/j.ijmedinf.2010.10.012.
[40] S. Barbagallo et al., "Optimization and planning of operating theatre activities: an original definition of pathways and process modeling.," BMC Medical Informatics & Decision Making, vol. 15, no. 1, pp. 38–53, 2015.
[41] K. H. Sung et al., "Application of clinical pathway using electronic medical record system in pediatric patients with supracondylar fracture of the humerus: a before and after comparative study.," BMC Medical Informatics & Decision Making, vol. 13, no. 1, pp. 1–8, 2013.
[42] M.-M. Bouamrane and F. S. Mair, "A qualitative evaluation of general practitioners' views on protocol-driven eReferral in Scotland.," BMC Medical Informatics & Decision Making, vol. 14, no. 1, pp. 1–24, 2014.
[43] M. J. Husain et al., "HERALD (health economics using routine anonymised linked data)," BMC medical informatics and decision making, vol. 12, no. 1, p. 24, 2012.
[44] N. Platt, M. Tarafdar, and R. Williams, "The complementary roles of Health Information Systems and Relational Coordination in alcohol care pathways: The case of a UK hospital," ECIS 2019 Proceedings, 2019.
[45] E. Øvrelid, T. Sanner, and A. Siebenherz, "Creating Coordinative Paths from admission to discharge: The role of lightweight IT in hospital digital process innovation," 2018.
[46] K. Eason and P. Waterson, "The implications of e-health system delivery strategies for integrated healthcare: Lessons from England," International Journal of Medical Informatics, vol. 82, no. 5, pp. e96–e106, 2013, doi: https://doi.org/10.1016/j.ijmedinf.2012.11.004.
[47] M. Schriek, O. Türetken, and U. Kaymak, "A Maturity Model for Care pathways.," in Proceedings of ECIS 2016, 2016, p. ResearchPaper127, [Online]. Available: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1112&context=ecis2016_rp.
[48] M. Juhrisch, G. Dietz, and H. Schlieter, "Towards Compliance in Organizational Engineering - A Case Study," in 20th Proceedings of European Conference on Informations System, Barcelona, 2012, p. Paper 168, [Online]. Available: https://aisel.aisnet.org/ecis2012/168/.
[49] M. Juhrisch, H. Schlieter, and G. Dietz, "Model-Supported Business Alignment of IT-Conceptual Foundations.," 2011, [Online]. Available: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1215&context=amcis2011_submissions.
[50] J. Greenwood-Lee, G. Wild, and D. Marshall, "Improving accessibility through referral management: setting targets for specialist care," Health Systems, vol. 6, no. 2, pp. 161–170, Jul. 2017, doi: 10.1057/hs.2015.20.
[51] K. Baker et al., "Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy," International Journal of Medical Informatics, vol. 103, pp. 32–41, 2017, doi: https://doi.org/10.1016/j.ijmedinf.2017.03.011.
[52] A. R. Cochran, K. M. Raub, K. J. Murphy, D. A. Iannitti, and D. Vrochides, "Novel use of REDCap to develop an advanced platform to display predictive analytics and track compliance with Enhanced Recovery After Surgery for pancreaticoduodenectomy," International Journal of Medical Informatics, vol. 119, pp. 54–60, 2018, doi: https://doi.org/10.1016/j.ijmedinf.2018.09.001.
[53] E. Aspland, D. Gartner, and P. Harper, "Clinical pathway modelling: a literature review," Health Systems, pp. 1–23, Sep. 2019, doi: 10.1080/20476965.2019.1652547.
[54] P. Richter, "Bringing Care Quality to Life: Towards Quality Indicator-Driven Pathway Modelling in Health Care Networks," ECIS 2019 Proceedings, 2019.
[55] E. Demir, M. M. Gunal, and D. Southern, "Demand and capacity modelling for acute services using discrete event simulation," Health Systems, vol. 6, no. 1, pp. 33–40, Mar. 2017, doi: 10.1057/hs.2016.1.
[56] G. Berntsen, F. Strisland, K. Malm-Nicolaisen, B. Smaradottir, R. Fensli, and M. Røhne, "The Evidence Base for an Ideal Care Pathway for Frail Multimorbid Elderly: Combined Scoping and Systematic Intervention Review," J Med Internet Res, vol. 21, no. 4, p. e12517, Apr. 2019, doi: 10.2196/12517.
[57] C. Hufnagl, E. Doctor, L. Behrens, C. Buck, and T. Eymann, "Digitisation along the Patient Pathway in Hospitals," ECIS 2019 Proceedings, 2019.
[58] K. Flott, R. Callahan, A. Darzi, and E. Mayer, "A Patient-Centered Framework for Evaluating Digital Maturity of Health Services: A Systematic Review," J Med Internet Res, vol. 18, no. 4, p. e75, Apr. 2016, doi: 10.2196/jmir.5047.
[59] F. Meng, C. K. Ooi, C. K. Keng Soh, K. Liang Teow, and P. Kannapiran, "Quantifying patient flow and utilization with patient flow pathway and diagnosis of an emergency department in Singapore," Health Systems, vol. 5, no. 2, pp. 140–148, Jun. 2016, doi: 10.1057/hs.2015.15.
[60] O. Korhonen and M. Isomursu, "Identifying Personalization in a Care Pathway: a Single-Case Study of a Finnish Healthcare Service Provider," ECIS 2017 Proceedings, 2017.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210094

Digital Health Intervention to Support Refugees in Switzerland

Patricia ROMAO[a,1], Stefanie NEUENSCHWANDER[a,1], Kerstin DENECKE[a,2] and Stephan NÜSSLI[a]
[a] *Bern University of Applied Sciences, Bern, Switzerland*

Abstract. Experiences of war and persecution often lead to mental health problems, resulting in post-traumatic stress disorders. In this work, we design a digital platform that aims at helping refugees coming to Switzerland by providing exercises for their mental health and information about daily life in Switzerland. In collaboration with the Swiss Red Cross (SRC), we collected requirements and developed a concept for information provision through in this platform. The architecture of a progressive web application (PWA) was identified as to best fulfill the given requirements. Based on the collected requirements mockups were created. In user interviews, we received feedback regarding the future system. We learned that the platform should include an avatar, which guides the user through the entire platform and asks questions. All texts should be accessible by a read-aloud function and exercises should be provided as videos. In summary, we learned that it is essential to involve the future user group in the development process since it is characterized by cultural diversity that has to be considered in the development and design. Enriched by this input, the next step is to realize the application in terms of a prototype.

Keywords. Mental health, mHealth, Progressive web application, Refugees mental health, Mental health care seekers

1. Introduction

In 2019, more than 79.5 million people worldwide were forced to flee their homes in the context of war and persecution [1]. The hardships experienced by those who flee their homes, both before and after, often lead to trauma-related disorders. These include post-traumatic stress disorder, depression, anxiety, and other psychosomatic symptoms [2]. In the incoming country, there is often a lack of adequate psychological services for these people. This leads to a reduced quality of life for those affected and a more difficult social and economic integration. Mental disorders among refugees and asylum seekers occur significantly more often than in the general population. Effective treatment options and cross-cultural, specialized treatment centers exist [7]. However, they tend to be overloaded by their target populations. General outpatient primary health care providers might be able to compensate for the lack of specialized treatment slots, but a study shows that barriers such as lack of funding of interpreters seem to hinder these providers [7].

In addition to the great personal suffering, there is also economic and social damage. This can be significantly reduced by early interventions [3]. The existing approaches to

[1] Contributed equally

[2] Kerstin Denecke, Corresponding author, Bern University of Applied Sciences. Biel / Bienne, Switzerland; E-mail: kerstin.denecke@bfh.ch

address these issues do not allow to support all affected individuals sufficiently due to the lack of resources. Therefore, new and innovative ways are required, for example by using digital psychological and psychosocial support. Nowadays, the smartphone serves as an indispensable tool that fulfills the needs for information provision, orientation, communication, and social contacts during and after the escape from the home country. Andersson et al. show in a review paper that psychosocial care via a digital channel has a promising effect [4]. Several mindfulness and meditation apps (e.g. MindSpace https://www.headspace.com/de) or even systems for online psychotherapy are available (MindDoc, https://www.minddoc.de). Besides the online psychotherapy, the MindDoc app offers exercises and an emotion diary. However, these approaches do not address the specific needs of refugees.

Step-by-Step [9] is a digital health intervention designed for refugees and migrants. It is built on evidence-based cognitive behavioral therapy (CBT) techniques: i) stress management, ii) behavioral activation, iii) positive self-talk, iv) promotion of social support, and v) relapse prevention. Fictional characters tell their story of how they overcame psychological distress. An illustrated doctor provides psychoeducation and introduces the interactive exercises, e.g., a breathing exercise (audio) or planning positive activities using input fields and a calendar [9]. This digital health intervention only addresses the mental health issue, but is not expected to additionally support the process of integration in the new country.

In this work, we aim at designing an internet-based platform that will support refugees in both, in psychological issues and in general questions about integration. It is still unclear which functionalities such an application should provide and which design issues have to be considered. In this work, we will collect requirements towards such application in collaboration with the Swiss Red Cross (SRC). We will generate design examples in form of mockups and collect feedback from refugees. The overall objective behind this effort is to develop in a next step such application achieving excellent user experience.

2. Methods

Requirements regarding the platform were collected in discussions with employees of the Swiss Red Cross (SRC) that are normally concerned by supporting refugees. Based on these requirements a mockup was created using Microsoft Powerpoint. Additionally, we considered best practices reported in literature [11-16]. Required images were selected from Freepik [5]. We also identified a system architecture that fits best with the collected requirements. Following human-centered design principles, interviews with five refugees from Syria were conducted to evaluate the mockup. The aim was to get their feedback on different design types of the mockup and the planned application in general. We conducted the interviews as A/B test, which is a simple controlled experiment. Two versions of the mockup (A and B) are compared. A/B tests are widely considered the simplest form of controlled experiment. They are useful for understanding user engagement and satisfaction regarding features.

Since this study was performed during the COVID-19 pandemic and physical meetings were impossible, interviews were conducted via Skype or Microsoft Teams. Interview partner (5 refugees from Syria that were already living in Switzerland for some time) were confronted with different variants of the mockup and they were asked to

decide and justify, which variant they prefer. Additionally, they were asked to provide their personal ideas.

3. Results

In this section, we summarize the collected requirements and provide an overview on the system architecture and mockup designs. Furthermore, results from the mockup testing are described.

3.1. Requirements

We identified basically 5 main requirements regarding platform functionalities and characteristics in our interviews.

- *Audiovisual presentation of the contents*: Content such as exercises targeting mental health should be presented in an audiovisual manner. It should be examined which form of presentation is best suited for this purpose.
- *Two types of content:* The platform should offer general information on integration of refugees as well as psychological support.
- *Logging status of a user:* A status log is needed to enable a user to continue where he left off when he uses the platform again.
- *Motivation*: The platform should motivate the user to come back and use it regularly.
- *Accessibility for people with low reading skills*: The content of the platform should also be accessible to people with low reading skills.

The application should be accessible by refugees anytime and anywhere via smartphone to provide information and psychological and psychosocial support in a digital manner. The digital health intervention will comprise different modules through which refugees will get support as needed.

3.2. Mockup

In the following, we are describing the different parts of the mockup that form the basis of the user interviews. The application is equipped with an avatar that can be selected by the user when launching the application for the first time. The avatar will guide him through the platform as seen in Figure 1. Additionally, the application provides information regarding integration as content. We created different versions for avatars and content presentations that were assessed by potential future users.
The mockup in Figure 1 shows three different versions of an avatar to find whether the target group is more attracted by avatars depicted as humans, by animals or by personified animals.

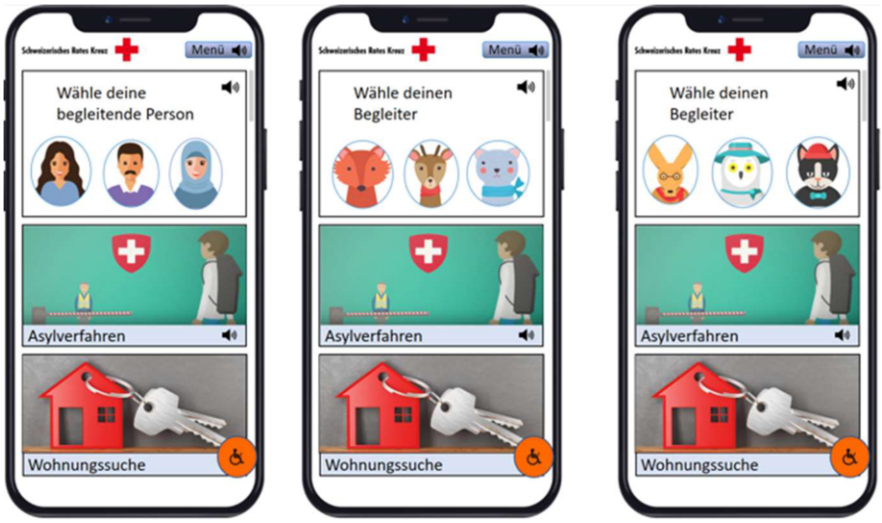


Figure 1 Mockup representing the different avatars.

Figure 2 shows different forms of content presentation. In the mockup, an example exercise provided by the SRC is used to test which way of presenting such content is most appealing to the potential user. There are three options to choose from: Text, audio file or video. Given these options, the requirement "audiovisual presentation of the contents" is considered. In addition, all texts, such as the statements of the avatar, have small loudspeakers that contain a read-aloud function.

Figure 2. Mockup representing different forms of representation of psychological exercises.

According to the requirements, the content of the platform should be accessible and understandable by persons with reduced reading skills. To address this requirement, we integrated the EqualWeb accessibility tool into the mockup. It is accessible by clicking on the wheelchair icon (see Fig. 2). The user interviews will show whether such tool is useful without explanation and introduction how to use it.

3.3. Feedback on the mockups

When selecting the avatar, the interviewees found it positive that the avatars depicted as human beings offer the option of a woman with or without a headscarf. Nevertheless, there were two interviewees who wished for an even broader spectrum regarding the appearance of the avatar. Icons of human beings were perceived as more serious than those of animals. In the case of the animal avatars, there was criticism that this could appear too childish and that users could feel that they were not being taken seriously. Nevertheless, two interviewees found the animals amusing and claimed they would like to choose an animal as avatar.

The fact that the avatar asks if it can address questions to the user was found to be sympathetic. All five interviewees were motivated to continue after an introduction dialog. Opinions were divided on the type of questions. Three persons prefer short and concise questions; the others stated to prefer longer sentences in which the avatar shows his understanding and tells something about himself.

In the case of accessibility, the orange button with the wheelchair pictogram was not entirely understood. One person even felt that such a button was unnecessary, since it would remind people of deficits, for example if they cannot read. In his opinion, the help should be built much more directly into the platform, as it was already done in the mockups with the speaker button indicating the read aloud function for each text.

In part, the functions of the accessibility tool were unclear. One person felt that it is necessary to be able to read Arabic very well to understand the functions. The option to choose the color of the background and the caption was positively evaluated by two people, as they would personally like to use it.

Four of the five interviewees would rather search for information in the beginning, when they were new to Switzerland, and later, when they were already settled, they would use the psychological support on the platform.

All interviewees would prefer to do exercises supported by a video, as they find that most exciting. One person would also use the text describing the exercise, depending on the situation, as they can get distracted by reading. Exercises provided only by audio was considered too boring by all interviewees.

3.4. System architecture

We decided to develop the application as progressive web application (PWA) since it corresponds to the ideas of the SRC and well addresses the requirements. A PWA supports offline mode, can display the content and save user data without internet connection. Furthermore, a PWA has a better performance than a common web application and therefore the loading time is short even with a slow internet connection. It can also send push notifications to the user to attract his attention to newly added content. However, the disadvantage of a PWA is that its functionalities depend on the browser and the operating system. Push notifications are for example currently unavailable for iOS systems [6].

A PWA comprises 4 main components:

- Web App Manifest: Allows the PWA to be added to the home screen and behave like a native application. The manifest contains data such as the name of the application, icons or default language.
- Application Shell: Is the so-called "skeleton" of the application. It is the part that does not have to be reloaded or, in other words, it displays the minimum required content that should be shown immediately when the application is started. The application shell contains only static content. As soon as this is shown, the dynamic content can be loaded. This component is not absolutely necessary for the implementation of a PWA, but it shortens the loading times. The prerequisite is that static and dynamic data are clearly separated in the application.
- IndexedDB: Is a JavaScript-based object-oriented database for structured data. Data that is needed for offline use is stored in this database as objects. However, large files such as videos are problematic to store. Not all platforms support the storage of video data objects. In addition, large files have an impact on the performance of the application that should not be underestimated.
- Service Worker: Is an intermediary between the application and the internet. The main task of the Service Worker is to use the browser API to preload content and display it to the user. It runs under a separate browser thread and enables additional functions that are normally known in native applications. This refers to functions such as push notifications, the offline function and caching of data.

4. Discussion

The characteristics of a PWA best meet the SRC's requirements for realizing the platform for refugees. Above all, offline operation is an important feature since refugees do not have a continuous internet connection. When implementing the PWA, it is important to ensure that only features that are supported by all browsers, or at least the most common browsers, are implemented. Another difficulty is the offline support for videos, since videos contain large amounts of data that can overload the storage capacity and slow down the application.

The interviews with the refugees from Syria showed that for such a project a close cooperation with the people concerned, i.e., the future users of the platform, is of great importance. A similar conclusion was drawn by Goodman et al. They collaborated with refugee populations, placed them at the center of the design process for digital mental health interventions, and in this way developed the most useful tools for this target group [8].

In the interviews with the participants, it became clear that addressing different cultures can be challenging. This was especially an issue in the selection of the avatar. It is essential to select an avatar as guide through the app that appeals to a large group of people, regardless of their cultural background. Offering a selection of personified avatars and animal avatars would probably be the solution that would pick up the most users. Cultural diversity of refugees impact also on what kind of support is needed and in which way the support can be provided best. This will impact on the design of digital

health interventions targeting this particular user group [8]. For this reason, there are first attempts to consider cultural diversity in digital health interventions [10].

Providing a platform that offers digital support to refugees in Switzerland has great potential. The interviews showed that the refugees are very interested and are always available for further interviews or usability tests. Furthermore, the feedback has shown that they are more motivated to obtain information and perform mental health exercises with this solution than if this only happens by means of text on paper. We developed our app designs by user-centered design, i.e. involving the user group in the development process. Similar developments confirmed that user-centered design is useful when developing applications for this particular user group [17].

It should be noted, however, that the people interviewed do not belong to the actual target group, as they have already been in Switzerland for a longer period and have already found an apartment and a job. Even if this does not rule out psychological problems, they have a certain stability in life that could have an influence on the assessment of the app. Whether refugees who have just arrived in Switzerland are just as responsive to it cannot be confirmed with certainty after these tests. Nevertheless, the people interviewed are very close to the target group and provided very helpful feedback.

5. Conclusion

In this paper, we described the design of an internet-based platform for refugees that provides support in psychological issues and in general questions about integration. We learned that adaptability is required in such application; interactive elements such as a guided interaction by an avatar or provision of video content helps in increasing user experience. Complexity of language should be adaptable as well. Since migration and integration of refugees is a time-consuming process, a digital intervention should provide support depending on the individual situation of the refugee or migrant.

In a next step, we will develop a prototype of the application, again in collaboration with the SRC and refugees to ensure that the application is useful and useable for the target user group.

References

[1] Gr F, Refugees i UHC for. UNHCR - Global Trends 2019: Forced Displacement in 2019 [Internet]. UNHCR Global Trends 2019. [cited 2020 Oct 29]. Available from: https://www.unhcr.org/globaltrends2019/

[2] Silove D, Ventevogel P, Rees S. The contemporary refugee crisis: an overview of mental health challenges. World Psychiatry. 2017;16(2):130–9.

[3] Trautmann S, Rehm J, Wittchen H-U. The economic costs of mental disorders. EMBO reports. 2016 Sep 1;17(9):1245–9.

[4] Andersson G, Titov N, Dear BF, Rozental A, Carlbring P. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry. 2019;18(1):20–8.

[5] Freepik | Graphikressourcen für alle [Internet]. Freepik. [cited 2021 Jan 29]. Available from: https://de.freepik.com

[6] Sheppard D. Beginning Progressive Web App Development - Creating a Native App Experience on the Web. apress; 2017.

[7] Kiselev, N., Morina, N., Schick, M. et al. Barriers to access to outpatient mental health care for refugees and asylum seekers in Switzerland: the therapist's view. BMC Psychiatry 20, 378 (2020).

[8] Goodman R, Tip L, Cavanagh K. There's An App For That: Context, Assumptions, Possibilities And Potential Pitfalls In The Use Of Digital Technologies To Address Refugee Mental Health, Journal of Refugee Studies, 2020;
[9] Carswell K, Harper-Shehadeh M, Watts S, van't Hof E, Abi Ramia J, Heim E, Wenger A, van Ommeren M. Step-by-Step: a new WHO digital mental health intervention for depression. Mhealth. 2018;4.
[10] Shala M, Morina N, Burchert S, et al. Cultural adaptation of Hap-pas-Hapi, an internet and mobile-based intervention for the treatment of psychological distress among Albanian migrants in Switzerland and Germany, Internet Interventions, Volume 21, 2020.
[11] Chan S, Godwin H, Gonzalez A, Yellowlees PM, Hilty DM. Review of Use and Integration of Mobile Apps Into Psychiatric Treatments. Curr Psychiatry Rep. 30. Oktober 2017;19(12):96.
[12] Golchert J, Roehr S, Berg F, Grochtdreis T, Hoffmann R, Jung F, u. a. HELP@APP: development and evaluation of a self-help app for traumatized Syrian refugees in Germany – a study protocol of a randomized controlled trial. BMC Psychiatry. 30. April 2019;19(1):131.
[13] Zehetmair C, Nagy E, Leetz C, Cranz A, Kindermann D, Reddemann L, u. a. Self-Practice of Stabilizing and Guided Imagery Techniques for Traumatized Refugees via Digital Audio Files: Qualitative Study. J Med Internet Res. 2020;22(9):e17906.
[14] Heim E, Burchert S, Shala M, Kaufmann M, Cerga Pashoja A, Morina N, u. a. Effect of Cultural Adaptation of a Smartphone-Based Self-Help Programme on Its Acceptability and Efficacy: Study Protocol for a Randomized Controlled Trial. 3. August 2020 [zitiert 3. Januar 2021]; Verfügbar unter: https://www.psycharchives.org/handle/20.500.12034/2768
[15] Maercker A. Niedrigschwellige und innovative Interventionen. In: Maercker A, Herausgeber. Traumafolgestörungen [Internet]. Berlin, Heidelberg: Springer; 2019 [zitiert 3. Januar 2021]. S. 299–310. Verfügbar unter: https://doi.org/10.1007/978-3-662-58470-5_15
[16] Wagner B. Online-Therapie - eine neue Perspektive in der Psychotherapie für Flüchtlinge und Asylbewerber? In: Psychotherapie Forum. 2016. pp. 124–31.
[17] Burchert S, Alkneme MS, Bird M, et al. User-Centered App Adaptation of a Low-Intensity E-Mental Health Intervention for Syrian Refugees. Frontiers in Psychiatry, 2019:9; 663

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210095

Developing an App for Cardiovascular Prevention and Scientific Data Collection

Lara Marie REIMER[a,1], Fabian STARNECKER[a,b,c], Heribert SCHUNKERT[a,b,c] and Stephan JONAS[a]
[a] *Technical University of Munich, Munich, Germany*
[b] *German Heart Centre Munich, Munich, Germany*
[c] *German Centre for Cardiovascular Research, Berlin, Germany*

Abstract. Background: Mobile apps may encourage a lifestyle that avoids unhealthy behaviors, such as smoking or poor nutrition, which promotes cardiovascular diseases (CVD). Yet, little data is available on the utilization, perception, and long-term effects of such apps to prevent CVD. Objectives: To develop a mobile app concept to reduce the individual CVD risk and collect information addressing research questions on CVD prevention while preserving data privacy and security. Methods: To validate the concept, a prototype will be built, and usability studies will be performed. Results: We expect to determine whether it is possible to reach a broad user base and to collect scientific information while protecting user data sufficiently. Conclusion: To address CVD prevention, we propose a mobile coaching app. We expect high acceptance rates in validation studies.

Keywords. mHealth, Mobile Applications, Cardiovascular Diseases

1. Introduction

Cardiovascular diseases (CVD) remain the leading cause of death in the Western world [1]. Apart from conventional risk factors, such as smoking or hypertension, unhealthy lifestyle, including poor nutrition and psychological stress, are known to increase CVD risk. Modern technologies inform about risk factors and coach their users to adhere to a healthier lifestyle. Studies have shown that mobile medical apps can positively impact daily behavior in the short term [2,3]. Lifestyle apps target a broader user base while focusing on a subset of risk factors [4]. To evaluate the impact of both application types on CVD risk, long-term research data is necessary. Yet, such data collections are limited by the high costs of continuous data acquisition. Moreover, there is a need to evaluate the effects of mobile apps in primary prevention. Previous trials mainly included subjects with diagnosed CVD or other high-risk related diseases [2,3].

2. Methods

This paper presents a mobile app concept, which aims to provide a lifestyle application that targets relevant risk factors for CVD. Moreover, the concept allows collecting

[1] Corresponding Author: Lara Marie Reimer, Technical University of Munich, Germany, E-Mail: reimer@tum.de

scientific data through the app on how lifestyle changes impact CVD risk. We will build a prototype to validate the concept in usability studies.

3. Results

The system architecture consists of multiple subsystems. Health information and study questionnaires are distributed via a content delivery network to a web platform and a mobile app. The web platform serves as an entry point into the app. The app provides personalized content for the user and gathers individual health information. All data is processed locally on the device, ensuring maximum data security and protection. Health data, collected via smartphone sensors and connected wearables, is automatically imported into the app. Additional information can be entered manually. Both data types include parameters used for the calculation of the individual risk for CVD.
The mobile app decides based on prerequisites whether a user is eligible for a study. Studies can specify additional data to be collected. Only upon user consent, the app transmits anonymized health data to the study database. The data transmission is secured using end-to-end encryption between the study database and the user's device. Each study is examined for suitability of conduction in the app and requires an ethical vote.

4. Discussion and Future Work

To address CVD prevention, we propose a mobile app that allows the app user to lower their individual risk. For gaining insights into prevention, the app offers an anonymous collection of clinical data from users. The development of the prototype will be finished in 2021. Usability studies will be performed, emphasizing whether we can establish a user base broad enough for trials and whether the app protects user data sufficiently. Due to the focus on data protection, we expect high acceptance rates of the app.

Funding

Funded by the Bavarian State Ministry of Health and Care through the research project DigiMed Bayern (www.digmed-bayern.de)

References

[1] P. Bhatnagar, K. Wickramasinghe, E. Wilkins, N. Townsend, Trends in the epidemiology of cardiovascular disease in the UK. *Heart*, **102**(24) (2016), 1945-1952.
[2] G.M. Coorey, L. Neubeck, J. Mulley, J. Redfern, Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. *Eur J Prev Cardiol* **25**(5) (2018), 505-521.
[3] L. Pfaeffli Dale., R. Dobson., R. Whittaker, R. Maddison, The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: A systematic review. *Eur J Prev Cardiol* **23**(8) (2016), 801-817.
[4] G. Flores Mateo, E. Granado-Font, C. Ferré-Grau, X. Montaña-Carreras, Mobile Phone Apps to Promote Weight Loss and Increase Physical Activity: A Systematic Review and Meta-Analysis. *J Med Internet Res* **17**(11) (2015), e253.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210096

Secure Peer Review and Feedback on Medical Findings via IHE XDS

Emmanuel HELM[a,1], Jacqueline SCHWEBACH[a,b], Andreas POINTNER[a], Anna LIN[a], Valentin ROTHENSTEINER[b], Dietmar KEIMEL[b] and Andreas SCHULER[a]
[a] *University of Applied Sciences Upper Austria, Research Department for Advanced Information Systems and Technology, 4232 Hagenberg, Austria*
[b] *CAS - Computer Anwendungs Systeme GmbH., 1090 Wien, Säulengasse 6/9*

Abstract. Background: There is a lack of secure official communication channels for peer review and peer feedback on medical findings. Objectives: We aimed to utilize the existing Austrian eHealth infrastructure to enable review and feedback processes. Methods: We extended the IHE XDW workflow document to enable the exchange of text messages (i.e., comments on documents or images) over an XDS infrastructure. Results: The workflow enables the exchange of comments on specific sections of CDA documents or radiological images and was verified in an XDS test environment. Conclusion: The presented solution is a proof of concept and the potential basis for the specification of a new IHE workflow definition.

Keywords. Workflow [L01.906.893], Radiology [H02.403.740]

1. Introduction

We developed a system for discussing findings and obtaining expert opinions and feedback via secure medical networks based on the Integrating the Healthcare Enterprise (IHE) Cross-Enterprise Document Sharing (XDS) integration profile. In this paper we describe the approach and discuss the results.

In 2018, Helm et al. [1] published a study on cross-enterprise communication and data exchange in Austrian radiology practices and departments. They conducted interviews with 4 experts each from hospitals and practices to identify the use cases and technology used for communication and data exchange. A key finding was that radiologists mainly use *informal* channels to obtain expert opinions, peer reviews, and peer feedback, e.g., online messengers [1].

This leads to several issues, e.g.: (1) under data protection law, the exchange of patient information via, e.g., WhatsApp is questionable, (2) the resolution of photos in these messenger apps is not suitable for a qualified assessment of radiological images, and (3) the communication is not available for follow-up visits and other physicians.

Thus, our motivation was to provide the means to enable secure communication of findings, expert opinions, and feedback over the existing Austrian eHealth infrastructure. To this end, we utilize the standards-based, centrally available services of the Austrian national electronic health record ELGA (Elektronische Gesundheitsakte) [2].

[1] Corresponding Author: Emmanuel Helm, FH OÖ F&E GmbH, Hagenberg, Austria, E-Mail: emmanuel.helm@fh-hagenberg.at

1.1. Background

The Austrian national electronic health record ELGA enables patient-centric document sharing across different healthcare providers. The architecture is based on the interplay of *actors* and *transactions* specified in the *integration profiles* of the non-profit organization IHE. These integration profiles define how to utilize established standards like Health Level Seven (HL7) or Digital Imaging and Communications in Medicine (DICOM) to solve common interoperability problems, e.g., the document exchange between healthcare providers [3]. In ELGA, the structure of the documents is based on the HL7 Clinical Document Architecture (CDA) standard.

In simplified terms, three transactions of the IHE XDS integration profile [4] are relevant to exchange clinical reports: (1) *ITI-18 Registry Stored Query*, to receive the metadata for a set of documents of a given patient, (2) *ITI-41 Provide and Register Set-b*, to upload a document, and (3) *ITI-43 Retrieve Document Set*, to receive a document. For the exchange of medical images, the Cross-enterprise Document Sharing for Imaging (XDS-I) integration profile [2], an extension of XDS, is relevant.

The IHE Cross-Enterprise Document Workflow (XDW) integration profile [4] states: "This profile defines an instrument, called a 'Workflow Document', to manage and track a shared workflow. It records the creation of tasks and maintains a historical record of tasks as they move through the associated workflow". While XDW defines form and function of this document, the actual sharing of the document is based on XDS. And while XDW is not part of ELGA yet, the existing XDS infrastructure already provides the means for future adoption of this integration profile.

1.2. Problem Statement

Previous research [1, 5, 6] described a lack of official communication channels for obtaining external feedback or expert opinions on reports or radiological imaging. In Austria, mainly *informal* channels are used, e.g., telephone exchanges with acquainted physicians or the use of online messengers with other peers [1]. In addition to quality, privacy, and security issues with online messengers, the lack of official channels for exchange makes business models for proven experts impossible. Currently, ELGA lacks active workflow support. To overcome these problems, we developed a workflow mechanism that enables cross-enterprise collaboration based on the existing eHealth infrastructure. We show its applicability in the radiology domain.

1.3. Relevance and Related Work

In 2014, Austria introduced the nationwide peer review program "Früh erkennen" [7] for breast cancer screening. To reduce the number of false negative and false positive diagnoses, radiologists are obliged to obtain a second opinion on their Breast Imaging Reporting And Data System (BIRADS) assessment. Strickland [5] describes the potential for quality assurance in radiology via the introduction of peer review and peer feedback processes. However, he also describes a lack of integrated system software on the market. Gunn et al. [6] point out existing tools such as RADPEER (ACR) but also highlight problems of sole radiologist-to-radiologist peer review (e.g., lack of context, or practices and terminology confusing to end-readers). They found that including the referring physician in the process is a potential solution to overcome these problems.

2. Methods

To enable peer review and peer feedback processes (i.e., workflows) on the existing Austrian eHealth infrastructure, we extended the XDW Workflow Document [4] with an "annotation" mechanism. The workflow document basically contains a structured list of tasks, documenting the execution of a workflow. In our case, a peer review or peer feedback process constitutes a workflow document, with a task for the communication of the peers. The workflow document itself is exchanged via the XDS infrastructure, as is the respective report that is referred to.

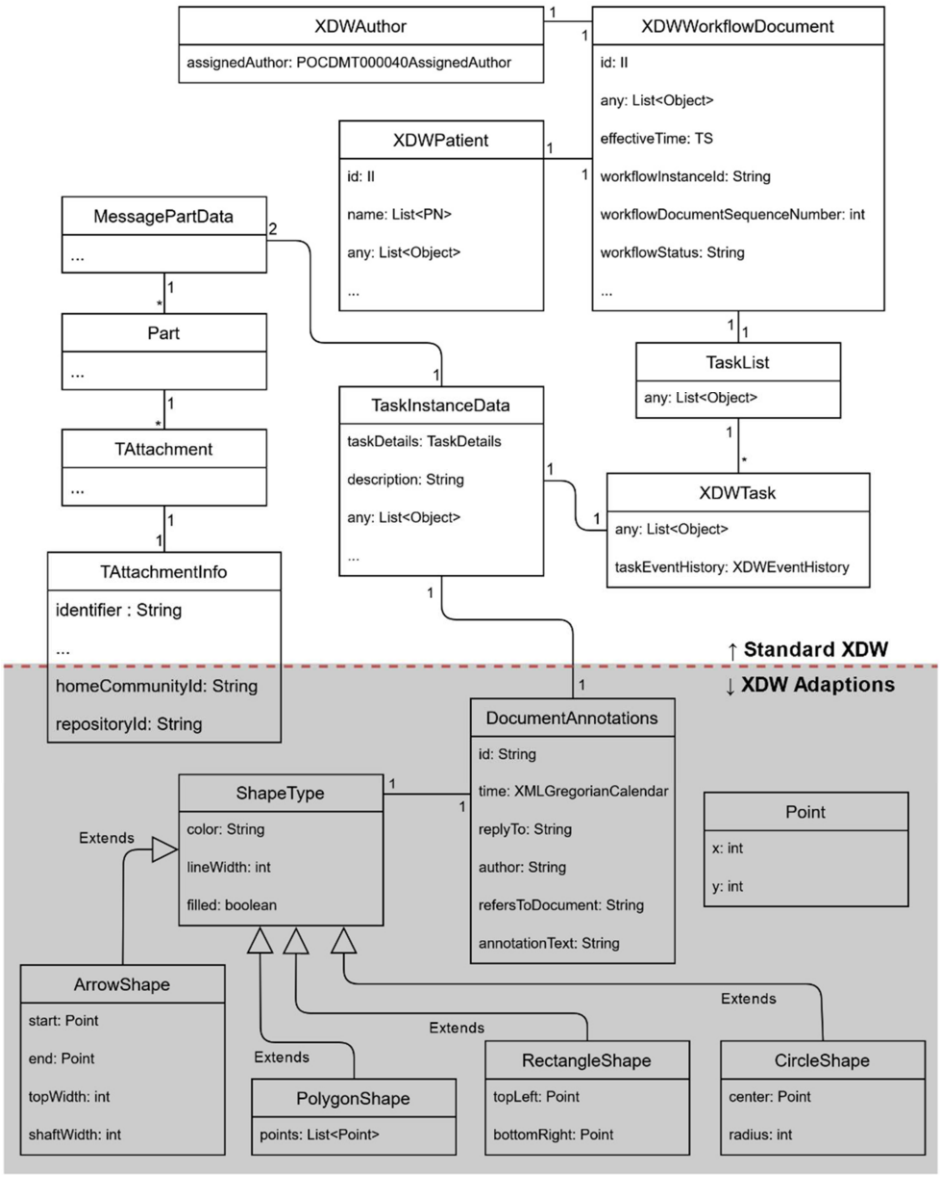


Figure 1: Schema outlining the adaptions to the XDW Document.

2.1. Adding Comments to XDW Workflow Documents

To utilize the XDW workflow document for comments, it was necessary to perform some adaptions (see Figure 1). We added objects that allow to write comments for specific parts of a referenced document or image. A *documentAnnotation* element was added to fulfill that requirement. This class contains the basic information required for a comment, namely an *author*, a date (*time*), and an *annotationText*. In addition, an *ID* to uniquely identify a comment inside an XDW document as well as a *replyTo* property – allowing to reply to specific comments – were added. For the rest of the adaptions, we must differentiate between commenting on CDA documents or DICOM objects (i.e., images).

For CDA documents it should be possible to comment on specific sections. For this purpose, a *refersToDocument* field was added, allowing to reference a section of a CDA document (cf. Listing 1). Furthermore, it is necessary to distinguish between top level comments and replies. For each top-level comment, the children are attached recursively matching the comment ID and the replyTo property.

```
<cas:documentAnnotation>
  <cas:id>1</cas:id>
  <cas:author>Dr. Tim Baader</cas:author>
  <cas:time>2020-11-28T08:01:12.000Z</cas:time>
  <cas:annotationText>
    Sollen noch weitere Abteilungen kontaktiert werden?
  </cas:annotationText>
  <cas:refersToDocument>1.2.40.0.34.11.5.2.9</cas:refersToDocument>
</cas:documentAnnotation>
```

Listing 1: Referring to a CDA-Section in an XDW-Document using *documentAnnotation*.

In case of DICOM, or more generally in the case of images, it should be possible to refer to an explicit position in the picture. We developed a concept to allow adding shapes with annotations as seen in Listing 2. For that purpose, *Markers* were added to the *DocumentAnnotation*. Each marker consists of a specific shape, e.g., *Arrow*, *Polygon*, *Rectangle* or *Circle*. These shapes do have some properties in common, i.e., a specific color, a line width, and a flag that indicates if the shape is filled or frame only. In addition to that, each shape type has properties to define its position, size, and form.

```
<cas:marker>
  <cas:rectangle>
    <cas:color>#00ff00</cas:color>
    <cas:lineWidth>1</cas:lineWidth>
    <cas:filled>false</cas:filled>
    <cas:topLeft>
      <cas:x>227</cas:x>
      <cas:y>304</cas:y>
    </cas:topLeft>
    <cas:bottomRight>
      <cas:x>247</cas:x>
      <cas:y>324</cas:y>
    </cas:bottomRight>
  </cas:rectangle>
</cas:marker>
```

Listing 2: Description of an image marker inside a *documentAnnotation*.

2.2. Referencing Reports or Images

An XDW document describes multiple tasks, where each task with its *TaskInstanceData* can define *input* and *output* to refer to relevant objects for the task (e.g., diagnostic reports). In our case, one XDW workflow document describes the workflow of a peer review or peer feedback to a single CDA document or DICOM object. The *input* of the task in the XDW document refers to the respective document or object.

However, the standard XDW input and output fields do not have all the information needed for retrieval of the object in a cross-community XDS environment (based on the XCA profile [4], e.g., ELGA [2]) via the IHE *ITI-43 Retrieve Document Set* transaction. This transaction needs at least a *sourceDocumentId*, a *homecommunityId* as well as a *repositoryId* [2]. To add those fields, we extended the *AttachmentInfo* class, by adding fields for the *homecommunityId* and *repositoryId*. The *identifier* field of the object was used to store the *sourceDocumentId*.

To retrieve DICOM objects, the IHE *RAD-69 Retrieve Imaging Document Set* was used as specified by [8]. Thus, *Study*-, *Series*-, *Instance*- and *HomeCommunityId* is required. These fields are again added as extensions to the *AttachmentInfo* class.

2.3. Testing Infrastructure

ELGA provides a test environment, which we used as a testing infrastructure. To avoid implementing the required SOAP transactions, we used Open eHealth Integration Platform (IPF) components [9]. These components provide an abstraction for accessing the IHE-based ELGA transaction endpoints. By using IPF, we can exchange CDA documents and image references via the ELGA test registries and repositories. In theory this would also enable us to upload XDW workflow documents, but by the time of this work this was not supported by the ELGA infrastructure.

For testing the XDW workflow, a custom XDW document store was implemented. A file storage providing a basic REST interface was hosted on our test server. This endpoint provides methods to query and store XDW workflow documents. In addition to that a simple client was developed, that allows to display CDA documents as well as images and provides the means to add comments via the described approach.

3. Results

The presented approach allows to comment on specific sections of CDA documents and images to enable the collaboration across different departments or even enterprises. The following results show the communication in an example client application. We were able to test and simulate different scenarios on the ELGA-based testing infrastructure and provide two examples here.

The comment section has the same functionality and design for both scenarios. It shows the different comments, that are either assigned to a specific section for CDA documents or to a specific region in the image. Each comment consists of three parts: (1) the author, (2) a date when the comment was submitted, and (3) the comment text. There are two different types of comments: Top-level comments that belong to a specific document section or image region, and answer comments, that reply on another comment. They can be distinguished by their indentations (cf. comment section in Figures 2 and 3).

3.1. Comments on Reports

The first example, cf. **Figure 2**, shows the communication between two physicians in a finding represented by a CDA document. The image is split into two sections. On the left-hand side, a rendered form of the CDA document is displayed, whereas the right-hand side shows the comment section. To provide a uniform representation across platforms, ELGA [2] defines a default layout for CDA documents and provides the respective Extensible Stylesheet Language Transformation (XSLT) file. This stylesheet can be applied on the XML representation of the CDA document, and transforms it into the HTML format, which can be displayed as shown in our example.

In addition to that, we added the comment concept. This consists of two parts, on the one hand there are the comments with the authors on the right side, on the other hand there is a circle with number in it for every specific section in the CDA document (cf. Figure 2). These numbers represent the comments that are written in a specific section.

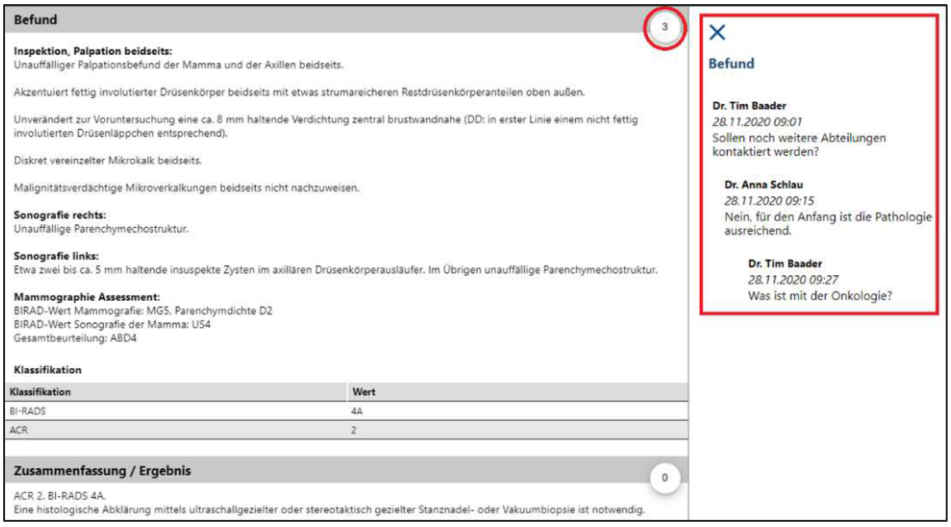


Figure 2: CDA document rendered with the ELGA stylesheet (left). The red circle highlights the number of comments in this specific section. The comments are visible on the right side.

3.2. Comments on Images

The second example, cf. **Figure 3**, shows the communication between two physicians but in this case for a single image. The left side shows the menu bar on top as well as the image, whereas on the right side the comments are placed. We utilize the RAD-69 transaction to query for a rendered JPEG image using the respective *Transfer Sytax* property. The image can be annotated with different forms, e.g., circles or arrows. Each of the shapes can have a different colour as well as different line width. In addition to the comments section on the right, each top-level comment is extended with the used shape as well as the colour to highlight the connection. Each new shape automatically creates a new top-level comment as well.

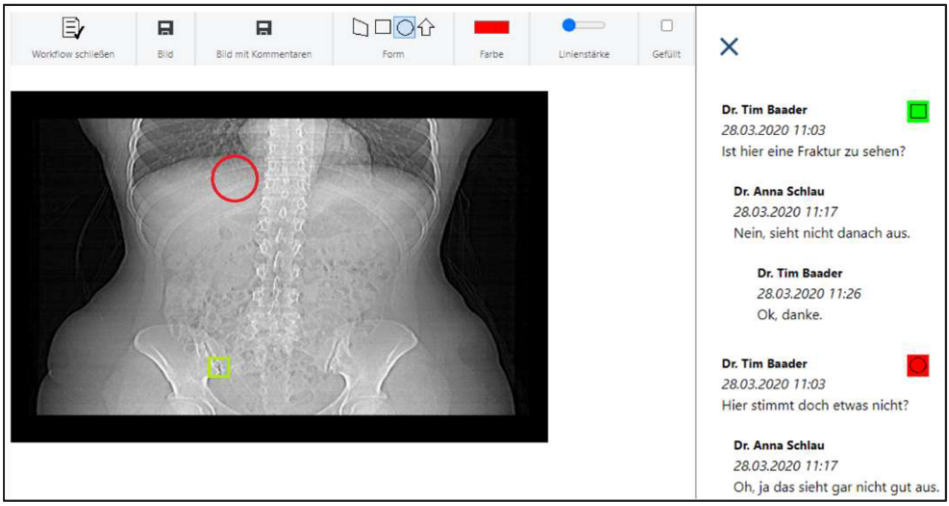


Figure 3: A radiology image with annotating shapes in the image and respective comments on the right side.

4. Discussion

The presented work is a first proof of concept on how to utilize XDW workflow documents for establishing secure and integrated communication channels in context of medical reports and images. There are still open issues and threats to the validity of the presented approach:

1. There is no IHE *workflow definition* document describing the use of XDW workflow documents for this purpose. Thus, the presented approach can only be seen as a predecessor of future standardized solutions.
2. ELGA does not support XDW workflow documents. These documents cannot be exchanged via the Austrian eHealth infrastructure (yet). However, the approach could already be implemented in local (XDS) environments.
3. The image approach only works on single rendered JPEGs. We did not implement a full DICOM viewer. However, the concept to store comments on DICOM objects in XDW workflow documents and exchange them via an XDS/ XCA infrastructure could be of interest for DICOM viewer or PACS vendors.
4. The described approach does not allow to address a specific peer with a request for feedback or expert opinion. Currently, the XDW document with the questions and comments, linked to a patient's CDA document or DICOM image, would be available to all healthcare providers with a valid treatment assertion for the patient. Thus, the approach uses ELGA as a platform, but would require additional peer-to-peer functionality for direct communication.
5. Due to the document-based nature of this approach, every exchange of comments or annotations requires the retrieval of the complete XDW workflow document. This leads to a certain redundancy in the communication. However, the XDW workflow document does not require significant bandwidth, e.g., the document in the examples above has 5 kB with about 100 lines of code.

5. Conclusions and Outlook

Our motivation was to provide the means to enable secure and integrated communication of findings, expert opinions, and feedback over the existing eHealth infrastructure. We presented an approach to utilize XDW workflow documents to enable cross-enterprise peer review and peer feedback. The approach was developed with respect to the Austrian eHealth infrastructure ELGA [2, 8] and is mostly based on international standards.

In practice, the developed workflow is either document centric or image centric. Comments can be made on sections of documents if they are structured based on the HL7 CDA standard. Alternatively, regions in images can be highlighted and commented.

In IHE the concept of a *workflow definition* can be used to specify how an XDW workflow document should be utilized to enable a certain workflow. Common examples are the Cross-enterprise Tumor Board Workflow Definition (XTB-WD) and the Cross-enterprise Basic eReferral Workflow Definition (XBeR-WD). Given the interest of the Austrian eHealth community, we plan to take a similar approach and prepare a first draft for a Cross-Enterprise (or even Cross-Community) Peer Review Workflow Definition (e.g., XPR-WD) document. Of course, this effort must be coordinated with multiple stakeholders, including the different chapters in IHE.

References

[1] E. Helm, A. Schuler and H. Mayr, "Cross-Enterprise Communication and Data Exchange in Radiology in Austria: Technology and Use Cases", *Studies in Health Technology and Informatics*, vol. 248, pp. 64-71, 2018

[2] ELGA GmbH, "ELGA-Gesamtarchitektur", [Online]. Version 2.30b, Available: https://www.elga.gv.at/fileadmin/user_upload/Dokumente_PDF_MP4/Technisches/ELGA_Gesamtarchitektur_2.30b.pdf, 2017, [Last Accessed 18.01.2021].

[3] T. Benson und G. Grieve, *Principles of Health Interoperability*, Springer, 2016.

[4] IHE International, Inc., "IHE IT Infrastructure (ITI) Technical Framework", vol. 1 (ITI TF-1) Integration Profiles, rev. 17.0, July 20, 2020

[5] N. H. Strickland, "Quality assurance in radiology: peer review and peer feedback", *Clinical radiology,* 70(11), pp. 1158-1164, 2015

[6] A. J. Gunn, C. I. Alabre, S. E. Bennett, M. Kautzky, T. Krakower, K. Palamara und G. Choy, "Structured feedback from referring physicians: a novel approach to quality improvement in radiology reporting", *American Journal of Roentgenology,* 201(4), pp. 853-857, 2013.

[7] K. Jeitler, T. Semlitsch, N. Posch, A. Siebenhofer und K. Horvath, "Breast cancer screening in Austria: Key figures, age limits, screening intervals and evidence", *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen,* Bd. 109, Nr. 4, pp. 363-370, 2015

[8] ELGA GmbH, " Anbindung von DICOM Ressourcen in ELGA", [Online]. Version 1.65, Available: https://www.elga.gv.at/fileadmin/user_upload/Dokumente_PDF_MP4/Technisches/AnbindungBilddaten_Gesamtarchitektur.pdf, 2020, [Last Accessed 18.01.2021].

[9] Open eHealth Foundation, "Open eHealth Integration Platform - IPF", [Online]. Available: https://github.com/oehf/ipf. [Last Accessed 18.01.2021].

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210097

Introducing a Navigation Algorithm for Reducing the Spread of Diseases in Public Transport Networks

Jens KLINKER[a,1], Mohamed Hechem SELMI[a], Mariana AVEZUM[a] and Stephan JONAS[a]
[a] *Professorship for Digital Health, Technical University Munich, Munich, Germany*

Abstract. Reducing passenger flow through highly frequented bottlenecks in public transportation networks is a well-known urban planning problem. This issue has become even more relevant since the outbreak of the SARS-CoV-2 pandemic and the necessity for minimum distances between passengers. We propose an approach that allows to dynamically navigate passengers around dangerously crowded stations to better distribute the passenger load across an entire urban public transport network. This is achieved through the introduction of new constraints into routing requests, that enable the avoidance of specific nodes in a network. These requests consider walks, bikes, metros, subways, trams and buses as possible modes of transportation. An implementation of the approach is provided in cooperation with the Munich Travel Corporation (MVG) for the city of Munich, to simulate the effects on a real city's urban traffic flow. Among other factors, the impact on the travel time was simulated given that the two major exchange points in the network were to be avoided. With an increase from 26.5 to 26.8 minutes on the average travel time, the simulation suggests that the time penalty might be worth the safety benefits.

Keywords. Physical Distance, Social Control, Social Distancing, COVID-19 Pandemic, Risky Health Behavior, Outbreaks, Prevention and Control

1. Introduction

Conventional navigation systems like Google Maps offer a number of fastest or shortest routes to get to a destination. While these options were sufficient in the past, the outbreak of the SARS-CoV-2 pandemic led to a new requirement. Since then, travelers have also been looking for safe routes through cities, that could lower the risk of an infection with the virus. In the case of SARS-CoV-2, one effective measure to reduce this risk is to keep a safe distance to other humans [1, 2].
In Germany, the minimal required safety distance in public places was set to 1.5 meters by the Federal Ministry of Health [3]. Any distance below that threshold is considered unsafe from a social distancing perspective.

Fulfilling these social distancing requirements turns out to be a challenging task for passengers in urban public transportation networks. This is mostly due to overcrowded transportation vehicles and exchange stations during rush-hour times that force people to break social distancing rules. Relatively in the beginning of the outbreak, public

[1] Corresponding Author: Jens Klinker, Technical University Munich, Germany, E-Mail: jens.klinker@tum.de

authorities in Germany took action to prohibit any events with more than 100 people. However, when considering that the average Munich subway has capacity for over 900 people, any public transportation ride operating at over 11% of its total capacity can already be considered a large event [4]. This issue is even more critical at the central exchange points in Munich's public transportation network.

In exchanges with city planners form the Munich Transport Corporation (MVG), it was explained that congestions rarely occur at most public transport stations. Instead, there is a small number of stations in Munich's public transportation network that get overcrowded reliably and predictably on a daily basis, due to their centrality in the network. Such stations, like "Sendlinger Tor" or "Hauptbahnhof", that are mainly used to change between public transport lines, pose a constant infection risk. At "Sendlinger Tor" for example, an average of 250.000 passengers board, alight or change public transport every day [5]. While this is an unusual threat during a pandemic, it is a constant threat to people in high-risk groups.

To solve this issue, new solutions have to be implemented that reduce the passenger load on central traffic nodes, so that safety distances can be established again [6]. The goal of such an optimization problem is to fulfill the safety requirement while causing the minimal necessary impact on other travel goals such as minimizing overall travel time to a destination.

2. Methods

To tackle this issue, we introduce a new routing algorithm that offers a safer route to a traveler's destination by avoiding highly congested stations in the public transport network. As the congestion level of a node is not constant, but depends on a large number of factors, a collaboration with the Munich Transport Corporation (MVG) was established. Using insights provided by the MVG, public transport stations can be dynamically toggled on and off before they reach a dangerously high congestion level in which safe distances can no longer be maintained. Currently, the assessment whether a station is considered safe or overcrowded is delegated to the MVG. To make this assessment as transparent and scientific as possible, it is advisable to take a closer look at parameters influencing the spreading risk in public transport stations. Due to its complexity, we haven't included this additional step in the scope of this publication.

The avoidance algorithm then uses the information about the available and unavailable stations to calculate a safer route with less congested exchange points, giving the control for social distancing back to the travelers. To simulate the impact on the travel time and other important features, the MVG provided a dataset containing the scheduled departure and arrival times of all public transport lines in Munich. Both the data and the congestion knowledge are used to simulate the results produced by an avoidance algorithm for safer routes.

To evaluate the performance the avoidance algorithm is compared to the results of the same algorithm without the avoidance feature. Core features like the change in travel time are then used to measure how the convenience of a route is impacted by the feature. As the convenience of a route is highly subjective, no final measure for the convenience could be produced. Instead, some underlying factors that partially correlate with a route's perceived convenience are presented.

As the proposed solution was developed with data from- and in cooperation with public authorities in Munich, the researched results and examples use cases are limited to the city of Munich.

2.1. Foundations

Public transportation has long been suggested as an alternative for road bottlenecks. During rush-hour, however, these are already overcrowded, and some cities thus prefer to limit passenger flow in bottleneck stations. In order to motivate passengers to consider going around an overcrowded station, two main reviews are presented: (1) Providing passengers with personalized multi-modal routing options and (2) calculating routes which have further advantages than a fastest travel time.

Review (1) is based on the work of Bucher *et al.* who describe how to integrate personal constraints on calculation of multi-modal routes [7]. To incorporate more travelling options into the multi-modal route calculation, the authors save a series of constraints in the form of a user profile.

As far as multi-modal routing goes, the state-of-the-art implementation is provided by OpenTripPlanner [8]. The open-source project uses public transportation schedules formatted as General Transit Feed Specification (GTFS) to calculate multi-modal routes that, among others, can combine cycling and public transportation segments. Since the consideration of multi-modal routes can be seen as an interesting approach to reduce the load of overcrowded public transportation intersection, OpenTripPlanner contains a "bannedStop" parameter that can be passed through a route request.

While the calculation of routes is a technical challenge, actually ensuring that the passengers use that route is a social one. Review (2) mentioned above relates to providing route advantages other than simply travel time. The basis for this is the work of Bunds *et al.* who analyzed how different route attributes were perceived by travelers and how these affected the choice of route [9]. Bunds et al. showed how air pollution, traffic, and noise level are the determining factors when deciding which route to walk through [9]. This allows us to infer that if a traveler can control for these variables during a route calculation, they can be more likely to accept longer walking segments. Furthermore, the fact that traffic and noise levels are directly correlated with overcrowded regions (even in public transportation), means that presenting this information to the traveler can serve as an important motivation for them to avoid these regions.

2.2. Constraint-based route personalization

OpenTripPlanner calculates multi-modal routes in the network based on a routing request that consists of a list of query parameters. In order to define the context of the route search, the request must specify the following information:

- **fromPlace:** Latitude and longitude of the start location.
- **toPlace:** Latitude and longitude of the end location
- **date:** Date on which the trip should depart.
- **time:** Time when the trip should depart.
- **mode:** Set of modes that a commuter is willing to use. The main modes supported by the system are walk, bike, car, and transit (buses, trains, trams).

In addition, the existing system supports multiple optional parameters that can be used to further manipulate the results. The most useful parameters for the problem at hand are:

- **bannedRoutes:** A comma-separated list of banned public transportation lines.
- **bannedStops:** Banned stations cannot be used to board or alight from a public transportation mode, but it is still possible to travel through them. This is achieved by blocking the pre-board and pre-alight edges that connect the transit network to the street network.
- **bannedStopsHard:** Stations that are removed from the network. It is no longer possible to board, alight or travel through these stations.

While the bannedStops parameter realizes the avoidance that the system aims to achieve for crowded stations, the time aware usage of this parameter to automatically avoid the stations during rush hour periods is still to be implemented.

2.3. Constraint integration in multi-modal routing

Similar to most state-of-the-art route planning services, OpenTripPlanner uses the A* algorithm to search for routes in the transportation network [10]. This algorithm keeps track of an ordered list of tentative routes and during each iteration the one with the smallest weight is extended. In order to achieve multi-modal routes, the algorithm is modified to loop over the available transportation modes during each iteration. For all outgoing edges of the last node, each mode that matches the type of the edge is used to traverse it. For instance, edges from the street network can be used for walking, biking, or driving, whereas edges of the transit network are restricted to a specific public transportation mode.

3. Results

3.1. Example use case

The following use case exemplifies how the avoidance algorithm calculates routes. We consider a trip in the Munich transit network and compare the route generated by the system with and without the automatic avoidance during rush-hour periods. For this demonstration, the experts from the Munich Transport Corporation (MVG) recommended to use the Sendlinger Tor, as it is one of the most congested stations in the network. In this case the routing request sent to the trip planner could, for example, use the following parameters:

- **fromPlace:** Nordfriedhof station with latitude 48.17312 and longitude 1.59686.
- **toPlace:** Theresienstraße station with latitude 48.15139 and longitude 11.56444.
- **date:** May 5th 2020.
- **time:** 08:00 am.
- **mode:** Walk, transit (buses, trains, trams)

This trip starts near a dorm for students and ends at a station used to access the technical university in the city, which makes it a realistic trip that students take on a daily

basis. The time of the request is within the morning rush-hour period from 07:00 am to 09:00 am.

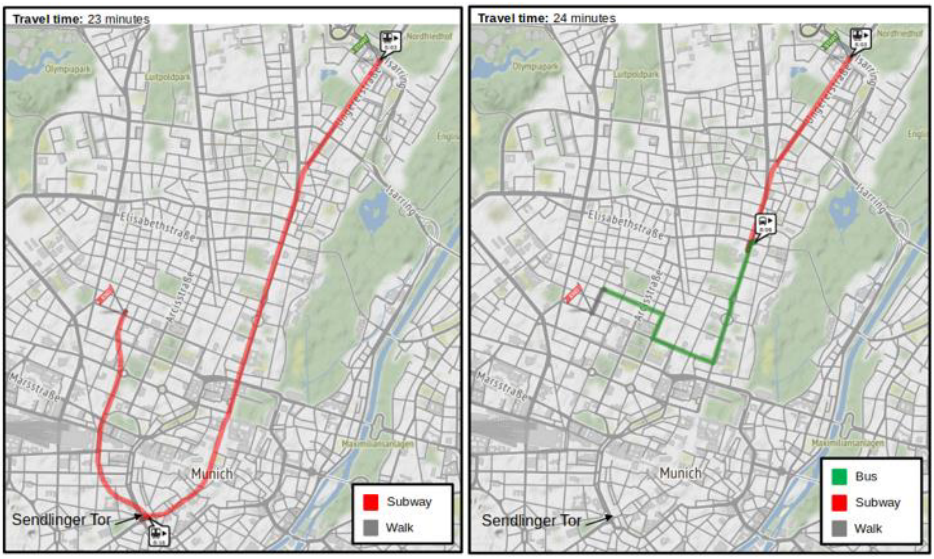


Figure 1. Route from Nordfriedhof to Theresienstraße. Left: Without automatic rush-hour avoidance. Right: With automatic rush-hour avoidance.

The route generated without avoidance can be seen in Figure 1 (left). It uses the subway from the origin, marked with a green flag, to get to Sendlinger Tor station. From there, a different subway line is used to get to the destination marked with a red flag. Including the time-aware avoidance results in a route that successfully avoids changing lines at Sendlinger Tor station as shown in Figure 1 (right). The new route cuts the subway part short before reaching banned stations and instead uses a bus to get to the destination. With activated avoidance, the travel time increases by one minute from 23 to 24 minutes.

3.2. General effect of avoidance on line changes

In this section we analyze how avoiding one station would affect the travel time and passenger distribution. The goals are to avoid overcrowding other stations in an attempt to scatter passengers to other stations and to maintain a reasonable travel time. To do so, we analyze the number of routes that use these stations to board, alight, or change lines. This can be considered as an estimation of the number of passengers at the stations. The analysis is based on a set of 1000 random routes located in the city of Munich. For each route, coordinates for the origin and the destination are sampled from an area centered at Sendlinger Tor with a radius of 4 kilometers. In addition, the following parameters are used for all routes:

- **date:** May 5th 2020
- **time:** 08:00 am.
- **mode:** Bicycle, walk, transit (buses, trains, trams)

The coordinates of the stops used to **board**, **alight**, or **change** lines are then extracted from each route and used to generate two heat maps. The first heat map shown in Figure 2 (left) summarizes the routes where no stops were avoided. The heat map in Figure 2 (right) considers the routes where Sendlinger Tor and Hauptbahnhof were automatically avoided during rush hour periods.

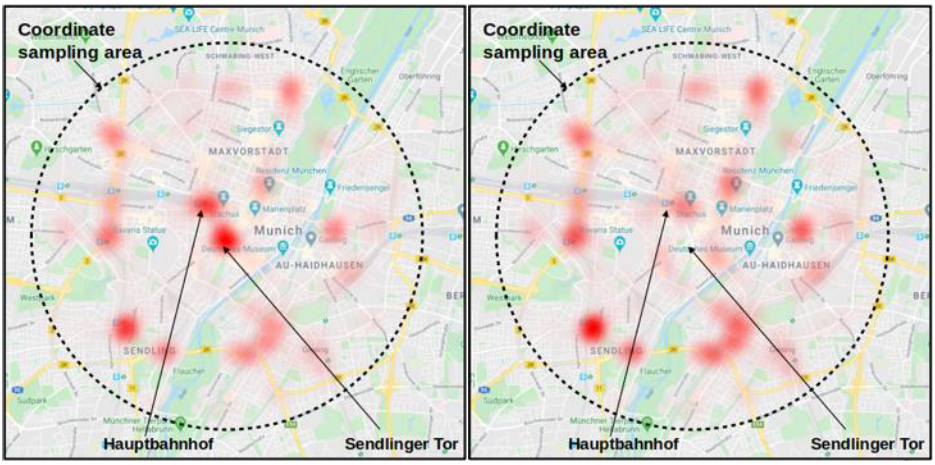


Figure 2. Heat maps showing how often transit stations are used to board or alight public transportation vehicles based on 1000 random routes. Left: Without activated avoidance feature. Right: With activated avoidance feature.

The left image in Figure 2 visualizes how Hauptbahnhof and Sendlinger Tor are originally the most used stations in the network, indicated through the dark red color of theses hot spots. While some areas on the outskirt of the network have similar size and darkness, they represent stations where commuters mostly board or alight a transit line but do not change them. With the automatic avoidance feature activated, the heat signature at the Sendlinger Tor and Hauptbahnhof both brighten up, while the direct areas around them darken slightly. This effect visualizes the distribution of the passenger load from these two stations to other nearby stations in the network. Instead of overcrowding a new station in the network, the load was distributed rather evenly over multiple surrounding stations located in close vicinity of the banned ones.

Aside the distribution of line transfers, the avoidance did not have notable effects on other route characteristics. The mean biking distance per route rose from 959 to 976 meters and the average walking distance slightly shrank from 469 meters to 468 meters. The mean waiting time at public transportation stations also remained unchanged at around 3.8 minutes per route. The mean runtime required to calculate a route decreased from 4.7 seconds to 3.7 seconds when the avoidance was activated. Finally, we consider the travel time of the routes used for the creation of the heat maps. The mean travel time increased from 26.5 minutes without avoidance to 26.8 minutes with avoidance.

3.3. Effect of avoidance on travel time

In this section we use isochrones to visualize the effect of the avoidance on the travel time in Munich's traffic network. Isochrones are graphs that measure location reachability from a specific origin. They consist of curves with equal travel time (Figure

3). OpenTripPlanner provides a service for generating isochrones out of the box by sending a request to the system similar to how routes are generated. In Figure 3 Sendlinger Tor was picked as the origin parameter for calculating the isochrones.

The generation of an isochrone starts by calculating a shortest path tree. OpenTripPlanner then builds a regular grid of samples covering the whole shortest path tree area. Finally, the sample points are connected based on their travel time to form the curves of the isochrones. The isolines are computed with the help of the Delaunay Triangulation Algorithm [11].

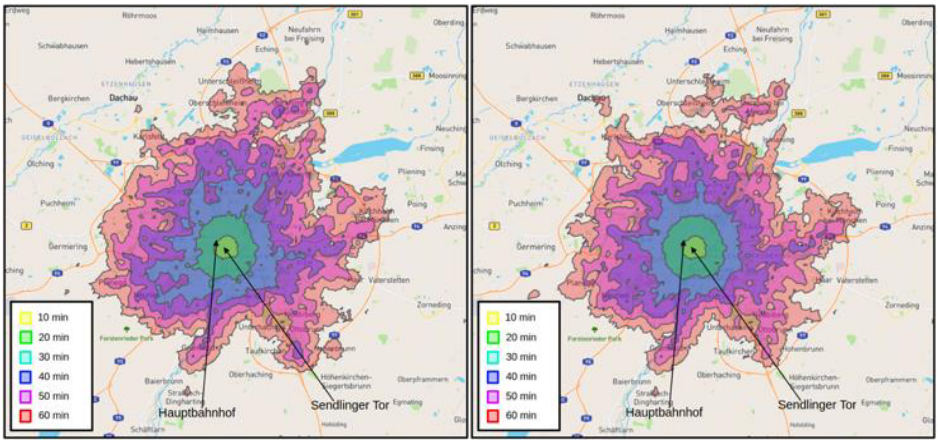


Figure 3. Isochrone of the travel time. Left: Without avoidance of Sendlinger Tor and Hauptbahnhof. Right: With avoidance of Sendlinger Tor and Hauptbahnhof.

Comparing the yellow and green areas in Figure 3 (left) and Figure 3 (right), shows that the avoidance of Hauptbahnhof and Sendlinger Tor did not have a noticeable effect on the area reachable within 20 minutes of the origin Sendlinger Tor. However, the curves that were affected the most are the ones from 30 to 40 minutes (blue and purple areas). When the crowded intersections are avoided, these areas became noticeably smaller. This reflects an increase in travel time for the destinations located within these areas. The last two curves that represent 50 and 60 minutes were also affected by the avoidance. In general, the areas reachable within 50 and 60 minutes are very similar in both isochrones, with a small decrease in reachability when avoidance is included.

4. Discussion

4.1. Result interpretation

Deleting two critical stations in the transportation network results in an overall increase in travel time, particularly for medium long routes. However, this increase is to be expected since the deleted nodes represent important connections in the transit network. Also, for most routes, the increase was so small that it is neglectable. This is specifically true for the random test set presented in Figure 2, where the average travel time for 1000 randomly generated routes only increased by 0.3 minutes. For longer routes, avoiding the central exchange points in Munich's traffic network usually results in less direct

routes with more line changes and an increase in waiting time. Regarding the travel time, the increase is rather small, when compared to the potential crowd reduction benefits presented in Figure 2. In contrast to our own expectations, passengers were evenly distributed by the algorithm, thereby preventing neighboring stations from overcrowding. However, this method could still pose a risk, if those neighboring stations are much smaller than the avoided station and therefore could be overloaded even by a comparably small number of passengers. Assuming that this is not the case, the changes to the overall travel time, as well as the stable physical activity level (walk/bike distance) required for a route, would be small enough to compensate for the potential health benefits. This is specifically the case for high-risk groups with a higher need for a safe passage.

4.2. Limitations

The greatest threat lies in the theoretical conception of the motivational aspects. Without real data from travelers, the true threshold for individuals to ignore a safe route can only be estimated. Also, a safety increase through the navigation around congested bottlenecks, can only be assumed but not measured yet. Hence, we do not know how effective such a measure could be unless it has been tested. Another bias lies in the selection of one single city to serve as a prove of concept. Even though the simulation produces promising results in Munich, there is currently no foundation on which the results could be compared to other cities.

4.3. Outlook

We suggest repeating the same evaluation with other cities in Germany, or even Europe, to understand how different transportation networks respond to the avoidance feature. With this measure, it could be determined whether the simulated results of this research are generalizable or whether they merely occurred due to the specific layout of Munich's traffic network. To reproduce these results in other cities, the public transport schedule data and position of traffic hot spots would be required. The format and accessibility of the public transport schedule data may vary between cities. During first investigations, we were able to obtain similar data sets for Nuremburg, Berlin and Duesseldorf. The second requirement for the extension of this work is to obtain the city specific knowledge about the most frequented nodes. For this knowledge a cooperation with local traffic experts would have to be established.

The results presented in the previous sections leave room for a variety of subsequent research areas. As hinted before, our research so far focused on simulations with planned trips. Consequently, the next step could be a comparison of these results to the real travel behavior of Munich's citizens. Currently such a data set does not exist, as there is no technological solution yet that can reliably track the lines, change points and modes of transport a traveler used in a route. Given such a solution was implemented, it could be used to compare the schedule data to the real travel behavior of passengers who are given the chance to test the avoidance router.

Finally, this router could not only be used to prevent infections with contagious diseases, but also to individualize routes to fit to the needs of physically impaired passengers. Firsts test with our algorithm have shown that it is possible to tailor the amount of physical activity through steps in a route to the settings of a passenger. Given that a passenger struggles with walking, it would be possible to generate routes that minimize the number of steps. Vice versa it is also possible to create routes that contain

a minimum number of steps for passengers who need or want to include more physical activity in their daily routines.

5. Conclusion

The presented approach simulated how safer routes that avoid overcrowded nodes in the traffic network of Munich could be generated, without causing the travel time to increase significantly. Even though it could not be determined how much safer such routes are, our simulations suggest that it could dissolve existing congestions without causing new ones at neighboring stations. This would allow passengers to maintain a safe distance to other passengers while they are in a public transport station. Determining when a node is at risk of being too congested was established through a collaboration with city specific transportation flow experts. In the here presented use case for the city of Munich, this knowledge was provided by the MVG who already monitor the public transportation network but required additional tools to steer the passenger flow and prevent congestions.

References

[1] C. Courtemanche, J. Garuccio, A. Le, J. Pinkston, A. Yelowitz, Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States, *Health Affairs* **39**(7) (2020), 1237-1246.

[2] K. Prather, C. Wang, R.T. Schooley, Reducing transmission of SARS-CoV-2, *Science* **368**(6498) (2020), 1422-1424.

[3] German Federal Ministry of Health, Novel Coronavirus SARS-CoV-2: Information and practical advice, https://assets.ctfassets.net/eaae45wp4t29/29lyMngJAvYypiWvqOTtVr/5707e9a6f5ea1681c4d0d1d5e09019f4/BMG_Corona_Information_englisch.pdf, last access: 31.01.2021

[4] Munich Transport Corporation, Aktuelle Zahlen und Fakten, https://www.mvg.de/dam/mvg/ueber/unternehmensprofil/mvg_in_zahlen_s, last access: 02.01.2021

[5] Munich Transport Corporation, Modernisierung U-Bahnhof Sendlinger Tor, https://www.mvg.de/ueber/mvg-projekte/bauprojekte/modernisierung-sendlinger-tor.html, last access: 13.03.2021

[6] T.R. Frieden, C.T. Lee, Identifying and interrupting superspreading events-implications for control of severe acute respiratory syndrome coronavirus 2, *Emerging infectious diseases* **26**(6) (2020), 1059.

[7] D. Bucher, D. Jonietz, M. Raubal, *A heuristic for multi-modal route planning*, in: Progress in Location-Based Services 2016, Springer, Germany, 2017. Pp. 211-229.

[8] C. Narboneta, K. Teknomo, OpenTripPlanner, OpenStreetMap, General Transit Feed Specification: Tools for Disaster Relief and Recovery, *CoRR*, arXiv:1609.01472 (2016).

[9] K.S. Bunds, J.M. Casper, J.A. Jonathan, J. Koenigstorfer, Recreational walking decisions in urban away-from-home environments: The relevance of air quality, noise, traffic, and the natural environment, *Transportation research part F: traffic psychology and behaviour* **65** (2019), 363-375.

[10] P.E. Hart, N.J. Nilsson, B. Raphael, A Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics* **4**(2) (1968), 100-107.

[11] M. Duckham, L. Kulik, M. Worboys, A. Galton, Efficient Generation of Simple Polygons for Characterizing the Shape of a Set of Points in the Plane, *Pattern Recogn.* **41**(10) (2008), 3224-3236.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210098

Health Pandemic and Social Media: A Content Analysis of COVID-Related Posts on a Telegram Channel With More Than One Million Subscribers

Simin MEHDIPOUR[a,b], Nazanin JANNATI[c], Mozhgan NEGARESTANI[a], Saber AMIRZADEH[d], Sareh KESHVARDOOST[e], Farzaneh ZOLALA[a], Atiyeh VAEZIPOUR[f], Maryam HOSSEINNEJAD[a,b,1] and Farhad FATEHI[g,h]

[a] *Social Determinants of Health Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran.*
[b] *Nursing and Midwifery Department, Islamic Azad University, Kerman Branch, Iran.*
[c] *Health Services Management Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran.*
[d] *Modeling in Health Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran.*
[e] *Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran.*
[f] *RECOVER Injury Research Center, The University of Queensland, Brisbane, Australia.*
[g] *School of Psychological Sciences, Monash University, Melbourne, Australia.*
[h] *School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran.*

Abstract. Background**:** Mobile-based social media play an important role in the dissemination of information during public health emergencies. Objectives: This study aimed to analyze the contents and trends of public messages posted on Telegram during Coronavirus Disease 2019 (COVID-19) pandemic. Methods**:** A content analysis of the 1781 messages, posted in a public Telegram channel with more than one million subscribers performed over 9-weeks. The messages were categorized into seven categories. Results**:** In total, 39% (n=703) of all messages were related to COVID-19. With the official confirmation of the case of COVID-19 in Iran, the number of COVID-related massages started to rise. Overall, the most frequent messages were of joke and humor (n=292, 41.5%), followed by educational messages (n=140, 19.9%). Conclusion**:** Our study showed that the most popular messages during first weeks of COVID pandemic were satirical, indicating that people may not had taken the risks of this pandemic seriously. It is crucial for health organizations to develop strategies for dissemination of reliable health information through social media.

Keywords. COVID-19, Coronavirus, Social media, Social networking, Mobile health.

[1] Corresponding Author, PhD Student in Health in Disaster and Emergencies, Social Determinants of Health Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran. E-mail: maryam.hosseininejad@gmail.com

1. Introduction

Evidence suggests that dissemination of the health information is a key to manage and control a pandemic like COVID-19 [1]. In particular, in emerging infectious disease outbreaks where people are drive to seek information from various channels due to the limited information available about the disease, source of infection, and consequences of the disease, the health-related content of social media can be critical [2-8].

Social media can also serve as a space for community to actively participate in the rapid sharing of information. They can lead to increasing awareness and reducing the risk of spreading the disease [1]. Some benefits of social media are increasing in people interaction, availability of information, accessibility of health information, exchanging social and emotional support, facilitating surveillance for public health, and informing future policy and planning [9]. However social media can provide an environment for the spread of information and correction of false information, they can also broadcast misinformation in health or other subjects [10-13].

Iran with a 23 million year-on-year hike has great growth in the number of mobile social media users [14]. Telegram is one of the most popular instant messaging mobile applications in Iran [15] with an estimated 45-50 million users [14, 16]. This messenger that was launched in 2013 [17], is an instant messaging service in which users can send and receive any type of files and multimedia, along with text messages. A Telegram channel is used for broadcasting various information and messages [18].

To date, there is little published research systematically analyzed the content and trends of health information in social media during health emergencies in Iran. The objectives of this study were to analyze the contents of the COVID-19 related messages posted on a popular Iranian Telegram channel during the first 9-weeks of COVID-19 pandemic and to determine their content, trend and changes over the time. This study is the first to provide insight into the coverage of health information on a messaging mobile app with over 45 million users in Iran.

2. Methods

We performed a retrospective content analysis of available public posts shared on Telegram. Content analysis is a methodology for quantitative analysis of social media messages and characterizing them in themes and subthemes. Content analysis has been frequently used in the analysis of websites and social media content [19-21].

We used Telegram analytics website [22] to identify the most popular public Telegram channels among Iranians. From the top five channels, we chose “Giz-Miz” channel based on Telegram analytics and experts’ opinion. The selection criteria were having more than 1,000,000 subscribers, publishing content for the general public and not having specific or professional contents. Giz-Miz with over 1.2 million subscribers is a Telegram channel for the general public that publishes posts about various social topics, including COVID-19. The language of the channel is Persian.

2.1. Data collection and coding

We retrieved all the posts on the selected Telegram channel through Telegram application programming interface (API), during a 9-week period from 23 January to 25

March 2020. After excluding the posts not related to Coronavirus or COVID-19, we extracted the images, videos, and texts, posting date, and the number of views for posts.

Two researchers read and analyzed the contents of a sample of 50 randomly selected posts independently and divided them into various categories and subcategories. Afterwards, the authors discussed all identified categories and agreed on the categories and subcategories. Seven categories identified. Then they independently coded all the remaining posts based on latent and manifest contents. In case of any disagreement, a third researcher was consulted. The coefficient of agreement (Cohen kappa) between two researchers was 0.82. Table 1 outlines the coding criteria and categories of the posts.

Table 1. Coding criteria and categories of the posts

Categories	Subcategories	Definition	Example
News	Reported Cases and Deaths	-	-Number of positive cases -Number of dead patients
Education	Prevention methods	Posts that teach people about disease prevention	- How to use a mask - How to wash hands
	Food and nutrition	Posts about recommended nutrition to strengthen the immune system during Coronavirus	- Healthy diet
	Corona Characteristics	Posts related to the characteristics of Coronavirus	-Survival of Coronavirus in the environment at different temperature
	Symptoms	Posts which teach different symptoms of COVID-19	-Fever, tiredness, dry cough
	Modes of transmission	Posts which discuss how Coronavirus could transmit among people	-Contact transmission
COVID-19 impact on society	Economic	Posts about how Coronavirus has affected people's incomes and the economies of countries	-The price of oil - Market crash
	Social	Posts about how Corona has influenced people's relationships and social life	-Closing schools, sports competitions and religious - Distribution free masks by people
Health care system	Diagnosis and treatment measures	Posts about measures have been taken to diagnose and treat COVID-19.	-Detection Kit -Treatment protocols
	Health care workers	Posts about Problems and challenges health care workers face.	-Lack of Personal protective equipment
Joke & humor	Social	Posts about something funny about Coronavirus which was said or done by people in the society	-Traditional Iranian Medicine
	Political	Posts about Something funny about Coronavirus which was said or done by government politicians.	-Elections during the spread of Corona
Combating coronavirus	Travel	Posts that discuss actions to prevent people from travelling	-Traffic restrictions
	Cleaning and Disinfecting Public Spaces	Posts about actions have been taken to disinfect cities.	-Disinfection of buses, gas stations and banks
	Providing Online Services	Posts about online services aimed at reducing people's presence in crowded places	-E-Banking, online shopping and E-learning
	Social distancing	Posts about actions have been taken in the community to maintain social distance	-Keep distance from other people -Work from home
Criticism	Social	Posts that criticized people's behavior in society	-Mask holders
	Political	Posts that criticized the government's performance on Coronavirus's actions	- No quarantine of cities - Iran's Mahan Air continue Flights to China

2.2. Quantitative data analysis of the posts

We performed descriptive statistics with Microsoft Excel 2019 and SPSS version 26. We calculated frequency and percentage to describe the characteristics of the posts including the post format, topics covered and the number of views. We presented weekly trends of the posts based on the identified categories. The weeks are numbered according to the timeline of the COVID-19 outbreak so that the week that the first case of disease was officially confirmed in Iran was labelled as week 0. Three weeks before that was labelled as week-3 to -1. The five weeks after the announcement of the first case were labelled as week 1 to 5. It's noticeable that the number of disease cases had a considerable increase over time and reached to 9656 cases in week 5.

3. Results

A total of 1781 posts were retrieved, reviewed, and analyzed. From these, 39% (n=703) were related to COVID-19 or coronavirus. Other posts were fun and humor, political and social, etc. The average number of views of posts was 526,970 (140,000~ 2,000,000). Coronavirus related posts were three types included of videos (n=293, 41.7%), photos (n=291, 41.4%) and text messages (n=119, 16.9%).

3.1. Content of the posts

Table 2 indicates the number of posts by categories on the channel. Most frequent post category was joke and humor with 41.5% (n=292); the number of social joke posts was seven times as many as the number of political joke posts. Education with 140, had the second highest number of posts. In this category, the prevention methods had more posts (63.5%) than other subcategories. Ranked in the third place was criticism, in which the percentage of social criticism was two times more than political criticism with 63.5% and 36.5%, respectively. Coronavirus's impact on society had the lowest number of posts with just 3.3%.

Table 2. The number of the posts by categories and subcategories

Categories	Subcategories	N (%)
News	Reported Cases and Deaths	54 (7.7)
Education	Prevention methods	89 (12.7)
	Food and nutrition	11 (1.6)
	Corona Characteristics	18 (2.6)
	Symptoms	9 (1.3)
	Modes of transmission	13 (1.8)
COVID-19 impact on society	Economic	11 (1.6)
	Social	12 (1.7)
Health care system	Diagnosis and treatment measures	26 (3.7)
	Health care workers	28 (4)
Joke & humor	Social	256 (36.4)
	Political	36 (5.1)
Combating COVID-19	Travel	24 (3.4)
	Cleaning and Disinfecting Public Spaces	6 (0.9)
	Providing Online Services	5 (0.7)
	Social distancing	31 (4.4)
Criticism	Social	47 (6.7)
	Political	27 (3.8)
Total		**703(100)**

3.2. The trend of the Posts

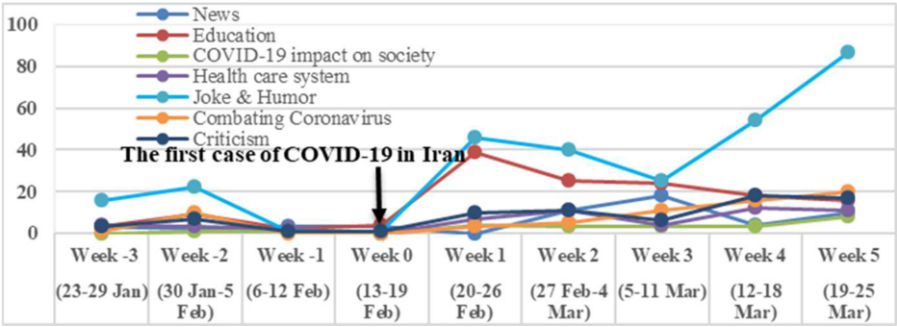


Figure 1. The trend of the messages by category during the first 9 weeks of COVID-19 pandemic

Figure 1 shows the trend of COVID-19 posts over time. All categories declined in the week -1 and 0. Afterwards, with an exception to education and joke posts, all posts showed a modest steady increase over time. With regards to education posts, there has been a high increase in week 3 before declining over time. However, jokes and humor post soared and reached to their max at the end of the study.

4. Discussion

Based on our findings, there was a rise in the number of humor posts as the number of cases increased. Although this finding is counterintuitive, the literature shows that during past public health crises, humors were spread through social media, too. For example, in the analysis of the content of twitter messages during H1N1 Influenza epidemic, the highest proportion (12.7%) of tweets was of humor type [6]. Also in another study about Ebola, the highest proportion (23%) of the posts, which were published on two image sharing platforms, were jokes [23]. Moreover, research shows that Jokes were the most common posts on twitter during the Zika outbreak [24]. This could be illuminated through different explanations. First, humors including social media jokes could be as a tool for stress reduction with a protective role in mental health [25] pandemic with such large extent as COVID-19 can affect mental health and develop stress-related disorder [26]. This unprecedented situation with social distancing and uncertainty about the disease takes a toll on the mental health of the population. This will be exacerbated when the population have been already affected by preexisting problems. For example, in Iran economic stagnation preceded the pandemic had already caused considerable stress and economic hardship in people [27]. Therefore, viewing humor posts help people better cope with adversity. The second reason that jokes are popular during the crises could be chasing a safe and subtle alternative way of criticizing authorities particularly when there might be consequences for the serious criticism [28, 29]. Based on Sulistyaningtyas et al, humors can appear as a discourse and social mechanisms in response to a critical situation like COVID-19 pandemic [1]. Finally, many useful information can be emitted through jokes. People can benefit from learning in a relaxing and fun atmosphere [1]. Moreover, in our study, a significant rise in the number of posts in week 5 can be due to approaching the Iranian New Year. Recommendation by health authorities for adhering to social distancing rules could motivate people to spread more posts on jokes through social media to create a more relaxing and joyful atmosphere.

The second most popular group of posts was education. Up to the time of drafting this manuscript, due to lack of vaccine and effective treatment, the best strategy to deal with COVID-19 has been prevention, which is well presented in the educational posts (63.6% of the educational posts) [30]. Another noticeable point of our findings was that despite the huge social effects of coronavirus pandemic on daily lives and social occasions, they were rather failed to address adequately. It seems that in similar studies, a larger proportion of social media messages contained the social and economic consequences of the disease epidemic. In one study, for example, 39 of 100 messages expressed fear/anxiety in the community and 15 of them mentioned the need for medical or economic assistance. These can reflect the social and economic effects of the disease [5]. In another study, 23% of the messages contained personal experiences of the disease, which can be considered a reflection of the social impact of the disease [6].

Health care system was another category which drew little attention. Having preexistence of human and financial limitations in place, as with any pandemic, over the study period health care system was overloaded with the number of patients. However, this was not addressed through messages which were posted (less than 8%). In a study of coverage Ebola epidemic on YouTube, 18% of posts discussed dangers for health care personnel, 11% was about medical help/resources in Africa and 2% discussed need for health care training; that is 31% totally, and it's a noticeable proportion of the posts [5]. Our findings may be an indication of low awareness of the critical role of the health system in the crises among the social media users.

Regarding the trend of the posts, in general, about 30% of posts (30 out of 103) were relevant to COVID-19 before the official announcement of the first case of COVID-19 in Iran (from 23 January to 18 February 2020). Despite this fact that coronavirus was declared as a public health emergency in the world, the greatest number of posts was joke and humor in the first two weeks. This shows that COVID-19 might not be perceived as a serious health issue in Iran. During this time, all flights from China were still in place, and no practical action was taken to prevent arriving the Coronavirus in Iran [31]. During the third week, the number of all COVID-19 posts reached its lowest level, maybe due to coinciding the Father's Day celebration in Iran which made people concentrate on it.

The number of posts increased and peaked a week after the official announcement of the first case of COVID-19 in Iran. During this week, the number of confirmed cases increased from 2 to 137. This sharp increase could be due to disregarding social distancing directives by the people who wanted to participate in the parliament election at the beginning of the week, despite the announcement of the disease. As a result of this increase in the number of COVID-19 cases, the number of education posts reached its peak because people needed information and became more aware of the disease risks and severity. Studies have shown that because of concerns in emergencies, people demand and search for information [2, 3, 25], and become more interested in gaining information about the disease [32]. However results of a study on Twitter during influenza H1N1 pandemic has shown that the number of H1N1-related Tweets declined while the number of cases of the disease was increasing [33]. This difference could be due to the fact that COVID-19 is a more contagious, widespread and new unknown disease that has caused great concern in the world.

In the first week after the onset of COVID-19 in Iran, joke and humor had the highest number of the posts; it may be because of stress control as mentioned earlier; but during the second and third weeks, along with the increase in the number of patients, the number of joke decreased. This show that the public knew that the situation had become more

critical. In the assessment of Twitter during 2009 H1N1 epidemic, the proportion of Tweets containing jokes declined over time [6]. In addition, the number of educational posts declined in these two weeks. A hypothesis can because of decreasing perceived severity over time, as had shown in a study on H1N1 pandemic [34]. The posts related to combating Coronavirus was the only category that has been on the rise from the 6th to the 9th week, reflecting that the government has taken measures such as disinfecting cities and traffic bans to prevent further spread of the disease.

In this study, we only retrieved posts related to coronavirus and COVID-19 and did not examine the trustworthiness of the posts. In addition, the messages analyzed were retrieved from one public Telegram channel and our findings may not be generalizable to channels with a specific topic of interest. However, the selected Telegram channel is one of the most popular public Telegram channels in Iran with diverse users of different ages and thus can be regarded as having a relatively high impact on the general public.

4.1. Conclusion

We analyzed the content of social media posts related to COVID-19, which were published on Telegram. Different categories of the posts have emerged from our data. Among them, jokes and educational posts were the most common. Posting a considerable amount of posts related to COVID-19 crisis highlighted the prominent role of social media in public emergencies. This could play a critical role in public education in such emergencies where people need and search for information. Reliable and valid information could help to control the widespread public health crises and their consequences. We also found that even with such an extensive crisis with many casualties, people could be distracted by social issues such as an election or traditional ceremonies. Therefore, appropriate knowledge and education should be made available by health organizations to the public to prevent the spread of the disease during outbreaks.

References

[1] Sulistyaningtyas T, Jaelani J, Suryani Y. Power of Knowledge and Community Social Class above Covid-19 Pandemic Information on Social Media. *Jurnal Komunikasi Ikatan Sarjana Komunikasi Indonesia* **5**(1) (2020), 52-62.

[2] Chen B, Shao J, Liu K, Cai G, Jiang Z, Huang Y, et al. Does Eating Chicken Feet With Pickled Peppers Cause Avian Influenza? Observational Case Study on Chinese Social Media During the Avian Influenza A (H7N9) Outbreak. *JMIR public health and surveillance* **4**(1) (2018), e32.

[3] Gu H, Chen B, Zhu H, Jiang T, Wang X, Chen L, et al. Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *Journal of medical Internet research* **6**(1) (2014), e20.

[4] Barrelet C, Bourrier M, Burton‐Jeangros C, Schindler M. Unresolved issues in risk communication research: the case of the H 1 N 1 pandemic (2009–2011). *Influenza and other respiratory viruses* **7**(2013), 114-9.

[5] Basch CH, Basch CE, Ruggles KV, Hammond R. Coverage of the Ebola virus disease epidemic on YouTube. *Disaster medicine and public health preparedness* **9**(5) (2015), 531-5.

[6] Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*.**5**(11) (2010), e14118.

[7] Jones JH, Salathe M. Early assessment of anxiety and behavioral response to novel swine-origin influenza A (H1N1). *PLoS one***4**(12) (2009), e8032.

[8] Tuarob S, Tucker CS, Salathe M, Ram N, editors. Discovering health-related knowledge in social media using ensembles of heterogeneous features. *Proceedings of the 22nd ACM international conference on Information & Knowledge Managemen*, 2013.

[9] Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research* **15**(4) (2013), e85.
[10] Al Khaja KA, AlKhaja AK, Sequeira RP. Drug information, misinformation, and disinformation on social media: a content analysis study. *Journal of public health policy* 39(3) (2018), 343-57.
[11] Bode L, Vraga EK. See something, say something: Correction of global health misinformation on social media. *Health communication*. 33(9) (2018), 1131-40.
[12] Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine* **240** (2019), 112552,4-12.
[13] Waszak PM, Kasprzycka-Waszak W, Kubanek A. The spread of medical fake news in social media–the pilot quantitative study. *Health policy and technology* **7**(2) (2018), 115-8.
[14] Latest Data on Iran: Surge in Social Media Use 2018, https://financialtribune.com/articles/sci-tech/81536/latest-data-on-iran-surge-in-social-media-use, updated February 06, 2018.
[15] Bastani P, Bahrami MA. COVID-19 Related Misinformation on Social Media: A Qualitative Study from Iran. *Journal of medical Internet research* **22**(6) (2020), e20021.
[16] Iqbal M. Telegram Revenue and Usage Statistics (2020) , https://www.businessofapps.com/data/telegram-statistics/ updated June 23, 2020..
[17] Number of monthly active Telegram users worldwide from March 2014 to April 2020 , https://www.statista.com/statistics/234038/telegram-messenger-mau-users/ , 2020.
[18] Dargahi Nobari A, Reshadatmand N, Neshati M, editors. Analysis of Telegram, an instant messaging service. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management;* 2017.
[19] Cherian R, Westbrook M, Ramo D, Sarkar U. Representations of Codeine Misuse on Instagram: Content Analysis. *JMIR Public Health and Surveillance* **4**(1) 2018, **e22**.
[20] Liao Q, Yuan J, Dong M, Yang L, Fielding R, Lam WWT. Public Engagement and Government Responsiveness in the Communications About COVID-19 During the Early Epidemic Stage in China: Infodemiology Study on Social Media Data. *Journal of Medical Internet Research* **22**(5) (2020), e18796.
[21] Massey PM, Kearney MD, Hauer MK, Selvan P, Koku E, Leader AE. Dimensions of Misinformation About the HPV Vaccine on Instagram: Content and Network Analysis of Social Media Characteristics. *Journal of Medical Internet Research*, **22**(12) (2020), e21451.
[22] Telegram analytics: Telegram channels 2020, https://tgstat.com/.
[23] Seltzer EK, Jean N, Kramer-Golinkoff E, Asch DA, Merchant R. The content of social media's shared images about Ebola: a retrospective study. *Public health* **29**(9) (2015) 1273-7.
[24] Gui X, Wang Y, Kou Y, Reynolds T, Chen Y, Mei Q, et al., editors. Understanding the Patterns of Health Information Dissemination on Social Media during the Zika Outbreak. *AMIA Annual Symposium proceedings* AMIA Symposium; 2017.
[25] Boerner M, Joseph S, Murphy D. The association between sense of humor and trauma-related mental health outcomes: Two exploratory studies. *Journal of Loss and Trauma* **22**(5) (2017), 440-52.
[26] Horesh D, Brown AD. Traumatic stress in the age of COVID-19: A call to close critical gaps and adapt to new realities. *Psychological Trauma: Theory, Research, Practice, and Policy* **12**(4) (2020), 331.
[27] Zandifar A, Badrfam R. Fighting COVID-19 in Iran; economic challenges ahead. *Archives of Iranian medicine* **23**(4) (2020), 284.
[28] Hart Mt. Humour and social protest: An introduction. *International Review of Social History* **52**(S15) (2007), 1-20.
[29] Oring E. Risky business: Political jokes under repressive regimes. *Western folklore*.**63**(3) (2004), 209-36.
[30] Wilder-Smith A, Freedman DO. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. *Journal of travel medicine* **27**(2) (2020) ,taaa020.
[31] TheNewArab. Mahan Air's return flights to China 'fuelled coronavirus outbreak in Iran' 2020, https://english.alaraby.co.uk/english/news/2020/3/31/iranian-airlines-return-flights-to-china-fuelled-coronavirus-outbreak, updated March 31, 2020..
[32] Eysenbach G. Infodemiology and Infoveillance: Framework for an Emerging Set of Public Health Informatics Methods to Analyze Search, Communication and Publication Behavior on the Internet. *Journal of Medical Internet Research* **11**(1) (2009), **e11**.
[33] Signorini A, Segre A, Polgreen P. The Use of Twitter to Track Levels of Disease Activity and Public Concern in the US during the Influenza A. *PLoS One* **6**(5) (2011), 10.
[34] Bults M, Beaujean DJ, de Zwart O, Kok G, van Empelen P, van Steenbergen JE, et al. Perceived risk, anxiety, and behavioural responses of the general public during the early phase of the Influenza A (H1N1) pandemic in the Netherlands: results of three consecutive online surveys. *BMC Public Health*, **11**(2) **(**2011), 1-13.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210099

A Mechanism for Verifying the Integrity and Immutability of Tuberculosis Data Using IOTA Distributed Ledger Technology

Vinícius LIMA[a,b,1], Filipe BERNARDI[a,b], Rui RIJO[c], Jó UEYAMA[d] and Domingos ALVES[e]

[a] *Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil*
[b] *Bioengineering Postgraduate Program, São Carlos School of Engineering, University of São Paulo, São Carlos, Brazil*
[c] *School of Technology and Management, Polytechnic Institute of Leiria, Leiria, Portugal*
[d] *Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil*
[e] *Department of Social Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil*

Abstract. Background: Intensified research and innovation and rapid uptake of new tools, interventions, and strategies are crucial to fight Tuberculosis, the world's deadliest infectious disease. The sharing of health data remains a significant challenge. Data consumers must be able to verify the consistency and integrity of data. Solutions based on distributed ledger technologies may be adequate, where each member in a network holds a unique credential and stores an identical copy of the ledger and contributes to the collective process of validating and certifying digital transactions. Objectives: This work proposes a mechanism and presents a use case in Digital Health to allow the verification of integrity and immutability of TB electronic health records. Methods: IOTA was selected as a supporting tool due to its data immutability, traceability and tamper-proof characteristics. Results: A mechanism to verify the integrity of data through hash functions and the IOTA network is proposed. Then, a set of TB related information systems was integrated with the network. Conclusion: IOTA technology offers performance and flexibility to enable a reliable environment for electronic health records.

Keywords. Electronic Health Records, Health Informatics, Tuberculosis, Computer Security

1. Introduction

Measured by the number of people who die each year, tuberculosis (TB) is the world's deadliest infectious disease [1]. To achieve the milestones for reductions in cases and deaths, the World Health Organization (WHO) has defined the global strategies and targets for TB prevention, care, and control, as part of The End TB Strategy, which defines 3 pillars to reduce TB burden [2]. One of these pillars consists of incorporating

[1] Corresponding Author: Vinícius Lima, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil, E-mail: viniciuslima@usp.br

an environment supported by digital health to promote intensified research and innovation and rapid uptake of new tools, interventions, and strategies.

In Brazil, the Tuberculosis Ecosystem (SISTB) for monitoring tuberculosis treatment via the Directly Observed Treatment (DOT) strategy and patients' follow-up has been developed [3], consisting of several modules. Information about the treatment may be recorded online or offline, as well as remote self-registration of medication intake by the patient through the upload of a video for a health team (VDOT) [4]. Also, the SISTB is able to communicate and share health data with other national and international repositories through a semantic interoperability layer [5].

TB records are composed of a variety of health data, such as demographics, treatment, exams, daily registration of medication intake, and contacts tracking. These pieces of data are linked together and visualized according to the objective of the health professional. Therefore, TB health information systems are segmented in several sections, where each one presents data within their context.

Despite these advances in the attendance of TB patients, the interinstitutional sharing of health data remains a significant challenge. The ineffective process of sharing data in the health sector results, in part, from the lack of trust between providers and data interoperability between healthcare IT systems and applications [6]. Data consumers must be able to verify the consistency and integrity of data. Trust relationships often exist between providers in the network and/or healthcare organizations but are particularly difficult to maintain due to the lack of consensus that the data can generate if the parties do not use the same healthcare system with a shared provider directory [7].

To address such challenges and enable patient and health stakeholders' control and autonomy, solutions using distributed ledger technologies (DLT) may be adequate. DLT's also offer opportunities for clinical research, real-time access to individualized data and the ability to set permissions for accessing and auditing data to ensure their integrity and security [8].

A DLT is based on the principles of peer-to-peer (P2P) network and cryptographic primitives. Each member in the network holds a unique credential and stores an identical copy of the ledger and contributes to the collective process of validating and certifying digital transactions for the network [9]. Information is encrypted and digitally-signed to guarantee authenticity and accuracy. Blockchain is one example of a DLT, which became famous with the digital currency Bitcoin in 2008 [10], but there are other approaches with distinct operation.

In this sense, this work proposes a mechanism and presents a use case in Digital Health, in which the main goal is to allow the verification of integrity and immutability of TB electronic health records using a DLT-based approach.

2. Methods

This research seeks to address an open issue in security of health data. Based on the relevance of the theme, a research question was defined to drive the development of the solution, as follows: "How to deliver a mechanism to allow systems' users to verify the integrity and immutability of a given TB health information?".

Due to the sensitivity of data, one key challenge to be faced is to provide guarantees that a given record, e.g., a video or an electronic health record, was not corrupted. An approach deployed in the SISTB Ecosystem is proposed to enable the possibility of

consumer systems/users to check the integrity of a health record by using hash functions and a DLT.

Hash functions are algorithms that transform a message of random length and generate a fixed-length output that acts as a unique identifier of a message. Any minimal changes to the original message will lead to an entirely different hash value [11]. In this sense, the solution is relevant because it may increase reliability of stored data.

2.1. IOTA

An attractive alternative to traditional blockchain architectures for DLTs is the Directed Acyclic Graph (DAG). DAG ledgers can provide many gains over traditional blockchain, including performance, scalability and transaction costs. In this work, we adopted the IOTA, developed in 2015, with no trading fees, blocks, or mining. Based on the M2M (Machine to Machine) principle, IOTA was designed specifically for the Internet of Things (IoT) industry with an interconnected architecture through a tangled network, the Tangle network [12].

All the data flow management can happen in channels with the exchange of Masked Authenticated Messages (MAM). MAM protocol is one of the most remarkable features of IOTA. It allows data to be shared securely via an encrypted channel. In this way, it is possible to ensure that the recipient of a transaction receives data with integrity through a trusted source [13]. To decrypt data stored in transactions, the key must be known, so only authorized parties can access the data.

The IOTA community maintains two public networks, and each one has its own Tangle to which nodes can attach transactions: the Mainnet and the Devnet. In addition, a private node can be deployed and attached to the network.

3. Results

The main outcomes of this work consist in the following: i) the deployment of a private node in the IOTA Devnet public network; ii) a mechanism to perform verifications of integrity of data stored in the SISTB and in the VDOT App. The SISTB Ecosystem relied on IOTA technology to increase trust in data accessed by multiple actors.

The private node was deployed in a virtual server. The software, called Hornet [14], is provided by the IOTA community. Although being connected to a public network, owning a private node allows an organization to securely interact with the Tangle, with specific permissions and without depending on the availability of third-party nodes.

The IOTA node automatically exposes an Application Programming Interface (API) to enable interoperability with external applications. The API allows writing and reading operations, i.e., sending and retrieving transactions, as well as the communication through confidential channels with encrypted data (MAM).

To communicate with the node, an additional auxiliary API was developed in Node.js that directly connects to the node's API using programming libraries [15] maintained by the IOTA community. This API was able to authenticate and send/retrieve transactions to/from the Tangle and work as a bridge to the network. Finally, through specific endpoints of the interoperability layer available in the SISTB Ecosystem, authorized systems were able to connect to the auxiliary API to interact with the Tangle. Figure 1 summarizes the communication flow and the integration between components.

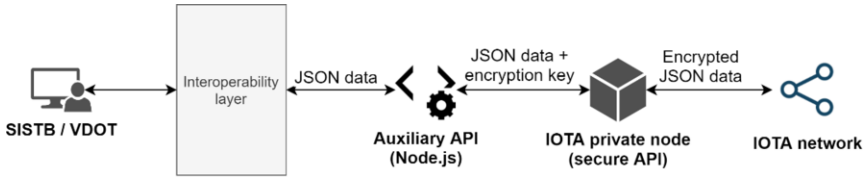


Figure 1. Communication flow

The SISTB Ecosystem needed the auxiliary API to interact with the Tangle. When an electronic health record was created or updated, or a video was submitted through the VDOT App, a data entry trigger calculated a hash for the data and called the auxiliary API, supplying the record ID, the hash and a timestamp as a JSON object. Then, a communication channel was established with the IOTA node, which effectively submitted the transaction (with data) to the Tangle. After successfully committing it, the transaction's ID was returned to the auxiliary API, which was stored it in the history of transactions for further reading operations. Figure 2 presents an example of data stored in the Tangle.

MESSAGE JSON


```
{
   "cod": "11245875",
   "hash": "E9E16315EAE1308D458BCBE079000C0632F5F2C35100DAF6B14330646E241536",
   "timestamp": "2020-12-07 12:34:05"
}
```


Figure 2. Data in JSON format stored in a transaction. This transaction can be verified at: https://explorer.iota.org/devnet/transaction/MLJOTLKQSVVXHJZBQCVMNBYSVMKCKCZBOMWYWGPEAERZZFDMQBHRCPPRXREHKIWFCQPOKANNGVIPEF999

Considering that this solution uses a shared network, it is important to keep data confidential in the Tangle. Therefore, the MAM protocol was used, which receives the message and a symmetric key known only by auxiliary API to encrypt and decrypt data. Systems from the SISTB Ecosystem do not know the key, because this API is the single point that writes and reads data from the Tangle.

Finally, a consumer system (or user) that wants to verify the integrity of a specific data can calculate the hash of the data by itself and compare it with the hash stored in the Tangle. If hashes match, it means that the data was not changed. When disclosing data for users, interfaces of the systems usually present the hash calculated in real-time, that is, when the user navigates to the page that contains the data, the transaction ID associated with that data and the hash extracted from the Tangle. If there is a mismatch, the user is alerted about a possible data corruption. Also, with the transaction ID in hands, the user can manually double-check the hash in the Tangle through an explorer tool provided by the IOTA community.

4. Discussion

IOTA technology is a free and open-source tool that enables a reliable environment for electronic health records underpinned by DLT features, such as data immutability, traceability and tamper-proof characteristics. Most existing studies on IOTA applications in health information systems have focused on conceptual designs and systems for verifying the integrity and immutability of electronic health records or on proposals with

managerial aspects of information [16]. Unlike the literature, a functional prototype was implemented to demonstrate the feasibility of the solution proposed in this work.

The main challenges of traditional blockchain-based solutions have been gradually overcome by IOTA Tangle and MAM. Despite being a constantly evolving solution, afforded by your active community, solid aspects in terms of cost, efficiency, scalability and flexibility in managing data access have been identified as advantageous in relation to their predecessors. Moreover, in health information systems, where the information is sensitive, the solution must comply with the latest regulations (for example, LGPD, a Brazilian GDPR). The immutability, traceability and tamper-proof characteristics reduce the risk of irregular processing and storage [17].

In this work, it is assumed that the party (SISTB) storing and sharing TB data can be trusted. One may say that, in this case, using a DLT tool becomes obsolete. However, the proposed solution seeks to protect a data consumer (other systems or users) from obtaining malicious or altered data in case of an unexpected security breach.

Finally, although a private node was deployed, a public network was used. The private node guarantees the connection with the Tangle with a dedicated and protected API. To increase availability, additional nodes can be configured for redundancy. Ideally, migrating for a private network is desirable, allowing a full control of nodes and participants.

5. Conclusion

This work presented the use of IOTA technology to deliver a mechanism to enable the verification of integrity of tuberculosis related data stored in the SISTB Ecosystem, enhancing trust on data in a reliable, safe and controllable way. IOTA offers performance and flexibility to enable a reliable environment for electronic health records.

Although the approach presented in this work was motivated by the TB scenario, its application is suitable in other health areas. The only requirement is to be able to calculate hashes for pieces of data in a given health information system to, then, write it into the IOTA network.

As future work, it is expected to establish a whole private network based on IOTA technology, so confidentiality and access control could be increased, which will allow the use of the network for additional purposes.

References

[1] Goldrick B. Once Dismissed, Still Rampant: Tuberculosis, the second deadliest infectious disease worldwide. AJN, Am J Nurs. 2004;104(9):68–70.

[2] WHO. A Global Action Framework for TB Research in Support of the Third Pillar of WHO´s End TB Strategy. 2015;75. Available from: https://www.who.int/tb/publications/global-framework-research/en/

[3] Crepaldi NY, Costa Lima V, Andrade Bernardi F, Albano dos Santos LR, Yamaguti VH, Pellison FC, et al. SISTB: an ecosystem for monitoring TB. Procedia Comput Sci [Internet]. 2019;164:587–94. Available from: https://doi.org/10.1016/j.procs.2019.12.224

[4] Albano dos Santos LR, Andrade Bernardi F, Prado GCS, Costa Lima V, Crepaldi NY, Marçal MA, et al. The perception of health providers about an artificial intelligence applied to Tuberculosis video-based treatment in Brazil: a protocol proposal. Procedia Comput Sci [Internet]. 2019;164:595–601. Available from: https://doi.org/10.1016/j.procs.2019.12.225

[5] Pellison FC, Rijo RPCL, Lima VC, Crepaldi NY, Bernardi FA, Galliez RM, et al. Data Integration

in the Brazilian Public Health System for Tuberculosis: Use of the Semantic Web to Establish Interoperability. JMIR Med Informatics. 2020;8(7).

[6] Cichosz SL, Stausholm MN, Kronborg T, Vestergaard P, Hejlesen O. How to Use Blockchain for Diabetes Health Care Data and Access Management: An Operational Concept. J Diabetes Sci Technol. 2019;13(2):248–53.

[7] Zhang P, Walker MA, White J, Schmidt DC, Lenz G. Metrics for assessing blockchain-based healthcare decentralized apps. 2017 IEEE 19th Int Conf e-Health Networking, Appl Serv Heal 2017. 2017;2017–Decem:1–4.

[8] Liang X, Zhao J, Shetty S, Liu J, Li D. Integrating blockchain for data sharing and collaboration in mobile healthcare applications. IEEE Int Symp Pers Indoor Mob Radio Commun PIMRC. 2018;2017–Octob:1–5.

[9] Wüst K, Gervais A. Do you need a Blockchain? IACR Cryptol ePrint Arch [Internet]. 2017;(i):375. Available from: https://eprint.iacr.org/2017/375.pdf

[10] Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008; Available from: https://bitcoin.org/bitcoin.pdf

[11] Teh J Sen, Tan K, Alawida M. A chaos-based keyed hash function based on fixed point representation. Cluster Comput [Internet]. 2019;22(2):649–60. Available from: https://doi.org/10.1007/s10586-018-2870-z

[12] Bhandary M, Parmar M, Ambawade D. A Blockchain Solution based on Directed Acyclic Graph for IoT Data Security using IoTA Tangle. Proc Fifth Int Conf Commun Electron Syst (ICCES 2020). 2020;(ICCES):827–32.

[13] Bhandary M, Parmar M, Ambawade D. Securing Logs of a System - An IoTA Tangle Use Case. Proc Int Conf Electron Sustain Commun Syst ICESC 2020. 2020;(ICESC):697–702.

[14] Community TI. HORNET - The IOTA community node [Internet]. Available from: https://github.com/gohornet/hornet

[15] Community TI. iota.js GitHub repository [Internet]. Available from: https://github.com/iotaledger/iota.js

[16] Zheng X, Sun S, Mukkamala RR, Vatrapu R. Accelerating Health Data Sharing : A Solution Based on the Internet of Things and Distributed Ledger Technologies Corresponding Author : 21.

[17] Hawig D, Zhou C, Fuhrhop S, Fialho AS, Ramachandran N. Designing a distributed ledger technology system for interoperable and general data protection regulation–compliant health data exchange: A use case in blood glucose data. J Med Internet Res. 2019;21(6):1–13.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210100

Machine Learning Based Risk Prediction for Major Adverse Cardiovascular Events

Michael SCHREMPF[a,1], Diether KRAMER[a], Stefanie JAUK[a,b], Sai P. K. VEERANKI[a], Werner LEODOLTER[a] and Peter P. RAINER[b]
[a] *Steiermärkische Krankenanstaltengesellschaft m. b. H., Graz, Austria*
[b] *Medical University of Graz, Graz, Austria*

Abstract. Background: Patients with major adverse cardiovascular events (MACE) such as myocardial infarction or stroke suffer from frequent hospitalizations and have high mortality rates. By identifying patients at risk at an early stage, MACE can be prevented with the right interventions. Objectives: The aim of this study was to develop machine learning-based models for the 5-year risk prediction of MACE. Methods: The data used for modelling included electronic medical records of more than 128,000 patients including 29,262 patients with MACE. A feature selection based on filter and embedded methods resulted in 826 features for modelling. Different machine learning methods were used for modelling on the training data. Results: A random forest model achieved the best calibration and discriminative performance on a separate test data set with an AUROC of 0.88. Conclusion: The developed risk prediction models achieved an excellent performance in the test data. Future research is needed to determine the performance of these models and their clinical benefit in prospective settings.

Keywords. Major Adverse Cardiovascular Events, Cardiovascular Diseases, Myocardial Infarction, Stroke, Machine Learning, Cardiovascular Risk, Electronic Medical Records, Risk Assessment

1. Introduction

1.1. Problem Statement

Cardiovascular diseases are one of the major death causes. In Austria, almost 40% of deaths are due to major adverse cardiovascular event (MACE) such as myocardial infarction or stroke [1]. During their lifetime, patients with cardiovascular diseases suffer from frequent hospitalizations and treatments, and their outcome depends often on 24-hour care after a severe cardiovascular event [2]. By identifying the risk at an early stage, many of these cases could have been prevented [3].

In many cases, a MACE is caused by advanced atherosclerotic vascular diseases. Atherosclerosis is strongly influenced by smoking, poor diet and lack of exercise. Since atherosclerotic vascular diseases develop over a long period of time, an early risk identification and preventive actions can decrease the risk of MACE [3].

[1] Corresponding Author: Michael Schrempf, Steiermärkische Krankenanstaltengesellschaft m.b.H., Graz, Austria, E-Mail: michael.schrempf@kages.at

One way of identifying patients at risk is the use of clinical prediction models. Common risk scores for cardiovascular diseases are the Framingham Risk Score [4], QRISK3 [5], SCORE [6] or ACC/AHA Risk Score [7].

The advantage of these scores is the use of a few predictors for the risk estimation. QRISK3 uses 22 predictors including age, smoking and prevalent diseases to predict the 10-year risk of a cardiovascular event. In a validation cohort, the score achieved a high discriminative performance with a C-statistic of 0.88 [5].

Compared to this, the Framingham Risk Score achieved an Area Under the Receiver Operating Characteristic (AUROC) of 0.79 in women. However, the Framingham Risk Score is using seven parameters to calculate the 10-year risk of cardiovascular disease. Furthermore, it should not applied to patients with diabetes.

The ACC/AHA Risk Score achieved an AUROC of 0.818 for African-American women. The poorest AUROC performance for non-Hispanic white men was 0.713.

Similar to the before mentioned score, the SCORE AUROC ranged from 0.71 to 0.84. It is using similar feature as the Framingham Risk Score and ACC/AHA Risk Score and was developed for European clinical practice.

However, all of the mentioned scores require manual input for risk assessment. The manual scoring often presents an additional burden for healthcare professionals instead of optimizing care. In addition, most risk scores are based on linear methods which are not able to account for non-linear relationships in highly complex data [8].

Therefore, the use of machine learning (ML) techniques is an alternative. ML algorithms can be trained on already available electronic medical records (EMRs) and account for numerous predictions in order to achieve high predictive performance.

Weng et al. [9] demonstrated that machine learning based approaches are able to predict cardiovascular events. However, their best performing model, a neural network, achieved only a performance of an AUROC of 0.764, which was comparable to the ACC/AHA score. Krittanawong et al. [10] demonstrated that the performances of the risk score could be outperformed by machine learning. This meta-analysis of prediction models for cardiovascular diseases showed that machine learning models can achieve AUROCs between 0.88 and 0.93.

1.2. Objectives

The aim of this study was to develop a prediction model based on electronic medical records to predict the 5-year risk of MACE. A machine learning approach was used in order to estimate the risk for each patient without additional efforts by clinicians and to account for various predictors available in the EMR system.

2. Methods

2.1. Data Overview

Data for modelling was provided by the public hospital provider Steiermärkische Krankenanstaltengesellschaft m.b.H. (KAGes). KAGes covers approximately 90% of acute care beds in the province of Styria, Austria, and provides ambulatory care services. Therefore, longitudinal health records of 2.4 million patients in Styria are available starting from 2006. The extracted EMR data included among others demographic data, ICD-10 coded diagnoses, nursing assessment, LOINC codes for laboratory data,

procedures, transfer data and ATC classes for medication. Available EMR data was included within the time span of 01/01/2009 and 12/31/2019. All patients under the age of 18 were excluded from modeling.

2.2. Label Definition

In order to identify patients with the prediction outcome MACE, transfer data, diagnoses and procedures were used. First, all patients with a coded diagnosis of angina pectoris (ICD-10 code I20) with a manual description associated to an acute coronary syndrome were included. In addition patients with one ICD-10 code of the diagnoses shown in Table 1 were included, if a matching procedure or transfer to a coronary catheterization laboratory was recorded for the hospital stay.

Table 1. ICD-10 Codes for outcome definition

ICD-10 code	Description	Type
I20	Angina Pectoris	Coronary heart disease
I21	Myocardial Infarction	
I24	Other acute ischemic heart disease	
I46	Cardiac arrest	
I63	Cerebral infarction	Cerebrovascular disease
I64	Stroke	
I71	Aortic aneurysm and dissection	Other cardiovascular disease
I74	Arterial embolism and thrombosis	

All patients with the following procedures were included in the cohort of MACE patients: coronary bypass surgeries, percutaneous coronary catheterization and systemic thrombolysis for myocardial infarction or stroke. In addition, magnetic resonance and computed imaging procedures for cerebrovascular diseases were used.

In case of a MACE, there is a high probability that the patient will receive cardiovascular associated procedures in one of three coronary catheterization facilities in the State of Styria. Thus, patients with transfers to the coronary catheterization laboratories in combination with a defined diagnosis or procedure were included as MACE patients.

Furthermore, the death data from Austria's Federal Statistical Office (Statistik Austria) enriched the definition of MACE patients: KAGes patients with death due to myocardial infarction (I21), cerebral infarction (I63) or stroke (I64) were included as MACE patients. Patients with the following ICD 10 codes were excluded, because these disease are mostly not caused by atherosclerosis [11]: intracerebral haemorrhage (I61, I62), aneurysm and dissection (I72).

After applying all inclusion and exclusion criteria, 86,365 admissions of 29,262 patients were included in the MACE group. The overlap of the identification of MACE patients using the four data modalities is shown in Fig. 1.

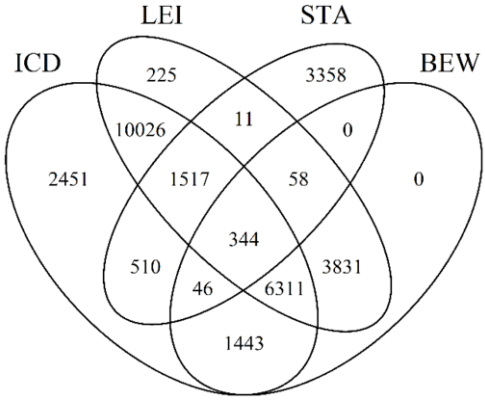


Figure 1. Venn diagram of the final selection of MACE patients using four modalities of identification: Diagnosis (ICD), procedures (LEI), death data (STA), and transfer data (BEW).

Finally, 181,150 admissions of 98,405 patients without MACE identification within 5 years after admission were extracted. For patients with multiple stays in the specified time period, all stays were treated separately. This resulted in 267,515 admissions for modeling.

2.3. Feature Extraction & Modelling

As a reference date for risk prediction, the discharge date of the admission included in the data set was set. Until this reference date, various features were created using already available data from the patient. Out of the data modalities described in section 2.1, 9,270 features were created.

Following, all constant features were deleted, before feature selection was applied. As a method of univariate feature selection, Chi-squared tests and ANOVA were applied, which removed 2,425 features from the data set. For further selection, a Generalized Linear Model with Lasso regularization and a Random Forest model were trained with the remaining features. Using only features with its coefficients greater than zero in or greater than the average feature importance in Random Forest, a feature size of 826 was achieved.

The final data of 267,515 admissions and 826 features was split into a 70% training and 30% test data set. Model training was performed only on the training data using a 5-fold cross-validation. Four different methods were used for training: Generalized Linear Model with Elastic Net regularization (GLM), Random Forest (RF), Gradient Boosting Machines (GBM) and Linear Discriminant Analysis (LDA).

To optimize the hyper-parameters of each method, the GridSearchCV function from scikit-learn [12] was used. GridSearchCV performs a 5-fold cross-validation for every hyper-parameter combination.

Finally, the performances of the models were evaluated on the separate test data set. The discriminative performance was measured using the AUROC value with 95% confidence intervals and ROC plots. The confidence intervals were calculated using the DeLong method [13]. Based on the closest topleft threshold, sensitivity, specificity, and precision were calculated. In order to analyse the calibration of the models, calibration plots were used showing the percentiles of risk probabilities in the x-axis and relative frequency of MACE on the y-axis. Additionally, an ethical vote was issued by the

Medical University of Graz for the development of these predictive models (30-146 ex 17/18). Furthermore, the TRIPOD statement [14] was used for model development, validation and reporting.

3. Results

In Table 2, the results of the evaluation on the test data are presented. Although the discriminative performance of all methods was excellent with AUROC above 0.86, the Gradient Boosting Machines achieved a slightly better performance than the others. Furthermore, a plot with the ROC curves of all models is shown in Fig. 2.

Table 2. 5-Year MACE prediction using four machine learning algorithms.

Method	AUROC [95%-CI]	Sensitivity	Specificity	Precision
GLM	0.867 (0.848 - 0.886)	0.798	0.767	0.619
RF	0.879 (0.860 - 0.897)	0.808	0.776	0.632
GBM	0.896 (0.879 - 0.913)	0.808	0.820	0.681
LDA	0.867 (0.848 - 0.886)	0.785	0.780	0.630

Fig. 3 presents the calibration plots of the models. Results of Gradient Boosting Machines showed the poorest performance among the models, with an underestimation of the risk for MACE at the lower percentiles and an overestimation of the risk for higher percentiles. Random forest and Generalized Linear Model showed a good calibration for all risk percentiles.

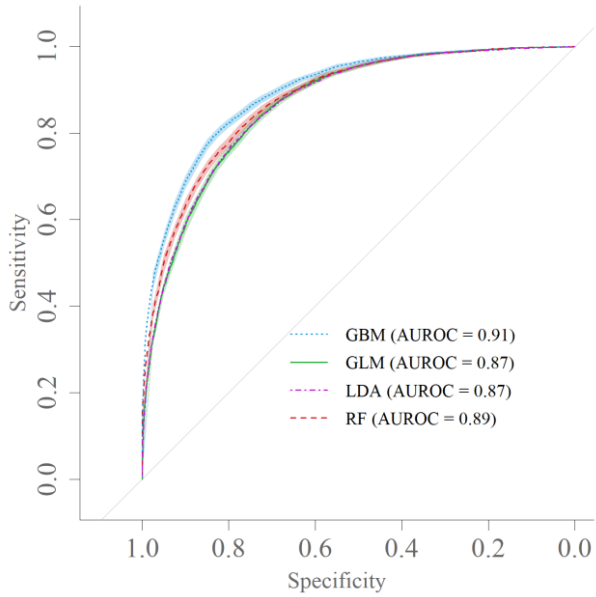


Figure 2. ROC curves for the trained models when applied on the test data (Generalized Linear Model (GLM), Random Forest (RF), Gradient Boosting Machines (GBM), Linear Discriminant Analysis (LDA))

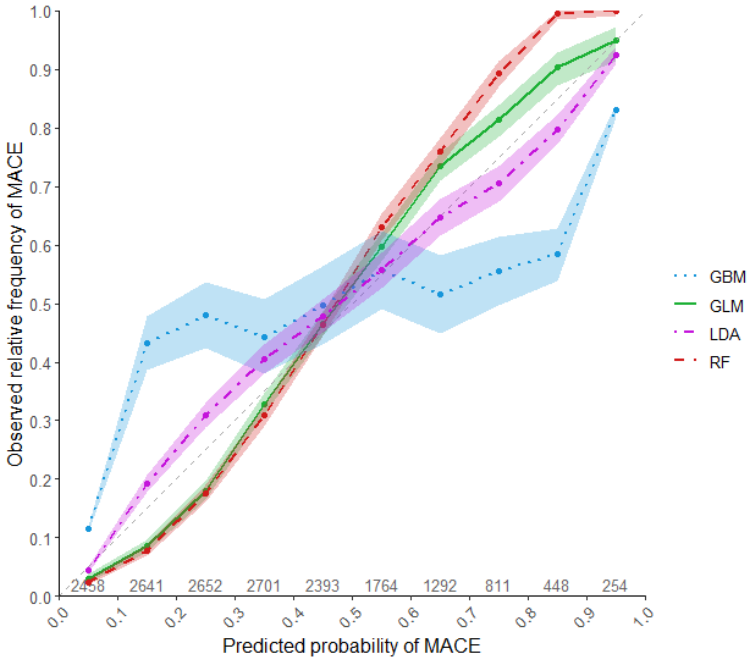


Figure 3. Calibration plots of the models when applied on the test data.

In addition, the 17 most important features of the Random Forest were investigated (see Fig. 4). Similar to QRISK3 or the other cardiovascular diseases related risk scores, age had a very large impact on risk prediction. An increased value for urea in serum or plasma, and potassium in serum or plasma were features with the highest influence, among the others. These features are associated with renal disorders, leading to an increased risk for cardiovascular diseases [5] [15] [16].

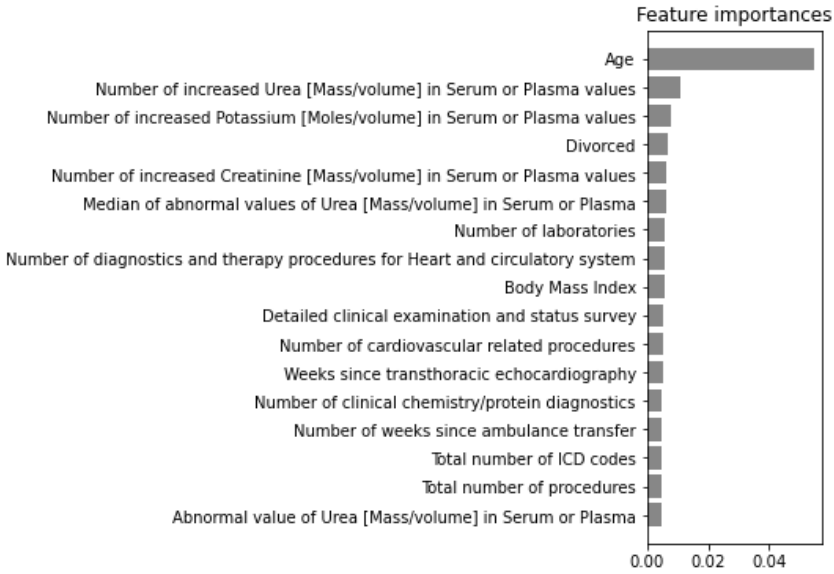


Figure 4. Gini Importance based Feature Importance of the Random Forest model.

4. Discussion

Our study presents the development of various machine learning models predicting the 5-year risk of a major cardiovascular event (MACE) of in-patients. EMRs of more than 127,667 patients were used to train the risk prediction models using different methods. Although GBM proved to be the best discriminator, the results of calibration curve were not satisfying. The three remaining models showed equally good results when considering both discrimination and calibration. This could be due to a comprehensive and effective feature selection.

With an AUROC of 0.879, the Random Forest model outperforms established cardiovascular risk scores such as QRISK3 or the ACC/AHA. However, the comparison has to be considered with care, as the models developed in our study predict the 5-year risk of MACE, while the published scores mostly predict a 10-year risk.

Further limitations need to be discussed. A main advantage of clinical scores over machine learning models is their generalizability. The use of few predictors for a risk prediction is independent of the availability of EMRs. As EMR systems and their records vary between hospitals, machine learning models developed in one hospital might not be applicable to other hospitals. External validation of machine learning-based models is crucial for a broad deployment.

KAGes covers almost 90% of the in-hospital stays in the region, and thus the EMR system provides a comprehensive population-based view of patient histories. However, the identification of patients with MACE can be biased if the event was recorded in another hospital outside the KAGes network. Using different modalities for labelling, including death records from the official Austrian statistical office, presents a major advantage of this study.

Although risk prediction by machine learning models might be highly accurate, more actions are needed in order to achieve a clinical benefit. In a first step, clinicians need to be informed about patient's individual risk factors using transparent visualization methods embedded in the hospital information systems. In a second step, it is essential to inform patients about preventive measures to modify the risk. Thus, future work should focus on how a predicted risk can be changed by staging preventive actions or treatment for patients with high risk of MACE.

The prediction models achieved an excellent performance in this study. In previous work, we demonstrated the high performance of machine learning model predicting delirium in a clinical setting [17]. Future work needs to focus on the prospective evaluation of the risk prediction models for MACE and determine the overall clinical benefit for patients.

References

[1] STATISTIK AUSTRIA, 'Todesursachen'. http://pic.statistik.at/web_de/statistiken/menschen_und_gesellschaft/gesundheit/todesursachen/index.html (accessed Jan. 28, 2021).

[2] L. Bergman, J. H. P. van der Meulen, M. Limburg, and J. D. F. Habbema, 'Costs of Medical Care After First-Ever Stroke in the Netherlands', *Stroke*, vol. 26, no. 10, pp. 1830–1836, Oct. 1995, doi: 10.1161/01.STR.26.10.1830.

[3] M. F. Piepoli *et al.*, '2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice', *Eur. Heart J.*, vol. 37, no. 29, pp. 2315–2381, Aug. 2016, doi: 10.1093/eurheartj/ehw106.

[4] R. B. D'Agostino *et al.*, 'General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study', *Circulation*, vol. 117, no. 6, pp. 743–753, Feb. 2008, doi: 10.1161/CIRCULATIONAHA.107.699579.
[5] J. Hippisley-Cox, C. Coupland, and P. Brindle, 'Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study', *BMJ*, p. j2099, May 2017, doi: 10.1136/bmj.j2099.
[6] R. Conroy, 'Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project', *Eur. Heart J.*, vol. 24, no. 11, pp. 987–1003, Jun. 2003, doi: 10.1016/S0195-668X(03)00114-3.
[7] D. C. Goff *et al.*, '2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines', *Circulation*, vol. 129, no. 25 suppl 2, pp. S49–S73, Jun. 2014, doi: 10.1161/01.cir.0000437741.48606.98.
[8] L. Breiman, 'Statistical Modeling: The Two Cultures', Stat. Sci., vol. 16, pp. 199–215, 2001.
[9] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, 'Can machine-learning improve cardiovascular risk prediction using routine clinical data?', *PLOS ONE*, vol. 12, no. 4, p. e0174944, Apr. 2017, doi: 10.1371/journal.pone.0174944.
[10] C. Krittanawong *et al.*, 'Machine learning prediction in cardiovascular diseases: a meta-analysis', *Sci. Rep.*, vol. 10, no. 1, p. 16057, Dec. 2020, doi: 10.1038/s41598-020-72685-1.
[11] K. Ritz, N. P. Denswil, O. C. G. Stam, J. J. van Lieshout, and M. J. A. P. Daemen, 'Cause and Mechanisms of Intracranial Atherosclerosis', *Circulation*, vol. 130, no. 16, pp. 1407–1414, Oct. 2014, doi: 10.1161/CIRCULATIONAHA.114.011147.
[12] F. Pedregosa *et al.*, 'Scikit-learn: Machine Learning in Python', *J. Mach. Learn. Res.*, vol. 12, p. 2825−2830, Oct. 2011.
[13] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, 'Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach', *Biometrics*, vol. 44, no. 3, p. 837, Sep. 1988, doi: 10.2307/2531595.
[14] G. S. Collins, J. B. Reitsma, D. G. Altman, and K. Moons, 'Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement', *BMC Med.*, vol. 13, no. 1, p. 1, 2015, doi: 10.1186/s12916-014-0241-z.
[15] D. Pandya, 'Assessment and Correlation of Urea and Creatinine Levels in Saliva and Serum of Patients with Chronic Kidney Disease, Diabetes and Hypertension– A Research Study', *J. Clin. Diagn. Res.*, 2016, doi: 10.7860/JCDR/2016/20294.8651.
[16] J. R. Montford and S. Linas, 'How Dangerous Is Hyperkalemia?', *J. Am. Soc. Nephrol.*, vol. 28, no. 11, pp. 3155–3165, Nov. 2017, doi: 10.1681/ASN.2016121344.
[17] S. Jauk *et al.*, 'Risk prediction of delirium in hospitalized patients using machine learning: An implementation and prospective evaluation study', *J. Am. Med. Inform. Assoc.*, vol. 27, no. 9, pp. 1383–1392, Sep. 2020, doi: 10.1093/jamia/ocaa113.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210101

Enabling FAIR Discovery of Rare Disease Digital Resources

Rajaram KALIYAPERUMAL[a,1], Núria QUERALT ROSINACH[a], Kees BURGER[a], Luiz Olavo BONINO DA SILVA SANTOS[a,b], Marc HANAUER[c] and Marco ROOS[a]
[a] *Leiden University Medical Center, Leiden, Netherlands*
[b] *University of Twente, Enschede, Netherlands*
[c] *Inserm, Paris, France*

Abstract Background: Integration of heterogenous resources is key for Rare Disease research. Within the EJP RD, common Application Programming Interface specifications are proposed for discovery of resources and data records. This is not sufficient for automated processing between RD resources and meeting the FAIR principles. Objective: To design a solution to improve FAIR for machines for the EJP RD API specification. Methods: A FAIR Data Point is used to expose machine-actionable metadata of digital resources and it is configured to store its content to a semantic database to be FAIR at the source. Results: A solution was designed based on *grlc* server as middleware to implement the EJP RD API specification on top of the FDP. Conclusion: *grlc* reduces potential API implementation overhead faced by maintainers who use FAIR at the source.

Keywords. rare disease, metadata, data management, information storage and retrieval, FAIR, semantic web

1. Introduction

Integrating sensitive, sparse and heterogenous datasets is key for accelerating research and patient care in the Rare Disease (RD) field. RD resources such as biobanks, patient registries and omics datasets are often dispersed and hard to find in order to integrate. The European Joint Programme on Rare Diseases (EJP RD) [1] brings together researchers across the RD field to address data integration problems by implementing the FAIR (Findable, Accessible, Interoperable and Reusable) data principles [2]. One of the goals of the EJP RD project is to build a so-called 'virtual platform' (VP) via which centralized and federated RD resources can be found and interrogated. Importantly, the underlying infrastructure needs to enable automated processing between the resources. A common API specification has been proposed that programmers can use in their code to discover resources and data records [3]. However, documented APIs are not sufficient for a dynamic, AI ready VP that is also interoperable outside of the EJP RD project: resources and their content need to be self-describing with global, machine readable models, in line with the FAIR principles. This can be achieved by using global ontologies defined in the Resource Description Framework (RDF). For many system designs, including the EJP RD VP, it will be a requirement that common APIs can work

[1] Corresponding Author: Rajaram Kaliyaperumal, Leiden University Medical Center, Leiden, Netherlands, E-Mail: r.kaliyaperumal@lumc.nl

seamlessly with this technology. In this paper, we propose a design to implement common APIs as a complement to resources that apply RDF to implement FAIR principles. The Orphanet data catalogue was used as an example.

2. Methods

The FAIR Data Point (FDP) is general-purpose software that helps users expose metadata of their resources in a FAIR manner, particularly to meet the FAIR requirement of machine readability [4]. An FDP therefore relies on Semantic Web technologies by design, including the Resource Description Framework (RDF), triple store (an RDF database) and the SPARQL Protocol and RDF Query Language (SPARQL) a W3C recommended standard to manipulate and query RDF [5]. FAIR for machines in the FDP is achieved in two ways. First, the FDP applies widely used RDF-based vocabularies such as the Data Catalog Vocabulary and Dublin Core to describe the resource in machine readable terms. Secondly, the FDP uses triple stores to store and serve the metadata. The REST API of the FDP also supports the Linked Data Platform (a W3C recommended specification to manipulate RDF documents via HTTP) GET call specification [6].
Within the EJP RD project, a set of common API specifications is also being defined for programmers, including an API for discovery of resources and data records. Version 0.2 of this API aims to address discovery of resources at a metadata level. The proposed APIs are based on the OpenAPI specification to implement REST API calls that responds with JSON documents. A requirement for the EJP RD Virtual Platform is that the APIs and RDF-based FAIR data points can work together. This is achieved by *grlc*, a server application that converts SPARQL queries into API calls and lets users transform the API call's response to JSON documents [7].

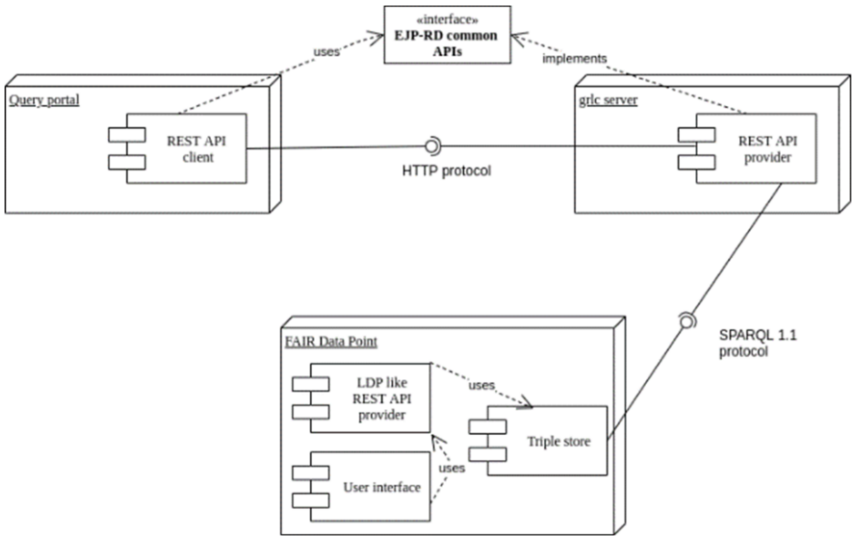


Figure 1. Design to implement EJP-RD API specification on top of a FAIR Data Point.

3. Results

We implemented an FDP that exposes metadata of biobank and patient registries captured by the Orphanet catalog[2] and configured it to store its RDF documents in a triple store. We propose a design where the EJP RD common APIs specification is implemented on top of the FDP (Figure 1). The *grlc* server is used as middleware that converts EJP RD specific common API calls into SPARQL queries and the response of the SPARQL queries into JSON documents. We created a reference implementation for the "Search resource" API call[3]. This API call searches biobanks or patient registries for a given disease code from the Orphanet catalog's FDP. Figure 2 shows the request URL and response for the disease code 'Orphanet_586' for our implementation.

Figure 2. *grlc* REST API's request URL and response to find resources for disease code `Orphanet_586'

4. Discussion

In our work we showed how generic FAIR software solutions such as the FAIR Data Point and triple stores can be used to achieve specific findability and interoperability goals. We have presented a straightforward solution to accommodate common APIs for programmers without compromising FAIR principles. For implementors who maintain data sources FAIR at the source, implementing a common API might be a substantial overhead. With our solution to use middleware components such as *grlc* we show how this overhead can be reduced. A drawback in our reference implementation is that we only converted successful API call responses (http 200). Converting other API call responses such as 'not found' (http 404) and 'unauthorized' (http 401) might require further extension to the *grlc* server.

References

[1] "EJP RD." [Online]. Available: https://www.ejprarediseases.org. [Accessed: 10-Feb-2021].

[2] M. D. Wilkinson *et al.*, "Comment: The FAIR Guiding Principles for scientific data management and stewardship," *Sci. Data*, vol. 3, p. 160018, Mar. 2016, doi: 10.1038/sdata.2016.18.

[3] "EJP RD Query Builder API specs:" [Online]. Available: https://github.com/ejp-rd-vp/query_builder_api. [Accessed: 10-Feb-2021].

[4] L. O. Bonino da Silva Santos *et al.*, "FAIR Data Points Supporting Big Data Interoperability," in *Enterprise Interoperability in the Digitized and Networked Factory of the Future*, M. Zelm, G. Doumeingts, and J. P. Mendonça, Eds. ISTE Press, 2016.

[5] "SPARQL 1.1 Protocol." [Online]. Available: https://www.w3.org/TR/sparql11-overview/. [Accessed: 10-Feb-2021].

[6] "Linked Data Platform 1.0." [Online]. Available: https://www.w3.org/TR/ldp/. [Accessed: 10-Feb-2021].

[7] A. Meroño-Peñuela and R. Hoekstra, "Grlc makes github taste like linked data APIs," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9989 LNCS, pp. 342–353, doi: 10.1007/978-3-319-47602-5_48.

[2] Orphanet: catalog https://www.orpha.net/consor/cgi-bin/index.php

[3] Reference implementation: http://grlc.io/api-git/rajaram5/SWAT4HCLS-hackathon-2021-ejprd-grlc

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210102

Visual Analytics in Delirium Management

Werner O. HACKL[a,1], Michael NETZER[a], Renate NANTSCHEV[a], Michael SCHALLER[a] and Elske AMMENWERTH[a]
[a] *UMIT - Private University for Health Sciences, Medical Informatics and Technology, Institute of Medical Informatics, Hall in Tirol, Austria*

Abstract. *Background*: Delirium is a patient safety issue that often occurs within the population of elderly people. As delirium may be characterized by fluctuating progress, the aim of this work is to find methods to visualize the occurrence of delirium over time in different patient stays in gerontopsychatric settings. *Methods*: We analyzed current data mining visualization techniques for clinical research using a delirium data set collected in a gerontopsychatric setting. *Results*: We identified heatmaps and dendrograms resulting from hierarchical clustering as a suitable visualization method. *Conclusion:* Heat maps with hierarchical clustering are a suitable data mining tool or visualization technique to study delirium cases in the time course of patient stays.

Keywords. Delirium management, patient safety, data visualization

1. Introduction

Gerontopsychatric wards in hospitals represent unique challenges in patient safety requirements. Delirium is one of these specific patient safety issues that often occur within the population of older people. It is associated with a massive burden to the affected people and the health care system [1,2]. Reliable risk prediction plus rapid and precise detection of cases are crucial in delirium management. Several screening instruments for detecting delirium are proposed, such as the Delirium Observation Screening Scale (DOSS) or the Confusion Assessment Method (CAM) [3]. However, their predictive ability might be limited in the psychiatric setting due to frequent side diagnoses such as dementia or depression. Therefore, in a recent study, we assessed these instruments in the gerontopsychatric setting and used different machine learning models to analyze their predictive ability. We found that CAM clearly outperformed DOSS in the specific setting and that some of our machine learning approaches were even a notch better than CAM [4]. A very important lesson learned of our study was that a horizontal view should be taken over time in addition to the pure detection of cases, as delirium can be a fluctuating condition.

Thus, this work aimed to find a suitable method to get an overview on and visualize the onset and course of delirium in the observed patient-stays.

[1] Corresponding Author: Dr. Werner Hackl, UMIT TIROL, Hall in Tirol, Austria E-Mail: werner.hackl@umit-tirol.at

2. Methods

We performed a literature analysis to identify current visualization techniques in clinical research. We found a very recent systematic overview of data mining visualization techniques including tables, charts, trees, maps, scatter plots, diagrams and graphs [5]. We evaluated these methods according to our requirements (i.e., visualization of time information). We used the statistical software *R* [7] and the *gplots* package for visualization.

As dataset we used the retrospectively collected data from two gerontopsychatric wards as described in [4] (Ethical approval granted by the Ethical Committee of the Medical University of Innsbruck - EK Nr: 1032/2019). In total, 732 patient stays from 593 distinct patients were available. During these stays DOSS (n=2.423), CAM (n= 2.368) and ICD-10 diagnoses of dementia (n=2.213) had been collected on a regular, weekly basis [4]. The data were pre-processed to match the results of the CAM/DOSS risk assessment surveys (yes/no/maybe) with the actual conditions according to the ICD-10 diagnosis as gold standard in the time sequence of each stay.

3. Results and Discussion

Based on the overview, we identified heatmaps as a suitable tool for visualization of delirium occurrence over time. In principle, heatmaps are tables where the color of each cell represents the magnitude of a certain value. In particular, rows represent patient-stays and columns represent the time steps. The colors of each cell represent the occurrence of delirium (yes/no) according to the gold standard or according to the result of CAM (yes/maybe/no) and DOSS (yes/no or zero to 13 points when analyzed in more granularity). The rows of the heatmaps representing patients are additionally sorted using a hierarchical clustering approach. The result of the hierarchical clustering is visualized by a dendrogram, which is a binary tree where each branch represents a cluster [6].
We could clearly identify different groups and subgroups of stays: i) stays with deliruim at admission with subgroups i.1) improvement or i.2) constant condition; ii) stays were delirium occurred during the stay, with same subgroups; iii) stays without delirium cases.
Next steps will be to analyze these groups and subgroups in more detail using process mining techniques.

Acknowledgement
This research was funded by the Austrian Science Fund (FWF) [PATIS: P 29076].

References

[1] S.K. Inouye, R.G.J. Westendorp, J.S. Saczynski, Delirium in elderly people. *Lancet*. **383**(2014), 911–22.
[2] National Clinical Guideline Centre. DELIRIUM: diagnosis, prevention and management. Clinical Guideline **103** (2010). https://www.ncbi.nlm.nih.gov/books/NBK65558/pdf/Bookshelf_NBK65558.pdf
[3] W. Hasemann, R. Kressig, D. Ermini-Fünfschilling, M. Pretto, R. Spirig, *S*creening, Assessment und Diagnostik von Delirien. *Pflege.***20**(4) (2007),191–204.
[4] M. Netzer, W.O. Hackl, M. Schaller, L. Alber, Evaluating Performance and Interpretability of Machine Learning Methods for Predicting Delirium in Gerontopsychiatric Patients. *IOS Press - dHealth* (2020)
[5] Q. Li , *Overview of Data Visualization,* in: Embodying Data. Springer, 2020 pp. 17–47
[6] M. Schonlau, The clustergram: A graph for visualizing hierarchical and nonhierarchical cluster analyses, *The Stata Journal* **2**(4) (2002), 391-402
[7] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/, 2021

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210103

Co-Designing a Primary Care Breathlessness Decision Support System: General Practitioners Requirements Analysis, Workflow Assessment and Prototype Development

Anthony Paulo SUNJAYA[a,b,1], Allison MARTIN[a,b] and Christine JENKINS[a,b]
[a] *Respiratory Division, The George Institute for Global Health, Sydney, Australia*
[b] *Faculty of Medicine, University of New South Wales, Sydney, Australia*

Abstract. Clinical decision support systems (CDSS) have been shown in a variety of diseases to lead to improvements in care. The aim of this study is to design a CDSS to assist GPs to assess and manage breathlessness, a highly prevalent symptom in practice. A focus group is conducted to explore the needs of general practitioners (GPs), assess current workflow to identify points for intervention and develop early prototypes for testing. Five GPs took part in the focus group elucidating 248 relevant data points which were then qualitatively analyzed using the Technology Acceptance Model as the theoretical framework. In general, there was a positive attitude towards the use of CDSS for breathlessness with various proposed features from the participants. Twelve high level workflow steps were identified with 5 as key points for intervention. Several proposed features such as reporting likelihood of causes of breathlessness in a patient, link with evidence-based recommendations, integration with clinical notes and patient education materials were translated into a prototype. Mixed-method studies are planned to assess its usability to inform subsequent iterations of the CDSS development.

Keywords. clinical decision support system, primary care, breathlessness

Introduction

Breathlessness defined as "a subjective experience of breathing discomfort that consists of qualitatively distinct sensations that vary in intensity"[1] is found in about 11% of Australians.[2] Changes to the biosphere and environment due to climate change[3], sequelae of respiratory infections like coronavirus disease 2019 (COVID-19), all of which negatively affects cardiopulmonary health will likely further increase the burden of breathlessness in Australia and globally. With 87.8% of Australians visiting their primary care physician at least once a year[4], most breathless patients are likely to present to primary care even when it is not their main or only presenting symptom.

The multidimensional aspect and myriad possible causes for chronic breathlessness, ranging from respiratory to cardiac to metabolic diseases, mental health and

[1] Corresponding Author: Anthony Paulo Sunjaya, The George Institute for Global Health, Sydney, Australia, E-mail: asunjaya@georgeinstitute.org.au

deconditioning, create a major diagnostic challenge for this very common problem. In the primary care setting, a study of patients referred for breathlessness reported that less than 30% had a fully concordant referral diagnosis with the final diagnosis.[1] An analysis of the online British Lung Foundation Breath Test also reported that even after seeking medical advice, 58% stated that their breathlessness had not improved.[5] The researchers hypothesised that this sub-optimal care for breathlessness may be due to the underutilization of effective evidence-based interventions for breathlessness.

Clinical decision support systems (CDSS) have been shown in a variety of diseases to lead to improvements in care. Even so, the systematic review also concluded that among the trials included, usage and compliance was generally low. Furthermore, compliance with CDSS recommendations was also reported to be low.[6]

Hence, while designing a user friendly and validated CDSS is of great importance, its implementation into practice relies on how clinician end users and practices respond to the integration of a CDSS with existing workflows.[6] Understanding the needs of clinicians who will be the end-users and having them participate from the start in a co-design process will improve uptake of digital health interventions.

The aim of this phase in the development of the breathlessness algorithm and CDSS is to explore the needs among general practitioners who care for people with breathlessness. Specific aims were to understand the burden of breathlessness in their practices, current assessment pathways and points for referrals, identify the most challenging point in care, their current use of CDSS, response to using CDSS for breathlessness and development of an early CDSS prototype for further testing.

1. Methods

1.1. Participants and Data Collection

A focus group (FGD) was conducted with five general practitioners (3 males and 2 females) from varying local health districts. Exploratory open-ended questions were used to enquire about the current state and challenges in breathlessness assessment and management. Questions were specifically developed to assess what could help GPs in this field, current use of other CDSS, facilitators and barriers to use, and their openness to using a breathlessness CDSS. The discussion lasted about 90 minutes in length and was conducted in February 2018. Clinicians were invited to participate through the local primary care health network and from doctors that have previously referred patients for breathlessness to the investigator's clinic. Average monthly prevalence of patients identifying breathlessness as a problem was between 5% to 12%. Workflow assessment was conducted using a business processes workflow diagram and prototypes developed using Balsamiq. Ethical approval was granted by the University of Sydney Human Research Ethics Committee. Informed consent forms were sent to participants prior to the FGD.

1.2. Data Analysis

The audio recording of the FGD was transcribed for analysis. Coding was performed using NVivo 12 using an approach as described by Terry et al.[7] With a focus on the intrinsic behavior intention of end users (doctors) to a CDSS, the Technology Acceptance Model (TAM) was used as framework for analysis and not others which has

substantial focus on external influences.[8] Direct quotations from participants were reported between single quotes.

2. Results

A total of 248 relevant data points was collected from the FGD. Data were analyzed into 2 main themes – perceived usefulness and perceived ease of use. We examined these themes from the view of current CDSS usage, related this to proposed features to improve the breathlessness CDSS and in brackets link this to the prototype translation.

2.1. Perceived Usefulness of a CDSS for Breathlessness

To elucidate the usefulness of a CDSS for breathlessness, we aimed to understand the current care pathways and challenges clinicians faced in practice when managing patients with breathlessness. The challenges described by GPs are summarized in Table 1.

Table 1. Current breathlessness care challenges reported by clinicians

Topic	Key points and quotes	Summary
Current practice challenges and tests	'Frustrating' 'moving diagnosis. 'Multifactorial' causes, 'multimorbidity' and 'uncertainty' found especially in the elderly. Common tests in practice are electrocardiography, complete blood count, spirometry and pulse oximetry.	Diagnosis is the most challenging component in the care spectrum and 'most important'. Among common tests spirometry is one which they can do 'okay' but am 'not an expert'.
Points for referral	'Concerned patients', 'personal worries', 'severity' and 'length of time with symptoms'. 'The reward for sending to a specialist is low'. 'Compartmentalized medicine'.	Gut feelings described as 'internal anxiety' and 'litmus test' as driver of deciding on a diagnosis, diagnostic tests done and referral.
Manage-ment	Issue with treatment 'adherence' in some patients including asthmatics. Patients often come with requests for testing after visiting 'naturopaths' and other alternative practitioners.	Once diagnosis is made it becomes possible to 'formulate a management plan'. Time needed to compare the utility of further tests.

Considering diagnosis to be the main challenge, most clinicians proposed a diagnosis focused CDSS, hence our focus on the diagnosis stages in the prototype (Figure 2A-C). The clinicians mention the need of an 'algorithm' for breathlessness. The algorithm should help with 'ruling things in but also ruling things out', importantly to help cast away the organic issues that if missed 'could kill the patient' and lead to determination of inorganic causes like 'deconditioning' and 'anxiety'. GPs could see the utility of a CDSS that can provide the 'percentages e.g. 10% deconditioning, 10% cardiac disease etc' and 'chance' for the causal disease of breathlessness in a patient (Figure 2C top). A CDSS would need to be based on the most recent 'evidence' with quick access to 'guidelines'. (Figure 2D left) Whereas another clinician proposed a CDSS that can recommend the most 'important decisions' early on. This was translated by presenting differentials from page one of the CDSS (Figure 2A). The clinician imagined that knowing these recommendations, he would be more inclined to prioritize them over other tests especially with the time limitation in practice where consults typically last for only '15 minutes'. The CDSS should also guide decisions based on the patient's 'previous results and current results' before recommending a 'referral' e.g. to a cardiologist.

Even so, one clinician pointed out that his problem with 'pathway algorithms is that sometimes it stops you thinking', a solution proposed for this by the group is for a CDSS that is 'simple but takes you down a diagnostic path as opposed to trying to establish a diagnosis'. This is with a note from one clinician that this only stands true 'till you're absolutely certain what it is'. Additionally, when this happens, risk scores and management oriented CDSS' can help to educate patients (Figure 2B – links to validated questionnaires).

One clinician described the usefulness of the cardiovascular (CVD) disease risk score in 'scaring patients' by seeing the impact of their activities on their health. Another possible use is if a CDSS can automatically provide the above 'side-by-side comparison' of the utility of diagnostic tests (Figure 2C). It can also be useful to provide 'management plans' with reference to red flags that patients need to be reminded to 'come back sooner rather than later' if particular symptoms appear. Another use for the CDSS in patients with inorganic causes is to provide guides on 'how fit patients can get' and 'improving their conditioning'. This is currently all done manually.

2.2. Perceived Ease of Use

Individuals' past actions have been shown to be a good reflection of their future behaviour.[9] Considering this, perceived ease of use is elucidated based on how easy or challenging it is to use current systems. Furthermore, as with usefulness, we will relate this with clinicians' proposed features to improve ease of use.

The clinicians reported current use of a variety of CDSS tools namely the 'CHAD score', 'AUSDRISK', 'MMSE score', 'DASS-21', 'Cervical screening flowchart', 'Health Tracker', 'FRAX score' and 'CVD risk calculators'. 'Embedded' or 'integration' with current electronic medical records system (EMR) is a feature reported to improve ease of use. One clinician described how she calculates CVD risk on the 'background' for all patients of appropriate age. Another describes CDSS automatically incorporating available data such as blood pressure and smoking to the score. Even so, integration with family history remains a caveat.

Others reported the use of more basic CDSS being 'a set of questions you answer and get the score'. On the other hand, another clinician reported some algorithms being more 'complicated than what he can handle', all clinicians did concur that the CDSS must be 'simple' and 'short' to fit with practice timing constraints. CDSS use is expected to 'save energy' through communicating ideas simply and providing only 'the most important points' in an algorithm.

The clinicians proposed something 'smaller than an algorithm' presented on screen that 'sums up the recommendation' based on the history and previous results. However, it should also provide the option of seeing the 'more complex' 'main algorithm' either by 'right clicking for further details' or through other buttons where they can bounce between the algorithms (Figure 2D). This also addresses the need of some of the clinicians to know 'how did they get there (the recommendation)' before deciding whether they are to agree or disagree. Especially one of them who preferred to have 'the computer pop the algorithm in from of me' and make their own decisions.

In addition, information from the EMR, integration with results from equipment (e.g. spirometry) was proposed. Furthermore, all clinicians concur that ease of use would be improved if the result of the CDSS could be directly 'incorporated into clinical notes' and even replace 'some of the history'. (Translated into ability to copy summary of CDSS recommendation into notes) Another idea was to incorporate patient's self-reported

history before 'actually going to the doctor' as part of the system. Lastly, one clinician suggested the analogy that the CDSS should be as easy to use as 'Google' where asking a specific question would lead you the most relevant guidelines and answers.

2.3. Attitude and Acceptance of CDSS for Breathlessness

In general, there was a positive attitude towards the use of CDSS for breathlessness. This is especially true when a 'diagnostic quandary' is reached or in providing help in deciding to send patients for 'something simple or should I escalate'. The clinicians find CDSS acts as a 'question channel', help 'validate my decision' and 'builds confidence I am in the right ballpark' for diagnoses that is at times made based on gut feelings.

Another aspect is the current 'swarm' of decision support tools, one clinician noted 'there's so many tools, there's so many things' that even when CDSS exist they are often forgotten and remain unutilized. The group suggest the use of appropriate reminders to remind clinicians that the CDSS exists.

An important aspect brought forward was whether the interventions including CDSS for breathlessness are positively affecting patients' 'quality of life' and their 'life expectancy'. Only when this is expected would one clinician said they "be open to expanding the number of patients with breathlessness they see", translating to screening more patients for breathlessness and routine CDSS use for this diagnosis purpose.

2.4. Clinical Workflow Assessment

Using a business-processes-workflow diagram, twelve high-level steps were identified in the process of assessment and management of breathlessness by GPs based on the results of the FGD question on how breathlessness patients are currently managed in practice and reflection from the clinician members of the team. Several points for intervention identified were at history taking, reviewing patient medical records, formulating a differential and personalized management plans, patient education and reporting into clinical notes as shown in Figure 1. Currently most of the steps in the workflow are undertaken manually by clinicians and involves their own clinical judgement based on varying levels of education and personal experience.

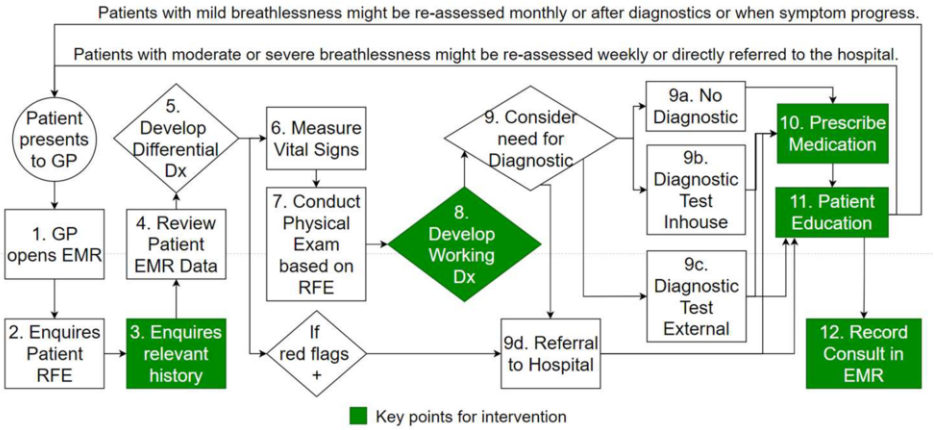


Figure 1. High level workflow overview and key points for CDSS intervention (green)
Dx – Diagnosis, EMR – Electronic Medical Record, RFE – Reason for Encounter

2.5. CDSS Prototype based on Clinician Requirement and Proposed Features

Based on the focus group and recent review on developing successful CDSS interfaces[10], we proposed the development of an expert system with integrated information display, order facilitators and documentation support which was then prototyped. Several screens from the prototype can be viewed in Figure 2.

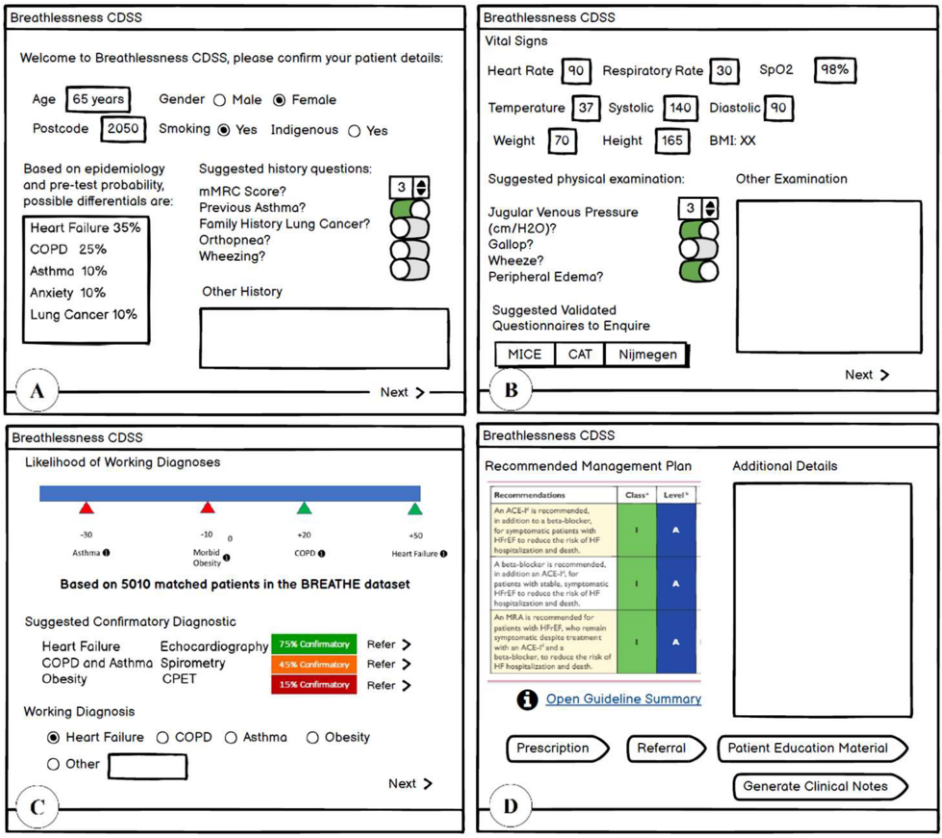


Figure 2. Breathlessness CDSS screen prototypes based on features proposed. *A: High yield history suggestions; B: Prioritising key physical exams; C: Diagnosis likelihoods, diagnostic utility of tests; D: Evidence based management suggestions*

3. Discussion

This early co-design with general practitioners who will be end users of the future breathlessness CDSS has provided a reflection of the current reality of breathlessness care, workflow and CDSS use in primary practice. It has also provided inputs for design of the CDSS prototype which will be tested in future mixed methods studies.

Some of the findings elucidated were similar to a recent focus group discussion conducted by Trinkley and colleagues among general practitioners to explore physicians preferred features of a CDSS. Features proposed included being clinically relevant and

providing customisable support, presentation of pertinent clinical information e.g. labs and vitals, and improving their workflow as proposed by our focus group participants.[11] Reminders to use the CDSS were another point proposed by our participants and previous studies have also reported interruptive alerts may be more effective if well-designed.[11]

The barriers to current CDSS use do not seem to be as extensive as reported from another focus group in the Netherlands among 24 primary care physicians which reported insufficient knowledge of the CDSS, irrelevant and high intensity alerts, lack of flexibility and CDSS update, effects on patient communication and additional time as well as effort to utilise it.[12] This will be further explored in our subsequent focus groups.

The utility of CDSS for educating patients on CVD risk described by clinicians is one that is also supported by existing international literature. In another study on the use of CDSS in primary care for cardiovascular disease, 18 months after CDSS implementation 98% of physicians reported the CDSS improved cardiovascular risk factor control in their patients and helped initiate cardiovascular risk discussion.[13] These results demonstrate that when doctors have tackled the learning curve post CDSS implementation, these tools can provide benefits to the patient-doctor consultation.

With 96% of general practitioners[14] in Australia using computer for clinical purposes, the use of CDSS for managing breathlessness has potential to bring great improvements in care. However, as a recent systematic review on identification and assessment of breathlessness in clinical practice which included 97 studies showed a lack of study in primary care.[15]

Previous studies on CDSSs' effectiveness related to breathlessness have been on disease oriented systems once a diagnosis has been reached such as in patients with asthma, COPD[16] and heart failure[17] which reported improvements in outcomes and even reduce costs. However, as the focus group has shown, the clinicians want to have a precursor to such systems, one that helps in guiding them to an appropriate diagnosis before moving to more specific disease oriented CDSS.

Furthermore, the group recognised the potential benefit of a CDSS to support interpretation of simple tests – such as spirometry, where typically they report greatest difficulty. This is in line with a study which reported only 57.8% of COPD patients diagnosed in primary care having post-bronchodilator spirometry results consistent with COPD or asthma.[18] These suggest lack of confidence in interpretation although it might also relate to the recording quality, both of which can be supported by CDSS.

There were mixed views among participants on when CDSS recommendations should be provided. Previous studies by Kostopoulou et al. has shown that physicians have an emerging judgement early in the consult and that 56% of physicians remained committed to initial diagnostic leaning even after receiving conflicting cues.[19] Therefore an early CDSS as proposed by a participant might be an option to counter this diagnostic inertia and make clinicians more open to new differentials.

While this study is limited to the contexts of the participating clinicians, in general it suggests a positive attitude towards use of a breathlessness CDSS system by general practitioners. Five key points for intervention were identified and with clinician proposed features were translated into a prototype. This first iteration of the co-design process would support planning of future mixed-method studies to assess its usability to inform subsequent iterations of the CDSS development. Further studies are also planned to assess the accuracy of the expert system.

Acknowledgements

This study was supported by a grant from Telstra Health and SPHERE Maridulu Budyari Gumal. APS is supported by a Scientia PhD scholarship from UNSW Sydney.

References

[1] Y.C. Huang, O.R. Ferry, S.C. McKenzie, R.V. Bowman, M. Hamilton, P.J. Masel, K.M. Fong, and I.A. Yang, Diagnosis of the cause of chronic dyspnoea in primary and tertiary care: characterizing diagnostic confidence, *J Thorac Dis* **10** (2018), 3745-3756.

[2] J.A. Bowden, T.H. To, A.P. Abernethy, and D.C. Currow, Predictors of chronic breathlessness: a large population study, *BMC Public Health* **11** (2011), 33.

[3] G. D'Amato, L. Cecchi, M. D'Amato, and I. Annesi-Maesano, Climate change and respiratory diseases, *European Respiratory Review* **23** (2014), 161-169.

[4] Department of Health, Annual Medicare Statistics, in, 2020.

[5] A.F. Elbehairy, J.K. Quint, J. Rogers, M. Laffan, M.I. Polkey, and N.S. Hopkinson, Patterns of breathlessness and associated consulting behaviour: results of an online survey, *Thorax* **74** (2019), 814-817.

[6] P. Matui, J.C. Wyatt, H. Pinnock, A. Sheikh, and S. McLean, Computer decision support systems for asthma: a systematic review, *NPJ Prim Care Respir Med* **24** (2014), 14005.

[7] G. Terry, N. Hayfield, V. Clarke, and V. Braun, Thematic Analysis, in: *The SAGE Handbook of Qualitative Research in Psychology T,* C. Willig and W.S. Rogers, eds., 2017.

[8] V. Fontil, K. Radcliffe, H.C. Lyson, N. Ratanawongsa, C. Lyles, D. Tuot, K. Yuen, and U. Sarkar, Testing and improving the acceptability of a web-based platform for collective intelligence to improve diagnostic accuracy in primary care clinics, *JAMIA Open* **2** (2019), 40-48.

[9] D. Albarracín and R.S. Wyer, Jr., The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions, *Journal of personality and social psychology* **79** (2000), 5-22.

[10] K. Miller, D. Mosby, M. Capan, R. Kowalski, R. Ratwani, Y. Noaiseh, R. Kraft, S. Schwartz, W.S. Weintraub, and R. Arnold, Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support, *J Am Med Inform Assoc* **25** (2018), 585-592.

[11] K.E. Trinkley, W.W. Blakeslee, D.D. Matlock, D.P. Kao, A.G. Van Matre, R. Harrison, C.L. Larson, N. Kostman, J.A. Nelson, C.-T. Lin, and D.C. Malone, Clinician preferences for computerised clinical decision support for medications in primary care: a focus group study, *BMJ Health & Care Informatics* **26** (2019), 0.4-0.

[12] M. Lugtenberg, J.-W. Weenink, T. Van Der Weijden, G.P. Westert, and R.B. Kool, Implementation of multiple-domain covering computerized decision support systems in primary care: a focus group study on perceived barriers, *BMC Medical Informatics and Decision Making* **15** (2015).

[13] J.M. Sperl-Hillen, A.L. Crain, K.L. Margolis, H.L. Ekstrom, D. Appana, G. Amundson, R. Sharma, J.R. Desai, and P.J. O'Connor, Clinical decision support directed to primary care patients and providers reduces cardiovascular risk: a randomized trial, *Journal of the American Medical Informatics Association* **25** (2018), 1137-1146.

[14] Australian Institute of Health and Welfare *Australia's health 2018*, Canberra, 2018.

[15] H.L. Elliott-Button, M.J. Johnson, U. Nwulu, and J. Clark, Identification and Assessment of Breathlessness in Clinical Practice: A Systematic Review and Narrative Synthesis, *J Pain Symptom Manage* (2019).

[16] K. Lisspers, G. Johansson, C. Jansson, K. Larsson, G. Stratelis, M. Hedegaard, and B. Ställberg, Improvement in COPD management by access to asthma/COPD clinics in primary care: Data from the observational PATHOS study, *Respiratory Medicine* **108** (2014), 1345-1354.

[17] A.I. Vitry, T.A. Nguyen, E.N. Ramsay, G.E. Caughey, A.L. Gilbert, S. Shakib, P. Ryan, A. Esterman, R.A. McDermott, and E.E. Roughead, General practitioner management plans delaying time to next potentially preventable hospitalisation for patients with heart failure, *Internal Medicine Journal* **44** (2014), 1117-1123.

[18] N.A. Zwar, G.B. Marks, O. Hermiz, S. Middleton, E.J. Comino, I. Hasan, S. Vagholkar, and S.F. Wilson, Predictors of accuracy of diagnosis of chronic obstructive pulmonary disease in general practice, *The Medical journal of Australia* **195** (2011), 168-171.

[19] O. Kostopoulou, J.E. Russo, G. Keenan, B.C. Delaney, and A. Douiri, Information distortion in physicians' diagnostic judgments, *Medical decision making : an international journal of the Society for Medical Decision Making* **32** (2012), 831-839.

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210104

Natural Language Processing for Free-Text Classification in Telehealth Services: Differences Between Diabetes and Heart Failure Applications

Fabian WIESMÜLLER[a,b,1], Dieter HAYN[a,c], Karl KREINER[a], Bernhard PFEIFER[d,e], Gerhard PÖLZL[f], Peter KASTNER[a], and Günter SCHREIER[a]
[a] *AIT Austrian Institute of Technology GmbH, Graz, Austria*
[b] *FH Joanneum, Graz, Austria*
[c] *Ludwig Bolzmann Institute for Digital Health and Prevention, Salzburg, Austria*
[d] *AIT Austrian Institute of Technology GmbH, Hall in Tirol, Austria*
[e] *Landesinstitut für Integrierte Versorgung – LIV Tirol, Innsbruck, Austria*
[f] *Department of Internal Medicine III, Cardiology and Angiology, Medical University Innsbruck, Austria*

Abstract. Telehealth services for long-term monitoring of chronically ill patients are becoming more and more common, leading to huge amounts of data collected by patients and healthcare professionals each day. While most of these data are structured, some information, especially concerning the communication between the stakeholders, is typically stored as unstructured free-texts. This paper outlines the differences in analyzing free-texts from the heart failure telehealth network *HerzMobil* as compared to the diabetes telehealth network *DiabMemory*. A total of 3,739 free-text notes from *HerzMobil* and 228,109 notes from *DiabMemory*, both written in German, were analyzed. A pre-existing, regular expression based algorithm developed for heart failure free-texts was adapted to cover also the diabetes scenario. The resulting algorithm was validated with a subset of 200 notes that were annotated by three scientists, achieving an accuracy of 92.62%. When applying the algorithm to heart failure and diabetes texts, we found various similarities but also several differences concerning the content. As a consequence, specific requirements for the algorithm were identified.

Keywords. Heart failure, diabetes, natural language processing, telehealth

1. Introduction

1.1. Background

A rising prevalence of chronic diseases like heart failure and diabetes mellitus is a major concern for healthcare systems, since they are not only accompanied by a high burden for the patients but also by high costs [1,2]. To counteract these medical and financial

[1] Corresponding Author: Fabian WIESMÜLLER, AIT Austrian Institute of Technology GmbH, Graz, Austria, E-Mail: fabain.wiesmueller.fl@ait.ac.at

issues, different disease management programs were implemented in recent years. E.g., in 2012, a heart failure telehealth system called *HerzMobil* was launched in Tyrol, Austria, by the healthcare provider Tiroler Landeskrankenanstalten GmbH [3]. In 2010, a diabetes management program called *DiabMemory* was started by the Versicherungsanstalt für öffentlich Bedienstete, Eisenbahnen und Bergbau (BVAEB). Both systems are based on a telehealth platform developed by the AIT Austrian Institute of Technology [4]. In both programs, patients are equipped with a specific smartphone app for submitting data concerning their chronic disease to a backend system, where healthcare professionals review the data. While in *HerzMobil*, the focus is set on cardiac parameters such as blood pressure and heart rate [3], participants of the *DiabMemory* program primarily upload data like blood sugar values, insulin administration and their daily food intake [5]. Both systems store these data in a structured format. However, it is possible for the participants (patients and healthcare professionals) of the *HerzMobil* as well as the *DiabMemory* network, to upload additional information in free-text format. This function is used as an additional communication channel between participants. In the *HerzMobil* network, this interaction is mostly used between physicians, nurses and other healthcare professionals, whilst in the *DiabMemory* network, most free-text messages derive from the communication between patients and healthcare professionals. Since these free-texts often contain important information, analyzing them in a structured way, based on natural language processing techniques, would provide additional value.

1.2. State-of-the-art

The amount of unstructured clinical notes in modern healthcare systems is already very high and increases rapidly [6]. Therefore, previous studies have already made efforts towards an analysis of free-text clinical notes. A study by Hebal et al. [7] compared methods for automated and manual data extraction from clinical notes. However, the scope of this study was rather small with 149 individual notes and, additionally, these notes had to follow a specific template for the automated extraction. Another approach for analyzing clinical notes is to detect named entities by using a domain ontology [8]. Amongst other techniques it is also possible, to use systems based on Support Vector Machines or Statistic Language Modeling [9,10].

1.3. Objectives

In this paper, automated classification of clinical free-text notes from a heart failure telehealth network and a diabetes telehealth network are compared. The paper focuses on a) differences in the frequency of free-text categories in the two corpuses and b) on implications on the natural language processing algorithm when applying tools developed for a heart failure application to free-texts from a diabetes telehealth service.

2. Methods

2.1. Corpus

The datasets for the analysis stem from the two telehealth systems: *HerzMobil* and *DiabMemory*. The *HerzMobil* dataset contained 1,564 notes with 55,737 tokens. These

1,564 notes were tokenized so that every resulting text snippet represented a separate sentence, resulting in 3,739 notes on a sentence layer.

In the *DiabMemory* network, the clinical notes can be further divided into *comments*, which were written by the patients themselves, and *feedbacks* which represent a response from a physician. The physicians' respond must not necessarily concern a patient's free-text comment, but it can also refer to an uploaded measurement or any other issue. For the evaluation of the *DiabMemory* system, 71,241 comments written by patients and 156,868 free-text feedbacks from physicians were available.

The clinical notes from both systems were exclusively written in German language, which had to be considered during keyword and regular expression creation in chapter 2.3.

Ethics approval for our analyses was granted by the Ethikkommission der Medizinischen Universität Innsbruck (vote nr. AN2015-0131 35/4.2 374/5.10(4092a)) and the Niederösterreichischen Ethikkommission (vote nr. GS1-EK-4/534-2018).

2.2. Natural language processing

Our algorithm was based on an approach developed by Gruber et al. [11, 12] in 2015, based on *HerzMobil* notes. Adaptions of these algorithms and of the categories used for classification were applied. The final algorithm consisted of pre-processing, stemming, and filtering based on regular expressions, which are described in detail in the following. Details concerning adaptions as compared to Gruber et al. are described in chapter 3.

2.2.1. Pre-processing

Since the notes contained personal information about patients and healthcare professionals, all personal data had to be de-identified using a rule-based algorithm during preprocessing, including data like names, age, address, etc., Additionally, in order to prevent linkage of the de-identified data with telehealth data especially in case of abnormal values mentioned in the free-texts, all numeric values were masked, resulting in e.g. *<Digit> km of running* instead of an actual distance, or *hba1c <Digit>*.

A previously developed algorithm by Wiesmueller et al. was applied to tag all notes containing time-related information like e.g. *gestern (yesterday)* [13]. Even though this algorithm was developed using the notes from the *HerzMobil* network, it was not yet available during the work by Gruber et al.

2.3. Regular expressions and keywords

A set of keywords which were analyzed by regular expressions was applied to all free-text notes. Each expression corresponded to one category. E.g., free-texts containing the word *krank (sick)* were classified as category *state of health.* The keywords for each category were stored in separate plain text files for further use in a Python algorithm.
In general, most regular expressions had the following structure

(?:^|\s)expression(?:\s|$).

The enclosing expressions were start and stop parameters which escaped the expression in the case of e.g. a blank character or a new line. As compared to Gruber et al., the number of keywords was reduced by simply using the stem of a word and appending a

leading and trailing *\w**. An example, where this technique was useful, is the expression *(?:^|\s)\w*schlecht\w*(?:\s|$)*. This single expression covered all cases of *schlecht (bad), schlechter (worse), verschlechtern (worsen), Verschlechterung (worsening),* etc.

2.4. Filter process

A Python[2] script was developed to analyze the notes. Therefore, all the prebuilt keyword files were transformed into a Python list object, in which each element represented an expression. Python's own library for regular expressions was used to filter the input notes. Afterwards, the results for each individual note were aggregated and grouped by patients and physicians, based on each note's sender and recipient ID, respectively.

To normalize the results for each participant, the overall number of free-text elements for each category were subsequently divided by the overall number of free-text elements of the respective participant.

2.4.1. Classification

We slightly adapted the 28 classes used by Gruber et al. and merged similar classes, leading to a smaller set of ten different categories as described in Table 1.

The developed regular expressions categorized each of the free-text elements into none, one or more of these categories.

Table 1 – Categories used for classifying the free-text elements, including short name, description, some examples and the number of keywords that were used in the regular expressions.

Short name	Description	Example	Number of keywords
State of health	Physical and mental well-being, pain, illness, etc.	Had a headache today.	91
Measurements	Measurements of physical activities and vital parameters	*DIGIT* km *DIGIT* kcal	10
Nutrition	Meals, food/drink intake	Had bread and ham for breakfast.	42
Activities	Physical activities, chores, errands	*DIGIT* h cycling	34
Medication	Medication and medication forms (pills, capsules, etc.)	Insulin injected at 10 pm.	245
Technical issues	Technical malfunctions/outages, test messages and phone-related notes	I could not transmit my measurements.	26
Medical appointment	Stationary and ambulatory visits at hospitals, rehab-institutes, physicians, etc.	I am going to be at the hospital next week.	11
Absence	Vacations and travelling	I am going abroad for two weeks.	6
Relatives	Relatives, their activities and visits	Yesterday was my daughter's birthday.	22
Timestamps	Time-related information	Today, tomorrow, at 10 pm, on Monday…	-

[2] http://www.python.org

2.5. Statistical analyses and evaluation

2.5.1. Reference annotation and validation

To validate our algorithm, a sample of 200 free-text notes were randomly selected from the *DiabMemory* dataset. These 200 notes consisted of an equal amount of 100 comments from the patients to the healthcare professionals and 100 feedbacks from the healthcare professionals to the patients. An annotation guideline was developed and provided, based on which three scientists independently annotated the whole set of 200 notes, resulting in three annotations per note. These annotators already had experience with classification of free-text clinical notes. The inter-annotator reliability was calculated via the Fleiss' kappa measure which is part of Pythons *numpy* library [14]. After the first annotation phase, divergent notes were analyzed jointly by the three annotators and an update of the guidelines was made to be more specific even in case of unclear free-text classifications. Afterwards, one of the three annotators applied the finale guideline on the in-sample set of 200 notes to create a *gold standard,* which was used for validating our algorithm.

2.5.2. Comparison of diabetes and heart failure free-text elements

The number and ratio of free-text elements per category as achieved for the *HerzMobil* and the *DiabMemory* corpus were compared with one another. The results of this comparison are depicted in chapter 3.3.

3. Results

3.1. Adaptions of the algorithm

During the development of the *DiabMemory* algorithm, two major differences between the *HerzMobil* and the *DiabMemory* system were identified. The first important difference was the origin of the notes. Whilst both datasets originated from a telehealth system, the *HerzMobil* notes were used for communication between various healthcare professionals like physicians and nurses. In *DiabMemory,* however, a healthcare professional directly communicated with the patient, which might have resulted in more inexpertly and personal expressions and conversations.

The second aspect was the different nature of the diseases heart failure and diabetes. Even though the cores of the algorithms were similar, some keyword categories had to be extended or changed, to fit the profile of both diseases. Nutrition is, for example, a very important aspect in the management of diabetes and, therefore, the nutrition-related keywords were vastly extended. Instead of the initially eight words from Gruber et al., the final dictionary for the *DiabMemory* system had 42 expressions for the nutrition category.

The list of medications had to be extended as well. Diabetes patients need different medication than heat failure patients, which had to be considered during the creation of the keywords. Insulin was especially important, since it is one of the most important medications of patients with diabetes. Activity-related notes represented an entirely new category. Since physical activities are a key factor in the management of diabetes, it was necessary to filter end extract this information. The importance of this category is also reflected by its relative frequency of 16.53% as shown in Table 3.

3.2. Statistical analyses and evaluation

3.2.1. Inter-observer-variability

The inter-observer-variability of the three annotators for the subset of *DiabMemory* notes was calculated based on the annotations prior the annotation guideline update. Based on Feiss' kappa [15], a slight agreement was identified (Fleiss' kappa = 0.017). The comparison of the results of each individual annotator with the gold standard resulted in an accuracy of 91.95%, 90.65% and 87.58%, respectively.

3.2.2. Accuracy of the analyzing algorithm for DiabMemory

To calculate the accuracy of the newly developed algorithm for *DiabMemory*, the manually annotated gold standard was compared to the results of the Python script. Using the previously annotated set of 200 notes resulted in an accuracy of 92.62% for the automated extraction algorithm.

Table 2. Contingency table as achieved for 200 notes, each of which could be positive or negative according to the ten categories, leading to 2,000 decisions as a whole.

	Positive (Annotation)	Negative (Annotation)	Total
Positive (Algorithm)	150	43	193
Negative (Algorithm)	112	1,695	1,807
Total	262	1,738	2,000

3.3. Comparing the relative frequency of categories

Table 3 shows the relative frequency of specific categories normalized by the overall number of notes.

Table 3. The relative frequency of notes per category for the *HerzMobil* and *DiabMemory* system in percent. Note: Relative frequencies do not sum up to 100 %, since 0, 1 or >1 categories can be applied to each note.

	HerzMobil	DiabMemory
State of health	**37.26**	25.33
Measurements	15.97	**27.43**
Nutrition	5.64	**19.8**
Activities	5.03	**16.53**
Medication	**38.05**	4.25
Technical Comments	**31.15**	3.04
Medical appointment	**22.90**	0.42
Absence	**4.09**	2.65
Relatives	**5.63**	3.57
Time references	**63.14**	23.55

4. Discussion

When trying to apply natural language processing methods developed within the *HerzMobil* telehealth program for heart failure patients on data from the *DiabMemory* diabetes monitoring program, we identified that several adaptions were required. These adaptions primarily related to the different focusses addressed in the two programs, i.e. hospital stays, medication adaptions, etc. for the (rather severely ill) heart failure patients, as compared to nutrition, physical activities, and vital parameters for the diabetes patients.

Although some of the differences identified related to the different type of communication (heart failure: between healthcare professionals, diabetes: patient to healthcare professionals), this effect had less influence than the different focus of the diseases.

All categories except three show higher frequencies in the *HerzMobil* notes than in the *DiabMemory* notes. This is because *HerzMobil* notes tend to be longer and, therefore, it is more likely for one note to have multiple categories. The three categories *measurements, nutrition* and *activity* seem to be of great importance for diabetes patients, since they are more common in the *DiabMemory* notes than in *HerzMobil*.

The inter-observer-variability with a Fleiss' Kappa result of 0.017 in between the three annotators shows, that annotating clinical notes is, in some cases, not quite clear and the annotator has room for interpretations. Comparing the accuracy of the three annotators with the newly developed algorithm shows, that the algorithm works with a slightly higher accuracy, even though the reference annotations significantly influenced the golden standard. However, since the annotation guidelines were adapted after the results used to calculate the inter-observer-variability and prior finalizing the golden standard, inter-observer-variability must be interpreted with care.

State of health is a very common category in *HerzMobil* as well as in *DiabMemory*. However, due to the nature of the keywords from Gruber et al., this category is predestined to contain false-positive results. Due to the very general expressions like *gut (good/well),* phrases which have no relation to the state of health might be considered as relevant. For example, the common greeting phrase *Guten Morgen (Good morning)* would be considered as a state of health expression. This is probably the reason for the high relative frequencies of 37.26% and 25.33% for the respective systems.

The difference in the number of medication-related notes between *HerzMobil* and *DiabMemory* might be explained by the origin of the notes. In an expert-to-expert communication like in *HerzMobil*, it is apparently more common to talk about the patient's medication, whilst in a patient to physician communication, the medication is a less frequent topic. Similarly, the rate of activity-related notes might be higher in *DiabMemory* because activities are not so much in the focus of the healthcare professionals (as compared to the patients), although activities are known to have a huge impact on the outcome in both, heart failure and diabetes management.

Due to the large number of 228,109 available notes, using an in-sample set of notes for the evaluation should have a neglectable impact on the results of this paper.

4.1. Outlook

Up to now, we have applied our classification algorithm to the telehealth data retrospectively, to structure the information included in the free-text clinical notes. These analyses have already helped us and will further help us to identify, which important information need to be recorded in a structured way.

Currently, we apply various machine learning and artificial intelligence algorithms to the data derived within our telehealth services to predict events like hospitalizations or dropouts. As a next step, we will include not only structured data, but also the classes derived from the free-text data in these models, which might further improve the model accuracy.

5. Conclusion

Structuring free-text notes from telehealth services via natural language processing methods provides valuable information, which can complement structured data. Even though methods derived for a specific telehealth program can be a good starting point for analyzing data from other programs, various adaptions might be necessary, especially regarding disease specific aspects.

Acknowledgement

The authors would like to thank the Versicherungsanstalt für öffentlich Bedienstete, Eisenbahnen und Bergbau (BVAEB) and the Landesinstitute für Integrierte Versorgung Tirol for supporting this scientific work.

References

[1] K. Barnett, S.W. Mercer, M. Norbury, G. Watt, S. Wyke, B. Guthrie, Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study, *The Lancet* **380**(9836) (2012 Jul), pp. 37-43.

[2] K. Guha and T. McDonagh, Heart Failure Epidemiology: European Perspective, *Current Cardiology Reviews* **9**(2) (2013 May), pp. 23-127.

[3] A. Von der Heidt, E. Ammenwerth et al., HerzMobil Tirol network: rationale for and design of a collaborative heart failure disease management program in Austria, *Wien klinische Wochenschrift* **126** (2014) 2014 Nov, pp. 734-41.

[4] AIT, DiabMemory, https://www.ait.ac.at/themen/telemedical-solutions/projects/diabmemory/, last access: 04.01.2018, URL, last access: 05.01.2021.

[5] V. Schusterbauer, D. Feitek, P. Kastner, H. Toplak, Two-Stage Evaluation of a Telehealth Nutrition Management Service in Support of Diabesity Therapy. *Studies in health technology and informatics* **248** (2018), pp. 314–321.

[6] H. J. Kong, Managing Unstructured Big Data in Healthcare System. *Healthcare informatics research*, ***25***(1) (2019), pp. 1–2.

[7] F. Hebal et al, Automated data extraction: merging clinical care with real-time cohort-specific research and quality improvement data., *Journal of Pediatric Surgery* **52**(1) (2017 Jan), pp. 149-152.

[8] E. Yehia et al, Ontology-based clinical information extraction from physician's free-text notes, *Journal of Biomedical Informatics* **98** (2019), pp. 103276.

[9] H. J. Lee et al, Identifying direct temporal relations between time and events from clinical notes, *BMC Med Inform Decis Mak* **18** (2018 Jul), pp. 49.

[10] R. Zhang et al, Detecting clinically relevant new information in clinical notes across specialties and settings, *BMC Med Inform Decis Mak* **17**(2) (2017 Jul), pp. 86.

[11] K. Gruber, R. Modre-Osprian, K. Kreiner, P. Kastner, G. Schreier, Development of text mining based classification of written communication within a telemedical collaborative network, *eHealth* (2015 Jun), pp. 35-42.

[12] Gruber K., Schreier G. Dashboard und Klassifikatoren in einer telemedizinischen Software zur Unterstützung der kollaborativen Herzinsuffizienz-Versorgung, *Technischen Universität Graz* (2015 Apr).

[13] Wiesmüller et al, Automated Extraction of Time References From Clinical Notes in a Heart Failure Telehealth Network (in press)

[14] Compute Fleiss' kappa using numpy, 2021 GitHub, Inc. https://gist.github.com/skylander86/65c442356377367e27e79ef1fed4adee, last access: 25.01.2021.

[15] Kappa statistics and Kendall's coefficients, 2019 Minitab, LLC, https://support.minitab.com/en-us/minitab/18/help-and-how-to/quality-and-process-improvement/measurement-system-analysis/supporting-topics/attribute-agreement-analysis/kappa-statistics-and-kendall-s-coefficients/, last accessed: 08.02.2021

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

Subject Index

Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

Author Index