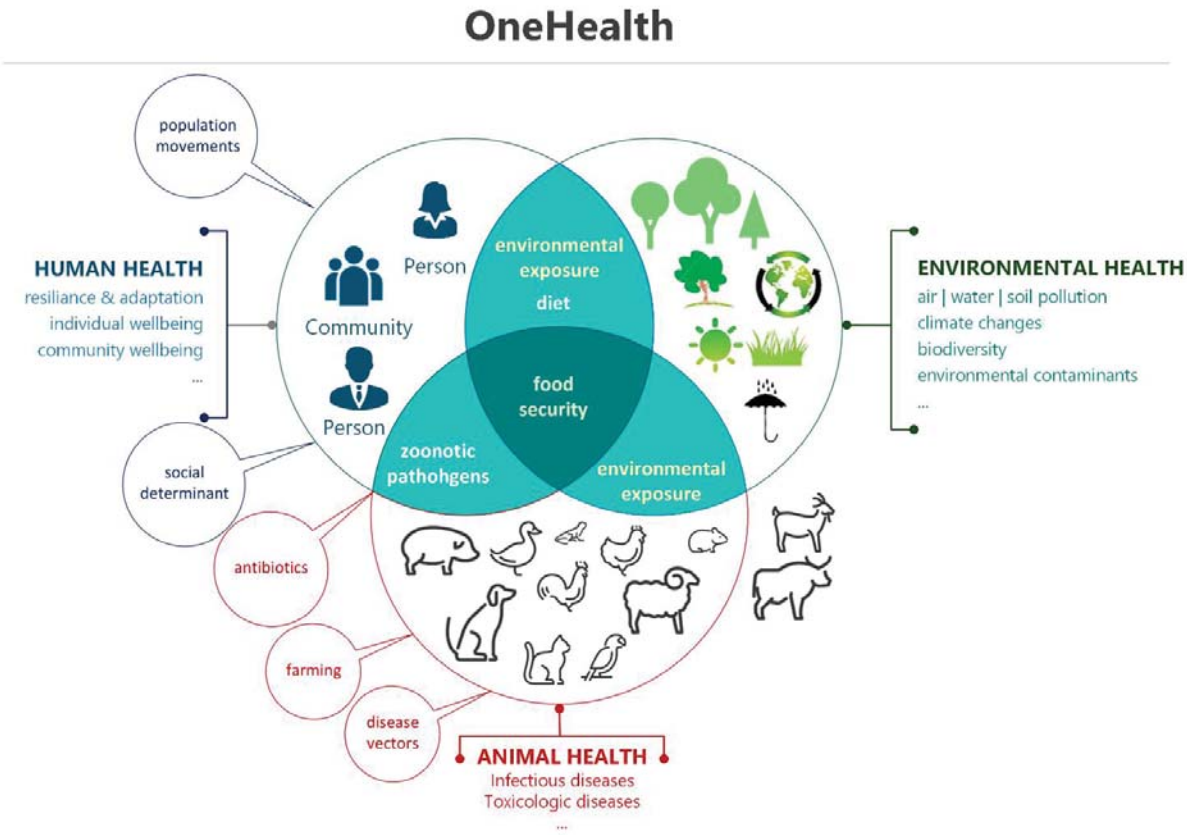
Special Issue

Collaborative Health Applications Across Disciplines: People, Animals, and Ecosystems



Applied Medical Informatics
46 (Suppl. 2) 2024

Editors: Lăcrămioara STOICU-TIVADAR, Parisis GALLOS and Sorana D. BOLBOACĂ

CONTENTS

Editorial

The current special issue of the Applied Medical Informatics journal comes out with 14 papers as extended abstracts of the posters presented during the EFMI-STC 2024 conference "Collaboration across disciplines for the health of people, animals, and ecosystems", 27-29 November 2024, Timişoara, Romania. This issue content has been peer-reviewed and edited by Medical Informatics scientists and researchers across the world. This event was organized by the Romanian Society of Medical Informatics, as a national member of the European Federation for Medical Informatics and hosted by University Politehnica Timişoara. West University Timişoara and "Victor Babeş" University of Medicine and Pharmacy Timişoara were co-organizers.

One Health is a concept that recognizes the interconnectedness of human, animal, and environmental health. It emphasizes that a holistic approach is essential to address health challenges effectively.

Digitalization, with its advancements in technology and data science, plays a crucial role in enhancing One Health efforts. By leveraging digital tools, we can collect, analyze, and share data across various sectors, leading to better insights and informed decision-making. Digital technologies facilitate real-time monitoring of diseases, early warning systems, and rapid response measures. They also enable the development of innovative solutions, such as telemedicine and remote sensing, to improve access to healthcare and environmental monitoring. However, the integration of digital technologies into One Health also presents challenges, including data privacy, cybersecurity, and the digital divide. Addressing these issues is essential to ensure the equitable and sustainable implementation of One Health initiatives in the digital age.

The papers cover specific topics, such as *Practical Implementation of Artificial Intelligence (AI)*, *Data and Standards in Healthcare*, *Human Aspects in Digital Health*, *Education in Healthcare - New Paradigms and Tools*, *Interprofessional Digital Communication and Collaboration in Healthcare*, and *Healthcare Information Systems*.

Lăcrămioara STOICU-TIVADAR
Faculty of Automation and Computers, University Politehnica Timişoara
Fellow of the International Academy of Health Sciences Informatics

Parisis GALLOS
Computational Biomedicine Research Lab, Department of Digital Systems, University of Piraeus
EFMI Publications Officer

Sorana D. BOLBOACĂ
Iuliu Haţieganu University of Medicine and Pharmacy Cluj-Napoca
Romanian Society of Medical Informatics, President

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2/2024, pp: S1-S4

Can Radiomics of Dynamic PET Imaging with 11C-methionine Predict EGFR Amplification Status in Glioblastoma?

Gleb DANILOV[1,*], Andrey POSTNOV[2], Diana KALAEVA[2], Nina VIKHROVA[2], Tatyana KOBYAKOVA[2], and Igor PRONIN[2]

[1] Laboratory of Biomedical Informatics and Artificial Intelligence, National Medical Research Center for Neurosurgery named after N.N. Burdenko, 4th Tverskaya-Yamskaya Str. 16, Moscow, Russian Federation
[2] Department of Neuroimaging, National Medical Research Center for Neurosurgery named after N.N. Burdenko, 4th Tverskaya-Yamskaya Str. 16, Moscow, Russian Federation
E-mails: gdanilov@nsi.ru; apostnov@nsi.ru; dkalaeva@nsi.ru; nvikhrova@nsi.ru; kobyakovata@nsi.ru; pronin@nsi.ru

* Author to whom correspondence should be addressed

Received: 28 June 2024/Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
Introduction: Epidermal growth factor receptor (EGFR) amplification predicts poor survival in patients with brain gliomas. *Purpose:* This study aimed to evaluate whether EGFR amplification status can be predicted using radiomics data from dynamic PET scanning with 11C-methionine. *Materials and Methods:* We analyzed 31 PET/CT scans from 31 patients (7 men 22.6% and 24 women 77.4%, mean age 59 ± 10 years). Three datasets were used to predict EGFR amplification status via machine learning: 1) Radiomic features calculated as time series for each image biomarker; 2) Dynamic tumor-to-normal brain ratio (T/N) of radiopharmaceutical uptake - time series of T/N peak for 26 frames; 3) Static T/N - peak, max, and average T/N for static images. *Results:* Radiomics-based models achieved an average accuracy of 1.0 using k-nearest neighbors across thirty subsampling experiments. Despite this promising result, we approach it critically, considering significant methodological limitations of our study and similar works. These include a small sample size, lack of standardized regions of interest, and absence of reproducibility tests for the selected imaging biomarkers and resulting models. *Conclusion:* Further research should focus on reproducibility, which is crucial for ensuring the non-randomness, generalizability, and real-world value of our findings.

Keywords: Glioblastoma; Radiomics; Positron Emission Tomography Computed Tomography; Artificial Intelligence; Epidermal growth factor receptor

Introduction

Epidermal growth factor receptor (EGFR) amplification predicts poor survival in patients with brain gliomas. In this study, we aimed to evaluate whether EGFR amplification status can be predicted using radiomics data from dynamic positron emission tomography (PET) scanning with 11C-methionine.

Materials and Methods

We analyzed preoperative PET scans with 11C-methionine from adult patients diagnosed with supratentorial glioblastoma (isocitrate dehydrogenase (IDH) wildtype) treated at the N.N. Burdenko National Medical Research Center for Neurosurgery between 2018 and 2020. The scanning continued for 20 minutes after intravenous

radiotracer injection. We used two methods for PET image reconstruction: dynamic and static. The dynamic images were obtained over 26 scanning time intervals consisting of 6 frames of 10 seconds each during the first minute, followed by 6 frames of 20 seconds, 6 frames of 30 seconds, 4 frames of 60 seconds, and finally, 4 frames of 150 seconds each. Static images were generated using data acquired between the 10th and 20th minutes (corresponding to frames 23 to 26). 3D OSEM algorithm with 5 iterations and 8 subsets and Gaussian filter of 5 mm was used for image reconstruction, incorporating attenuation correction from low-dose computed tomography.

Data preprocessing and generation of time-activity curves for 11C- methionine were performed using PMOD software (version 4.2). An experienced radiologist with 5 years of expertise delineated the metabolic tumor volume (MTV), which was subsequently used as the volume of interest for mask creation. To assess the relative metabolic activity of 11C-methionine in glioblastoma, the values of the standardized uptake value (SUV) were summarized in the most active tumor region (SUVt) and in normal brain tissue (SUVn). Then the tumor-to-brain ratio was derived as T/N = SUVt/SUVn. In our study, we applied three different methods for calculating the T/N ratio:

1) T/Npeak: SUVt was measured as the average uptake value in the most metabolically active 1 cm^3 of the tumor
2) T/Nmax: SUVt was measured as the maximum uptake value within the MTV
3) T/Nmean: SUVt was measured as the mean uptake value within the MTV

For all three methods, SUVn was calculated as the average uptake value in a spherical volume of interest (~10 mm in radius) placed in a mixed white and gray matter area of healthy tissue in the contralateral hemisphere.

In line with our dynamic and static image reconstruction methods, we analyzed two types of T/N data: static and dynamic. For static T/N data, we calculated T/Npeak, T/Nmax, and T/Nmean using the static image. Dynamic T/N data were represented as a time series of T/Npeak values for each of the 26 frames.

In addition to the standard T/N measurements, we calculated a wide range of radiomic biomarkers within the MTV for each frame of the dynamic images. To achieve this, we applied an MTV mask to the dynamic images and excluded all values outside the masked region. We then computed radiomic features from these masked images using the RIA library. The voxel values from the MTV were discretized into 2, 4, 8, 16, 32, 64, and 128 bins. Our analysis included first-order statistics, gray level co-occurrence matrix, gray level run length matrix, and geometry-based statistics. Each radiomic feature was represented as a time series derived from the 26 frames of the dynamic image. All calculations and data analysis were performed using R programming language (version 4.3.1) in the RStudio Server IDE (version 2023.09.0, build 463) on an NVIDIA DGX A100 supercomputer.

Thus, to achieve the goal of predicting EGFR amplification status based on PET data, we obtained three datasets:

1) Radiomic features - calculated as time series for each image biomarker
2) T/N (dynamic) - time series of T/N peak for 26 frames
3) T/N (static) - T/N peak, T/N max, and T/N average for static images

Using the feasts and fabletools R packages, we generated 43 features for each time series (including radiomic features and dynamic T/N) to use as predictors in machine learning. For radiomics data, we selected only those predictors that showed statistically significant differences between groups with positive and negative EGFR amplification ($p < 0.05$). We trained various machine learning (ML) models on each dataset, including k-nearest neighbors, naïve Bayes, decision trees, logistic regression with LASSO regularization, random forest, support vector machine, XGBoost, and CatBoost to predict the binary EGFR amplification status (0,1). Each ML experiment with a distinct model was repeated 30 times using subsampling. In each iteration, the original dataset was split into training (45%) and testing (55%) subsets at the patient level. The mlr3verse package ecosystem was used to execute the machine learning procedures.

Results

Our study included 31 PET scans from 31 patients (7 (22.6%) men and 24 (77.4%) women, avg. age 59 ± 10 years). In our sample, only 9 (29%) patients had a positive EGFR amplification status. The median T/N values for T/Npeak, T/Nmax, and T/Nmin were 3.72 [3.22; 4.32], 3.88 [3.22; 4.69], and 1.45 [1.28; 1.81], respectively.

The radiomics pipeline produced 4839 initial features (as time series for 26 frames) which were converted into 191,039 time series features. Only the 1,140 predictors that showed statistically significant differences between the EGFR amplification subgroups were considered for ML. Table 1 summarizes the results of ML experiments in

predicting EGFR amplification status across the three datasets.

Table 1. Performance metrics for the top three ML models in predicting EGFR amplification status for each dataset

Predictors	Model	BACC	ACC	SEN	SPE	F1	AUC
Radiomic features	KNN	1.000	1.000	1.000	1.000	1.000	1.000
Radiomic features	CB	0.983	0.990	0.967	1.000	0.982	1.000
Radiomic features	RF	0.893	0.937	0.787	1.000	0.866	0.999
T/N (dynamic)	LR	0.505	0.592	0.293	0.717	-	0.446
T/N (dynamic)	SVM	0.503	0.663	0.113	0.892	-	0.466
T/N (dynamic)	DT	0.500	0.706	0.000	1.000	-	0.500
T/N ratio (static)	CB	0.504	0.629	0.200	0.808	0.217	0.558
T/N ratio (static)	KNN	0.508	0.645	0.173	0.842	-	0.563
T/N ratio (static)	DT	0.500	0.706	0.000	1.000	-	0.500

BACC – balanced accuracy; ACC – accuracy; SEN – sensitivity; SPE – specificity; AUC – area under ROC-curve

Discussion

We found no publications on using radiomics to predict EGFR amplification based on dynamic and static PET images. MRI studies suggest that predicting EGFR amplification status is challenging. However, B. Sohn et al. (2023) showed promising results in this task (ROC AUC = 0.80) using radiomics on dynamic contrast-enhanced MRI data [1]. N. Vikhrova et al. (2024) revealed statistically significant differences in T/N for PET with 11C-methionine between glioblastoma patients with positive and negative EGFR amplification [2]. Additionally, Z. Li et al. (2022) demonstrated that radiomic features from dynamic O-(2-18F-fluoroethyl)-L-tyrosine PET images can enhance survival prognosis in patients with IDH-wildtype glioblastoma [3]. These findings also justify investigation into the diagnostic and prognostic value of radiomics obtained from dynamic 11C-methionine PET images.

In our study, T/N ratios from 11C-methionine PET in various forms proved to be poor predictors of EGFR amplification (Table 1). In contrast, dynamic radiomic features, transformed into time series features, were selected from a large pool and enabled us to distinguish between positive and negative EGFR classes with high accuracy. Despite these impressive results, we approach them critically, considering the significant methodological limitations including a small sample size, model overfitting risks, lack of standardized regions of interest, and absence of reproducibility tests for the selected imaging biomarkers and resulting models. Such reproducibility is crucial for ensuring the non-randomness, generalizability, and real-world value of our findings [4].

Conclusions

Radiomics enables the creation of a large feature space of quantitative biomarkers, allowing for the selection of predictors that effectively determine EGFR status based on dynamic PET images of glioblastoma with 11C-methionine. However, the reproducibility of these results is unknown and needs to be established in future studies.

List of Abbreviations: ACC – accuracy; AUC – area under ROC-curve; BACC – balanced accuracy; EGFR – epidermal growth factor receptor; IDH – isocitrate dehydrogenase; ML – machine learning; MTV – metabolic tumor volume; PET – positron emission tomography; ROC – receiver operating characteristic; SEN – sensitivity; SPE – specificity; SUV – standardized uptake value; T/N – tumor-to-normal brain ratio.

Author Contributions: GD, AP and IP defined the research's aim and the study's design. AP, DK, NV and TK collected the data. GD produced radiomic features and statistical analysis. All authors read and approved the final manuscript.

Funding: The study was supported by the Ministry of Science and Higher Education of the Russian Federation under agreement No. 075-15-2024-561.

Ethics Statement: This research adhered to the ethical guidelines outlined in the Declaration of Helsinki.

Data Availability Statement: Research data is unavailable publicly due to privacy and ethical restrictions.

Conflict of Interest: The authors declare no conflicts of interest.

References

1. Sohn B, Park K, Ahn SS, Park YW, Choi SH, Kang SG, et al. Dynamic contrast-enhanced MRI radiomics model predicts epidermal growth factor receptor amplification in glioblastoma, IDH-wildtype. J Neurooncol. 2023;164(2):341-351. doi: 10.1007/s11060-023-04435-y.
2. Vikhrova NB, Kalaeva DB, Batalov AI, Pronin IN. Phenotypical and genetic heterogeneity of gliobastomas, comparison of MRI and PET/CT parameters with molecular genetic characteristics of the tumor. IP Pavlov J High Nerv Act. 2024;74(1):48-59. doi: 10.1007/s11055-024-01672-0.
3. Li Z, Holzgreve A, Unterrainer LM, Ruf VC, Quach S, Bartos LM, et al. Combination of pre-treatment dynamic [18F]FET PET radiomics and conventional clinical parameters for the survival stratification in patients with IDH-wildtype glioblastoma. Eur J Nucl Med Mol Imaging. 2022;13;1-11. doi: 10.1007/s00259-022-05988-2.
4. Danilov G, Shevchenko A, Afandiev R, Batalov A, Pogosbekyan E, Zakharova N, et al. Reproducibility of Radiomic Features in Glial Brain Tumors. Stud Health Technol Inform. 2024;316:1165-1166. doi: 10.3233/SHTI240617.

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2 /2024, pp: S5-S8

Is it Beneficial to Use Different Thresholds Over Time for Early Prediction Model?

Sungsoo HONG[1,*], Hyunwoo CHOO[1], Kyung Hyun LEE[1], Sungjun HONG[2], Ki-Byung LEE[1,3], and Chang Youl LEE[3]

[1] AITRICS, Inc., 218 Teheran-ro, Gangnam-gu, 06221 Seoul, Republic of Korea
[2] Medical AI Research Center, Research Institute for Future Medicine, Samsung Medical Center, 81 Irwon-ro, Gangnam-gu, 06351 Seoul, Republic of Korea
[3] Division of Pulmonary, Allergy and Critical Care Medicine, Hallym University Chuncheon Sacred Heart Hospital, 77 Sakju-ro, 24253 Chuncheon, Republic of Korea
E-mails: sshong@aitrics.com; hwchoo@aitrics.com; lkh256@aitrics.com; hsj2864@skku.edu; hasej@aitrics.com; doclcy@hallym.or.kr

* Author to whom correspondence should be addressed

Received: 12 July 2024/Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
In production settings, deep learning models often rely on fixed thresholds. This study investigates whether using varying thresholds over time enhances predictive accuracy and clinical utility, especially for early sepsis prediction. We retrospectively analyzed EMR data from Hallym University Chuncheon Sacred Heart Hospital (2018-2022), excluding patients aged under 18 or without vital signs. Utilizing the AITRICS-VC SEPS deep learning model, which predicts sepsis using six vital signs, eleven lab results and patient information, we examined prediction thresholds at one-hour intervals before sepsis onset. Optimal thresholds for each interval were identified using the Youden index. Net benefit and decision curve analysis compared the performance of time-varying versus global thresholds. Results show interval-specific thresholds yield higher net benefits and increased true positive detections: 456 (0-1 hour), 122 (1-2 hours), 41 (2-3 hours), and 29 (3-4 hours) before sepsis onset. This suggests dynamically adjusting thresholds over time can improve early sepsis detection and patient outcomes.

Keywords: Deep learning; Early prediction; Sepsis; Threshold adjustment; Net benefit

Introduction

It is common to use the area under the receiver operating characteristic curve to present or evaluate the performance of deep learning models. However, in production settings, a specific threshold is typically set to evaluate performance. This raises the question: is this the best convention? For time-series prediction models, such as those used for early prediction of sepsis, it may be beneficial to use different thresholds over time. The model's prediction values may be lower when far from the onset time and relatively higher when close to the onset time. In this study, we explore whether using different thresholds over time is beneficial. We apply the concepts of net benefit and decision curve analysis to evaluate this approach.

Materials and Methods

Study Design

Our objective is to demonstrate the benefits of applying different thresholds over time in a deep learning model focused on the early prediction of sepsis onset. We hypothesize that the optimal threshold varies over time intervals before the onset, which can be particularly beneficial for early detection and intervention.

To determine if the optimal threshold varies over the specified time intervals, we split each episode into the intervals of 0-1 hour, 1-2 hours, 2-3 hours, and 3-4 hours before the onset. For negative episodes (those without an onset time), we used the timestamp of the last observation recorded instead. The Youden index, defined as the maximum of the sum of sensitivity and specificity scaled by subtracting 1, is a useful for finding the optimal threshold considering the balance. It is used to identify the optimal threshold for each interval [1].

Data Collection and Processing

We retrospectively collected EMR data in Hallym University Chuncheon Sacred Heart Hospital in period of from 2018-01-01 to 2022-12-31. Patients who are below 18 years of age or who have no vital signs reported were excluded. Collected data was refined into episode per each patient. Episodes were segmented into one-hour intervals relative to the sepsis onset time.

Deep Learning Model for Early Prediction of Sepsis

In this study, we utilized a deep learning model called AITRICS-VC SEPS, which has been approved by the Korea's Ministry of Food and Drug Safety as a medical AI solution that predicts future sepsis occurrence [2]. The model gets six vital sign, eleven lab results, and patient information, refer to the past in time-series manner, and then returns future sepsis risk probability.

Net Benefit and Decision Curve Analysis

To evaluate whether a model does more good than harm when used in clinical practice, we utilize the concept of net benefit [3-4]. Net benefit provides a measure that balances the true positive outcomes against the false positive outcomes, adjusted by the odds of a given threshold probability. The formula for net benefit is defined as:

$$Net\ benefit = \frac{True\ Positive}{N} - \frac{False\ Positive}{N} \times \frac{p_t}{1 - p_t} \quad (1)$$

where N is the total sample size and p_t is a threshold probability to define when a patient is positive. The number of true positives and false positives are determined based on the threshold probability p_t. Additionally, Standardized net benefit is more interpretable when compared with net benefits, and the outcome prevalence of each group is used to weight.

The decision curve is plotted with threshold probabilities on the x-axis and their corresponding net benefit values on the y-axis.

Results

Demographic Information

During the period, we collected 46,842 patients. Of which 22,317 (47.6%) were female and 24,525 were male (52.4%). Mean age of all patients was 58.2 (Standard deviation 18.3). The prevalence of sepsis was 6.7% (n=3,120).

Decision Curve Analysis

Figure 1 illustrates the decision curves for interval-group-specific cutoffs alongside the global cutoff. For all interval groups, the optimal thresholds calculated using the Youden index correspond to higher net benefit values, compared to those derived from the global cutoff.

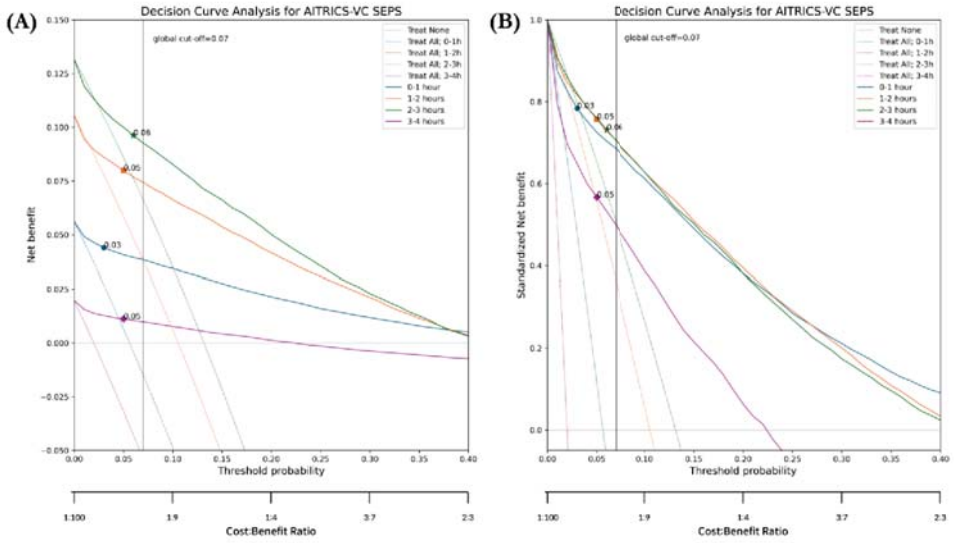


Figure 1. (A) Net benefit curves for each interval group. **(B)** Standardized net benefit curves. The global threshold (0.07) is represented by a vertical line. Optimal thresholds for each benefit line are indicated by numbers and distinct markers. The dashed line represents a policy of treating all patients regardless of their score.

Group-wise optimal thresholds enhanced the detection of true positives across time intervals. Increases were observed in the number of true positives: 456 (9.4%) within the 0-1 hour interval, 122 (4.1%) within 1-2 hours, 41 (1.8%) within 2-3 hours, and 29 (3.1%) within 3-4 hours.

Discussion

In this study, we explored that using varying threshold over time would be beneficial than using a single global threshold when deploying a deep learning model, especially for early sepsis prediction. To achieve this, we utilized the net benefit concept and decision curve analysis.

Our retrospective study highlights the benefits of using varying thresholds. However, a key limitation is the reliance on known onset times, which allows for grouping data by hour-based intervals. In real-world settings, the onset time cannot be predicted in advance. To address this, we suggest including informative messages when alarms are triggered using a global cutoff, indicating the observed findings.

Conclusions

This study demonstrates that applying time-varying thresholds in early sepsis prediction models significantly enhances detection accuracy and clinical decision-making benefits. The improved net benefits and increased true positive rates highlight the potential for this approach to be adopted in clinical settings, offering a more responsive and precise method for early intervention.

List of Abbreviations: Not applicable.

Author Contributions: SSH, HC defined the research's aim and the experiments' design. SJH participated in the design of the study. SSH carried out the experiments and performed the statistical analysis. SSH makes figure. HC and SSH write the manuscript. SSH and KHL collected the data and clean them. KHL, SJH critically reviewed the draft and helped manuscript to be completed.

Funding: This research received no funding.

Ethics Statement: This study is approved by the institute review board of Hallym University Chuncheon Sacred Heart Hospital (IRB No: CHUNCHEON 2023-03-007-002). The informed consent is waived as this is a retrospective study that has minimal risk to patients.

Data Availability Statement: The data utilized in this study were obtained under the supervision and with the grant support of Hallym Chuncheon Sacred Heart Hospital. For access to the data, individual contact with the hospital is required.

Conflict of Interest: HC, SSH, KHL and KBL are employee of AITRICS.

References

1. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J. 2005;47(4):458-472. doi: 10.1002/bimj.200410135.
2. Sung M, Hahn S, Han CH, Lee JM, Lee J, Yoo J, et al. Event Prediction Model Considering Time and Input Error Using Electronic Medical Records in the Intensive Care Unit: Retrospective Study. JMIR Med Inform. 2021;9(11):e26426. doi: 10.2196/26426.
3. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-574. doi: 10.1177/0272989X06295361.
4. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. J Clin Oncol. 2016;34(21):2534-2540. doi: 10.1200/JCO.2015.65.5654.

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2 /2024, pp: S9-S12

Beyond the ROC Curve: Activity Monitoring to Evaluate Deep Learning Models in Clinical Settings

Hyunwoo CHOO[1,*], Kyung Hyun LEE[1], Sungsoo HONG[1], Sungjun HONG[2], Ki-Byung LEE[1,3], and Chang Youl LEE[3]

[1] AITRICS Inc., 218 Teheran-ro, Gangnam-gu, 06221Seoul, Republic of Korea
[2] Medical AI Research Center, Research Institute for Future Medicine, Samsung Medical Center, 81 Irwon-ro, Gangnam-gu, 06351 Seoul, Republic of Korea
[3] Division of Pulmonary, Allergy and Critical Care Medicine, Hallym University Chuncheon Sacred Heart Hospital, 77 Sakju-ro, 24253 Chuncheon, Republic of Korea
E-mails: sshong@aitrics.com; hwchoo@aitrics.com; lkh256@aitrics.com; hsj2864@skku.edu; hasej@aitrics.com; doclcy@hallym.or.kr

* Author to whom correspondence should be addressed

Received: 12 July 2024/Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
We evaluated 'VITALCARE-SEPS', a deep learning model for sepsis prediction, using the activity monitoring operator characteristics curve with two different scoring algorithms. This evaluation is crucial as the AMOC curve addresses the time-dependent nature of predictions, providing a more nuanced performance assessment than traditional ROC metrics. Our findings demonstrate that the AMOC curve significantly enhances the evaluation of time-series predictions, enabling more accurate and continuous performance monitoring of machine learning models in clinical settings. This approach can improve model deployment and ultimately lead to better patient outcomes in healthcare.

Keywords: Deep learning; Sepsis; Activity monitoring; ROC (receiver operating characteristic) curve

Introduction

The area under the receiver operating characteristic (ROC) curve (AUC-ROC) is a well-established metric for evaluating machine learning models in binary classification tasks. It provides a single scalar value summarizing the trade-off between true positive and false positive rates across various thresholds. However, AUC-ROC may not be ideal for time-series contexts or dynamic environments, such as clinical settings, as it treats each inference point equally without considering temporal aspects or the sequence of predictions.

To address this limitation, time-dependent ROC curves were introduced [1,2] to evaluate model performance at different time points, showing how prediction accuracy evolves over time. Yet, this approach still falls short in capturing the timing of predictions relative to clinical outcomes, failing to reward early predictions or penalize late ones, which is crucial in time-sensitive scenarios like sepsis detection.

Tom Fawcett proposed the Activity Monitoring Operator Characteristics [3] (AMOC) curve to offer a more nuanced evaluation of time-series predictions. The AMOC curve extends traditional and time-dependent ROC analyses by incorporating the timing of predictions, rewarding early correct predictions and penalizing late or incorrect ones. This approach provides a more comprehensive assessment, especially where timely interventions are critical.

Despite its advantages, the AMOC curve remains underutilized for evaluating time-series predictions. This study introduces the AMOC curve to assess the deep learning model 'AITRICS-VC SEPS,' used for early sepsis detection in clinical settings. By comparing the AMOC curve with traditional and time-dependent ROC curves, we show that the AMOC curve captures both accuracy and timing of predictions, highlighting its effectiveness and advocating for its broader use in scenarios where timely predictions are crucial.

Materials and Methods

Study Design

We evaluated the deep learning model 'VITALCARE-SEPS', deployed in clinical settings. The model uses five vital signs, twelve lab results, and demographic information to generate a pseudo-probability of sepsis (0 to 1) upon new events (e.g., new lab results). Model inference outputs were retrospectively collected from EMR data of patients with sepsis infected from Hallym University Chuncheon Sacred Heart Hospital (2018-2022). Each model output was binarily labeled based on whether sepsis occurred within four hours. AMOC curves were drawn from these labeled inference results.

Activity Monitor Operator Characteristics

To illustrate a typical curve, we plot the false alarm rate against the average score. We evaluate each model inference output against threshold values from 0 to 1, at 0.01 intervals. This systematic quantization allows assessment across all thresholds. Scoring each inference across all patients and aggregating the results yields counts of true and false alarms and the average score.

Scoring Functions Used to Draw AMOC Curve

To evaluate model performance and construct the AMOC curve, we employed two scoring functions that account for both prediction accuracy and the timing of predictions:

1) Simple Score Calculation:

This function assigns a binary score where a true positive (score = 1) is recorded if the predicted probability (PredictionScore) exceeds a defined threshold and the event occurs (ObservedOutcome = 1). All other cases receive a score of 0.

2) Time-Sensitive Scoring:

This advanced function rewards early and moderate correct predictions differently while penalizing missed events and false positives. Specifically:

- Early Correct Predictions: Receive a higher reward if the event is predicted before a pre-defined time threshold.
- Moderate Correct Predictions: Receive a lower reward if predicted after this threshold.
- Penalties: Applied for missed events and false positives, scaled based on the time to event occurrence.

Results

From 2018 to 2022 in Hallym University Chuncheon Sacred Heart Hospital, we retrospectively collected EMR data from 3,460 sepsis positive patients. From those patients, we generated 93,922 model inference outputs. We illustrated two AMOC curves applying different scoring algorithms on Figure 1.

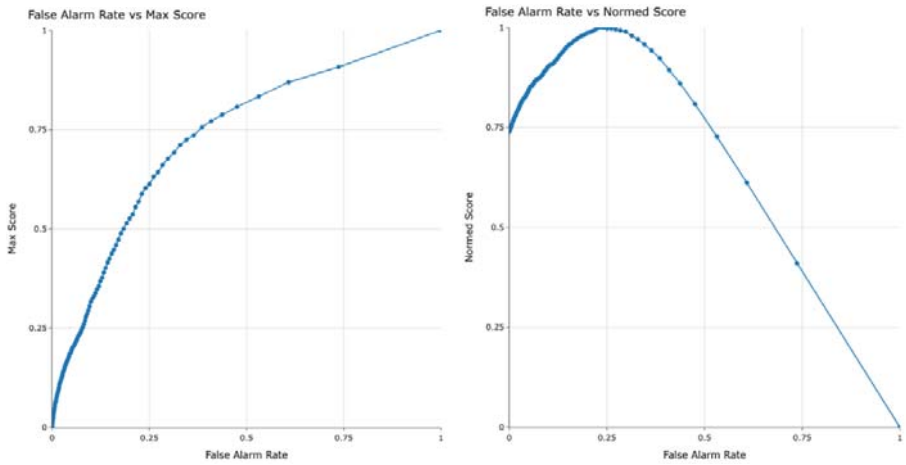


Figure 2 (Left) AOMC curve drawn using simple scoring function **(Right)** AOMC curve drawn using time-sensitive scoring function

Discussion

In this study, we used the Activity Monitoring Operator Characteristics (AMOC) curve to evaluate the 'VITALCARE-SEPS' deep learning model for sepsis prediction, applying two different scoring functions. This analysis highlights each approach's strengths and limitations in capturing the timing of predictions, crucial in clinical settings.

The AMOC curve using the simple scoring function (Figure 2, left) treats all correct predictions equally, regardless of timing. This method results in a curve where increases in the false alarm rate often correspond to proportional increases in the score. While straightforward, it does not account for the timing of predictions, potentially overestimating model performance in time-sensitive contexts.

Conversely, the AMOC curve with the time-sensitive scoring function (Figure 2, right) addresses this by incorporating prediction timing into the evaluation. It rewards early correct predictions and penalizes missed events and false positives based on their timing. This results in a more nuanced curve that better reflects the practical needs of clinical decision-making, distinguishing between models that perform well in a timely manner and those that do not.

The time-sensitive scoring function enhances the AMOC curve's intuitiveness for real-time monitoring and evaluation. By accounting for the timing of predictions, it provides a more accurate assessment of model effectiveness in critical settings like sepsis detection. This refined approach could improve the deployment and performance of machine learning models in healthcare, potentially leading to better patient outcomes.

Further research is needed to explore optimal reward and penalty parameters for the time-sensitive scoring function and their impact on model evaluation.

Conclusions

The AMOC curve enhances the evaluation of time-series predictions, leading to more accurate and continuous monitoring of clinical machine learning models.

List of Abbreviations: activity monitoring operator characteristics (AMOC)

Author Contributions: HC defined the research aim, designed the experiment, created the figures, and wrote the manuscript. HSS HSJ, and KL participated in the research design, conducted the experiments, and critically reviewed the manuscript.

Funding: This research received no funding.

Ethics Statement: This study is approved by the institute review board of Hallym University Chuncheon Sacred Heart Hospital (IRB No: CHUNCHEON 2023-03-007-002). The informed consent is waived as this is a retrospective study that has minimal risk to patients.

Data Availability Statement: The data utilized in this study were obtained under the supervision and with the grant support of Hallym Chuncheon Sacred Heart Hospital. For access to the data, individual contact with the hospital is required.

Conflict of Interest: HC, SSH, KHL and KBL are employee of AITRICS.

References

1. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56(2):337-344. doi: 10.1111/j.0006-341x.2000.00337.x.
2. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Med Res Methodol. 2017;17(1):53. doi: 10.1186/s12874-017-0332-6.
3. Fawcett T, Provost F. Activity monitoring: Noticing interesting changes in behavior. In: KDD '99: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining. 1999:53-62. doi: 10.1145/312129.312195.

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2 /2024, pp: S13-S16

Relationship between In-Hospital Sepsis Prediction Score and Prevalence of Community-Onset Sepsis: Triage for Sepsis Risk Management

Kyung Hyun LEE[1,*], Hyunwoo CHOO[1], Sungsoo HONG[1], Sungjun HONG[2], Ki-Byung LEE[1,3], and Hochan CHO[4]

[1] AITRICS. Inc, 218 Teheran-ro, Gangnam-gu, 06221 Seoul, Republic of Korea
[2] Medical AI Research Center, Research Institute for Future Medicine, Samsung Medical Center, 81 Irwon-ro, Gangnam-gu, S06351 eoul, Republic of Korea
[3] Division of Pulmonary, Allergy and Critical Care Medicine, Hallym University Chuncheon Sacred Heart Hospital, 77 Sakju-ro, 24253 Chuncheon, Republic of Korea
[4] Department of Internal Medicine, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, 42601 Daegu, Republic of Korea
E-mails: lkh256@aitrics.com; hwchoo@aitrics.com; sshong@aitrics.com; hsj2864@skku.edu; hasej@aitrics.com; hochan3632@gmail.com

* Author to whom correspondence should be addressed

Received: 12 July 2024/Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
Early diagnosis of sepsis is crucial in clinical practice. Several studies have proposed sepsis prediction models to forecast the onset of sepsis in hospitals. However, validation of prediction models for community-onset sepsis, which is sepsis developed before admission to the hospital, is insufficient. This study investigates the relationship between the in-hospital prediction model scores and community-onset sepsis. We used hierarchical logistic regression analysis to explore the relationship between sepsis prevalence and AITRICS-VC SEPS tertile categories while adjusting for potential confounders. The low-SEPS group was used as the reference group. The odds ratio (ORs) of sepsis comparing the moderate versus low SEPS group are 1.198 (95%, 1.075-3.654), and the high versus low VC-SEPS group are 8.683 (95%, 4.995-15.095). Even though the sepsis prediction model was designed to predict in-hospital sepsis, high prediction scores are related to the prevalence of community-onset sepsis. This result implies that SEPS scores can stratify sepsis risks and be considered a patient assessment tool for triage.

Keywords: Sepsis; Prediction; Deep learning; Risk stratification; Regression analysis

Introduction

Identifying sepsis early is essential in clinical settings [1]. Numerous prediction models have been developed to identify in-hospital sepsis [2]. However, research validating these models' effectiveness on patients who develop sepsis before hospital admission, known as community-onset sepsis, is insufficient. This study examined the connection between community-onset sepsis and the prediction score of AITRICS-VC SEPS, a model designed to predict in-hospital sepsis.

Materials and Methods

Study Participants and Definition of Sepsis

We collected patient data from the electronic medical records (EMR) of Keimyung University Dongsan Hospital (KUDH) from June 5th, 2023, to January 31st, 2024. Patients were identified as having sepsis using Rhee et al.'s Sepsis Clinical Surveillance Definition criteria [3]. Sepsis was classified into community-onset sepsis (COS) and hospital-onset sepsis (HOS), depending on whether it was identified within or after 48 hours of hospital admission. We excluded patients who identified as HOS in this analysis.

Sepsis Prediction Model (AITRICS-VC SEPS)

The sepsis prediction model was developed based on deep learning [4] and has been approved by the Korean Ministry of Food and Drug Safety for use in general wards as an AI-based clinical decision support system (CDSS). This model primarily aims to generate early warning scores to help medical staff screen patients before they develop septic shock. This model generates a score to identify sepsis in patients by providing an alert 4 hours before a formal diagnosis. It can handle data from clinical practice in time series and requires systolic blood pressure, diastolic blood pressure, pulse rate, respiratory rate, and body temperature to calculate scores. Additional laboratories related to sepsis can also be used to improve the model's predictive performance. It achieved the area under the receiver operating characteristic (AUROC) curves for 0.894 (95% CI, 0.832-0.937) on identifying COS in KUDH.

Statistical Analysis

We investigated whether AITRICS-VC SEPS could be used to stratify the risk of developing sepsis. Since age and gender are closely linked to sepsis prevalence [5], we grouped VC-SEPS scores into three categories based on age-specific and gender-specific percentiles: (Low-SEPS / Moderate-SEPS / High-SEPS). Additionally, we assessed if AITRICS-VC SEPS could serve as an independent biomarker for predicting sepsis. We used hierarchical logistic regression analysis to explore the relationship between sepsis prevalence and AITRICS-VC SEPS tertile categories while adjusting for potential confounders such as age, gender, body mass index (BMI), systolic blood pressure (SBP), and respiratory rate (RR). The low-SEPS group was used as the reference group.

Results

Study Populations

A total of 6455 patient data were enrolled in the study, and 325 patients were diagnosed as having sepsis. In sepsis patients, 229 patients were identified as COS, and 96 patients were identified as HOS. In this study, 6130 non-sepsis patients and 229 COS patients were used for analysis.

Hierarchical Logistic Regression Analysis

Table 1 shows the results of hierarchical regression analysis when incrementally adding variables related to sepsis. In the crude analysis, the odds ratios (ORs) of sepsis comparing the moderate versus low SEPS group are 2.516 (95% CI, 1.377-4.598), and the ORs of the high versus low SEPS group are 13.225 (95% CI, 7.779-22.484) for sepsis. After adjustments for age, gender, BMI, SBP, and RR, the ORs of sepsis comparing the moderate versus low SEPS group are 1.198 (95% CI, 1.075-3.654), and the high versus low VC-SEPS group are 8.683 (95% CI, 4.995-15.095).

Table 1. Results of the hierarchical regression analysis when incrementally adding variables related to sepsis. All values in each cell are reported in the form of Odds Ratio (Confidence Interval)

Variable		Regression 1	Regression 2	Regression 3	Regression 4
SEPS	Low	ref	ref	ref	ref
	Moderate	2.516 (1.377-4.598)	1.973 (1.070-3.637)	1.971 (1.069-3.633)	1.982 (1.075-3.654)
	High	13.225 (7.779-22.484)	10.212 (5.923-17.609)	10.127 (5.850-17.532)	8.683 (4.995-15.095)
Age			1.060 (1.049-1.072)	1.060 (1.049-1.072)	1.059 (1.048-1.071)
Gender	Male		ref	ref	ref
	Female		0.323 (0.240-0.433)	0.324 (0.241-0.435)	0.341 (0.253-0.459)
BMI			0.963 (0.926-1.000)	0.963 (0.927-1.001)	0.967 (0.930-1.005)
SBP				0.999 (0.994-1.005)	0.997 (0.992-1.003)
RR					1.125 (1.076-1.175)

Discussion

In this study, we investigated that SEPS score can be used as independent predictors to assess COS. ORs were exponentially increased when the SEPS score increased even after adjustment for confounding factors. The SEPS high group was highly related to the prevalence of COS compared to the SEPS low group. However, the limitation of our study is that we focused only on COS, which is identified near admission. The relationship between SEPS scores and HOS could be different compared to COS. So, further analysis is necessary to make guidance for interpreting SEPS scores.

Conclusions

Even though AITRICS-VC SEPS were designed to predict HOS, SEPS scores are highly related to the prevalence of COS. This result implies that SEPS scores can stratify sepsis risks and be considered a patient assessment tool for triage.

List of Abbreviations: Odds Ratio (ORs), Body Mass Index (BMI), Systolic Blood Pressure (SBP), Respiratory Rate (RR), Electronic Medical Records (EMR), Confidence Interval (CI), Clinical Decision Support System (CDSS), Community-onset Sepsis (COS), Hospital-onset Sepsis (HOS), Keimyung University Dongsan Hospital (KUDH), area under the receiver operating characteristic (AUROC).

Author Contributions: KHL and KBL defined the research's aim and the experiment's design. KHL, SSJ, HWC, SSH, KBL, and HCC participated in the study's design. KHL carried out the experiments and performed the statistical analysis. KHL made a table. KHL wrote the manuscript. KHL and HCC collected the data and cleaned them. HWC, SJH, KBL, and HCC critically reviewed the draft and helped complete the manuscript.

Funding: This research received no funding.

Ethics Statement: The Institutional Review Board (IRB) of KUDH approved this study, and a waiver of consent was obtained (IRB No. 2020-08-045).

Data Availability Statement: The data utilized in this study were obtained under the supervision and grant support of Keimyung University Dongsan Hospital. Individual contact with the hospital is required for access to the data.

Conflict of Interest: HWC, SSH, KBL, and KHL are AITRICS employees.

References

1. van der Poll T, van de Veerdonk FL, Scicluna BP, Netea MG. The immunopathology of sepsis and potential therapeutic targets. Nat Rev Immunol. 2017;17(7):407-420. doi: 10.1038/nri.2017.36.

2. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med. 2020;46(3):383-400. doi: 10.1007/s00134-019-05872-y.
3. Rhee C, Dantes R, Epstein L, Murphy DJ, Seymour CW, Iwashyna TJ, et al; CDC Prevention Epicenter Program. Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009-2014. JAMA. 2017;318(13):1241-1249. doi: 10.1001/jama.2017.13836.
4. Sung M, Hahn S, Han CH, Lee JM, Lee J, Yoo J, et al. Event Prediction Model Considering Time and Input Error Using Electronic Medical Records in the Intensive Care Unit: Retrospective Study. JMIR Med Inform. 2021;9(11):e26426. doi: 10.2196/26426.
5. Ko RE, Kang D, Cho J, Na SJ, Chung CR, Lim SY, et al.; Korean Sepsis Alliance (KSA) investigators. Influence of gender on age-associated in-hospital mortality in patients with sepsis and septic shock: a prospective nationwide multicenter cohort study. Crit Care. 2023;27(1):229. doi: 10.1186/s13054-023-04515-5.

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2 /2024, pp: S17-S20

Let's not Miss the Opportunity to Improve Rare Disease Reporting Through ICD-11 Implementation

Miroslav ZVOLSKÝ[1,2,*], Šárka DAŇKOVÁ[1], Marie VIKDAL[3], Carine ALSOKHN[4], Robert JAKOB[4], and Milan BLAHA[1]

[1] Institute of Health Information and Statistics of the Czech Republic, Palackého náměstí, 375/4, 12800, Prague, Czech Republic
[2] First Faculty of Medicine, Charles University, Kateřinská, 32, 12108, Prague, Czech Republic
[3] Code Systems Department, Norwegian Directorate of Health, PO Box 220, Skøyen, 0213, Oslo, Norway
[4] Data Analytics and Delivery for Impact Division, World Health Organization, Avenue Appia, 20, 1211, Geneva, Switzerland
E-mails: miroslav.zvolsky@uzis.cz; sarka.dankova@uzis.cz; marie.vikdal@helsedir.no; alsokhnc@who.int; jakobr@who.int; milan.blaha@uzis.cz

* Author to whom correspondence should be addressed;

Received: 25 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
The identification of rare diseases in health data is crucial for their epidemiological definition. Beyond the existing terminology system ORPHAcodes, we describe the possibilities of the 11th revision of the International Classification of Diseases. We analyzed the linkages of ORPHAcodes to ICD-11 items with only 376 items lacking unique representation in ICD-11. ICD-11 can be used to identify most rare diseases in administrative data.

Keywords: Rare diseases; Classification; Terminology; Administrative data

Introduction

Rare diseases (RD) are defined differently in different parts of the world, but the basic parameter for defining rarity is most often the epidemiological characteristic of prevalence. In order to determine prevalence RD have to be identified in the collected data.

In Europe, the definition of a rare disease is "one affecting fewer than 5 out of 10,000 members of the general population", while it must be a serious, chronic and often life-threatening condition [1-4].

An important information resource in the field of rare diseases is Orphanet, a multilingual information portal on rare diseases and orphan drugs [5].

Rare disorders defined in Orphanet nomenclature (ORPHAcodes) are cross-referenced with other international terminologies, classifications, and reference databases (including OMIM, ICD-10, ICD-11, SNOMED-CT, MedDRA, UMLS, MeSH, and GARD), in order to enable interoperability between different clinical information systems [6].

The content of Orphanet is published and regularly updated on Orphanet portal, where a full-text search system and an interactive presentation of rare disease classification hierarchies are available. Its full content can be downloaded at Orphadata as the Orphanet Nomenclature Pack, available in nine languages including Czech.

Based on Orphanet database and knowledge resources several projects expanding the original work took place in recent years: 1) Joint European project RD-ACTION in 2014-2020 with the main objectives to support the development and sustainability of the Orphanet database and ensure an appropriate codification of RD in health information systems, 2) European project RD-CODE in 2019-2021 with objectives to develop rules and structured

documentation for ORPHAcoding needs and pilot implementations in target European countries (Czech Republic, Malta, Romania, Spain, Italy/Veneto Region), 3) European projects OD4RD and OD4RD2 in 2022-2025 with main objective to contribute to the generation of standardized, interoperable data on RD diagnosis, through maintenance of the Orphanet nomenclature of RD.

Materials and Methods

In addition to specific rare disease data collections, rare disease cases can also be found in existing administrative data sources.

Administrative health data refers to the information that is collected within the health-care system for reasons other than clinical care. It may include claims for reimbursement, records of health services provided, medical procedures carried out, prescribed medication and information about conditions for which the health care service was provided.

Different coding systems are used in particular administrative databases to identify disease or health condition. Examples of existing coding systems with the ability to capture rare diseases are: ICD-10 (International Classification of Diseases, 10th Revision) and derived modifications [Classification and coding system supplemented by a terminology layer], ICD-11 (International Classification of Diseases, 11th Revision) [Classification, terminology and coding system], MedDRA (Medical Dictionary for Regulatory Activities)[Medical terminology dictionary], MeSH (Medical Subject Headings) [Controlled vocabulary for literature indexing], OMIM (Online Mendelian Inheritance in Men) [Catalogue of human genes and phenotypes], ORPHAcodes [Terminology and coding system], SNOMED CT (Systemized Nomenclature of Medicine, Clinical Terms) [Terminology and ontology].

However, the main classification/coding system used in general administrative data collections for the purpose of standardized Health status information is the ICD-10. In some countries the ICD-10 or modified ICD-10 coding schemes are enriched by specialised clinical terminologies. A good example is Alpha-ID system in Germany. The Alpha-ID is a sequential and stable identification number, which is allocated to each entry in the alphabetical index of ICD-10.

11th revision of ICD is the result of an unprecedented collaboration with clinicians, statisticians, coders, classification and IT experts from around the world, to ensure interoperability and comparability of digital health data. ICD-11 provides access to 17,000 diagnostic categories, with over 100,000 medical diagnostic index terms [7]. World Health Assembly adopted ICD-11 in May 2019, and it came into effect on 1st January 2022 [8].

ICD-11 includes detailed and frequently used information and combinations. However, it is not restricted to these coumpounds; an infinite number of meaningful code combinations can be created. All entities in the ICD-11 are rendered in the Foundation, which is an acyclic graph (meaning no entity can ever be its own descendant as or parent) of all entities and their relationship trees. Unlike in previous revisions ICDs, the Foundation may have multiple inheritances, where a single term may have one or more, sometimes many more, parents [9].

As a part of the ICD revision, collaboration with Orphanet resulted in including all rare disease terminology in the terminology layer of ICD-11. The ongoing WHO - Orphanet collaboration was set to enable regular updates of rare diseases content in the ICD-11 [10].

There are several possibilities of identification RD in ICD-11.

Diseases or health conditions that are more common, where greater public health importance requires statistical outputs, have their own ICD-11 for Mortality and Morbidity Statistics (MMS) code (e.g. CA25.0 Classical cystic fibrosis).

All rare diseases, although it may not have its own MMS code, can be identified in the ICD-11 terminology layer, referred to as it's Foundation with a Uniform Resource Identifier (URI) assigned to each entity and are therefore identifiable in the data. The highly detailed identification allows subsequent aggregation into general categories for statistical outputs.

A rare condition can also be captured by combining multiple ICD-11 codes or URIs. A disease that does not individually meet the prevalence criteria for rarity can be classified as a rare disease when additional case-specific

information is captured via code combination. For example, an uncommon cancer can be considered rare when detailed information about its localization, histopathology, or other specific clinical parameters is provided.

ORPHAcodes are updated continuously in the database, while the official release of Orphanet Nomeclature Pack in July is valid for following year. ICD-11 releases are published in February valid for following year. The publication of each system releases differ which can be one of the reasons for the difference RD listing in ICD-11. An ongoing effort is underway to facilitate the timely updating cycles with Orphanet.

ICD-11 offers multilingual web-based digital tools for conversion of text based clinical information into coded health data [9].

Results

ICD-11 includes nearly 5,500 rare diseases and their synonyms in the Foundation and aggregated under the same nonspecific ICD-11 code [11].

Based on the Orphanet Terminology Pack, we analyzed the number of ORPHAcodes rare disease terminology concepts (version available for download from 1 July 2023) in relation to ICD-11 (release 2024-01).

The ORPHAcodes system contains a total of 11,136 concepts in the current Nomenclature pack. Of these, 1,230 have the property 'OBSOLETE' or 'NON RARE IN EUROPE', these concepts should not be used for active coding of rare diseases in Europe. A further 3,308 have the level 'group of disorders' or 'subtype of disorder' in ORPHAcodes (see in Table 1).

Table 1. Distribution of 9,906 active ORPHAcodes terminology concepts by level of detail

Level of detail	**Number of items**
Disorders	6,598
Groups of disorders	2,264
Subtypes of disorders	1,044
Total	**9,906**

Table 2 shows further breakdown of the 6,598 rare diseases identified by the ORPHAcode mapped to ICD-11 MMS categories.

Table 2. ORPHAcodes terminology concepts mapped to ICD-11 MMS according to the type of linkage

BTNT (ORPHAcode is broader than the targeted code used to represent it)	225
E (Exact mapping: the two concepts are equivalent)	1,367
NTBT (ORPHAcode is narrower than the targeted code used to represent it)	3,618
Total aligned rare diseases	**5,210**

In 376 cases is the rare disease defined in ORPHAcodes not separately identifiable using ICD-11.

A closer breakdown of the 3,618 ORPHAcodes terminology concepts that are classifiable in the ICD-11 MMS but do not have their own unique identification in it shows that in most cases the ICD-11 Foundation terminology layer provides the unique identification (3,242 cases).

Conclusions

Compared to ICD-10, ICD-11 identifies almost all rare disease entities. Rare disease reporting through ICD-11 can standardize rare disease data (in terms of quality and completeness) and enable linkage to comprehensive clinical information on episodes of care, comorbidities, and other health status factors (signs, symptoms, phenotype, ICD-11 terminology features, such as extension codes). By achieving standardization, it can improve diagnosis, treatment, and quality of life of patients with rare diseases.

List of Abbreviations: Not applicable.

Author Contributions: MZ defined the research's aim drafted the idea of the article, coordinated the manuscript and derived the final version. ŠD, MV and CA discussed and helped to create the first version of the manuscript. RJ and MB reviewed and corrected the manuscript. All authors read and approved the final manuscript.).

Funding: This research received no funding.

Ethics Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

References

1. European Union. Regulation (EC) N°141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products. [online] 2000 [accessed on 16 June 2024]. Available from: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:018:0001:0005:EN:PDF
2. European Union. Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. [online] 2011 [accessed on 16 June 2024]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011L0024
3. COUNCIL RECOMMENDATION of 8 June 2009 on an action in the field of rare diseases. [online] 2009 [accessed on 16 June 2024]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:151:0007:0010:EN:PDF
4. Operational descriptions of rare diseases. Rare Diseases International. [online] [accessed on 16 June 2024]. Available from: URL: https://www.rarediseasesinternational.org/description-for-rd/
5. Rath A, Olry A, Dhombres F, Brandt MM. Representation of rare diseases in health information systems: The orphanet approach to serve a wide range of end users. Hum Mutat 2012;33:803-808. doi: 10.1002/humu.22078
6. Procedural document: Orphanet nomenclature and classification of rare diseases. [online] June 2023 [accessed on 16 June 2024]. Available from: https://www.orpha.net/pdfs/orphacom/cahiers/docs/GB/eproc_disease_inventory_R1_Nom_Dis_EP_05.pdf
7. ICD-11 Fact Sheet. World Health Organization. [online] 2020 [accessed on 16 June 2024]. Available from: https://icd.who.int/en/docs/icd11factsheet_en.pdf
8. Eleventh revision of the International Classification of Diseases, WHA72.15. [online] 28 May 2019 [accessed on 16 June 2024]. Available from: https://cdn.who.int/media/docs/default-source/classification/icd/icd11/a72_29-en_icd-11-adoption.pdf
9. Chute CG, Çelik C. Overview of ICD-11 architecture and structure. BMC Med Inform Decis Mak. 2021;Suppl 6:378. doi: 10.1186/s12911-021-01539-1.
10. Aymé S, Bellet B, Rath A. Rare diseases in ICD11: making rare diseases visible in health information systems through appropriate coding. Orphanet J Rare Dis. 2015;10:35. doi: 10.1186/s13023-015-0251-8.
11. Mazzucato M, Pozza LVD, Facchin P. ORPHAcodes use for the coding of rare diseases: comparison of the accuracy and cross country comparability. Orphanet J Rare Dis. 2023;18:267. doi: 10.1186/s13023-023-02864-6.

Applied Medical Informatics
Review Vol. 46, No. Suppl. 2 /2024, pp: S21-S24

Questionnaires for Evaluation the Nutritional Status of Cancer Patients

Claudia Raluca BALASA VIRZOB[1,*], Melania Camelia FIZEDEAN[1], and Corina VERNIC[2]

[1] Department of Clinic Nursing, University of Medicine and Pharmacy "Victor Babeş", Eftimie Murgu Street, No. 2, 300041, Timişoara, Romania
[2] Medical Informatics and Biostatistics, University of Medicine and Pharmacy "Victor Babeş", Eftimie Murgu Street, No. 2, 300041, Timişoara, Romania
E-mails: virzob.claudia@umft.ro; fizedean.camelia@umft.ro; cvernic@yahoo.com

* Author to whom correspondence should be addressed

Received: 27 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
The evaluation of the nutritional status of the cancer patient is necessary during the oncological follow-up. The use of assessment tools becomes reliable in the nutritional status of these patients, providing accurate information on the patient's situation. The aim of this study was to identify valid and reliable tools in the assessment of the nutritional status of cancer patients. A scope assessment was performed to search for medical articles in the last four years. A total of 552 articles were initially identified, which were examined, reaching a final selection of 13 studies. The most commonly used tool for nutritional assessment was the Patient-Generated Subjective Global Assessment. Questionnaires for evaluation the nutritional status vary from qualitative assessments to quantitative measurements, and can be subjective or objective.

Keywords: Cancer; Nutritional status; Nutritional assessment; Questionnaire

Introduction

Cancer represents the most important cause of death and morbidity in Europe after cardiovascular diseases, with more than 3.7 million new cases and 1.9 million deaths each year [1]. The nutritional assessment for these patients is important, because malnutrition in cancer patients has physical and psychologically consequences, and is associated with a worse prognosis of the disease and with decrease in health-related quality of life [2]. This study is related to digital epidemiology, and aims to identify valid and reliable tools in the assessment of the nutritional status of cancer patients.

Materials and Methods

This study was conducted using DECS-MeSH descriptors and Boolean operators. In addition, the PRISMA flow chart were used. The databases utilized was SCOPUS, Web of Science, Medline, and PubMed. The main descriptors used in the research were: "cancer", "neoplasia", "malignant tumors", "malignancies", "nutritional assessment", 'nutritional status", "questionnaire", "treatment", and "therapy".

Results

A total of 552 articles were identified in the selected databases, and the duplicates were eliminated. In order to

found out the valid and reliable methods for evaluation the nutritional status of cancer patients, 13 articles were included, following the PRISMA flow chart, which is shown in Figure 1.

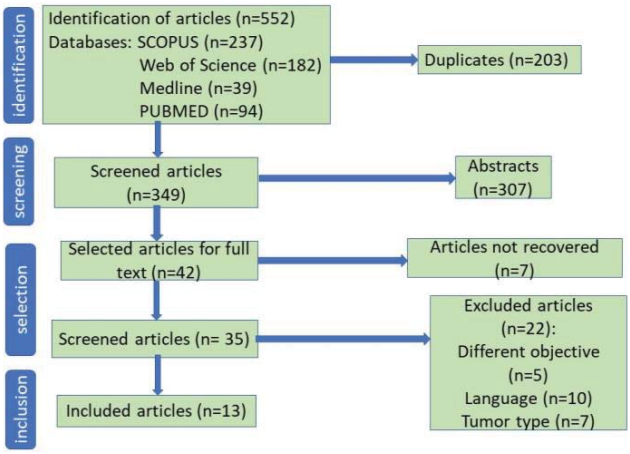


Figure 1. The PRISMA flow chart with the selected articles

It was found that the most frequent nutritional status assessment questionnaire for cancer patients was the Patient-Generated Subjective Global Assessment, followed by the Nutrition Risk Screening 2002 (Table 1).

Table 1. The main aspects of the articles included in the study

Authors	Year	Country	Type of study	No. Patients	Type of questionnaire
Kim et al. [3]	2019	Korea	Prospective	34	The Patient-Generated Subjective Global Assessment
Harada et al. [4]	2021	Japan	Retrospective	11	Short Nutritional Assessment Questionnaire
Badrasawi et al. [5]	2021	Palestine	Cross-sectional	100	The Subjective Global Assessment questionnaire
Hasegawa et al. [6]	2021	Tokio	Prospective cohort	38	The Subjective Global Assessment questionnaire
Sonneborn-Papakostopoulos et al. [7]	2021	Germany	Cross-sectional	109	The Mini Nutritional Assessment questionnaire
Beukers et al. [8]	2021	Netherlands	Systematic review	16	The Mini Nutritional Assessment, Nutritional Risk Index, and Patient-Generated Subjective Global Assessment
Tralubo et al. [9]	2022	Portugal	Observational	561	The Subjective Global Assessment questionnaire
Wang et al. [10]	2022	China	Cross-sectional	6685	The Nutrition Risk Screening 2002
Lee et al. [11]	2023	Korea	Observational	109	The dish-based semi-quantitative food frequency questionnaire
Vamvakari et al. [12]	2023	Greek	Observational	152	The Patient-Generated Subjective Global Assessment, Nutritional Risk Screening-2002, Simplified Nutritional Appetite Questionnaire, and Geriatric nutritional risk index
Boesenecker et al. [13]	2023	Germany	Observational	90	The anonymous questionnaire
Virtosu et al. [14]	2020	Republic of Moldova	Observational	100	The Romanian version for The Patient-Generated Subjective Global Assessment
Habina [15]	2022	Romania	Observational	46	Nutritional Risk Screening-2002

Discussion

The data showed that the most common nutritional status assessment questionnaire for cancer patients is the Patient-Generated Subjective Global Assessment. A Romanian language version was also identified for the Patient-Generated Subjective Global Assessment [14], and it was noted that in Romania, The Nutrition Risk Screening 2002 is the most used questionnaire [15].

Badrasaw et al. [5] and Kim et al. [3] conducted studies to observe the effect of the administration of nutritional supplements on the nutritional status of patients with pancreatic and biliary tract cancer undergoing oncological treatment [3,5].

Beukers et al. [8] in their systematic review, associated the results of this type of questionnaire with tolerance to systemic treatment, while Hasegawa et al. [6] used another version of this questionnaire, the Global Subjective Assessment, which is completed only by medical staff, focusing on weight deficit and physical assessment, and prioritizing these aspects over the symptoms manifested by the patients [6,8].

Sonneborn-Papakostopoulos et al. [7] used the Mini Nutritional Assessment questionnaire as a tool to evaluate the nutritional status of cancer patients, evaluating weight loss during defined periods of time and the presence of symptoms associated with the disease and treatment effect. This questionnaire does not require biochemical determinations or anthropometric parameters [7].

Conclusions

This study highlighted that questionnaires used in evaluation of the nutritional status vary from subjective assessments to objective measurements, emphasizing the need for an individualized approach. The questionnaires can improve the precision of nutritional interventions and also the life quality of the cancer patients.

List of Abbreviations: Not applicable.

Author Contributions: CRBV manuscript search, data collection, analysis, discussion, conclusions, and editing.

Funding: This research received no funding.

Ethics Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I would like to acknowledge Victor Babeş University of Medicine and Farmacy Timişoara.

Conflict of Interest: The authors declare no conflict of interest.

References

1. International Agency for Research on Cancer. Cancer Today [Internet] 2024. [accessed on 10 June 2024]. Available from: https://gco.iarc.fr/today/en].
2. García-Luna PP, Campos JP, Pereira Cunill JL. Causes and Clinical Impact of Malnutrition and Cachexia in Cancer Patients. Nutr Hosp. 2006; 21:10–16.
3. Kim SH, Lee SM, Jeung HC, Lee IJ, Park JS, Song M, Lee DK, Lee SM. The Effect of Nutrition Intervention with Oral Nutritional Supplements on Pancreatic and Bile Duct Cancer Patients Undergoing Chemotherapy. Nutrients 2019;11:1145. doi: 10.3390/nu11051145.
4. Harada K, Okagaki M, Neriya H, Yoshii K, Sekido K, Higashi A. Short Nutritional Assessment Questionnaire as a predictor of undernutrition in cancer patients receiving outpatient chemotherapy: A retrospective study. Eur J Oncol Nurs. 2021;54:102013. doi: 10.1016/j.ejon.2021.102013.

5. Badrasawi M, Al-Adhame A, Doufish A. Association of Malnutrition and Low Quality of Life among Cancer Patients Receiving Chemotherapy, Palestine. East Mediterr. Health J. 2021;27:459-466. doi: 10.26719/2021.27.5.459.
6. Hasegawa Y, Ijichi H, Saito K, Ishigaki K, Takami M, Sekine R, et al. Protein Intake after the Initiation of Chemotherapy Is an Independent Prognostic Factor for Overall Survival in Patients with Unresectable Pancreatic Cancer: A Prospective Cohort Study. Clin. Nutr. 2021;40:4792-4798. doi: 10.1016/j.clnu.2021.06.011.
7. Sonneborn-Papakostopoulos M, Dubois C, Mathies V, Heß M, Erickson N, Ernst T, Huebner J. Quality of Life, Symptoms and Dietary Habits in Oncology Outpatients with Malnutrition: A Cross-Sectional Study. Med. Oncol. 2021;38:20. doi: 10.1007/s12032-021-01460-7.
8. Beukers K, Voorn MJJ, Trepels R, van de Wouw Y, Vogelaar J, Havermans RC, et al. Associations between Outcome Variables of Nutritional Screening Methods and Systemic Treatment Tolerance in Patients with Colorectal Cancer: A Systematic Review. J. Geriatr. Oncol. 2022;13:1092-1102. doi: 10.1016/j.jgo.2022.06.010.
9. Trabulo C, Lopes J, da Silva Dias D, Gramaça J, Fernandes I, Gameiro R, et al.. Assessment of nutritional status of oncology patients at hospital admission: A Portuguese real-world study. Front Nutr. 2022;9:972525. doi: 10.3389/fnut.2022.972525.
10. Wang P, Tan Y, Soh KL, Soh KG, Ning C, Xue L, Lu Y, Yang J. Nutrition Risk Screening 2002 for Adult Cancer Patients: A Systematic Review and Meta-Analysis. Nutr Cancer. 2024;76(7):573-583. doi: 10.1080/01635581.2024.2352901.
11. Lee SA, Choi HK, Park SJ, Lee HJ. Development and Validation of a Food Frequency Questionnaire for Evaluating the Nutritional Status of Patients with Cancer. Nutrients. 2023;15(4):1009. doi: 10.3390/nu15041009.
12. Vamvakari K., Evangelou I, Kalafati IP, Kipouros M, Kosti RI, Kasti AN, Androutsos O. Assessment of nutritional status using objective and subjective methods in Greek patients with cancer. memo-Magazine of European Medical Oncology. 2023;16(3):200-207. doi: 10.1007/s12254-023-00877-2.
13. Boesenecker SJ, Mathies V, Buentzel J, Huebner J. Nutrition and physical activity in cancer patients: a survey on their information sources. J Cancer Res Clin Oncol. 2023;149(7):3823-3833. doi: 10.1007/s00432-022-04282-w.
14. Virtosu A, Harea D, Harea G, Morleano E. Evaluarea nutriţională la pacientul oncologic. Buletinul Academiei de Ştiinţe a Moldovei. Ştiinţe Medicale 2020;4(68):312-320.
15. Habina D. Assessing the risk of malnutrition in cancer patients during radiotherapy Medicine. Environmental Science. Dietetician.ro 2023. Available from: https://doi.org/10.26416/diet.4.4.2023.9094.

Applied Medical Informatics
Original Research Vol. 46, No. Suppl. 2 /2024, pp: S25-S28

Maximizing Research on Long COVID using FHIR and OMOP

Eugenia RINALDI[1,*], Lorenzo CANZIANI[2], Salvatore CAUTADELLA[3], Chiara DELLACASA[3], Anna GORSKA[2], Juan Mata NARANJO[3], Thomas OSMO[4], Miroslav PUSKARIC[5], Elisa ROSSI[3], and Sylvia THUN[1]

[1] Berlin Institute of Health at Charité-Universitaetsmedizin Berlin, Luisenstr. 65 10117 Berlin, Germany
[2] University of Verona, Via S. Francesco, 22, Verona, Italy
[3] Cineca Consorzio Interuniversitario, Bologna, Via Magnanelli, 6/3, 40033 Casalecchio di Reno BO, Italy
[4] Centre Informatique National de l'Enseignement Supérieur, 950 Rue de St - Priest, 34000 Montpellier, France
[5] High-Performance Computing Center, Nobelstraße 19, 70569 Stuttgart, Germany
E-mails: eugenia.rinaldi@bih-charite.de (*); lorenzomaria.canziani@univr.it; s.cautadella@cineca.it; c.dellacasa@cineca.it; anna.gorska@univr.it; j.naranjo@cineca.it; osmo@cines.fr; miroslav.puskaric@hlrs.de; e.rossi@cineca.it; s.thun@bih-charite.de

* Author to whom correspondence should be addressed

Received: 28 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
In 2020 The European commission funded the ORCHESTRA project with the aim to join the efforts of several European research centers in the research around the COVID-19 disease. One of the main challenges was to harmonize data across the different cohorts and countries. The introduction of standard terminologies such as SNOMED CT or LOINC helped establish a common language within the project. Over 3500 variables from several information categories were mapped to international codes from standard terminologies. After four years since the start of the pandemic, the study of long COVID seems to be of particular relevance due to the long-term effects that some people keep experiencing even after the infection has disappeared. To facilitate this research, we selected the ORCHESTRA variables that concerned long COVID and mapped them to the standards FHIR and OMOP to possibly support further data exchange with other research organizations.

Keywords: FHIR; OMOP CDM; Long COVID; Standard; Interoperability

Introduction

The project ORCHESTRA started in 2020 in the emergency of the COVID-19 pandemic. It reunites partners from several European and non-European countries with the purpose of developing new knowledge on the novel SARS-CoV-2 virus. One of the main objectives of the project was the harmonization of data across different protocols to be able to perform analysis across the different cohorts. A dedicated group within the project worked on semantic interoperability and mapped COVID-related variables to international codes stemming from standard terminologies [1,2] such as SNOMED CT, LOINC, NCIt, ICD and ATC. Four years after the start of the pandemics, international research is focusing on the long-term effects of the COVID-19 disease, the so called long COVID [3]. To facilitate research, we selected the ORCHESTRA variables relevant for long COVID and mapped them to standards such as FHIR [4] and OMOP [5] which are commonly used for data exchange.

Materials and Methods

We started from the data dictionary of the ORCHESTRA project [6], where over 3500 variables and their answer options have been mapped to international standard terminologies. The data dictionary is a file where the variables used for data collection are listed together with their answer options, and additionally, the information category assigned by the research group to the variable can be retrieved. Starting from this file, we selected the variables from the information categories "long COVID", "new medical events", "COVID severity". Additionally, we also extracted variables relating to the category "demographics" which included information on comorbidities. We removed cohort-specific information (e.g. "record Id") as well as doubles, and started mapping the selected variables to FHIR.

For some of the variables, such as for those relating to problems that persist in time after the infection, mapping was possible to both the FHIR resources "Observation" and "Condition". In these cases, we opted for the resource "Observation".

To proceed with the second step, and map the variables also to the OMOP standard, we utilized the online instrument ATHENA offered by OHDSI which helps identify the correct OMOP tables based on the terminology concepts. For this procedure, we utilized the mapping to the international standard terminology codes available within the ORCHESTRA project.

Results

The first result was the selection of the variables that could be particularly relevant for long COVID and that should be considered for the mapping to FHIR and to OMOP. We obtained a total of 104 variables across 5 different information categories. The 104 elements were mapped to the appropriate FHIR resources which turned out to be Observation and Condition as shown in Table 1.

Table 1. Distribution of the variables in the mapping to the FHIR resources Condition and Observation

Information category	FHIT Condition	FHIR Observation	Total
Comorbidities	2	7	9
COVID-19 severity	0	6	6
Demographics	0	7	7
Long COVID	5	18	23
New medical event	46	13	59
Total	***53***	***51***	***104***

The 104 variables were further mapped to the OMOP tables. As shown in Table 2, the tables resulted to be Condition, Observation and Measurement.

Table 2. Distribution of the variables in the mapping to the OMOP tables Condition_occurence, Measurement and Observation

Information category	OMOP Condition_occurence	OMOP Measurement	OMOP Observation	Total
Comorbidities	5	1	3	9
COVID-19 severity	0	0	6	6
Demographics	0	0	7	7
Long COVID	16	0	7	23
New medical event	47	4	8	59
Total	***68***	***5***	***31***	***104***

Discussion

Most of the selected ORCHESTRA variables concern examinations, symptoms and diseases and were mapped to the FHIR resources "Observation" and "Condition". However, there are cases where both resources could have been used. For example, symptoms, which are usually temporary problems, are generally modelled with the FHIR resource "Observation". However, in the case of long COVID, symptoms persist in time and could be handled as long-lasting problems and therefore mapped to the FHIR resource "Condition". However, we decided to map them to "Observation" to facilitate comparison with initial symptoms. OMOP distinguishes between a systematic examination with a standardized procedure (Measurement) and a general evaluation of fact (Observation) and this is why scores and indexes were mapped to the OMOP table "Measurement". Interesting is the case of the Child- Pugh score where some items belong to the OMOP domain Observations and others to the Measurement domain due to the different types of evaluation involved. All symptoms are mapped to the domain "Condition" in OMOP and this results in a higher number of variables mapped to the corresponding table as shown in Table 2.

The selection of relevant ORCHESTRA variables is not exhaustive and could certainly be extended to include, for example, further information about comorbidities or the questionnaires about the psychological effects of the pandemic. This study represents only an initial attempt to show how data exchange for long COVID could be supported using interoperability standards.

Conclusions

The need to efficiently exchange research information at international level to advance research has been highlighted by the COVID-19 pandemic. The broad adoption of interoperability standards for terminology and format of data in research communities would strongly support this goal. In Europe, the European Health Data Space [7] is expected to use FHIR to smooth the flow of data from primary care to secondary use.

List of Abbreviations: Not applicable.

Author Contributions: E.R. drafted the first version and following revisions of the manuscript under ST's supervision. LC and AG contributed to the selection of variables. All other authors contributed additional content, edits, and references. All authors approved the final version.

Funding: This research was funded by European Union's Horizon 2020 research and innovation program under Grant agreement no. 101016167, 6 November 2020

Data Availability Statement: The list of variables and relative mapping can be found at https://cloud.orchestra-cohort.eu/s/j2Fe4AxpHLysdJt

Acknowledgments: We would like to thank the entire ORCHESTRA Project Group for their support.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Bodenreider O, Cornet R, Vreeman DJ. Recent Developments in Clinical Terminologies - SNOMED CT, LOINC, and RxNorm. Yearb Med Inform. 2018;27(1):129-139. doi: 10.1055/s-0038-1667077.
2. International S. GlobeNewswire News Room. 2020. SNOMED International Joins COVID-19 Interoperability Alliance to Support Access to Terminology and Health Data Resources. Available from: https://www.globenewswire.com/en/news-release/2020/05/22/2037552/0/en/SNOMED-International-Joins-COVID-19-Interoperability-Alliance-to-Support-Access-to-Terminology-and-Health-Data-Resources.html
3. Mahase E. Covid-19: What do we know about "long covid"? BMJ. 2020;370:m2815. doi: 10.1136/bmj.m2815.

4. Benson T, Grieve G. FHIR Resources. In: Benson T, Grieve G, editors. Principles of Health Interoperability: SNOMED CT, HL7 and FHIR [Internet]. Cham: Springer International Publishing; 2016. p. 361–79. (Health Information Technology Standards). Available from: https://doi.org/10.1007/978-3-319-30370-3_20
5. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The Usage of OHDSI OMOP - A Scoping Review. Stud Health Technol Inform. 2021;283:95-103. doi: 10.3233/SHTI210546.
6. Rinaldi E, Stellmach C, Rajkumar NMR, Caroccia N, Dellacasa C, Giannella M, et al. Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic. NPJ Digit Med. 2022;5(1):75. doi: 10.1038/s41746-022-00620-x.
7. Marcus JS, Martens B, Carugati C, Bucher A, Godlovitch I. The European Health Data Space (December 12, 2022). IPOL | Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament Policy Department studies [Internet]. 2022 [cited 2024 Feb 16]. Available from: https://www.ssrn.com/abstract=4300393

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2/2024, pp: S29-S32

SERO Suicide Prevention App: A Preliminary Study of User Experiences in Real World

Kerstin DENECKE[1,*], François VON KAENEL[1], Caroline GURTNER[2], and Michael DURRER[3]

[1] Bern University of Applied Sciences, Department Engineering and Computer Science, Institute Patient-centered Digital Health, Quellgasse 21, 2502 Biel, Switzerland
[2] Bern University of Applied Sciences, School of Health Professions, Murtenstrasse 10, 3008 Bern, Switzerland
[3] Lucerne Psychiatry, Areal Kantonsspital 11, 6000 Luzern 16, Switzerland
E-mails: kerstin.denecke@bfh.ch; francois.vonkaenel@bfh.ch; caroline.gurtner@bfh.ch; michael.durrer@lups.ch

* Author to whom correspondence should be addressed

Received: 24 May 2024/Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
The study's objective is to explore users' experiences with the SERO app, a suicide prevention app. Its main functionalities include a self-assessment tool, safety plan, information and contact numbers that can be personalized. To collect the experiences, all registered users were contacted by E-mail and asked to fill a questionnaire with 3 open-ended questions and 2 questions with predefined answers. 74 persons answered the questionnaire. Generally, they confirmed that the app is easy to use and particularly the safety plan and self-assessment tool were very much appreciated. The responses were valuable to learn about the challenges of users in benefiting from the app usage, which is providing sufficient guidance for developing safety strategies.

Keywords: Usability; Suicide prevention; Testing; Digital health intervention

Introduction

There has been a growing interest in using digital tools to provide timely and accessible support in self-management to individuals at risk of suicide [1]. Research has shown the effectiveness of these digital interventions, particularly in regions with limited access to healthcare [2]. However, user non-adherence remains a challenge, often leading to high dropout rates [3] . A valuable source of information are observation studies to collect feedback from users that use the solution in real world without controlled settings. The focus of this study is to explore the real-world user experience of the SERO app [4], a digital intervention designed specifically for suicide prevention.

The SERO app is designed to support 1) people with suicidal behavior and 2) their families. People with suicidal behavior can use the app for self-assessment using the PRISM-S method [5], help-seeking and safety planning. They can access their personal safety plan and share it with relatives. Relatives can access the safety plan of an at-risk person who has chosen to share this information with them. They receive self-reflection support from the app (resource plan), as well as information on how to support the person seeking help. When the reminder feature is enabled, the application sends items retrieved from the safety plan or resource plan as pop-up messages to the user's phone. Data collected by the app is stored on MIDATA, a data storage ecosystem. 1597 users are registered by March 9, 2024 with 123 registered as relative. By understanding perceptions of usefulness and challenges in interacting with the app from a user's perspective, we aim to inform future iterations of digital interventions and ultimately improve support for those at risk of suicide and their relatives.

Materials and Methods

The link to the self-customized online survey was sent to all registered users on 5 February 2024. The consent management implemented in MIDATA allows to send emails to registered users while maintaining their anonymity. The survey was open to be filled for 4 weeks. We kept the questionnaire as short as possible and easy to answer. The survey contained 4 questions. None of them was mandatory: 1) In which role did you used the app? (4 options), 2) Which features of the app do you find useful? (open-ended question), 3) Which difficulties in usage do you had? (open-ended question), 4) Which suggestions for improvement do you have? (open-ended question). Additionally, we determined the Net Promoter Score (NPS) which is a customer loyalty metric that ranges from -100 to 100. NPS is calculated based on responses to the question: "On a scale of 0 to 10, how likely are you to recommend the SERO app to a friend or colleague?" Based on their scores, respondents are categorized into three groups: promoters, passives and detractors.

Results

The survey was answered by 74 persons. They answered the 5 questions in average in 5:16 minutes. 62.2% (n=46) claimed of having used the app in the role of a person with suicidality; 21.6% (n=16) used it in the role of a "relative"; 5.4% (n=4) selected both options (individual with suicidality and relative). 10.8% (n=8) claimed they do not know in which role they used the app. The NPS was 28 resulting from 40 promotors, 15 passives and 19 detractors.

Useful Features

We received 73 free text responses about the usefulness of the SERO app features. Some of these were explanations without any indication of useful features. The safety plan was mentioned most often as a useful feature with 37 mentions (50.7%). Self-assessment was the second most frequently mentioned feature (n=24, 32.9%). The availability of emergency numbers (n=6, 8.3%) and contact numbers for friends and family (n=2, 2.7%) in the app were less frequently mentioned. The resource plan was considered useful by 16 people (21.9%). Push notifications, information and sharing the safety plan were each mentioned by 1 participant (1.3%). 10 participants (13.7%) stated that everything in the app was useful. For example, one statement was "*I think all areas are very important. Sometimes you lose track of things or you can no longer access your resources*". (translated from German). 3 people (4.1%) stated that they had not used the app but had recommended it to health professionals (n=1) or patients (n=2, 2.7%). 1 person (1.3%) stated that he/she had been recommended the app but had not yet used it.

Difficulties in Interacting with the SERO App

Most of the 65 participants who answered the question about difficulties confirmed that they had not experienced any difficulties (n=41, 63%). 2 people (3%) complained about having to log in repeatedly to the app. Sharing the safety plan was difficult for 3 people (4.6%). Changing the order of contacts, adding new contacts, completing the self-assessment and logging out to change roles were each reported by 1 person (1.5%). Difficulties in filling in the safety plan were mentioned by 2 people (3%). Some participants reported challenges with the structure of the tools, i.e. in formulating the items to be included in the safety plan and answering the follow-up questions associated with each self-assessment (PRISM-S). Feedback suggests that the questions related to the reflections on the positioning of the intervertebral discs are difficult to answer. One comment was *"When assessing the risk (prism method), the follow-up questions are very similar and difficult to answer."*. Regarding the safety plan, participants found it difficult to formulate helpful beliefs; the differences between "distraction strategies" and "coping strategies" to be added by the user were not obvious.

Possibilities for Improvement

Possible improvements or confirmations of satisfaction are provided by 49 participants and include adding a dedicated pastoral care number alongside the existing emergency contacts. A clear overview of functionalities should be available on the starting screen of the app. The safety plan could be better structured, e.g. by allowing

users to place resources or motivational elements around an avatar representing themselves, potentially helping to motivate them to persevere. Highlighting the key elements of the safety plan would make navigation easier. To increase accessibility, suggestions were made for offline use and access without registration, allowing users to look up information provided by the app without restrictions. More comprehensive information on coping strategies and skills, including suggestions for dealing with self-harm or extreme stress, coping with difficult emotions and situations, crisis management exercises and push notifications to check for early warning signs were requested. One participant suggested the inclusion of a diary feature to track and evaluate their day, to monitor anxiety, stress, health and sleep patterns.

Discussion

From our study, we can confirm that the safety plan and self-assessment are perceived useful features of a suicide prevention app, even when it is used without being embedded in a therapeutic context. The reported interaction challenges are rather related to the self-management capabilities of an individual at risk. We assume that the reflection on the self-assessment could become easier when it has been practiced together with a therapist several times before using the app. In future work, we will find out how to improve the app to support in this more appropriately. Based on the detailed feedback and experiences shared by the participants of our survey, several suggestions for future iterations of digital interventions for suicide prevention can be derived. Users should be able to access some of the application's resources and features offline, ensuring that support is available without an internet connection. The SERO app provides already some information offline, but users have to be logged in which requires an internet connection if not done before. To lower the barrier for first-time or occasional users seeking immediate support, certain features should be even accessible without the need to register.

Our study comes along with some limitations. 74 out of 1597 registered users participated in the survey which cannot considered to be representative, and generalizations are impossible. However, it was not our intention to get a representative sample, but to get impressions from users for improving the app. In contrast to other papers who report on usability tests with their users conducted in controlled settings as part of clinical trials, our survey provides feedback from individuals that really used the app in real world. We expect that the collected feedback is more valuable than feedback collected in controlled settings with selected participants. We do not know whether the participants are aware of all features available in the app (e.g. sharing, push notifications). It remains an open issue whether this digital tool can support in acute crisis situations.

Conclusions

In conclusion, the perceived usefulness of the SERO app, specifically the safety plan and self-assessment tool was evidenced by the positive perception. The study underscores the need of establishing an ongoing user feedback mechanism to facilitate continuous app refinement and ensure alignment with the dynamic landscape of user requirements and real-world application. Although regulatory constraints preclude the integration of certain enhancements suggested by the participants, the recommendations to enhance the utility of the app underscore the potential to increase user engagement and personalization in the app.

List of Abbreviations: PRISM-S – Pictorial Representation of Illness Self Measure – Suicide, NPS – Net Promotor Score.

Author Contributions: KD defined the research's aim and the experiment's design; carried out the experiments and analyzed the data. FVK, CG, MD interpreted the results. All authors contributed to the discussion. KD coordinated and MD helped to draft the manuscript. All authors read and approved the final manuscript.

Funding: This research was funded by Gesundheitsförderung Schweiz, grant number PGV03.033.

Ethics Statement: The development and study were approved by the ethics committee North-western and central Switzerland under project number 2022-00870.

Data Availability Statement: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Larsen ME, Nicholas J, Christensen H. A Systematic Assessment of Smartphone Tools for Suicide Prevention. PLoS One. 2016;11(4):e0152285. doi: 10.1371/journal.pone.0152285.
2. Torok M, Han J, Baker S, Werner-Seidler A, Wong I, Larsen ME, et al. Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health. 2020;2(1):e25-e36. doi: 10.1016/S2589-7500(19)30199-2.
3. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med. 2015;45(13):2717-26. doi: 10.1017/S0033291715000665.
4. Meier L, Gurtner C, Nuessli S, Miletic M, Bürkle T, Durrer M. SERO - A New Mobile App for Suicide Prevention. Stud Health Technol Inform. 2022;292:3-8. doi: 10.3233/SHTI220310.
5. Ring M, Harbauer G, Haas S, Schuetz C, Andreae A, Maercker A, et al. Validierung des Suizidalitätseinschätzungsinstrumentes PRISM-S (Pictoral Representation of Illness Self Measure - Suicidality) [Validity of the suicidality assessment instrument PRISM-S (Pictoral Representation of Illness Self Measure - Suicidality)]. Neuropsychiatr. 2014;28(4):192-7. German. doi: 10.1007/s40211-014-0123-9.

Applied Medical Informatics
Original Research Vol. 46, Suppl. 2 /2024, pp: S33-S36

Usability Evaluation of a Digital Teleradiology Case Study Using the System Usability Scale in the WFDT Project

Anna BOEHM[1] and Thomas LUX[2,*]

[1] Faculty of Health Care, Niederrhein University of Applied Sciences, Reinarzstraße 49, 47805 Krefeld, Germany
[2] Competence Center eHealth, Niederrhein University of Applied Sciences, Reinarzstraße 49, 47805 Krefeld, Germany
E-mails: anna.boehm@hs-niederrhein.de; thomas.lux@hs-niederrhein.de

* Author to whom correspondence should be addressed

Received: 21 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
Background: The global pandemic has significantly increased the use of e-learning. In the project "Knowledge modules and case studies in the digital transformation" (Wissensbausteine und Fallstudien in der digitalen Transformation – WFDT), digital learning content is created in H5P and integrated into the Learning Management System (LMS) to be made available as Open Source Material. The teleradiology content is presented through case studies oriented towards case-based learning. *Aim*: The objective of usability testing is to identify potential areas for improvement and ensure these are incorporated into subsequent development and quality assurance processes. *Methods*: The usability test is carried out with the SUS to obtain an overview of usability within 10 items, complemented by questions on font size and qualitative feedback on the presentation. In an asynchronous setting, students work on the case study, model a teleradiology process with software, and answer the SUS and additional questions. *Results*: A total of 20 students from the Health Care Management and Medical Informatics degree programme participated in the usability test. The case study usability was rated as M = 59.63. Eleven students rated the font size as exactly right. The tasks, examples, and clarity of the presentation were positively rated. Further criticism was made of the task elements and the design of the presentation. *Conclusions*: The content structure, integration of tasks, and examples will be retained for this and other case studies. The design of the case studies will be fundamentally transformed into an H5P-proprietary format.

Keywords: Healthcare Education; Case-based Learning; System Usability Scale; Health Care Management

Introduction

E-Learning is the utilisation of digital media in a variety of educational settings. In recent years, due to the coronavirus pandemic, there has been a marked increase in the use of this technology [1]. To transform learning content into a digital format that is both appealing and freely accessible, the "*Knowledge modules and case studies in the digital transformation*" (German title: *Wissensbausteine und Fallstudien in der digitalen Transformation* – WFDT) project is developing digital case studies and learning modules utilizing H5P within a learning management system (LMS).

The case studies are based on the case-based learning approach, an active learning method focused on solving cases or problems under the guidance of a lecturer [2, 3]. The WFDT project deals with a case study in teleradiology. All necessary data for process modelling is provided within the case study. The findings of the usability tests will inform the indication of potential improvements to facilitate the development and quality assurance processes.

Materials and Methods

The System Usability Scale (SUS) is used to measure usability and was selected because it is an established and standardised tool providing a good overview of a system's usability. The SUS is designed so that positive and negative items are presented alternately to reduce response bias. The ten items request information regarding utilisation, technical support, integration of functions, structure, and learnability of the system [4]. A five-point Likert scale is used, with responses ranging from 0 (*complete rejection*) to 4 (*complete agreement*). The total score is calculated using the formula: *SUS = ((X-5) + (25-Y)) * 2.5*. X is the sum of all odd-numbered items, Y is the sum of all even-numbered items. The total score (SUS value) ranges from 0 to 100. Scores are defined as follows: scores >80.3 indicate excellent usability, 68-80.3 good usability, 68 acceptable usability, 51-68 poor usability and scores <51 terrible usability [5, 6]. The survey was supplemented with additional questions about font size, degree programme, age of the students and the time in minutes needed to complete the presentation and process modelling. Qualitative criteria were requested via free text: *What did you particularly like? What did you not like at all?*

In a blended learning setting, the case study is completed by students using a laptop, tablet, or smartphone. Students are asked to download the modelling software used for the teleradiology process modelling. They will work on this and then complete the SUS and additional questions. All steps, except downloading the modelling software, can be implemented in the LMS and are available to students asynchronously.

Results

Of 91 students enrolled in the Health Care Management and Medical Informatics degree programmes within the Process Management LMS course, 23 students completed the H5P teleradiology presentation, and 20 students participated in the evaluation.

The overall usability score is $\boldsymbol{M_1} = 59.63$. The *figure 1(a)* illustrates the SUS by end devices. The analysis by categories (*see fiugre 1(b)*)demonstrates that the utilisation, integration of functions, and learnability of the system are in the neutral range with a median and mean value close to 2. The categories of technical support and lack of structure have a median value of 1, indicating a near-complete rejection of these categories.

The font size appears to be optimal for 11 students, while 6 students find the font size rather too small. The mean processing time for the presentation is $\boldsymbol{M_2} = 50$ *minutes*, with a $\boldsymbol{SD_1} = 30$ *minutes*. The mean processing time for process modelling using EPC is $\boldsymbol{M_3} = 56$ *minutes*, with a $\boldsymbol{SD_2} = 31$ *minutes*. Due to technical constraints, the process model can only be processed on a laptop. Sixteen students provided a free text response to the question of what they particularly liked about the presentation (**+** in Table 2), while nine students provided a free text response to the question of what they did not like about it (**-** in Table 2).

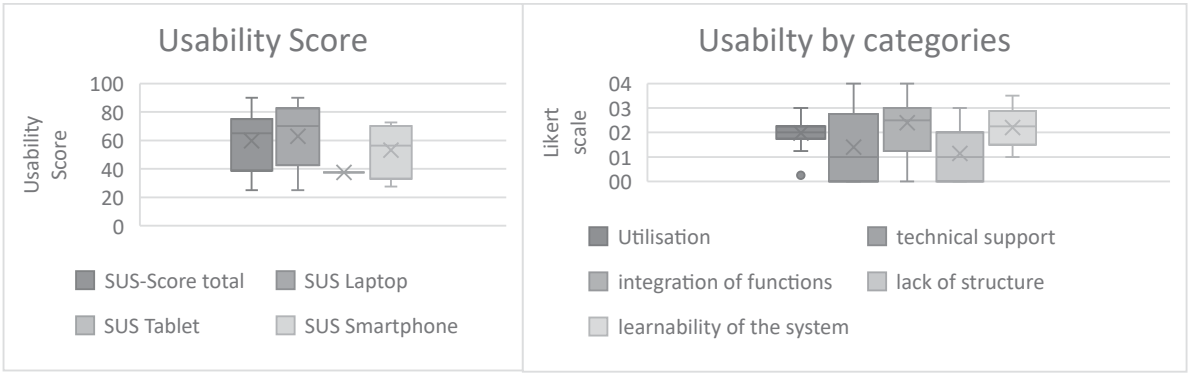


Figure 1. (a) SUS-Scores overall and end devices **(b)** Usability described in categories according

Table 2. Free text responses summarized in categories

categories	feedback	end devices
tasks	+ tasks in the presentation	6 laptop; 1 tablet
	+ examples	2 laptop; 1 smartphone
	- control	1 laptop; 2 smartphone
presentation	+ clear arrangement	3 laptop; 1 smartphone
	- design (font colour)	1 laptop

Discussion

The results of the usability test yielded an overall score of $M_1 = 59.63$, with a range of *25 to 90*. According to the SUS rating scale (see [5,6]), the overall is rated as poor usability. The free text responses indicate that the students found the clear arrangement, the tasks, and examples in the presentation satisfactory. This will therefore be adapted for this and other case studies, also emphasized by the almost complete rejection of the category structureless. The control of the task elements and the design of the presentation are perceived as negative aspects.

The results of the free text responses are used to further development and customisation. A fundamental change is the transition from a presentation-based format to an H5P-proprietary format. Another modification is the adaptation of the navigation elements for the drag-and-drop task type. Font size and colour will be implemented directly in H5P to support different scenarios. One limitation of the survey is the low response rate, with only 22% of enrolled students providing answers. Nevertheless, 86% of the students who completed the tasks are also rated the usability, so the results are valid and useful for further development. Moreover, the usability test is based on a single case study. As the project progresses, the usability of further content and changes to existing content will be re-evaluated. Another limitation is the asynchronous realisation. It would be beneficial to conduct further usability tests in a synchronous matter, as this would facilitate the prompt resolution of any queries.

Conclusions

Although the usability score is generally considered poor, the students' ratings of the case study in H5P are notably positive, particularly in the free text statements. The tasks and examples presented in the presentation are particularly noteworthy. Potential improvement is evident in the design of the presentation. These will be revised in the future to reassess usability and compare the results presented here with those of subsequent assessments. As a further research question, it should be critically examined whether the SUS method is suitable and which other methods could provide even more detailed findings.

List of Abbreviations: WFDT: Wissensbausteine und Fallstudien in der digitalen Transformation; LMS: Learning Management System; SUS: System Usability Scale; M: Mean; SD: Standard Deviation

Author Contributions: AB planned/performed usability test; TL research objective, coordinated, approved manuscript

Funding: This research was funded by Förderlinie "OERContent.nrw 2022": Produktion von digitalen Lehr-/Lernmaterialien für das Landesportal DH-NRW 214-76.01.08-37-2023-4344/31/05/2023 by the Ministry of Culture and Science of the State of North-Rhine Westphalia (Germany).

Data Availability Statement: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Kilian L. Handbuch E-Learning. Lehren und Lernen mit digitalen Medien. 5th ed. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG; 2018.

2. Das S, Das A, Rai P, Kumar N. Case-based learning: Modern teaching tool meant for present curriculum: A behavioral analysis from faculties' perspective. J Educ Health Promot. 2021;10:372. doi: 10.4103/jehp.jehp_1265_20.
3. Rezaee R, Haveshki F, Barati-Boldaji R, Mehrabi M. The effect of case-based e-learning on academic performance and problem-solving ability in nursing students: A pre- and post-test study. J Educ Health Promot. 2022; 11:302. doi: 10.4103/jehp.jehp_1613_21.
4. Brooke J. SUS: A 'quick and dirty' usability scale.1st ed. London: CRC Press; 1996. Chapter 21, SUS: A 'quick and dirty' usability scale; p. 189-194.
5. Will T. Measuring and Interpreting System Usability Scale (SUS). *UIUX* Trend 2017; [accessed on 13 June 2024]. Available from: https://uiuxtrend.com/measuring-system-usability-scale-sus/
6. Brooke J. SUS: A Retrospective. JUX - The Journal of User Experience 2014;8:29-40. [accessed on 13 June 2024] Available from: https://uxpajournal.org/sus-a-retrospective/

Applied Medical Informatics
Review Vol. 46, Suppl. 2 /2024, pp: S37-S40

A Digital Health Evaluation Framework

Elizabeth M. BORYCKI*, Amr FARGHALI, Amanda L. JOSEPH, Ryan KLETKE, Helen MONKMAN, Eleah STRINGER, and Andre W. KUSHNIRUK

School of Health Information Science, University of Victoria, Human and Social Development Building, Room 202, V8N 2G7, Victoria, Canada
E-mail (*): emb@uvic.ca

* Author to whom correspondence should be addressed

Received: 24 June 2024 /Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
Purpose: Digital health evaluation frameworks are needed to guide the development, implementation and evaluation of innovative, integrated systems of digital care. In this paper we describe a framework that was developed to evaluate the implementation of digital health technologies at a regional level. *Materials and Methods:* The framework was developed in a series of iterative phases beginning with a review of digital health frameworks used in Canada, which was followed by a scoping review focused on frameworks, models and theories used internationally to evaluate digital health technologies. Data extracted from articles were analyzed thematically to arrive at factors, concepts and measures that were incorporated in the final version of the framework following researcher discussions. *Results:* A range of themes and concepts emerged in the areas of: (1) organizational and context factors, (2) system, (3) use and process, and (4) outcomes. *Conclusions:* Several new themes and concepts were identified and incorporated into the new digital health evaluation framework.

Keywords: Digital Health Evaluation Framework; Health Informatics; Evaluation; Framework; Model; Theory

Introduction

A healthcare sector priority has been to connect primary, community, acute care and long-term care settings using digital health technologies. Over the past 50 years, considerable efforts have been made to digitize healthcare; for example, greater than 90% of physicians use electronic medical records (EMRs) in most OECD (i.e., Organization of Economic Cooperation and Development) countries. Yet, healthcare providers continue to record patient health information in their own EMRs without being able to digitally exchange this data with other healthcare organizations (e.g., regional health authorities, clinics, hospitals) [1]. This may lead to digital fragmentation of information, unnecessary interactions, and errors in the transmission of patient information across organizations. Researchers have identified a need to provide digital care in a meaningful way to prevent fragmentation of patient information [2, 3]. To address this issue, researchers and policy makers have identified a need for digital health evaluation frameworks to understand the effects of digital systems of care upon patients and health professionals [4-5]. In this paper we present a framework developed following a review of digital health frameworks in Canada [i.e., 6] and a scoping review. Our work aimed to create a framework for evaluating key aspects of digital health technologies and systems of care. The findings represent a contribution to the digital health and health informatics literature's as digital health frameworks have historically focused on individual health technologies, rather than digital health systems of care.

Materials and Methods

A scoping review was conducted using Arskey and O'Malley's approach [7]. The PubMed®, EBSCOhost®, CINAHL®, Web of Science®, and IEEE Xplore® databases were searched for the years 2010 to 2020 using the terms "evaluation" AND (framework or model or theory) AND "health information system". Identified articles were then uploaded to Covidence® [see 8], for title and abstract screening by three researchers using the criteria in Table 1.

Table 1. Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
• Written in English • Evaluated health information system(s) • A framework, model or theory used to develop, test, extend and/or evaluate a health information system(s)	• Not written in English • An opinion or review article • A medical device study • A description of a database that stores data • A study focused on a clinical problem

A final set of articles were identified and a full review of each paper took place. Identified articles and documents, which met the eligibility criteria were read, and relevant data were extracted (i.e., researcher name, date of publication, research questions, model/framework/theory and documented dimensions and concepts, subjects/participants, setting, methods, findings, and conclusions) [7]. Data were analyzed for themes and concepts to develop and extend the framework. This work involved researchers, reviewing the extracted scoping review data, including the themes and concepts and extending the framework through discussion [9].

Results

Three hundred sixty-three articles were identified; 78 duplicates were removed. The titles and abstracts of 285 articles were screened using the inclusion and exclusion criteria in Table 1. Two hundred fifteen items were excluded as they did not meet the criteria. Seventy articles were assessed for their eligibility with a full review of each article. Seventeen articles were excluded after a full review as they did not meet the criteria in Table 1. Fifty-three articles remained for data extraction. A thematic analysis of the data extractions was conducted. The following themes emerged: (1) organizational and contextual factors, (2) system, (3) use and process, and (4) outcomes. New concepts that emerged under the "system" theme were reliability, privacy and usability. Under the theme "use and process" new concepts included user characteristics, needs, and experience as well as workflow. Implementation emerged as an essential concept that included several important aspects of digital health implementation. Outcomes included new concepts such as quality of decision making, quality of care and others. Figure 1 shows the complete evaluation framework, including the set of factors, concepts, and measures that could be considered when evaluating digital health systems of care. We developed this framework through an iterative process. Initially, we drew on the previously accepted Canada Health Infoway's Benefits Evaluation Framework, that incorporates aspects of DeLone and McLean's work [10], and we also drew on Donabedian's system structure, process, and outcome quality model for healthcare [11]. We then integrated themes and concepts from our scoping review through researcher discussion (see Figure 1).

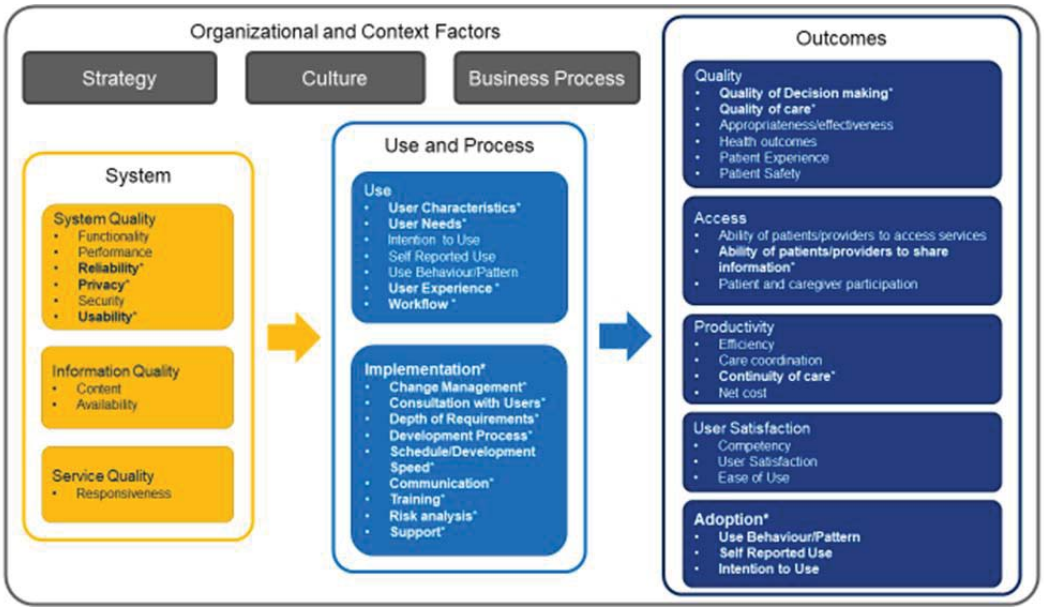


Figure 1. Digital health evaluation framework

Conclusions

The framework described in this paper was used to drive an evaluation of a regional digital health technology implementation. The framework offers an approach to studying digital health systems of care and guiding their evaluation. As technologies are integrated into digital health systems of care we need to consider past frameworks as well as new themes and concepts that have emerged and need to be integrated into developing evaluation frameworks. It should be noted that such frameworks can be designed to provide a superset of themes and concepts and could be considered as guidance for policy makers for inclusion in particular digital system of care evaluation projects. The development of evaluation protocols, such as questionnaires and interview prompts using the framework will vary depending on the themes and concepts selected.

List of Abbreviations: Not applicable.

Author Contributions: EB and AK worked on all aspects of the paper, developed the research protocol and oversaw the project. EB, AK, ALJ and ES worked on the scoping review portion. EB, AK, HM, AF and RK worked on the development of the framework, including its graphical depiction. EB and AK wrote the manuscript.

Funding: This research was funded by Mitacs. Grant number IT16333 with the support of Telus.

Ethics Statement: An ethics waiver was granted by the University of Victoria Ethics Board (publicly available data was used).

Data Availability Statement: The articles reviewed in the paper are publically available and the summary tables and references for all the articles identified in the scoping review can be obtained from the corresponding author.

Acknowledgments: We would like to acknowledge the administrative supports provided by Lois Holizki.

Conflict of Interest: The authors have no conflict of interest to report.

References

1. OECD. Health at a glance 2021: OECD indicators, OECD Publishing, Paris, 2021. [cited 03/11/2024] Available from: https://doi.org/10.1787/ae3016b9-en.

2. Senathirajah Y, Kaufman DR, Cato KD, Borycki EM, Fawcett JA, Kushniruk AW. Characterizing and visualizing display and task fragmentation in the electronic health record: Mixed methods design. JMIR Hum Factors. 2020;7(4):e18484. doi: 10.2196/18484.
3. Borycki EM, Farghali A, Kushniruk AW. Integrating human patterns of qualitative coding with machine learning: A pilot study involving technology-induced error incident reports. Stud Health Technol Inform. 2022;295:276-280. doi: 10.3233/SHTI220716.
4. Tran Ngoc C, Bigirimana N, Muneene D, Bataringaya JE, Barango P, Eskandar H, et al. Conclusions of the digital health hub of the Transform Africa Summit (2018): strong government leadership and public-private-partnerships are key prerequisites for sustainable scale up of digital health in Africa. BMC Proc. 2018;12(Suppl 11):17. doi: 10.1186/s12919-018-0156-3.
5. Joseph AL, Stringer E, Borycki EM, Kushniruk AW. Evaluative frameworks and models for health information Systems (HIS) and health information technologies (HIT). Stud Health Technol Inform. 2022;289:280-285. doi: 10.3233/SHTI210914.
6. Lau F, Hagens S, Muttitt S. A proposed benefits evaluation framework for health information systems in Canada. Healthc Q. 2007;10(1):112-116, 118.
7. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol. 2005;8:19-32. doi: 10.1080/1364557032000119616.
8. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available from: www.covidence.org.
9. Zhang Y, Wildemuth BM. Qualitative analysis of content. In: B. M. Wildemuth, Ed., Applications of Social Research Methods to Questions in Information and Library Science, Libraries Unlimited, 2009. pp. 1-12.
10. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst. 2003;19(4):9-30. doi: 10.1080/07421222.2003.11045748.
11. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743-8. doi: 10.1001/jama.260.12.1743.

Applied Medical Informatics
Review Vol. 46, Suppl. 2 /2024, pp: S41-S44

Challenges and Opportunities of Clinical Data Warehousing

Niklas GIESA[1,*], Anne Rike FLINT[1], Sedir MOHAMMED[1], Louis AGHA-MIR-SALIM[1], Sebastian Daniel BOIE[1], Fabian PRASSER[2], and Felix BALZER[1]

[1] Institute of Medical Informatics at Charité – Universitätsmedizin Berlin, Invalidenstr. 90, 10115 Berlin, Germany
[2] Medical Informatics, Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
Emails: niklas.giesa@charite.de; anne-rike.flint@charite.de; sedir.mohammed@charite.de; louis.agha-mir-salim@charite.de; sebastian-daniel.boie@charite.de; fabian.prasser@charite.de; felix.balzer@charite.de

* Author to whom correspondence should be addressed

Received: 27 June 2024 /Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
Clinical data warehousing (CDW) aims to provide an integrated database for secondary health analysis. We conducted a rapid review including 22 studies published between 2018 and 2022 that reported CWD implementation details in addition to individual experiences. Our results comprise technical details and reveal current opportunities and challenges of CDW. Many studies positively highlight standardized tools and models building on well-integrated clinical concepts. Enhanced tooling and data modeling were oftentimes hindered by bad data quality and organizational burdens. We call for enforcing synergies between clinical, technical, and organizational expertise to successfully roll out a CDW project.

Keywords: Clinical Data Warehousing; Electronic Health Records; Clinical Data Model; Rapid Review

Introduction

Clinical information systems (CIS) store high volumes of electronic health records (EHRs) [1]. To build a valuable integrated single point of data for secondary analysis, clinical data warehousing (CDW) facilitates extract, transform, and load (ETL) processes, harmonizing data from different source systems and file formats [2]. Traditional open-source projects, such as the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) or i2b2/tranSMART provide analytical tools in addition to database schemas [3]. OMOP CDM follows a normalized design principal suggested by Bill Imnon for analytical data stores while i2b2 is structured in de-normalized star-schemas inspired by Ralph Kimball [4].

Emerging big data infrastructures, like the Hadoop [5, 6], and advanced terminology systems, like SNOMED Clinical Terms [8], have been changing the CWD landscape over the past years. Besides relational data base management systems (RDBMS) document-oriented storages like NoSQL databases have gained popularity for exchanging data in interoperable formats like Fast Healthcare Interoperability Resources (FHIR) [6, 7].

We conducted a rapid review investigating technical implementations while highlighting resulting opportunities and challenges of CDW projects. Previous reviews have focused on specific clinical domains rather than analyzing these aspects of implementation projects [2, 8]. This study aimed to provide a general conspectus guiding an early conceptual phase before the concrete CWD rollout.

Methods

We searched PubMed for publications dealing with CDW implementations. Eligibility criteria for title and abstract were CDW related search terms, such as "clinical" in the combination of "data warehouse", "data repository", "data inventory", etc.[1], yielding a total of 136 publications. Exclusions were made for references with clinical trial search terms in their title, like "electronic data capture", "case report form", "clinical trial", etc.[2], reducing the number of references to 120. After applying inclusion- and exclusion criteria for published work between 2018 and 2022, 55 references remained for full-text screening. While screening we focused on technical reports that revealed details on the used CDM and the DBMS. Subsequently, we covered 22 publications in our rapid review that we conducted instead of a systematic review to reduce time and complexity while summarizing the most important results. We screened references during 2023 and see our rapid review as an initial exploration of the complex CDW research area.

Results

Table 1 displays applied review categories and the number of references (absolute frequencies) that fall into corresponding category items, e.g., Year Published. The underlying data table including all citations is provided in our GitHub repository [9]. 31% of the included studies (7 out of 22) were published in 2019 while 2021 and 2022 make up 18% (4 out of 22) each. The number of patients covered by the CDM ranged a lot between 100 and 70 million (M) while projects tend to either include little or vast number of records.

Table 1. Review categories for 22 references denoted as [absolute frequency] × [category item]. T: thousand, M: million, RDBMS: Relational Database Management System, OMOP: Observational Medical Outcomes Partnership

Year Published	Patient Volume	Clinical Data Model	Database Technology
5 × 2018	5 × [1M-70M]	6 × i2b2	16 × RDBMS
7 × 2019	4 × [100T-1M]	4 × OMOP	3 × RDBMS + NoSQL
2 × 2020	3 × [10T-100T]	2 × Dr. Warehouse	1 × Hadoop
4 × 2021	5 × [1T-10T]	1 × i2b2 + OMOP	1 × NoSQL
4 × 2022	5 × [100-1T]	8 × non-standard	1 × Hadoop + NoSQL

The most popular standard CDM design was i2b2 (6 out of 22) followed by OMOP (4 out of 22). Rinner et al. [3] combined these two approaches integrating standardized clinical vocabulary available on the data platform Athena. Garcelon et al. [10] introduced the project Dr. Warehouse providing ETL pipelines for the population of self-designed star-schemas. Plenty of CDW projects (8 out of 22) followed a non-standard CDM, five of them stored EHRs in de-normalized formats.

73% (16 out of 22) studies implemented a RDBMS that contributes to the seamless installation of i2b2 or OMOP. In contrast, Afshar et al. [5] and Artemova et al. [6] used novel infrastructure via implementing Hadoop solutions combining structured EHRs with unstructured data like images or texts. Only one project focused on document-based NoSQL data storages [11].

Authors positively highlighted vast data integration and harmonization capabilities of different modalities overcoming disjointed data silos [5, 6, 11, 12]. Curry et al. [11] presented the combination of traffic data and EHRs. The standardized vocabulary of OMOP was seen as a valuable base for the semantic interoperable data exchange between i2b2, FHIR, and additional export formats [3, 7, 12]. Projects also enhanced existing ETL pipelines by Notebooks for R or Python enabling data quality analysis and data exploration capabilities [6, 12]. Cossin et. al. built upon the i2b2 toolset developing a front-end for EHR annotations [13].

[1]Additional inclusion search terms: ("clinical") ("data") "store", "archive", "schema", "knowledge base", "lake", "base", "platform"
[2]Additional exclusion search terms: "clinical study", "digital documentation", "electronic scan", "electronic document", "randomized control study"

Data governance and data access patterns were identified as challenging when operating a CDW [12, 14, 15]. Walters et al. comprehensively described how they dealt with data requests in compliance with data privacy regulations [14]. Fleuren et al. explained their complex pseudonymization framework [16]. Although vendors provide powerful data integration tools, semantic integration was hampered by insufficient data quality [11, 12, 15, 16]. Record linkage, as the assignment of unique identifiers to each EHR across source systems, required a lot of data exploration [15, 16]. Complex toolsets were implemented at the expense of costly end-user training [13, 14].

Discussion

We found the usage of open-source CDMs as a major opportunity of CDW implementation projects. Although i2b2 seems to be favored by developers, semantic integration is especially enforced by the OMOP CDM [3]. De-normalized design principles, advocated by Ralph Kimball [4] in the field of data warehousing research, appeared to be suitable for the central integrated storage of EHRs. Surprisingly, most of the CDW were implemented on the basis of RDBMS, a few authors describe integration aspects by novel data systems like NoSQL or Hadoop. The document-based format FHIR [7] was seen as an export format instead of a suitable CDW structure. Data governance and privacy require synchronized organizational processes with data engineering tasks [14]. We also found that semantic interoperability cannot be solved with a standard data format but demands a certain level of data quality.

The results of this work are limited to a small number of references that were available between 2018 and 2022. We aimed to provide a first basis for a more complex literature analysis. This study would have benefited from a more comprehensive systematic review sharpening our claims and arguments. In the future, we plan to analyze current CDW studies with respect to interoperability aspects of i2b2 and OMOP in the combination with FHIR.

Conclusion

Results drawn from our rapid review call for deep synergies between technical, clinical, and organizational expertise addressing data quality and governance issues while profiting from advanced open-source CDW tools.

List of Abbreviations: EHR: Electronic Health Record, CDW: clinical data warehouse, CDM: Clinical Data Model, FHIR: Fast Healthcare Interoperability Resources, RDBMS: Relational Data Base Management Systems, OMOP: Observational Medical Outcomes Partnership, HDFS: Hadoop Distributed File System, DBMS: Data Base Management System, T: thousand, M: million

Author Contributions: NG, AF, SM extracted items, screened articles. NG wrote the manuscript, LA, SB, FB proofread.

Funding: This research received no funding.

Ethics Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: None.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Coorevits P, Sundgren M, Klein GO, Bahr A, Claerhout B, Daniel C, Dugas M, Dupont D, Schmidt A, Singleton P, De Moor G, Kalra D. Electronic health records: new opportunities for clinical research. J Intern Med. 2013;274(6):547-60. doi: 10.1111/joim.12119.

2. Gagalova KK, Leon Elizalde MA, Portales-Casamar E, Görges M. What You Need to Know Before Implementing a Clinical Research Data Warehouse: Comparative Review of Integrated Data Repositories in Health Care Institutions. JMIR Form Res. 2020 Aug 27;4(8):e17687. doi: 10.2196/17687.
3. Rinner C, Gezgin D, Wendl C, Gall W. A Clinical Data Warehouse Based on OMOP and i2b2 for Austrian Health Claims Data. Stud Health Technol Inform. 2018;248:94-99.
4. Yessad L, Labiod A. Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault. 2016 International Conference on System Reliability and Science (ICSRS), Paris, France, 2016, pp. 95-99, doi: 10.1109/ICSRS.2016.7815845.
5. Afshar M, Dligach D, Sharma B, Cai X, Boyda J, Birch S, et al. Development and application of a high throughput natural language processing architecture to convert all clinical documents in a clinical data warehouse into standardized medical vocabularies. J Am Med Inform Assoc. 2019;26(11):1364-1369. doi: 10.1093/jamia/ocz068.
6. Artemova S, Madiot PE, Caporossi A; PREDIMED group; Mossuz P, Moreau-Gaudry A. PREDIMED: Clinical Data Warehouse of Grenoble Alpes University Hospital. Stud Health Technol Inform. 2019 Aug 21;264:1421-1422. doi: 10.3233/SHTI190464.
7. Solbrig HR, Hong N, Murphy SN, Jiang G. Automated Population of an i2b2 Clinical Data Warehouse using FHIR. AMIA Annu Symp Proc. 2018 Dec 5;2018:979-988.
8. Hamoud A, Hashim AS, Awadh WA. Clinical data warehouse: a review. Iraqi Journal for Computers and Informatics 2018;44(2):16-26. doi: 10.25195/2017/4424.
9. Niklas G. Challanges and Opportunities of Clinical Data Warehousing. Available from: https://github.com/ngiesa/cdw_review (accessed 20/06/2024)
10. Garcelon N, Neuraz A, Salomon R, Faour H, Benoit V, Delapalme A, et al. A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse. J Biomed Inform. 2018;80:52-63. doi: 10.1016/j.jbi.2018.02.019.
11. Curry AE, Pfeiffer MR, Metzger KB, Carey ME, Cook LJ. Development of the integrated New Jersey Safety and Health Outcomes (NJ-SHO) data warehouse: catalysing advancements in injury prevention research. Inj Prev. 2021;27(5):472-478. doi: 10.1136/injuryprev-2020-044101.
12. Fruchart M, Quindroit P, Patel H, Beuscart JB, Calafiore M, Lamer A. Implementation of a Data Warehouse in Primary Care: First Analyses with Elderly Patients. Stud Health Technol Inform. 2022;294:505-509. doi: 10.3233/SHTI220510.
13. Cossin S, Lebrun L, Aymeric N, Mougin F, Lambert M, Diallo G, et al. SmartCRF: A Prototype to Visualize, Search and Annotate an Electronic Health Record from an i2b2 Clinical Data Warehouse. Stud Health Technol Inform. 2019;264:1445-1446. doi: 10.3233/SHTI190476.
14. Walters KM, Jojic A, Pfaff ER, Rape M, Spencer DC, Shaheen NJ, et al. Supporting research, protecting data: one institution's approach to clinical data warehouse governance. J Am Med Inform Assoc. 2022;29(4):707-712. doi: 10.1093/jamia/ocab259.
15. Visweswaran S, McLay B, Cappella N, Morris M, Milnes JT, Reis SE, et al. An atomic approach to the design and implementation of a research data warehouse. J Am Med Inform Assoc. 2022;29(4):601-608. doi: 10.1093/jamia/ocab204.
16. Fleuren LM, Dam TA, Tonutti M, de Bruin DP, Lalisang RCA, Gommers D, et al. The Dutch Data Warehouse, a multicenter and full-admission electronic health records database for critically ill COVID-19 patients. Crit Care. 2021;25(1):304. doi: 10.1186/s13054-021-03733-z.

Leveraging Cloud Technology for Personalized Multiple Sclerosis Care: A Comprehensive Data Management and Visualization Approach

Jose Manuel PINILLOS RUBIO[1,*] and Minerva VIGUERA MORENO[2]

[1] Universidad Internacional de La Rioja, Escuela Superior de Ingeniería y Tecnología (ESIT), 26006, Logroño, Spain
[2] Programa de Doctorado en Ciencias Biomédicas y Salud Pública UNED-IMIENS, Universidad Nacional de Educación a Distancia (UNED), 28015 Madrid, Spain
E-mails: pinillosrubio@gmail.com; mviguera7@alumno.uned.es

* Author to whom correspondence should be addressed

Received: 28 June 2024 /Accepted: 11 September 2024/ Published online: 21 November 2024

Abstract
This project develops a cloud-based solution for securely managing clinical data and patient-reported outcomes (PROMs) for multiple sclerosis (MS) patients. Utilizing REDCap for data collection, we incorporated clinical outcomes and PROMs from 300 MS patients over 18 months, supporting a machine learning (ML) based clinical decision support system. Our cloud architecture, featuring segregated data handling and enhanced security protocols using AWS, ensures robust data integrity and confidentiality. Key improvements include streamlined data ETL processes and an interactive online-based dashboard that facilitates the visualization of clinical data and PROMs, crucial for effective clinical decision-making. Initial results indicate a successful implementation in enhancing data management, with implications for personalized and predictive medicine. This framework not only elevates clinical data handling efficiency but also integrates PROMs into clinical practice effectively.

Keywords: Multiple Sclerosis; Data Visualization; Cloud Computing; Patient Reported Outcome Measures; Decision Support Systems

Introduction

Multiple sclerosis (MS) is an autoimmune, neurodegenerative disease with unpredictable progression, which complicates clinical decision-making and significantly impacts patient quality of life [1]. Addressing these challenges, digital solutions for clinical data management, as well as apps designed to collect PROMs, are common. Nevertheless, technological solutions that allow for the integration of these data and their appropriate visualization by healthcare professionals are scarce [2]. In this context, our project aimed to develop a comprehensive cloud-based solution to manage a wide array of data types involved in MS care. The primary objective of this initiative was to establish a robust platform that ensures thorough data checking, guarantees security and anonymity, and lays the groundwork for future integration of Machine Learning (ML) and Artificial Intelligence (AI) technologies. This platform will not only facilitate the advanced analysis needed to drive forward personalized medicine but also support more nuanced and data-driven clinical decision-making processes. The secondary objective revolved around the development of an interactive online-based dashboard, designed to enhance the visualization of both clinical data and patient-reported outcomes (PROMs), making it an essential component in the daily clinical management of MS [3]. By providing real-time data interaction capabilities, the dashboard serves as a critical support system for clinicians, enabling the effective incorporation of PROMs into routine clinical practice and ensuring that treatment decisions are informed by up-to-date and comprehensive patient data.

Materials and Methods

Data Collection and Management

Data for this study were systematically collected using the REDCap platform, tailored to gather both clinical outcomes (CROs) and patient-reported outcomes (PROMs) for a cohort of 300 multiple sclerosis (MS) patients over an 18-month period. We structured the data collection into two separate databases to maintain the anonymity and integrity of the data. A clinical database was developed to record baseline demographic information as well as clinical history, diagnostic details, and ongoing assessments of disease progression during biannual clinical visits. Separately, we created PROMs Database where patients could independently submit their responses to validated PROMs questionnaires through their electronic devices [4].

Data Anonymization and Security

Each participant was assigned a unique identifier at the point of inclusion, ensuring all data remained anonymous. This identifier linked the clinical and PROMs data longitudinally (captured in different time-points from every subject), allowing for comprehensive analysis without compromising patient privacy.

Cloud Architecture and Data Processing

The cloud architecture was designed for high security, operational efficiency, and scalability, as shown in Figure 1, using AWS [5]:

- Front-end and Back-end Segregation: The public front-end is protected with access controls via AWS Cognito and is separated from the backend, which is hosted in a private VPC. The interconnection between them is made through an API Gateway that controls access to the system's private services.
- User Interface: Developed in React and hosted on AWS Amplify, the interface provides secure access through authentication with AWS Cognito.
- Data Processing: The AWS Lambda function, which manages the ETL process, is responsible for importing data from S3 buckets into the Aurora MySQL database, where data validation and transformation tasks are also performed to ensure their coherence and prepare them properly for analysis.
- Data Storage and Backup: Uses dual instances of Aurora MySQL, one active and one standby, to ensure data availability and robustness against failures.

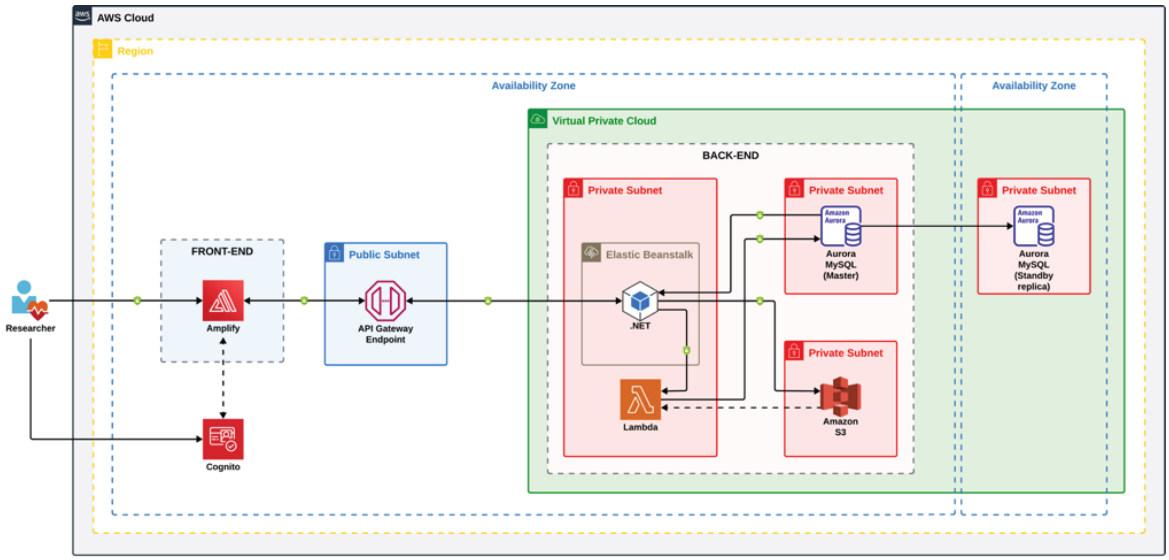


Figure 1. Schema of the cloud architecture developed for the implemented solution

Monitoring and Security Management

Continuous monitoring and advanced security measures, such as data encryption in transit and at rest and granular IAM policies, are implemented to protect data integrity and access. Tools like Amazon CloudWatch are used to oversee system performance and health.

Extraction, Transformation, and Loading Processes

Our ETL (Extraction, Transformation, and Loading) processes are designed to optimize the quality and structure of data for ML applications: during the Extraction and Preparation phase data from REDCap is extracted, checked for completeness, and validated. Transformation implies exporting data into a standardized CSV format, making it suitable for ingestion into a database. Finally, data is securely loaded into the cloud database using refined AWS Lambda processes to ensure integrity and alignment with our analytical needs.

Interactive Online-Based Dashboard

A key feature of our methodology is the development of an interactive dashboard hosted within the cloud infrastructure. This dashboard allows for real-time visualization and analysis of both clinical data and PROMs, enabling dynamic data interaction and clinical decision support by integrating data trends. Thus, we provide a powerful tool for clinical decision-making, facilitating the adoption of PROMs into routine practice.

Results

The implementation of our cloud solution has effectively managed comprehensive data collection and integration, ensuring high-quality data. Its scalable design supports robust data processing, enabling enhancements to our ML predictive models and facilitating ongoing research. The AWS-based platform is accessible to hospital staff via web authentication, allowing clinicians to view historical and updated patient information by simply knowing the assigned patient number. The interactive online dashboard has been especially impactful, offering real-time data visualization that aids clinicians and researchers in informed decision-making.

Discussion

Our solution represents a significant step forward in the management of MS, integrating clinical data with PROMs; other available solutions focus on clinical data or PROMs separately, which complicates the inclusion of patient perceptions in clinical practice. It supports rigorous data management with checks, security, and anonymity, enhancing the use of machine learning (ML) and artificial intelligence (AI) in clinical settings. The platform's interactive dashboard is vital for converting complex data into actionable insights, enabling real-time visualization of clinical data and PROMs. The major limitation of the platform is its design specific to a scientific study context, which will require modifications and adjustments for broader use.

Conclusions

The project has effectively met its objectives, establishing a secure and efficient cloud-based platform that enhances data management and lays the groundwork for future ML and AI applications in MS care. The interactive dashboard has proven useful in supporting clinical decision-making. As we advance toward a data-driven healthcare system, the technologies developed here are set to significantly improve patient management and care.

List of Abbreviations: MS - Multiple Sclerosis, ML - Machine Learning, AI - Artificial Intelligence, PROMs - Patient-Reported Outcome Measures, AWS - Amazon Web Services, ETL - Extraction, Transformation, and Loading, REDCap - Research Electronic Data Capture, VPC - Virtual Private Cloud, API - Application Programming Interface, IAM - Identity and Access Management.

Author Contributions: JMP designed and developed the solution. MV designed the data collection and study data capture.

Funding: This research received no funding.

Ethics Statement: This study was approved by ethical board of San Pedro Hospital (Logroño, Spain) in 2022.

Data Availability Statement: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Klineova S, Lublin FD. Clinical Course of Multiple Sclerosis. Cold Spring Harb Perspect Med. 2018;8(9):a028928. doi: 10.1101/cshperspect.a028928.
2. Matthews PM, Block VJ, Leocani L. E-health and multiple sclerosis. Curr Opin Neurol. 2020;33(3):271-276. doi: 10.1097/WCO.0000000000000823.
3. D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord. 2019;33:61-66. doi: 10.1016/j.msard.2019.05.019.
4. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al; REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95:103208. doi: 10.1016/j.jbi.2019.103208.
5. Fusaro VA, Patil P, Gafni E, Wall DP, Tonellato PJ. Biomedical cloud computing with Amazon Web Services. PLoS Comput Biol. 2011;7(8):e1002147. doi: 10.1371/journal.pcbi.1002147.

Applied Medical Informatics
Vol. 46, Suppl. 2 /2024, pp: S49-S52

The Case for Telemedicine from a Sustainability Perspective

Elisa SALA[1], Luisa BROGONZOLI[2], Marco DE BENEDICTIS[3], Maria Franca TOMASSI[4], Gabriella PAOLI[4], and Mauro GIACOMINI[5,*]

[1] Department of Political and Social Sciences, University of Pavia, Pavia, Italy
[2] Fondazione The Bridge, Milan, Italy
[3] Università Bocconi, Milan, Italy
[4] Azienda Ligure Sanitaria A.Li.Sa., Genoa, Italy
[5] Department of Informatics, Bioengineering, Robotics and System Engineering (DIBRIS), University of Genoa, Via all'Opera Pia 13, 16145, Genoa, Italy
E-mails: csfondazione@fondazionethebridge.it; centrostudi@fondazionethebridge.it; marco.debenedictis00@gmail.com; mft.tom@libero.it; gabriella.paoli@alisa.liguria.it; mauro.giacomini@unige.it

* Author to whom correspondence should be addressed

Received: 28 June 2024 /Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
This article examines telemedicine from a sustainability perspective, emphasising its alignment with the triple bottom line: social, environmental and economic benefits. Socially, telemedicine democratises healthcare by facilitating constant communication between patients and physicians, especially for the elderly, chronic patients and those with mobility difficulties. With regard to the environment, a life-cycle approach reveals the considerable potential of telemedicine to reduce greenhouse gas emissions by minimising patient travel. Furthermore, telemedicine also has the potential to decrease other environmental impacts, such as energy consumption in healthcare facilities and waste from personal protective equipment. From an economic perspective, telemedicine facilitates more efficient resource allocation, enhances system sustainability, and has thus the potential to reduce costs. It enables earlier and more optimal treatment of patients, improves adherence to treatment plans, and reduces hospitalisations and emergency room visits. The article examines the role of telemedicine in the transformation of the healthcare landscape, particularly in the context of Ministerial Decree 77 and the National Plan for Recovery and Resilience (PNRR), with a focus on the Liguria Region. The Liguria telemedicine plan, which was approved in 2023 and financed by the PNRR, has been designed to meet the health needs of an ageing population. The initial results indicate a growing adoption of telemedicine, particularly in the fields of endocrinology and diabetology. The discrepancies in data between sources signal the necessity for enhanced data integration.

Keywords: Telemedicine; Sustainability; Standard; Digital health

Introduction

Telemedicine presents advantages from a sustainability perspective, encompassing social, environmental and economic aspects.

Telemedicine has a social sustainability component, with the potential to equalise the healthcare system. Patients can utilise telemedicine to maintain regular contact with their general practitioners, as well as specialists, which is particularly beneficial for older [1] and chronic patients and those unable to travel. Telemedicine facilitates frequent monitoring [2] and enables patients to participate in remote clinical trials [3], to which they previously might have lacked the opportunity to take part in due to distance, thus decreasing the dropout rate. It serves as a tool for equal access to care and enables the delivery of more personalized and effective care for all patients.

It is challenging to draw a singular conclusion on the eco-advantages of telemedicine due to case-by-case variability and significant data gaps. Nevertheless, many studies proved that telemedicine has great potential to significantly reduce Greenhouse Gas emissions, primarily by eliminating patient commutes to hospitals [4, 5]. Additionally, telemedicine can lower other environmental impacts, such as energy demands for climate control in sanitary facilities and waste generation from personal protective equipment use during in person visits [6].

The concept of economic sustainability encompasses not only the reduction of costs [7, 8] but also the improved allocation of resources and the assurance of long-term system viability and endurance [9]. Early and optimal patient treatment and enhanced adherence are key elements in the optimisation of the economic impact of the system [10]. Digital health interventions can improve efficiency and convenience for patients and providers, reduce healthcare costs from decreased hospitalization and emergency room visits.

This paper will analyse the changes to the treatment landscape and the role of telemedicine, focusing on the analysis of the Liguria Region, presenting data on telemedicine usage in this area. In particular for Liguria, the Regional Health Authority (Alisa) initially implemented telemedicine with resolution 417 of November 2020.

Materials and Methods

The Ligurian population as of 1 January 2021 was just over 1.5 million residents.

The average age was 49.3 years and individuals over 65 years old were 28.7% of the total, compared to 23.5% at the Italian level. The Ligurian territory is characterized by a persistent gap between the coast and the hinterland with a prevalence of mountainous and hilly surfaces.

The continuous aging of the population - with the consequent greater incidence of chronic pathologies, such as heart diseases, chronic respiratory diseases, diabetes and tumors - combined with the conformation of the territory give rise to the necessity of using all possible tools for the management of complex medical care, with an accent on telemedicine.

The Liguria Region approved the Telemedicine Plan (2023) based on population health needs, taking into account already existing experiences - such as televisit and teleconsultation - and incorporating telemonitoring and teleassistance, which have yet to be fully developed. The plan will be developed in accordance with the National Telemedicine Platform, with funding provided by PNRR. It will benefit from the use of standardized technology and uniform measurement methods. Moreover, the sharing of health data is a crucial aspect of any telemedicine project, and will be facilitated by the Electronic Health Record project.

Results

Of the four possible telemedicine activities proposed and funded by the PNRR, televisits were chosen for evaluation. This choice depended on the fact that televisits were already formalised and employed during the COVID-19 period.

Televisits reporting can be considered a relatively novel activity, that will require additional time to become fully operational. For this reason, to ensure data consistency, two different information sources were cross-referenced. The first step was to corroborate the data from the telemedicine platform, according to which approximately 60 teleconsultations were conducted in 2020, about 1,250 in 2021, around 3,300 in 2022, and about 4,000 in 2023. In contrast, the Ligurian Data Warehouse (DWH), which provides data derived from Ambulatory Specialist Reporting, recorded no activity in 2020 and 2021 but reported nearly 6,700 televisits in 2022 and about 8,200 in 2023. The discrepancy between the two sources is mainly due to ASL3 Genovese, which reported a higher number of televists than what was recorded on the platform. Excluding ASL3 Genovese, the local healthcare authority that provided the most televisits is IRCCS Gaslini. However, both healthcare authorities primarily provide endocrinology and diabetology consultations remotely, as stipulated by the regional programming resolution (DGR 389/23). In 2023, there were almost 27 million specialist outpatient services for external patients, compared to nearly 25 million in 2022.

The following table shows the trend of televisits for the years 2023 and 2022. It is highlighted that the number of televisits is limited compared to the total number of follow up visits, but the prospect of reducing patients' travel and movement is of crucial relevance, as already highlighted by the data from the Gaslini Institute.

Table 1. Trends of televisits 2022-2023 in Liguria. In the analysis first visits, visits for rehabilitation plans, and emergency room visits were excluded, as they were considered inconsistent with the analysis

% of visits /total outpatient services		% of first visits and follow-ups / total visits		% of follow-up visits /total first visits and follow-up visits		% of video visits / follow-up visits		of which, out of region		Gaslini: of which, out of region	
2023	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**	**2022**
9.4%	9.5%	75%	75%	45.5%	45.3%	1.0%	0.8%	6.2%	4.0%	29.6%	26.3%

Discussion

A novel model of healthcare is emerging, centering on territorial medicine and the patient as the primary focus. One potential model is value-based healthcare [11, 12], which is founded on the pursuit of value, with a particular focus on the value derived for the patient. The model entails strategic and methodological processes that are also based on the incorporation of cutting-edge technology to enhance services and outcomes. The value for the patient represents the central axis around which services are organised and operated.

The growing utilisation of telemedicine, testified also by the Ligurian data, is a crucial step in the adoption of this model. However, another fundamental aspect of this process is data, as its measurement and evaluation facilitate the creation of more efficient care pathways, which in turn result in improved outcomes and reduced costs. It is therefore of the utmost importance to implement effective data collection mechanisms, such as the Electronic Health Record.

Conclusions

The implementation of telemedicine represents a viable model for the sustainable provision of healthcare services, whereby patient outcomes are balanced with cost efficiency and environmental responsibility. The effective collection and utilisation of data, as exemplified by the EHR, are of paramount importance to ensure the optimal impact of telemedicine. In a perspective of transition towards value-based care, strategic and technological advancements are fundamental. It is mandatory that telemedicine platforms are implemented to be able to bring together the many projects that have been financed in recent years, especially at a European level [13].

References

1. Bhatia R, Gilliam E, Aliberti G, Pinheiro A, Karamourtopoulos M, Davis RB, et al. Older adults' perspectives on primary care telemedicine during the COVID -19 pandemic. J Am Geriatr Soc. 2022;70(12):3480-3492. doi: 10.1111/jgs.18035.
2. Corbett JA, Opladen JM, Bisognano JD. Telemedicine can revolutionize the treatment of chronic disease. Int J Cardiol Hypertens. 2020;7:100051. doi: 10.1016/j.ijchy.2020.100051.
3. Crepaldi C, Venturi A, Brogonzoli L, Iardino MR. Decentralization of Clinical Trials: Opportunities, Risks and Development Paths. Stud Health Technol Inform. 2023;305:269-272. doi: 10.3233/SHTI230480.
4. Purohit A, Smith J, Hibble A. Does telemedicine reduce the carbon footprint of healthcare? A systematic review. Future Healthc J. 2021;8(1):e85-e91. doi: 10.7861/fhj.2020-0080.
5. Holmner A, Ebi KL, Lazuardi L, Nilsson M. Carbon footprint of telemedicine solutions--unexplored opportunity for reducing carbon emissions in the health sector. PLoS One. 2014;9(9):e105040. doi: 10.1371/journal.pone.0105040.

6. Thiel CL, Mehta N, Sejo CS, Qureshi L, Moyer M, Valentino V, Saleh J. Telemedicine and the environment: life cycle environmental emissions from in-person and virtual clinic visits. NPJ Digit Med. 2023;6(1):87. doi: 10.1038/s41746-023-00818-7.
7. AA.VV. La responsabilità sociale negli studi economico-aziendali. Milano: F. Angeli; 2007.
8. Olivotto L. Verso la trasformazione sostenibile dell'azienda. Manag Control 1 2022 [Internet]. 2022 [accessed on 21 August 2024]. Available from: http://digital.casalini.it/5213824
9. Airoldi G, Brunetti G, Coda V. Corso di economia aziendale. Bologna: Il mulino; 2010.
10. Nkyekyer J, Clifford SA, Mensah FK, Wang Y, Chiu L, Wake M. Maximizing Participant Engagement, Participation, and Retention in Cohort Studies Using Digital Methods: Rapid Review to Inform the Next Generation of Very Large Birth Cohorts. J Med Internet Res. 2021;23(5):e23499. doi: 10.2196/23499.
11. Porter PME. Redefining Health Care: Creating Value-Based Competition on Results. Harvard Business Press, 2006.
12. Porter ME. What is value in health care? N Engl J Med. 2010;363(26):2477-81. doi: 10.1056/NEJMp1011024.
13. Paleari L, Malini V, Paoli G, Scillieri S, Bighin C, Blobel B, Giacomini M. EU-Funded Telemedicine Projects - Assessment of, and Lessons Learned From, in the Light of the SARS-CoV-2 Pandemic. Front Med (Lausanne). 2022;9:849998. doi: 10.3389/fmed.2022.849998.

Applied Medical Informatics
Vol. 46, Suppl. 2 /2024, pp: S53-S56

Using ICT Tools for Exploring the Impact of Urban Blue-Green Spaces on Human Health and Well-Being

Sanja IVANKOVIC[1,*], Dragana JOVANOVIC[2], Predrag SAVIC[1], Vesna KARADŽIĆ[2], and Milena VASIĆ[2]

[1] Clinical Hospital Center "Dr Dragisa Misovic Dedinje", Ulica Sokobanjska 17, Belgrade, Centralna Srbija 11040, Serbia
[2] Institute for Public Health of Serbia, Dr. Subotića 5, Beograd, Grad Beograd 11000, Serbia
E-mail (*): projekti.sanja@gmail.com

* Author to whom correspondence should be addressed

Received: 28 June 2024 /Accepted: 11 September 2024 / Published online: 21 November 2024

Abstract
Worldwide, public health and well-being are significantly impacted by disruptions in physical, biological, and ecological systems stemming from urban environmental burdens. The H2020 project, "Healthier Cities through Blue-Green Regenerative Technologies: the HEART Approach," employs an innovative systemic methodology that integrates water (blue) and vegetated areas (green) infrastructure to address these challenges. The primary methodological approach of HEART clinical studies across three demonstration cities focuses on understanding how HEART project interventions relate to four chronic disease groups: cardiovascular, respiratory, mental, and metabolic conditions, with healthy participants as the control group. HEART study is using ICT tools as instruments in studying how blue-green spaces (BGS) impact public health and well-being in urban areas. 800 volunteers are equipped with smart-bands during their active time at demo sites in the cities of Belgrade, Athens and Aarhus. These wearable devices enable collection of biometric data, and project oriented mobile applications developed in collaboration between IT and medical experts guide participants on BGS visits, potentially monitor other health-related data and facilitate digital collection of socio-economic-medical questionnaires. Data collected from clinical measurements, laboratory tests, wearable devices, environmental sensors, and socio-economic-medical questionnaires will be analyzed using AI tools and statistical methods. These methods help to identify patterns and correlations between BGS visits and health outcomes. Overall, these ICT tools and methods are integral to the HEART project, enabling comprehensive data collection, real-time monitoring, and advanced analysis to evaluate the impact of BGS on public health and well-being in urban environments.

Keywords: Public health; Well-being; ICT tools; Chronic diseases; Blue-green spaces

Introduction

Globally, public health and well-being are affected by the disruptions of physical, biological, and ecological systems caused by the environmental burden in cities. The increase of non-communicable diseases morbidly and mortally has been recorded worldwide. The quantified negative impact of environmental pollution on health in the European Union resulted in 2021 with 253,000 deaths reported as attributed to exposure to fine particulate matter (PM2.5) above 5 $\mu g/m^3$, while 52,000 were attributed to nitrogen dioxide (NO_2) above 10 $\mu g/m$ [1]. Additionally, short-term exposure to ozone (O3) above 70 $\mu g/m^3$ was attributed to 22,000 deaths. Concerning specific causes of mortality, exposure to increased PM2.5 attributed at first place to deaths from ischemic heart disease, while NO2 harmed diabetes mellitus [2]. Oppositely, the quantified positive impact of spending more

than 30 minutes in blue-green spaces on human health in urban areas was less investigated and expressed in absolute numbers of people that directly benefited, but rather as lower odds ratios for a specific health outcomes such as depression or high blood pressure [3].

Public health and well-being are not standard urban planning criteria and are rarely integrated into that process, implicating the need for positioning citizens' health in the centre of urban planning and a good balance between social, environmental, and economic aspects [4]. Natural-based solutions (NBS) in urban areas should be designed in such a way as to enable healthy interaction between people and ecosystem services, which contributes to the improvement of public health and well-being. The relationship between urban planning and PH and WB enhancement is complex [5].

Based on the innovative systemic methodology of synergizing interactions of water (blue) and vegetated areas (green) infrastructure the H2020 project entitled Healthier Cities through Blue-Green Regenerative Technologies: the HEART Approach addresses these issues. Its principal aims are to significantly improve urban health and reduce health disparities by bringing nature to improve livability conditions and to alter the conventional approach to urban planning to focus on integrated nature-based solutions and concepts with emphasis on public health and citizens' well-being while considering sensitive societal and environmental aspects.

To achieve that the HEART Holistic Approach was built as a multidimensional framework consisting of four dimensions (individual health of study participants, blue-green solutions, environmental and community health) with the Health Centre Planning Matrix (HCPM tool) positioned in its centre.

Methodology

The main methodological approach to HEART clinical studies in all three demo cities was to understand the relationship of the HEART project interventions (recommendation) to four selected chronic diseases: cardiovascular, respiratory, mental and metabolic and healthy participants as the control group, assessing environmental (air quality, weather parameters, noise) distance to NBS, social, economic factors, as well as medical history of diseases and lifestyle factors. The HEART intervention/recommendation instructed participants in the following way: to perform 2-3 visits per week as a minimum, spend in the NBS/BGS at least 45 minutes to one hour and 15 minutes. Selected demo parks/locations with blue-green characteristics are in Belgrade (Serbia), Athens (Greece) and Aarhus (Denmark).

The goal is to engage 800 participants in all three counties until the end of the clinical study (December 2024). Recruited participants are randomly divided into study groups and their health will be monitored during the three-month study period, afterwards will be compared with those who suffer from the same diseases but without following instructions to visit the parks, just having their daily routine in the so-called grey areas (urban zones where they live and work).

Results

The ICT tools and methods play an important role in the HEART project by enabling comprehensive data collection, real-time monitoring, and advanced analytics to evaluate the impact of BGS on public health and well-being in urban settings.

ICT Tools Deployed

Mobile Applications: Medical experts were synergizing with IT experts to develop and test mobile applications tailored to fit the needs of each participant group. These applications are used to instruct participants on their visits to BGS, monitor health related data, track their emotional state, collect socio-economic-medical questionnaires digitally, and potentially monitor other health-related data.

Wearable Devices: Volunteers are equipped with 2 types of wearable devices (smart-bands), equipped with sensors to monitor various health metrics such as activity levels, heart rate, blood oxygen levels, stress levels, and

emotional states. Beside the clinical measurements in the hospital setting, these devices are crucial in assessing the impact of BGS on participants' health.

Hospital data base: Specially designed for the clinical study within the HEART projects, Hospital data base enables each Hospital involved in the study to keep the study medical records and provide connection with project developed applications.

Environmental Sensors: Environmental in-situ sensors are deployed in demo sites to gather real-time data on environmental parameters like air quality, weather conditions, and other factors relevant to the study. These sensors help in understanding the environmental context in which participants are exposed during their visits to BGS.

Satellite Images: Satellite images are used as a tool to potentially gather spatial data related to the BGS and their surroundings. This data will be utilized to analyze vegetation coverage, urban heat island effects, or other relevant environmental factors affecting public health.

Methods for Evaluation

AI Tools and Statistical Methods: Data collected from various sources including clinical measurements, laboratory tests, wearable devices, environmental sensors, and socio-economic-medical questionnaires are analyzed using AI tools and standardized statistical methods. These tools help in correlating different datasets to identify patterns, trends, and correlations between BGS visits and health outcomes.

Digital Data Collection: Apart from wearable devices, hospital database and mobile applications, digital data collection methods include self-reporting through the HEART applications and digital filling of socio-economic-medical questionnaires. This ensures efficient data collection and management throughout the study respecting data privacy.

Conclusion and Actions

Robust datasets that will be obtained within the clinical studies in all three demo locations will undergo statistical and advanced statistical analysis, thriving to confirm defined hypotheses and answer research questions. Concerning selected diseases, the HEART investigators are interested in qualifying benefits for the group of participants who visited BGSs according to the instructions in comparison to those who did not follow the recommendation, specifically focusing on the status of the respective disease (before and after blue-green intervention) and existence and type of the correlation between defined clinical and environmental indicators. In addition to that, holistic validation will include social and economic aspects to better characterize observed results and potential impact on public health and well-being. Furthermore, obtained results from the HEART studies should help decision-makers and urban planners in building new or renovating existing blue-green areas in towns and cities.

Deployment of the all ICT tools used to reach the study goals, the HEART consortium highlights importance of the strong bonds between technology, health and environment.

List of Abbreviations: BGS- blue-green spaces; NBS- natural-based solutions; HCPM- Health Centre Planning Matrix;

Funding: This research received no funding.

Data Availability Statement: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

References

1. European Environment Agency. European Union 8th Environment Action Programme. Monitoring report on progress towards the 8th EAP objectives - 2023 edition, 115 pp. Luxembourg: Publications Office of the European Union. doi: 10.2800/34224.

2. European Environment Agency. Briefing no. 23/2023, Harm to human health from air pollution in Europe: burden of disease 2023. EN HTML: TH-AM-23-026-EN-Q - ISBN: 978-92-9480-614-7 - ISSN: 2467-3196 - doi: 10.2800/721439.
3. Dumitru A, Wendling L. Evaluating the impact of nature-based solutions: A handbook for practitioners. European Commission EC. 2021. doi: 10.2777/244577.
4. Tefera Y, Soebarto V, Bishop C, Kandulu J, Williams C. A Scoping Review of Urban Planning Decision Support Tools and Processes That Account for the Health, Environment, and Economic Benefits of Trees and Greenspace. Int J Environ Res Public Health. 2023;21(1):48. doi: 10.3390/ijerph21010048.
5. Mouratidis K. Urban planning and quality of life: A review of pathways linking the built environment to subjective well-being. Cities 2021;115:103229. doi: 10.1016/j.cities.2021.103229.

AUTHOR INDEX

KEYWORD INDEX