Interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 13 (2024) ISSN 1929-073X

Contents

Original Papers

Predicting the Effectiveness of a Mindfulness Virtual Community Intervention for University Students: Machine Learning Model (e50982)
Christo El Morr, Farideh Tavangar, Farah Ahmad, Paul Ritvo, MVC Team. 4

Improvements in Neoplasm Classification in the International Classification of Diseases, Eleventh Revision: Systematic Comparative Study With the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision (e52296)
Yicong Xu, Jingya Zhou, Hongxia Li, Dong Cai, Huanbing Zhu, Shengdong Pan. 76

Evaluation of an mHealth App on Self-Management of Osteoporosis: Prospective Survey Study (e53995)
Magnus Bendtsen, Bodil Schönwandt, Mette Rubæk, Mette Hitz. 108

Intramural Health Care Through Video Consultations and the Need for Referrals and Hospital Admissions: Retrospective Quantitative Subanalysis of an Evaluation Study (e44906)
Katharina Schmalstieg-Bahr, Miriam Colombo, Roland Koch, Joachim Szecsenyi, Friedrich Völker, Eva Blozik, Martin Scherer. 119

Behavioral Insights from Vaccine Adoption in Nigeria: Cross-Sectional Survey Findings (e47817)
Sohail Agha, Ifeanyi Nsofor, Drew Bernard, Sarah Francis, Nandan Rao. 129

Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study (e43585)
Zhuoga Deji, Yuantao Tong, Honglian Huang, Zeyu Zhang, Meng Fang, M Crabbe, Xiaoyan Zhang, Ying Wang. 141

The Effect of Body Temperature Changes on the Course of Treatment in Patients With Pneumonia and Sepsis: Results of an Observational Study (e52590)
Domen Guzelj, Anže Grubelnik, Nina Greif, Petra Povalej Bržan, Jure Fluher, Žiga Kalamar, Andrej Markota. 156

Evaluation of the Accuracy, Credibility, and Readability of Statin-Related Websites: Cross-Sectional Study (e42849)
Eunice Ling, Domenico de Pieri, Evenne Loh, Karen Scott, Stephen Li, Heather Medbury. 179

Online Visibility and Scientific Relevance of Strabismus Research: Bibliometric Analysis (e50698)
Aleksander Stupnicki, Basil Suresh, Saurabh Jain. 191

XSL•FO
RenderX

Effectiveness of a Smartphone App to Promote Physical Activity Among Persons With Type 2 Diabetes: Randomized Controlled Trial (e53054)
Stephanie Bonn, Madeleine Hummel, Giulia Peveri, Helén Eke, Christina Alexandrou, Rino Bellocco, Marie Löf, Ylva Trolle Lagerros. 196

A Preliminary Checklist (METRICS) to Standardize the Design and Reporting of Studies on Generative Artificial Intelligence–Based Models in Health Care Education and Practice: Development Study Involving a Literature Review (e54704)
Malik Sallam, Muna Barakat, Mohammed Sallam. 210

Factors Associated With Worsened Mental Health of Health Care Workers in Canada During the COVID-19 Pandemic: Cross-Sectional Survey Study (e50064)
Ali AbdulHussein, Zahid Butt, Stanko Dimitrov, Brian Cozzarin. 233

A Web-Based Tool to Assess Social Inclusion and Support Care Planning in Mental Health Supported Accommodation: Development and Preliminary Test Study (e45987)
Sharon Eager, Helen Killaspy, Joanna C, Gillian Mezey, Peter McPherson, Megan Downey, Georgina Thompson, Brynmor Lloyd-Evans. . . .
2 4 4

Viewpoints

Designing mHealth Apps to Incorporate Evidence-Based Techniques for Prolonging User Engagement (e51974)
Rebecca Monachelli, Sharon Davis, Allison Barnard, Michelle Longmire, John Docherty, Ingrid Oakley-Girvan. 18

Using Routine Data to Improve Lesbian, Gay, Bisexual, and Transgender Health (e53311)
Catherine Saunders. 26

Ad Hoc Modifications to a High Dependency Psychiatric Unit for People With Dementia During the COVID-19 Period (e49618)
Thanita Pilunthanakul, Giles Tan. 35

Reviews

Strategies to Alleviate the Burden Experienced by Informal Caregivers of Persons With Severe Mental Disorders in Low- and Middle-Income Countries: Scoping Review (e48587)
Olindah Silaule, Daleen Casteleijn, Fasloen Adams, Nokuthula Nkosi. 39

Digital Methods for the Spiritual and Mental Health of Generation Z: Scoping Review (e48929)
Susanna Park, Bridgette Do, Jacqlyn Yourell, Janice Hermer, Jennifer Huberty. 64

Theoretical Perspectives Underpinning Research on the Physician-Patient Relationship in a Digital Health Practice: Scoping Review (e47280)
Damoun Nassehi, Birgitta Gripsrud, Ellen Ramvi. 97

Gamification and Oral Health in Children and Adolescents: Scoping Review (e35132)
Rui Moreira, Augusta Silveira, Teresa Sequeira, Nuno Durão, Jessica Lourenço, Inês Cascais, Rita Cabral, Tiago Taveira Gomes. 166

XSL•FO
RenderX

Original

Application of AI in Sepsis: Citation Network Analysis and Evidence Synthesis (e54490)
MeiJung Wu, Md Islam, Tahmina Poly, Ming-Chin Lin. 86

Corrigenda and Addenda

Correction: Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study (e59070)
Zhuoga Deji, Yuantao Tong, Honglian Huang, Zeyu Zhang, Meng Fang, M Crabbe, Xiaoyan Zhang, Ying Wang. 231

Case Reports

Treating Spontaneous Pneumothorax Using an Innovative Surgical Technique Called Capnodissection Pleurectomy: Case Report (e54497)
Ghaith Qsous, Prashanth Ramaraj, Sanjeet Avtaar Singh, Philip Herd, Nayandra Sooraj, Malcolm Will. 262

Radiological Progression of Degenerative Cervical Myelopathy in a Clinically Stable Patient: Case Report (e48212)
Rishi Umeria, Oliver Mowforth, Munashe Veremu, Benjamin Davies, Mark Kotter. 266

Research Letter

Visual “Scrollytelling”: Mapping Aquatic Selfie-Related Incidents in Australia (e53067)
Samuel Cornell, Amy Peden. 273

XSL•FO
RenderX

Original Paper

Predicting the Effectiveness of a Mindfulness Virtual Community Intervention for University Students: Machine Learning Model

Christo El Morr[1], PhD; Farideh Tavangar[2], MSc; Farah Ahmad[1], PhD; Paul Ritvo[3], PhD; MVC Team[4]

[1]School of Health Policy and Management, York University, Toronto, ON, Canada
[2]Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[3]Kinesiology & Health Science, York University, Toronto, ON, Canada
[4]See Acknowledgments

Corresponding Author:
Christo El Morr, PhD
School of Health Policy and Management
York University
4700 Keele Street
Toronto, ON, M3J 1P3
Canada
Phone: 1 426 736 2100 ext 22053
Email: elmorr@yorku.ca

Abstract

Background: Students' mental health crisis was recognized before the COVID-19 pandemic. Mindfulness virtual community (MVC), an 8-week web-based mindfulness and cognitive behavioral therapy program, has proven to be an effective web-based program to reduce symptoms of depression, anxiety, and stress. Predicting the success of MVC before a student enrolls in the program is essential to advise students accordingly.

Objective: The objectives of this study were to investigate (1) whether we can predict MVC's effectiveness using sociodemographic and self-reported features and (2) whether exposure to mindfulness videos is highly predictive of the intervention's success.

Methods: Machine learning models were developed to predict MVC's effectiveness, defined as success in reducing symptoms of depression, anxiety, and stress as measured using the Patient Health Questionnaire-9 (PHQ-9), the Beck Anxiety Inventory (BAI), and the Perceived Stress Scale (PSS), to at least the minimal clinically important difference. A data set representing a sample of undergraduate students (N=209) who took the MVC intervention between fall 2017 and fall 2018 was used for this secondary analysis. Random forest was used to measure the features' importance.

Results: Gradient boosting achieved the best performance both in terms of area under the curve (AUC) and accuracy for predicting PHQ-9 (AUC=0.85 and accuracy=0.83) and PSS (AUC=1 and accuracy=1), and random forest had the best performance for predicting BAI (AUC=0.93 and accuracy=0.93). Exposure to online mindfulness videos was the most important predictor for the intervention's effectiveness for PHQ-9, BAI, and PSS, followed by the number of working hours per week.

Conclusions: The performance of the models to predict MVC intervention effectiveness for depression, anxiety, and stress is high. These models might be helpful for professionals to advise students early enough on taking the intervention or choosing other alternatives. The students' exposure to online mindfulness videos is the most important predictor for the effectiveness of the MVC intervention.

Trial Registration: ISRCTN Registry ISRCTN12249616; https://www.isrctn.com/ISRCTN12249616

(Interact J Med Res 2024;13:e50982) doi:10.2196/50982

KEYWORDS

machine learning; virtual community; virtual care; mindfulness; depression; anxiety; stress; students; online; randomized controlled trial; Canada; virtual; artificial intelligence; symptoms; behavioral therapy; sociodemographic; mindfulness video; online video

XSL•FO
RenderX

Introduction

Students' mental health crises were recognized before the COVID-19 pandemic and deepened during the pandemic. University students are experiencing an increase in psychological distress on North American campuses. A student survey of 32 Canadian postsecondary institutions reported high anxiety (56.5%), hopelessness (54%), seriously depressed mood (37.5%), and overwhelming anger (42%) [1]. A similar survey in 2016 revealed higher distress levels [2]. In 2013, a study of 997 students at York University (site of this study) indicated that 57% reported depression scores sufficient for diagnosable clinical depression, while 33% reported anxiety scores in ranges typically indicative of panic disorder and generalized anxiety disorder [3]. The situation appears similar at universities in the United States [4,5] and worldwide; in 2018, the World Health Organization reported increasing mental disorders in college and university students worldwide [6]. Mental health challenges among university students demand attention. This is a vulnerable period, as 70% of mental health problems emerge before the age of 25 years. Without intervention, these problems can worsen and hinder students' personal and academic success [7]. COVID-19 has negatively impacted university students' mental health [8-10].

University student distress is both an individual and societal challenge. Losses in productivity during the study and at work due to distress and mental disorders are associated with indirect but significant economic burdens [11]. Canadian estimates show that mental disorders cost nearly US $37 billion yearly, with 9.8% due to direct medical costs, 16.6% and 18.2% due to long-term loss and short-term work loss, respectively, and 55.4% due to the loss of healthy function (ie, loss of the utilities of vision, hearing, speech, mobility, dexterity, emotion, cognition, and pain as assessed in the Health Utilities Index Mark 3 system) [12].

While mental distress and disorder are becoming more prevalent in students, the counseling offered in colleges and universities needs to catch up with demand. For example, from 2007 to 2012, full-time enrollment in the Ontario college system increased from 167,000 to 210,600 (a 26% increase), while the number of counselors employed in the college system increased from 146 to 152.7 (a 4.6% increase) [13]. This discrepancy leaves students underserved and counselors overwhelmed amid the increasing distress [14].

Mindfulness-based interventions have been demonstrated to positively impact psychological and physical health [15-17], with multiple meta-analyses demonstrating positive impacts in clinical and nonclinical populations [18-22]. However, with large numbers of students (50,000 to 60,000 on some campuses), there may not be enough trained personnel to convey helpful mindfulness-based practices directly. Instead, in the eHealth domain, virtual communities (VCs) [23], that is, online communities, have been used in health care to provide e-education tools and online support to empower active participants in health enhancement [24-26]. VCs can scale up mindfulness interventions at lower costs to a broader range of students, especially those restricted from attending clinics due to time-place discontinuities. VCs preserve anonymity (with reduced stigmatization) while promoting voluntary, supportive, interpersonal connections.

We developed a web-delivered mindfulness program (mindfulness virtual community [MVC]) to reduce symptoms of depression, anxiety, and stress in university students and conducted a randomized controlled trial (RCT) targeting university students at a Canadian university to examine its effectiveness. Following a successful RCT [26-29], we wanted in this secondary analysis (1) to develop a machine learning (ML) model to predict the effectiveness of the online mindfulness intervention on mental health outcomes using sociodemographic and self-reported features and (2) to investigate if exposure to mindfulness videos was highly predictive of the intervention's success.

Methods

Prediction Problem

This study aims to predict the effectiveness (ie, success vs nonsuccess) of the online mindfulness intervention on mental health outcomes; as such, this is a retrospective prognostic analysis of a classification problem per individual (ie, participants in the MVC mindfulness intervention).

Data Set Source

This is a retrospective analysis, where we analyzed an anonymized data set. The data were deidentified, and consent was obtained during the RCT; no further consent was sought for this secondary data analysis since nonidentifiable data were used. The data set was collected via an RCT described in detail elsewhere [28]. The parent study design consisted of a 2-arm parallel-design RCT, comparing a group assigned to the web-based MVC program to a waitlist control group. Participants in the study were students who were at least 18 years of age, reported English language fluency, self-reported high confidence in completing the study, and actively enrolled in an undergraduate program. This paper is based on the MVC intervention sample recruited in fall 2017, winter 2017, and fall 2018. The MVC intervention was an 8-week program and was comprised of three components: (1) 12 online videos for mental health education; (2) 3 anonymous discussion boards on depression, anxiety, and stress; and (3) anonymous, 20-minute group-based live videoconferences led by a mental health professional with training in mindfulness during which students could raise questions related to mindfulness (Figure 1).

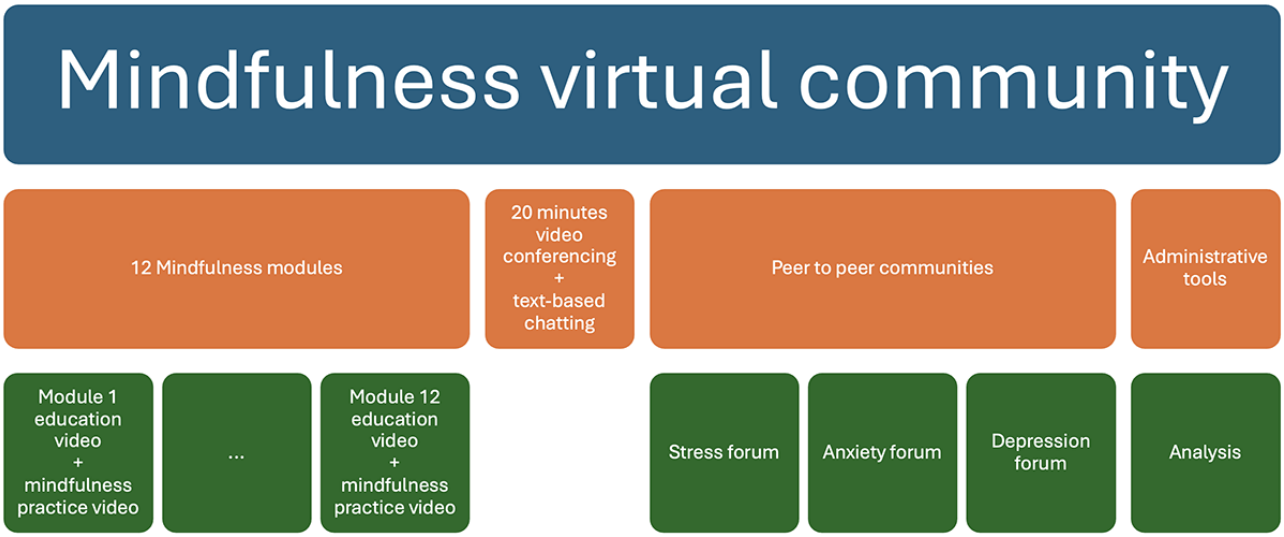
Each of the 12 mental health modules consisted of 1 educational content video and 1 mindfulness practice video recorded in both male and female voices and offered in both high and low resolution (a total of 8 videos per module); participants could choose the type of video they wanted to watch for each module. The videos were available for participants 24 hours a day to watch or listen to on computers, phones, or tablets at their convenience. The module scripts and audio recordings were created by one of the investigators with extensive experience as a clinical psychologist and researcher in mindfulness. They were based on mindfulness and cognitive behavioral therapy principles and informed by the prior student-based focus group

XSL•FO
RenderX

study [30,31]—the choice of moving and still images used in creating the videos involved collaborative work. The topics of the 12 modules included the following: overcoming stress, anxiety, and depression; mindfulness and being a student; mindfulness for better sleep; thriving in a fast-changing world; healthy intimacy; destigmatization; no more procrastination; pain reduction and mindfulness; healthy body image; healthier eating; overcoming trauma; and relationships with family and friends.

The primary RCT outcomes were depression, anxiety, and perceived stress, following hypotheses that symptom scores for depression, anxiety, and stress at T2 (after 8 weeks) would be significantly better in the MVC group when compared with the waitlist control group. The outcomes were measured with the following validated scales: Patient Health Questionnaire-9 (PHQ-9) [32], Beck Anxiety Inventory (BAI) [33], and Perceived Stress Scale (PSS) [34]. The secondary aim was to assess the impact of 3 elements of the MVC intervention on the outcomes. Participants also completed a sociodemographic questionnaire section at the T1 (baseline) survey.

Figure 1. The mindfulness virtual community design.



Ethical Considerations

The previous study received ethics approval from the Human Participant Research Committee (certificate e2016-345) of the York University. This ML study received ethics approval from the same committee (certificate e2023-012); the approval covers secondary analysis without additional consent. Participants in the original study had the option to receive an honorarium of CAD $50 (US $37.5) or 2% in course grade (for professors who gave this permission) or 3 credits (equivalent to 2% course grade) in the Undergraduate Research Participation Pool of the Department of Psychology. The participants' data were anonymized.

Participants

We aimed to build a model to predict who will likely benefit from the intervention, unlike the RCT study, where overall intervention effectiveness was determined (and supported by analysis) by comparing intervention and control groups. That is why we have analyzed intervention group data only to understand individual differences in response to the intervention.

Data Preparation

The data set consisted of 209 students who took the MVC intervention during fall 2017, winter 2018, and fall 2018. The effectiveness of the intervention was determined using the minimal clinically important difference (MCID), that is, the level of reduction in symptoms that psychologists consider clinically meaningful, for each of the mental health outcomes.

We adopted evidence from psychology that determines the MCID to be a 5-point reduction in PHQ-9 for depression [35,36], an 8.8-point reduction in BAI for anxiety [30,31], and an 11-point reduction in PSS for stress [37,38]. Any reduction equal to or above the MCID was labeled an effective intervention (label=1); otherwise, it was deemed ineffective (label=0).

To build a good prediction model from the training set, the data must be balanced. The class labels of the target variables, PHQ-9, BAI, and PSS, used in this study were not balanced. In our case, the percentage of instances with label=1 was extremely low: 50 (23%) for PHQ-9, 48 (24%) for BAI, and 8 (3.8%) for PSS, leading to a substantial imbalance. To alleviate the imbalanced data, we applied an oversampling method using the sklearn.resample function available in Python (version 3; Python Software Foundation).

Missing Data

Missing data in the outcomes were 12 (5.7%) for BAI and PHQ-9 and 13 (6.2%) for PSS of the 209 records. Missing data for the outcomes were dropped from the data set. There were no missing values in the predictors.

Labels and Features

The outcome variables were the 3 MCIDs associated with PHQ-9, BAI, and PSS being met or not for each instance. To investigate whether we can predict MVC's effectiveness using sociodemographic and self-reported features, the following features were used: sex (male and female), country of birth

XSL•FO
RenderX

(Canada and other), first language (English and other), education level (bachelor degree and other), ethnicity (White and non-White), marital status (married and other), age, number of weekly working hours, and self-rated health (poor, fair, good, very good, and excellent). To investigate the importance of exposure to mindfulness videos, in comparison with these features, in the prediction of intervention success, we added the total number of mindfulness videos watched to the previous data set.

Algorithms

Seven different classification algorithms, representing different learning paradigms, were used in this study: logistic regression (LR), support vector machine (SVM), random forest (RF), decision tree (DT), k-nearest neighbor (KNN), adaptive boosting (AdaBoost), and gradient boosting that showed good performance in previous studies that targeted depression, anxiety, and stress [35,39,40]. The implementations of the classification algorithms provided in the scikit-learn ML library [41] were used. The data set was split into 80% for training and 20% for testing. Hyperparameter tuning for each algorithm was performed using a grid search over a 10-fold cross-validation on the training data set. The optimal hyperparameters for the classification algorithms and their values for the data set without exposure to videos and the data set with exposure to videos are presented in Tables 1 and 2, respectively.

Each classifier's performance was compared with the best overall performance, leading to the selection of the best prediction model for the psychological outcomes. The classifiers' performances were assessed based on several evaluation metrics, including the percentage of correctly classified instances or the accuracy, sensitivity, specificity, and area under the curve (AUC) of the receiver operating characteristic curve. The best performance, as measured by the AUC score, was chosen for each algorithm.

To evaluate the features' importance in predicting intervention success, the data set with the total exposure to mindfulness videos was used to build predictive models. The RF algorithm was used to measure the features' importance. The hyperparameters used for the classification algorithms and their values that provided the optimal model are presented in Table 2.

XSL•FO
RenderX

Table 1. Algorithms and their corresponding optimal hyperparameters found by grid search (data set without videos).

Algorithm	Parameters
Logistic regression	
PHQ-9[a]	C=1, penalty=l1, solver=liblinear
BAI[b]	C=1, penalty=l1, solver=liblinear
PSS[c]	C=1, penalty=l1, solver=liblinear
Support vector machine	
PHQ-9	C=10, γ=0.1, kernel=rbf
BAI	C=10, γ=0.1, kernel=rbf
PSS	C=10, γ=0.1, kernel=rbf
Random forest	
PHQ-9	Max_features=auto, n_estimators=500, max_depth=8, criterion=entropy
BAI	Max_features=auto, n_estimators=500, max_depth=8, criterion=gini
PSS	Max_features=auto, n_estimators=500, max_depth=8, criterion=gini
Decision tree	
PHQ-9	Max_leaf_nodes=59, random_state=42, min_samples_split=2, criterion=entropy
BAI	Max_leaf_nodes=56, random_state=42, min_samples_split=2, criterion=entropy
PSS	Max_leaf_nodes=16, random_state=42, min_samples_split=2, criterion=entropy
K-nearest neighbor	
PHQ-9	N_neighbors=2, weight=distance, leaf size=27, P=1
BAI	N_neighbors=2, weight=distance, leaf size=1, P=1
PSS	N_neighbors=1, weight=dniform, leaf size=1, P=1
Adaptive boosting	
PHQ-9	n-estimators=5000, max_depth=3, learning rate=0.5
BAI	n-estimators=5000, max_depth=3, learning rate=0.9
PSS	n-estimators=500, max_depth=3, learning rate=0.9
Gradient boosting	
PHQ-9	Learning rate=0.05, max depth=6, n-estimators=100, subsample=0.9, max_features=none, min_samples_split=2
BAI	Learning rate=0.02, max depth=10, n-estimators=1000, subsample=1.0, max_features=none, min_samples_split=2
PSS	Learning rate=0.01, max depth=6, n-estimators=1000, subsample=0.9, max_features=sqrt, min_samples_split=2

[a]PHQ-9: Patient Health Questionnaire-9.

[b]BAI: Beck Anxiety Inventory.

[c]PSS: Perceived Stress Scale.

XSL•FO
RenderX

Table 2. Algorithms and their corresponding optimal hyperparameters found by grid search (data set with exposure to videos).

Algorithm	Parameters
Logistic regression	
PHQ-9[a]	C=1, penalty=l1, solver=liblinear
BAI[b]	C=0.1, penalty=l2, solver=newton-cg
PSS[c]	C=100, penalty=l2, solver=lbfgs
Support vector machine	
PHQ-9	C=10, γ=0.01, kernel=rbf
BAI	C=1, γ=1, kernel=rbf
PSS	C=1, γ=1, kernel=rbf
Random forest	
PHQ-9	Max_features=auto, n_estimators=500, max_depth=7, criterion=entropy
BAI	Max_features=auto, n_estimators=200, max_depth=8, criterion=gini
PSS	Max_features=auto, n_estimators=500, max_depth=8, criterion=gini
Decision tree	
PHQ-9	Max_leaf_nodes=39, random_state=42, min_samples_split=2, criterion=entropy
BAI	Max_leaf_nodes=53, random_state=42, min_samples_split=3, criterion=gini
PSS	Max_leaf_nodes=16, random_state=42, min_samples_split=2, criterion=gini
K-nearest neighbor	
PHQ-9	N_neighbors=1, weight=uniform, leaf size=14, P=1
BAI	N_neighbors=2, weight=distance, leaf size=1, P=1
PSS	N_neighbors=2, weight=uniform, leaf size=1, P=1
Adaptive boosting	
PHQ-9	n-estimators=500, max_depth=3, learning rate=0.5
BAI	n-estimators=500, max_depth=3, learning rate=0.7
PSS	n-estimators=2000, max_depth=3, learning rate=0.7
Gradient boosting	
PHQ-9	Learning rate=0.5, max depth=50, n-estimators=50, subsample=0.9, max_features=sqrt, min_samples_split=2
BAI	Learning rate=0.04, max depth=10, n-estimators=1000, subsample=0.5, max_features=none, min_samples_split=2
PSS	Learning rate=0.03, max depth=8, n-estimators=1000, subsample=0.5, max_features=none, min_samples_split=2

[a]PHQ-9: Patient Health Questionnaire-9.

[b]BAI: Beck Anxiety Inventory.

[c]PSS: Perceived Stress Scale.

Results

Demographics

Table 3 presents the demographic characteristics of participants at baseline. Of 209 students, 73.2% (n=153) were female, 8.1% (n=17) were married, and 21.1% (n=44) were White. Most participants were born in Canada, and English was their first language. The median (IQR) of age, work hours per week, and the total number of mindfulness videos watched were 21 (19-23) years, 10 (0-18), and 16 (9-30), respectively.

XSL•FO
RenderX

Table 3. Characteristics of participants at baseline (N=209).

Characteristics	Values
Sex, n (%)	
Male	56 (26.8)
Female	153 (73.2)
Marital status, n (%)	
Married	17 (8.1)
Other	192 (91.9)
Ethnicity, n (%)	
White	44 (21.1)
Non-White	165 (78.9)
Language, n (%)	
English	136 (65.1)
Other	73 (34.9)
Country of birth, n (%)	
Canada	119 (56.9)
Other	90 (43.1)
Education, n (%)	
High school diploma or General Education Development or college degree or certificate program	182 (87.1)
Bachelor degree	27 (12.9)
Self-reported general health, n (%)	
Poor or fair	43 (20.6)
Good or very good or excellent	166 (79.4)
Age (years), median (IQR)	21 (19-23)
Average number of hours at work per week, median (IQR)	10 (0-18)
Total number of mindfulness videos watched, median (IQR)	16 (9-30)

Objective 1: Predicting MVC's Effectiveness Using Sociodemographic and Self-Reported Features

Table 4 summarizes the evaluated models' performances: sensitivity, specificity, accuracy, and AUC, using 10-fold cross-validation.

The results showed that both gradient boosting (AUC=0.85 and accuracy=0.83) and DT (AUC=0.84 and accuracy=0.81) are slightly better compared to AdaBoost and KNN (AUC=0.82 and accuracy=0.80) as well as SVM (AUC=0.81 and accuracy=0.80) and outperformed the remaining classification algorithms for predicting a clinically significant reduction in PHQ-9. The best classifiers for predicting a clinically significant reduction in BAI were RF (AUC=0.93 and accuracy=0.93), followed by AdaBoost (AUC=0.92 and accuracy=0.92) and gradient boosting (AUC=0.87 and accuracy=0.87), which outperformed the remaining classifiers. Two classifiers, gradient boosting and DT, gained the perfect accuracy and AUC (AUC=1 and accuracy=1) for predicting a clinically significant reduction in PSS, followed by the near-perfect scores for SVM and AdaBoost (AUC=0.99 and accuracy=0.99). Meanwhile, LR had the lowest performance for PHQ-9, BAI, and PSS in terms of AUC (0.64, 0.75, and 0.73, respectively) and accuracy (0.66, 0.75, and 0.73, respectively).

The results were close to those found in the models built without video exposure. Gradient boosting (AUC=0.89 and accuracy=0.88) was the best predictor for a significant reduction in PHQ-9, followed closely by AdaBoost and DT (AUC=0.84 and accuracy=0.81), which outperformed the remaining classification algorithms. The best classifiers for predicting a clinically significant reduction in BAI were AdaBoost and SVM (AUC=0.93 and accuracy=0.93), followed closely by gradient boosting (AUC=0.91 and accuracy=0.92) and RF (AUC=0.90 and accuracy=0.90), which outperformed the remaining classifiers. Four classifiers, gradient boosting, AdaBoost, RF, and SVM, gained the perfect AUC and accuracy (AUC=1 and accuracy=1) for predicting a clinically significant reduction in PSS, followed by the near-perfect score for KNN (AUC=0.99 and accuracy=0.99) and DT (AUC=0.97 and accuracy=0.97). Meanwhile, LR had the lowest performance for PHQ-9, BAI, and PSS in terms of AUC (0.62, 0.60, and 0.79, respectively) and accuracy (0.63, 0.60, and 0.80, respectively).

Using the second data set (ie, enriched with the exposure to videos), RF was used to detect features' importance in relation

XSL•FO
RenderX

to the 3 outcomes. The most predictive feature for the PHQ-9, BAI, and PSS was the total exposure to the mindfulness videos, followed by the average number of working hours per week and age for PHQ-9 and BAI. In contrast, age and the average number of working hours per week were the second and third most important predictors for PSS, respectively.

Table 4. Classification report of the machine learning algorithms for outcomes.

Algorithm	AUC[a]	Accuracy	Sensitivity	Specificity
Logistic regression				
PHQ-9[b]	0.64	0.66	0.57	0.72
BAI[c]	0.75	0.75	0.75	0.75
PSS[d]	0.73	0.73	0.73	0.74
Support vector machine				
PHQ-9	0.81	0.80	0.90	0.75
BAI	0.77	0.77	0.79	0.75
PSS	0.96	0.96	1.0	0.91
Random forest				
PHQ-9	0.78	0.76	0.87	0.69
BAI	0.93	0.93	0.86	1.0
PSS	0.99	0.99	1.0	0.97
Decision tree				
PHQ-9	0.84	0.81	0.96	0.72
BAI	0.84	0.83	0.93	0.75
PSS	1.0	1.0	1.0	1.0
K-nearest neighbor				
PHQ-9	0.82	0.80	0.91	0.72
BAI	0.78	0.78	0.68	0.88
PSS	0.96	0.96	1.0	0.91
Adaptive boosting				
PHQ-9	0.82	0.80	0.91	0.72
BAI	0.92	0.92	0.89	0.94
PSS	0.99	0.99	1.0	0.97
Gradient boosting				
PHQ-9	0.85	0.83	0.91	0.78
BAI	0.87	0.87	0.86	0.88
PSS	1.0	1.0	1.0	1.0

[a]AUC: area under the curve.
[b]PHQ-9: Patient Health Questionnaire-9.
[c]BAI: Beck Anxiety Inventory.
[d]PSS: Perceived Stress Scale.

Objective 2: Importance of Exposure to Mindfulness Videos in Comparison With Sociodemographics and Self-Reported Features in Predicting Intervention Success

After the introduction of the total exposure to the mindfulness videos to the data set, new predictive models were built (Table 5).

XSL•FO
RenderX

Table 5. Classification report of the machine learning algorithms for outcomes (data set with exposure to videos).

Algorithm	AUC[a]	Accuracy	Sensitivity	Specificity
Logistic regression				
PHQ-9[b]	0.62	0.63	0.57	0.67
BAI[c]	0.60	0.60	0.61	0.59
PSS[d]	0.79	0.80	0.85	0.74
Support vector machine				
PHQ-9	0.78	0.75	0.96	0.61
BAI	0.93	0.93	0.86	1.00
PSS	1.00	1.00	1.00	1.00
Random forest				
PHQ-9	0.84	0.83	0.87	0.81
BAI	0.90	0.90	0.86	0.94
PSS	1.00	1.00	1.00	1.00
Decision tree				
PHQ-9	0.84	0.81	0.96	0.72
BAI	0.80	0.80	0.86	0.75
PSS	0.97	0.97	1.00	0.94
K-nearest neighbor				
PHQ-9	0.81	0.78	0.96	0.67
BAI	0.84	0.83	0.89	0.78
PSS	0.99	0.99	1.00	0.97
Adaptive boosting				
PHQ-9	0.84	0.81	0.96	0.72
BAI	0.93	0.93	0.89	0.97
PSS	1.00	1.00	1.00	1.00
Gradient boosting				
PHQ-9	0.89	0.88	0.96	0.83
BAI	0.91	0.92	0.86	0.97
PSS	1.00	1.00	1.00	1.00

[a]AUC: area under the curve.

[b]PHQ-9: Patient Health Questionnaire-9.

[c]BAI: Beck Anxiety Inventory.

[d]PSS: Perceived Stress Scale.

Discussion

Principal Results

The study investigated the predictability of the effectiveness of an MVC designed for undergraduate students to reduce symptoms of depression, anxiety, and stress as measured by PHQ-9, BAI, and PSS. The effectiveness was measured by the MCID for PHQ-9, BAI, and SPSS. Several algorithms were used to predict the MCID.

Predicting Intervention Success With Sociodemographic and Self-Reported Measures

We successfully built ML-based models that predicted the effectiveness of the MVC intervention. The highest AUC was achieved for gradient boosting to predict the intervention effectiveness for PHQ-9 and PSS (AUC=0.85 and AUC=1, respectively), followed closely by DT (AUC=0.84 and AUC=1, respectively) and AdaBoost (AUC=0.82 and AUC=0.99, respectively). The RF model had the highest AUC to predict BAI (AUC=0.93), followed closely by AdaBoost (AUC=0.92). AdaBoost might be the algorithm of choice for the 3 outcomes, as it is fairing a close second best for BAI and a close third best

XSL•FO
RenderX

for PHQ-9 and PSS. Gradient boosting and AdaBoost are both good choices to predict the intervention success for the 3 outcomes. It might be argued that AdaBoost might be preferable, given that it is usually less prone to overfitting than gradient boosting; however, there is no need to use the same algorithm to build the 3 predictors for the 3 outcomes.

We could not make a direct comparison with other studies that measured the 3 outcomes among university students using the same validated scales (PHQ-9, BAI, and SPSS). However, for PHQ-9, the performance of our model is higher than the one found in a previous study among adults in Korea using the Center for Epidemiologic Studies—Depression Scale 11 (AUC=0.87 and accuracy=0.86) [40] as well as the one found in a study in the United States that defined the success of the intervention as a 5-point reduction in PHQ-9 or a 4-point reduction in the General Anxiety Disorder screener-7 values (AUC=0.60 and accuracy=0.71) [35]. Regarding anxiety, the predictive model developed in this study had a higher performance (accuracy=0.92) than another study that used the Self-Rating Anxiety Scale, which did not report AUC but reported an accuracy of 0.84.

Feature Importance

Exposure to mindfulness videos was the most important factor in predicting the intervention's success. This study has demonstrated a link between the MVC intervention's success and exposure to mindfulness videos. It also confirms the results of the previous MVC pilot study that proved that exposure to mindfulness videos alone, without interaction between participants via an online discussion forum and without weekly videoconferencing with a coach, effectively reduced symptoms of depression, anxiety, and stress [26]. In other words, it indicates the ability of MVC to be deployed at a large scale without an increase in human resources. Scalability is a critical factor for eHealth intervention deployment in large populations. This finding suggests that scaling up an effective e-mental health MVC is possible in a cost-effective manner; scalability is one of the recognized failures in eHealth implementations [42].

Practical and Policy Implications

The MVC intervention does not provide clinical support; it is a platform that offers self-management of mental health symptoms (depression, anxiety, and stress). The MVC intervention proved to be effective [26-28] in reducing symptoms of depression, anxiety, and stress in university students. This study builds a predictive model that predicts intervention success using sociodemographic and self-reported measures; this will allow counseling services on university campuses to assess the usefulness of MVC for a particular student before taking the intervention and advise them accordingly to use MVC or to opt for another type of intervention. This will enable counseling services to personalize the advice to students' profiles and allow students to manage their symptoms with the most appropriate intervention.

The other finding related to videos being the most important factor in predicting intervention success confirms the ability of MVC to be deployed at a large scale without an increase in human resources. The number of working hours is another important predictor of the success of the intervention. Although the provincial governments in Canada support university education, students must pay for their education and bear the cost of living. Not surprisingly, they work long hours, especially if they belong to a marginalized community. Our findings align with other studies that suggest that longer working hours outside the university and difficulty paying bills were recognized as predictors of poor mental health among students [43]. In Ontario, where the sample was taken, Statistics Canada recently reported an increased reliance of academic institutions on students' fees in higher education to the extent that 54% of all college revenues in 2019/2020 were downloaded on students, which translates into an overall decline in public funding [44]. This situation pushed students to longer working hours; one can argue that since student debt has been recognized as negatively associated with mental well-being and academic outcomes [45,46], providing access to free higher education, supported by taxes such as in most of Europe, could enhance students' mental well-being as it would relieve them from the need for long working hours.

Strengths and Limitations

One of the strengths of this study is the ability to predict the intervention's success based on a few demographics and one question about self-rated health. Hence, the predictive model can be used in real life to indicate the suitability of online mindfulness intervention for specific individuals and possibly suggest alternatives if the model predicts noneffectiveness. The excellent AUC and accuracy measures make the models suitable for implementation and evaluation in real-life scenarios. However, the ML models must be monitored continuously if implemented for daily use (eg, a counseling service) [47,48].

A limitation of this study is that it relied on research done on 1 site; future research with larger samples with participants from multiple universities and colleges would better test the generalizability of results as it allows us to test the effectiveness of the models on external data.

Conclusions

Our results suggest that we can build high-performing models to predict MVC intervention effectiveness for depression, anxiety, and stress based on simple sociodemographics and self-reported features and that exposure to mindfulness videos is the most important predictor for the effectiveness of the intervention. Our findings provide evidence that scaling MVC can be done without additional cost for support and that the predictive models might be useful for professionals to advise students early enough on taking the intervention or choosing other alternatives.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the contribution of all students who spent valuable time participating in the study. The MVC Team members are Sahir Abbas, BSc; Yvonne Bohr, PhD; Manuela Ferrari, PhD; Wai Lun Alan Fung, MD, ScD, FRCPC; Louise Hartley, PhD; Amin Mawani, PhD; Kwame McKenzie, MD, FRCPC; and Jan E Odai, BA. They made contributions to several aspects of the project and the results' development. They approve the final version and agree to be accountable for all aspects of the submitted paper. The work reported in this paper was funded by the Canadian Institutes for Health Research and eHealth Innovations Partnership Program grant (EH1-143553). The project's principal investigators are CE, FA, and PR.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

CE, FA, and PR designed the original mindfulness virtual community study questionnaire, received the funds, and contributed equally. CE supervised FT, who performed and reported the analysis. CE verified the analysis and prepared the first draft. All authors provided critical feedback and revised it.

Conflicts of Interest

It is the understanding of the university and researchers that the Project Intellectual Property belongs to the CE, FA, and PR. The industry partner ForaHealthyMe.com owns all rights and titles to the copyrights of any computer source code software developed from this research project.

References

1. American College Health Association-National College Health Assessment II: Canadian reference group executive summary spring 2013. American College Health Association. 2013. URL: https://tinyurl.com/vxn2swse [accessed 2024-04-18]
2. American College Health Association-National College Health Assessment II: Canadian reference group executive summary spring 2016. American College Health Association. 2016. URL: https://tinyurl.com/bdfpv8b4 [accessed 2024-04-18]
3. Pirbaglou M, Cribbie R, Irvine J, Radhu N, Vora K, Ritvo P. Perfectionism, anxiety, and depressive distress: evidence for the mediating role of negative automatic thoughts and anxiety sensitivity. J Am Coll Health 2013;61(8):477-483 [FREE Full text] [doi: 10.1080/07448481.2013.833932] [Medline: 24152025]
4. Eisenberg D, Hunt J, Speer N, Zivin K. Mental health service utilization among college students in the United States. J Nerv Ment Dis 2011;199(5):301-308. [doi: 10.1097/NMD.0b013e3182175123] [Medline: 21543948]
5. Han B, Compton WM, Eisenberg D, Milazzo-Sayre L, McKeon R, Hughes A. Prevalence and mental health treatment of suicidal ideation and behavior among college students aged 18-25 years and their non-college-attending peers in the United States. J Clin Psychiatry 2016;77(6):815-824 [FREE Full text] [doi: 10.4088/JCP.15m09929] [Medline: 27232194]
6. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, et al. WHO World Mental Health Surveys International College Student project: prevalence and distribution of mental disorders. J Abnorm Psychol 2018;127(7):623-638 [FREE Full text] [doi: 10.1037/abn0000362] [Medline: 30211576]
7. The human face of mental health and mental illness in Canada 2006. Government of Canada. 2006. URL: https://www.phac-aspc.gc.ca/publicat/human-humain06/pdf/human_face_e.pdf [accessed 2024-04-18]
8. Allen R, Kannangara C, Vyas M, Carson J. European university students' mental health during Covid-19: exploring attitudes towards Covid-19 and governmental response. Curr Psychol 2022:1-14 [FREE Full text] [doi: 10.1007/s12144-022-02854-0] [Medline: 35194357]
9. Koelen JA, Mansueto AC, Finnemann A, de Koning L, van der Heijde CM, Vonk P, et al. COVID-19 and mental health among at-risk university students: a prospective study into risk and protective factors. Int J Methods Psychiatr Res 2022;31(1):e1901 [FREE Full text] [doi: 10.1002/mpr.1901] [Medline: 34932250]
10. Zapata-Ospina JP, Patiño-Lugo DF, Vélez CM, Campos-Ortiz S, Madrid-Martínez P, Pemberthy-Quintero S, et al. Mental health interventions for college and university students during the COVID-19 pandemic: a critical synthesis of the literature. Rev Colomb Psiquiatr (Engl Ed) 2021;50(3):199-213 [FREE Full text] [doi: 10.1016/j.rcpeng.2021.04.001] [Medline: 34479842]
11. Ungar T. The health care payment game is rigged. National Post. 2015 Apr 28. URL: https://nationalpost.com/opinion/thomas-ungar-the-health-care-payment-game-is-rigged [accessed 2024-05-03]
12. Lim KL, Jacobs P, Ohinmaa A, Schopflocher D, Dewa CS. A new population-based measure of the economic burden of mental illness in Canada. Chronic Dis Can 2008;28(3):92-98 [FREE Full text] [Medline: 18341763]
13. Bayram N, Bilgel N. The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. Soc Psychiatry Psychiatr Epidemiol 2008;43(8):667-672 [FREE Full text] [doi: 10.1007/s00127-008-0345-x] [Medline: 18398558]

XSL•FO
RenderX

14. Pfeffer A. Ontario campus counsellors say they're drowning in mental health needs. CBC. 2019. URL: https://www.cbc.ca/news/canada/ottawa/mental-health-ontario-campus-crisis-1.3771682 [accessed 2019-07-07]
15. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
16. Keng SL, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. Clin Psychol Rev 2011;31(6):1041-1056 [FREE Full text] [doi: 10.1016/j.cpr.2011.04.006] [Medline: 21802619]
17. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. A meta-analysis. J Psychosom Res 2004;57(1):35-43 [FREE Full text] [doi: 10.1016/S0022-3999(03)00573-7] [Medline: 15256293]
18. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009;15(5):593-600 [FREE Full text] [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
19. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. J Consult Clin Psychol 2010;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
20. Vøllestad J, Nielsen MB, Nielsen GH. Mindfulness- and acceptance-based interventions for anxiety disorders: a systematic review and meta-analysis. Br J Clin Psychol 2012;51(3):239-260 [FREE Full text] [doi: 10.1111/j.2044-8260.2011.02024.x] [Medline: 22803933]
21. Eberth J, Sedlmeier P. The effects of mindfulness meditation: a meta-analysis. Mindfulness 2012;3(3):174-189 [FREE Full text] [doi: 10.1007/s12671-012-0101-x]
22. Sedlmeier P, Eberth J, Schwarz M, Zimmermann D, Haarig F, Jaeger S, et al. The psychological effects of meditation: a meta-analysis. Psychol Bull 2012;138(6):1139-1171. [doi: 10.1037/a0028168] [Medline: 22582738]
23. Bender JL, Jimenez-Marroquin MC, Ferris LE, Katz J, Jadad AR. Online communities for breast cancer survivors: a review and analysis of their characteristics and levels of use. Support Care Cancer 2013;21(5):1253-1263 [FREE Full text] [doi: 10.1007/s00520-012-1655-9] [Medline: 23179491]
24. El Morr C. Mobile virtual communities in healthcare: managed self care on the move. 2007 Presented at: Telehealth '07: The Third IASTED International Conference on Telehealth; May 31, 2007-June 1, 2007; Montreal, Canada. [doi: 10.4018/978-1-59904-002-8.ch105]
25. Morr CE, Kawash J. Mobile virtual communities research: a synthesis of current trends and a look at future perspectives. Int J Web Based Communities 2007;3(4):386-403 [FREE Full text] [doi: 10.1504/ijwbc.2007.015865]
26. Ahmad F, El Morr C, Ritvo P, Othman N, Moineddin R, MVC Team. An eight-week, web-based mindfulness virtual community intervention for students' mental health: randomized controlled trial. JMIR Ment Health 2020;7(2):e15520 [FREE Full text] [doi: 10.2196/15520] [Medline: 32074061]
27. Ritvo P, Ahmad F, El Morr C, Pirbaglou M, Moineddin R, MVC Team. A mindfulness-based intervention for student depression, anxiety, and stress: randomized controlled trial. JMIR Ment Health 2021;8(1):e23491 [FREE Full text] [doi: 10.2196/23491] [Medline: 33264098]
28. El Morr C, Ritvo P, Ahmad F, Moineddin R, MVC Team. Effectiveness of an 8-week web-based mindfulness virtual community intervention for university students on symptoms of stress, anxiety, and depression: randomized controlled trial. JMIR Ment Health 2020;7(7):e18595 [FREE Full text] [doi: 10.2196/18595] [Medline: 32554380]
29. El Morr C. Virtual communities, machine learning and IoT: opportunities and challenges in mental health research. Int J Extreme Autom Connectivity Healthc 2019;1(1):4-11 [FREE Full text] [doi: 10.4018/ijeach.2019010102]
30. Carter S, Greenberg J, Funes CJ, Macklin EA, Vranceanu AM. Effects of a mind-body program on symptoms of depression and perceived stress among adults with neurofibromatosis type 2 who are deaf: a live-video randomized controlled trial. Complement Ther Med 2021;56:102581 [FREE Full text] [doi: 10.1016/j.ctim.2020.102581] [Medline: 33197677]
31. Eskildsen A, Dalgaard VL, Nielsen KJ, Andersen JH, Zachariae R, Olsen LR, et al. Cross-cultural adaptation and validation of the Danish consensus version of the 10-item Perceived Stress Scale. Scand J Work Environ Health 2015;41(5):486-490 [FREE Full text] [doi: 10.5271/sjweh.3510] [Medline: 26111225]
32. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. JAMA 1999;282(18):1737-1744 [FREE Full text] [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
33. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
34. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385-396. [Medline: 6668417]
35. Hornstein S, Forman-Hoffman V, Nazander A, Ranta K, Hilbert K. Predicting therapy outcome in a digital mental health intervention for depression and anxiety: a machine learning approach. Digit Health 2021;7:20552076211060659 [FREE Full text] [doi: 10.1177/20552076211060659] [Medline: 34868624]
36. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42(12):1194-1201 [FREE Full text] [doi: 10.1097/00005650-200412000-00006] [Medline: 15550799]

XSL•FO
RenderX

37. Goldsworthy S, Palmer S, Latour JM, McNair H, Cramp M. A systematic review of effectiveness of interventions applicable to radiotherapy that are administered to improve patient comfort, increase patient compliance, and reduce patient distress or anxiety. Radiography (Lond) 2020;26(4):314-324 [FREE Full text] [doi: 10.1016/j.radi.2020.03.002] [Medline: 32245711]
38. Leentjens AFG, Dujardin K, Marsh L, Richard IH, Starkstein SE, Martinez-Martin P. Anxiety rating scales in Parkinson's disease: a validation study of the Hamilton anxiety rating scale, the Beck anxiety inventory, and the hospital anxiety and depression scale. Mov Disord 2011;26(3):407-415 [FREE Full text] [doi: 10.1002/mds.23184] [Medline: 21384425]
39. Wang C, Zhao H, Zhang H. Chinese college students have higher anxiety in new semester of online learning during COVID-19: a machine learning approach. Front Psychol 2020;11:587413 [FREE Full text] [doi: 10.3389/fpsyg.2020.587413] [Medline: 33343461]
40. Na KS, Cho SE, Geem ZW, Kim YK. Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm. Neurosci Lett 2020;721:134804 [FREE Full text] [doi: 10.1016/j.neulet.2020.134804] [Medline: 32014516]
41. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830 [FREE Full text]
42. Moroz N, Moroz I, D'Angelo MS. Mental health services in Canada: barriers and cost-effective solutions to increase access. Healthc Manage Forum 2020;33(6):282-287 [FREE Full text] [doi: 10.1177/0840470420933911] [Medline: 32613867]
43. Roberts R, Golding J, Towell T, Reid S, Woodford S, Vetere A, et al. Mental and physical health in students: the role of economic circumstances. Br J Health Psychol 2010;5(3):289-297 [FREE Full text] [doi: 10.1348/135910700168928]
44. Trends in private and public funding in Canadian colleges, 2019/2020. Statistics Canada. 2023. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/220120/dq220120c-eng.htm [accessed 2023-03-26]
45. Pisaniello MS, Asahina AT, Bacchi S, Wagner M, Perry SW, Wong ML, et al. Effect of medical student debt on mental health, academic performance and specialty choice: a systematic review. BMJ Open 2019;9(7):e029980 [FREE Full text] [doi: 10.1136/bmjopen-2019-029980] [Medline: 31270123]
46. Boyles JD, Ahmed B. Does student debt affect dental students' and dentists' stress levels? Br Dent J 2017;223(8):601-606 [FREE Full text] [doi: 10.1038/sj.bdj.2017.876] [Medline: 29074930]
47. Gurevich E, El Hassan B, El Morr C. Equity within AI systems: what can health leaders expect? Healthc Manage Forum 2023;36(2):119-124 [FREE Full text] [doi: 10.1177/08404704221125368] [Medline: 36226507]
48. Kundi B, El Morr C, Gorman R, Dua E. Artificial intelligence and bias: a scoping review. In: Yampolskiy RV, editor. AI and Society: Tensions and Opportunities. New York: CRC Press Taylor & Francis; 2023:199-213.

Abbreviations

AdaBoost: adaptive boosting
AUC: area under the curve
BAI: Beck Anxiety Inventory
DT: decision tree
KNN: k-nearest neighbor
LR: logistic regression
MCID: minimal clinically important difference
ML: machine learning
MVC: mindfulness virtual community
PHQ-9: Patient Health Questionnaire-9
PSS: Perceived Stress Scale
RCT: randomized controlled trial
RF: random forest
SVM: support vector machine
VC: virtual community

Edited by T de Azevedo Cardoso; submitted 18.07.23; peer-reviewed by M Musker, D Carvalho; comments to author 24.11.23; revised version received 05.12.23; accepted 05.04.24; published 13.05.24.

Please cite as:
El Morr C, Tavangar F, Ahmad F, Ritvo P, MVC Team
Predicting the Effectiveness of a Mindfulness Virtual Community Intervention for University Students: Machine Learning Model
Interact J Med Res 2024;13:e50982
URL: https://www.i-jmr.org/2024/1/e50982
doi:10.2196/50982
PMID:38578872

XSL•FO
RenderX

©Christo El Morr, Farideh Tavangar, Farah Ahmad, Paul Ritvo, MVC Team. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 13.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Designing mHealth Apps to Incorporate Evidence-Based Techniques for Prolonging User Engagement

Rebecca Monachelli[1], BS; Sharon Watkins Davis[1], MPA; Allison Barnard[1], BA; Michelle Longmire[1], MD; John P Docherty[2], MD; Ingrid Oakley-Girvan[1,3], MPH, PhD

[1]Medable Inc, Palo Alto, CA, United States
[2]Weill Cornell Medical College, White Plains, NY, United States
[3]The Public Health Institute, Oakland, CA, United States

Corresponding Author:
Ingrid Oakley-Girvan, MPH, PhD
Medable Inc
525 University Ave
Palo Alto, CA, 94301
United States
Phone: 1 8778206259
Email: oakley@stanford.edu

Abstract

Maintaining user engagement with mobile health (mHealth) apps can be a challenge. Previously, we developed a conceptual model to optimize patient engagement in mHealth apps by incorporating multiple evidence-based methods, including increasing health literacy, enhancing technical competence, and improving feelings about participation in clinical trials. This viewpoint aims to report on a series of exploratory mini-experiments demonstrating the feasibility of testing our previously published engagement conceptual model. We collected data from 6 participants using an app that showed a series of educational videos and obtained additional data via questionnaires to illustrate and pilot the approach. The videos addressed 3 elements shown to relate to engagement in health care app use: increasing health literacy, enhancing technical competence, and improving positive feelings about participation in clinical trials. We measured changes in participants' knowledge and feelings, collected feedback on the videos and content, made revisions based on this feedback, and conducted participant reassessments. The findings support the feasibility of an iterative approach to creating and refining engagement enhancements in mHealth apps. Systematically identifying the key evidence-based elements intended to be included in an app's design and then systematically testing the implantation of each element separately until a satisfactory level of positive impact is achieved is feasible and should be incorporated into standard app design. While mHealth apps have shown promise, participants are more likely to drop out than to be retained. This viewpoint highlights the potential for mHealth researchers to test and refine mHealth apps using approaches to better engage users.

(Interact J Med Res 2024;13:e51974) doi:10.2196/51974

KEYWORDS

adherence; app design; attrition; mHealth; user engagement; user experience; proof-of-concept

Introduction

Smartphones have a global penetration estimated at 3.9 billion users [1], enabling mobile health (mHealth) apps to reach even low-resource areas and underserved populations [2,3]. mHealth apps have been developed to enable remote participation in clinical trials [4-7] and provide health education, health management, and other uses across the continuum from prevention through active treatment to palliative care [8]. Decentralized clinical trials using mHealth technologies promise faster participant accrual and a higher return on investment than traditional site-based trials [9]. mHealth apps have been shown to reduce inpatient readmission rates and decrease the length of hospital stay [10]. mHealth can increase knowledge and improve confidence and communication with health professionals [11]. However, while participants readily sign up for mHealth education and decentralized clinical trial apps, retention remains a major challenge [12-14]. For mHealth apps to succeed, users must consistently engage with them [15,16].

Engagement in mHealth apps has been conceptualized to include behavior, cognition, and affective components [17]. However, measures of patient engagement are underreported and lack consistency [18,19]. Participants are more likely to drop out than be retained despite app elements such as feedback,

XSL•FO
RenderX

reminders, in-app support, gamification, and participant compensation [20-23].

We developed a conceptual model to optimize patient engagement based on different phases of the engagement process [24]. Because digital literacy and anxiety have been shown to be negatively correlated with engagement [25], we established an approach to develop and test the educational components of our conceptual model to enhance app engagement by increasing health literacy, enhancing technical competence, and improving feelings about clinical trials. This Viewpoint aims to report on a series of exploratory mini-experiments, demonstrating the feasibility of testing our engagement conceptual model.

How We Conducted the Exploratory Mini-Experiments

Testing Design

We used a product testing approach rather than the traditional research evaluation approach. We used a group of existing product testers who are patients or caregivers working for Medable to rapidly test different iterations of our educational videos. Questionnaires were used both before and after participants viewed the videos, and semistructured interviews were also conducted.

Data Collection

We developed apps to collect specific data from participants over 1 week's duration through questionnaires available on their smartphones before and after exposure to videos, as shown in Table 1. The videos were based on a review of the literature defining and studying each of these 3 target areas: health literacy, technical competence, and feelings about participation in clinical trials [24]. Each concept area was tested separately, with questionnaires specific to the educational component.

Table 1. Schedule of tasks or questionnaires.

Task or questionnaire	Frequency	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Sign up	Once	Task 1	N/A[a]	N/A	N/A	N/A	N/A
Demographics (age, gender, education, housing, race, ethnicity, location, health condition, and notification)	Once	Task 2	N/A	N/A	N/A	N/A	N/A
Technology competence questions (TAM3[b] CANX,[c] worries about pressing the wrong button and device preference)	Twice	Task 3[d]	Task 2[e]	N/A	N/A	N/A	N/A
Technology competence video combined with practice questions (video on how to answer questions and practice questions)	Once	N/A	Task 1[d,e]	N/A	N/A	N/A	N/A
Health literacy questions (BRIEF[f], own health knowledge, clinical trial knowledge, and BMI)	Twice	N/A	N/A	Task 1[d,ed]	Task 2[e]	N/A	N/A
Health literacy video with knowledge check questions (St. Luke's University Health Network's video, "Wellness 101 – How to Improve Your Overall Health"; knowledge check questions)	Once	N/A	N/A	N/A	Task 1[d,e]	N/A	N/A
Clinical trials question (temperature scale regarding participant's feelings about study participation, Figure S1 in Multimedia Appendix 1)	Twice	N/A	N/A	N/A	N/A	Task 1[d,e]	Task 2
Clinical trials video (National Institute of Diabetes and Digestive and Kidney Diseases video: "Why Should I Join a Clinical Trial?")	Once	N/A	N/A	N/A	N/A	N/A	Task 1[d,e]
Study complete	Once	N/A	N/A	N/A	N/A	N/A	Task 3

[a]N/A: not applicable.

[b]TAM3: Technology Acceptance Model 3.

[c]CANX: Computer Anxiety.

[d]Set up with an automated morning reminder: "You have new tasks available today in the Patient Engagement app!"

[e]Set-up with an automated 8 pm reminder: "You have uncompleted tasks in the Patient Engagement app. Please finish them before midnight!"

[f]BRIEF: Brief Health Literacy Screening Tool.

XSL•FO
RenderX

In addition to collecting data through the questionnaires noted in Table 1, we conducted individual semistructured interviews with each participant at the end of the series of user evaluations using a video conferencing platform. Based on feedback, we revised specific videos and content and then conducted a second round of feedback. In this second round, after all the questions were answered in each section, participants were asked for immediate feedback with the open-ended question, "What did you think of this video?"

Recruitment

We used a product testing approach rather than a research evaluation approach. We recruited 6 individuals from the Medable Patient Care Network (PCN) who participated in this product development effort from May 2022 through July 2022. The PCN is a group of patients and caregivers who provide insights and user feedback from their perspective for a variety of apps being developed as products at Medable.

Ethical Considerations

This work was conducted and approved under the the Advarra IRB (Pro00062352). PCN members were paid an hourly rate of approximately US $150 by Medable for their work on behalf of the network in support of Medable product development efforts. Informed consent was obtained from all participants. All data were deidentified.

Assessment and Interventions

Health Literacy

The Brief Health Literacy Screening Tool [26] was selected to measure change before and after the health literacy intervention video. This questionnaire has four items that are rated on a 5-point Likert scale from "always" to "never": (1) How often do you have someone help you read hospital materials? (2) How often do you have problems learning about your medical condition because of difficulty understanding written information? (3) How often do you have a problem understanding what is told to you about your medical condition? (4) How confident are you filling out medical forms by yourself? Two questions were asked in addition to the Brief Health Literacy Screening Tool using a 10-point scale: "How much do you know about your own health," "How much do you know about clinical trials," and 1 true or false question: "Do you know your body mass index (BMI)?" (Multimedia Appendix 1).

Before showing a video, a module with a knowledge check portion was included to facilitate pre- and postvideo knowledge comparison. St. Luke's University Health Network's video, titled "Wellness 101 – How to Improve Your Overall Health" [27] was chosen for the content area of health literacy. The video provided 5 tips to improve an individual's overall health. Five knowledge-related questions based on the video were asked. A BMI calculator was included as the final task in this section. Participants were instructed to "Try calculating your BMI on this website" using the CDC BMI calculator [28].

Technology Competence

We measured internet skills using a section of the Technology Acceptance Model 3 [29]. Statements presented to the participant included "The study website does not scare me at all," "Working with the study website makes me nervous," "The study website makes me feel uncomfortable," and "The study website makes me feel uneasy." We also asked, "While using the study website, I'm worried that I might press the wrong button and make a mistake that crashes the program" and "I am most comfortable using my (multi-select) iPad/Tablet, Smart Phone (iPhone or Android), Computer, Other."

The video we used to increase technology competence was created in-house and showed participants examples of how to click on checkboxes, "radio button" response buttons, or move the cursor to a particular spot to answer different types of questions, including multiple choice, multiple selections, and a sliding scale. Participants were then asked to practice answering the same type of questions on their own (Multimedia Appendix 1 for the Technology Competence Questionnaire).

Clinical Trials

We also asked the question, "When it comes to your feelings about participating in this study, how do you rate your comfort?" using a scale ranging from "0, meaning no distress; totally relaxed" to "100, reflecting the highest anxiety/distress that you have ever felt" (Figure S1 in Multimedia Appendix 1). We showed the video from the National Institute of Diabetes and Digestive and Kidney Diseases, "Why Should I Join a Clinical Trial?" [30].

Semistructured Interviews

RM and AB conducted 45-minute interviews using Zoom (Zoom Technologies Inc) video conferencing with all 6 study participants using semistructured guides. Open-ended questions explored how participants felt about the process of using the app and how they felt about the questions that were asked through the app. A 5-point Likert scale was used to determine whether they agreed or disagreed with several statements focusing on how useful and informational they felt each of the videos and questionnaires were, for example, "I found the knowledge check questions to be useful" and "I felt less anxious about the idea of participating in a clinical trial after completing the knowledge check."

Data Analysis

Given the small sample size and our product testing approach, we used simple descriptive statistics to give us insights into the differences between the "before" and "after" questionnaire results. The semistructured interviews were reviewed for commonalities.

What We Found

User Demographics

Users testing the smartphone apps ranged in age from 54 to 69 years. Of the 6 participants, 3 identified as female and 3 as male. Five stated their race as White, 1 selected Black or African American, and none identified as Hispanic. Participants lived in the United States and Europe; 5 owned their homes and 1 rented. The majority (n=5) indicated they had 1 or more health conditions, and 5 had completed at least 4 years of college. All participants indicated they preferred to receive notifications

XSL•FO
RenderX

before noon, 4 preferred SMS text message notifications, and 2 preferred email notifications.

Health Literacy

The mean score and SD for each survey item before and after viewing the instructional video are listed in Table S1 in Multimedia Appendix 1. Lower scores indicated a more positive response for 2 of the questions, specifically, "How confident are you filling out medical forms by yourself?" and "Do you know your Body Mass Index (BMI)?" The mean score for each survey item was slightly more positive after viewing the instructional video, except for the item "How often do you have a problem understanding what is told to you about your medical condition?" However, the SD was greater than the change in mean scores, indicating that it could be due to chance. There were no changes in questionnaire responses before and after watching the instructional video for 3 of the 6 participants, improvement in 2 questions and a decline in 1 question for 1 participant, improvement in 4 questions and no change in 3 questions for 1 participant, and improvement in knowledge of their overall health and a decline in the clinical trial knowledge item for 1 participant.

Participants' feedback from the semistructured interviews revealed negative feelings toward the video "Wellness 101 – How to Improve Your Overall Health." One participant described the video as "juvenile," while another noted concerns that some participants might object to the health video if they already smoke or have a high BMI. All participants agreed or strongly agreed to the question, "I found the knowledge check questions to be useful." In total, 4 of the 6 participants neither agreed nor disagreed with the statement, "I felt less anxious about the idea of participating in a clinical trial after completing the knowledge check." Participants liked the alternative health literacy video, "5 Ways to Make the Most of Your Doctor Visit" [31].

Technology Competence

The mean scores for all items were more positive after viewing the video. The scores for 3 participants improved after watching the instructional video but declined for 2 participants. There was no change for 1 participant (Table S1 in Multimedia Appendix 1).

Participants (N=6) shared positive feedback about the video showing how to answer and practice questions. When asked, "Watching somebody else demonstrate how to answer questions made me feel like I knew what was expected of me in the study," participants answered, "agree" (2 participants) and "strongly agree" (4 participants). Additionally, participants answered "strongly agree" (4 participants) to the survey, "Practicing answering questions on my own made me feel less anxious about participating in the study." Some participants felt the questions in this section were redundant and thought all the questions could be combined into 1 question. In addition, participants thought the questions felt negative with the emphasis on the terms "anxious" and "nervous" and suggested changing the questions to make them seem more positive. One participant suggested making the technology anxiety section optional for those who feel more comfortable using the study website.

Clinical Trials

Mean scores were slightly higher prior to viewing the instructional video compared with after the video. One participant improved by 10 points, 1 decreased by 10 points, and the other 4 stayed the same.

Participants had positive things to say about the video "Why Should I Join a Clinical Trial?" [30] When asked, "I found the video to be useful," 2 participants answered "agree" and 4 answered "strongly agree."

Additional Overall User Feedback

The participants had several specific suggestions. One participant suggested making sure that the videos were clearly specific to diseases or therapy areas in the trial and gave specific information on the trial structure. Two participants suggested including additional content on participant safety. We did an ad hoc assessment of this suggestion and sent the National Cancer Institute's "Patient Safety in Clinical Trials" video for feedback [32]. The majority (n=5) participants liked the video. The National Institute of Mental Health's video, "What are the risks and benefits of participating in clinical research?" was also considered [33]. Most (n=4) participants liked the video and 2 did not. Some participants thought the video was juvenile and better for children or young adults. Others also thought the video did not explain concepts such as placebos well. In response to this feedback, the following resources from the National Institute on Aging were added: "What are Clinical Trials and Studies?" [34] and "Clinical Research Benefits, Risks and Safety" [35], after which we sought a second round of feedback. Most participants (n=5) liked the additional resources.

Discussion

Key Lessons

Maintaining continuous and complete use of mHealth apps has remained a persistent problem that has not yielded even sophisticated solutions such as timed and individualized user messaging. A newer and evolving understanding of the foundational importance of user engagement with mHealth suggests that this problem comes from a lack of appreciation by mHealth app designers of the complex and multicomponent structures behind user engagement. We have built on prior knowledge and work to develop a model of engagement that accounts for the complexity of engagement [24].

This viewpoint was an exploratory study to determine the feasibility of this approach and to guide the refinement of this interactive test strategy. We learned several key lessons: (1) Specificity—the participants endorsed the recommendation that the interventions should be specific to the educational needs of the target of the mHealth app. The most positive feedback was given to the video we developed de novo to teach participants the technical competence required to correctly and effectively use the app to report their evaluations. (2) Attention to inadvertent adverse affective variables—the participants noted the importance of avoiding or rephrasing medical terms that

XSL•FO
RenderX

could be seen as demeaning by some participants (eg, obesity, age, and infirmity). (3) Individualization—the participants clearly reflected different levels of need for improving their technical competence and health literacy. Our results indicate the potential importance of personalization in health app design addressing individuals' levels of need and cultural and personal sensitivities. For example, a way to allow more individualization is to allow users to potentially opt out of certain learning features if they do not think they need them.

Comparison With Prior Work

Other studies have assessed how to adjust apps to increase engagement. One study found positive effects on adherence from personalization or tailoring of the app content to users' needs, push notification reminders, user-friendly design, and personal support along with digital intervention [36]. However, the high dropout rate in app usage remains a major challenge [37]. A recent literature review found that despite factors such as appropriate reminders and feedback, app participants were more likely to drop out than be retained [20]. App literacy skills have been identified as a major factor in the uptake and engagement of smartphone apps [38]. Although we identified studies recommending web-based interventions to increase health literacy and technical skills [39,40], we have not found other studies testing approaches to increase those skills.

Limitations and Strengths

Our sample size was limited to 6 participants, 5 of whom were highly educated. We were unable to use any statistical significance measures because of the small sample size or draw conclusions that would apply to a larger population. Our highly educated sample is a limitation because these individuals may have better digital literacy than the general population. The main goal of our series of mini-experiments was to assess the feasibility of our approach and see whether we could retain the interest of the participants and obtain useful feedback on the interventions, which was successful. This type of testing is intended to evaluate tailored iterations of the app after gathering rapid participant feedback, with the ultimate goal of developing an app that will engage users. The next phase of our work will be to undertake the systematic testing of each component of this model in a larger and more diverse sample. We will then be able to refine the interventional enhancements for those components for a broader population. Ultimately, the functional use of this approach requires much larger, population-specific samples.

Conclusions

To fulfill the promise of mHealth apps to improve health outcomes, apps need to be improved so they reduce participant attrition. Health care apps do not work for people who do not use them. To date, app feedback, notifications and reminders, in-app support, gamification, and participant compensation have not been consistently successful in eliminating participant dropout. This study highlights the potential to develop and refine mHealth apps using evidence-based interventions derived from a broad range of behavioral and social science to increase engagement as a way of improving participant retention. This viewpoint highlights the potential for mHealth researchers to test and refine mHealth apps using approaches to better engage users. The preliminary experience reported in this viewpoint supports the feasibility of this iterative approach to create and refine engagement interventional enhancements for each element of the multidimensional, multicomponent theory of the engagement process.

Acknowledgments

We thank members of the Medable Patient Care Network for their valuable input and critical insights. No artificial intelligence products were used to develop questionnaires, interventions, analysis, or write any of this manuscript. Research reported in this publication was partially supported by the National Cancer Institute of the National Institutes of Health under contract numbers HHSN261201700030C and HHSN261201800010C. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the data release requirements of the Patient Care Network and the intellectual property associated with this work but are available from the corresponding author on reasonable request.

Authors' Contributions

RM, AB, JPD, and IO-G conceived the mini-experiments designed to test the model developed by JPD and IO-G. RM and AB developed the app to collect data, coordinated with the Patient Care Network, and collected and analyzed the data. RM, JPD, IO-G, and SWD wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflicts of Interest

RM, JPD, SWD, and AB were employed by Medable at the time the data collection and manuscript writing was done. ML and IOG are currently employed at Medable. Medable is a clinical trial software-as-a-service platform and evidence-generation company and supported the authors' conduct of this work without interference.

XSL•FO
RenderX

Multimedia Appendix 1
Study questionnaires and responses.
[DOCX File , 158 KB - ijmr_v13i1e51974_app1.docx]

References

1. Newzoo global mobile market report 2021. Newzoo. 2021. URL: https://newzoo.com/insights/trend-reports/newzoo-global-mobile-market-report-2021-free-version [accessed 2023-07-18]
2. Purnell JQ, Griffith J, Eddens KS, Kreuter MW. Mobile technology, cancer prevention, and health status among diverse, low-income adults. Am J Health Promot 2014;28(6):397-402 [FREE Full text] [doi: 10.4278/ajhp.120816-ARB-396] [Medline: 24200336]
3. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
4. Dorsey ER, Kluger B, Lipset CH. The new normal in clinical trials: decentralized studies. Ann Neurol 2020;88(5):863-866. [doi: 10.1002/ana.25892] [Medline: 32869367]
5. Sarraju A, Seninger C, Parameswaran V, Petlura C, Bazouzi T, Josan K, et al. Pandemic-proof recruitment and engagement in a fully decentralized trial in atrial fibrillation patients (DeTAP). NPJ Digit Med 2022;5(1):80 [FREE Full text] [doi: 10.1038/s41746-022-00622-9] [Medline: 35764796]
6. Sessa C, Cortes J, Conte P, Cardoso F, Choueiri T, Dummer R, et al. The impact of COVID-19 on cancer care and oncology clinical research: an experts' perspective. ESMO Open 2022;7(1):100339 [FREE Full text] [doi: 10.1016/j.esmoop.2021.100339] [Medline: 34953404]
7. Patel S, Goldsack JC, Cordovano G, Downing A, Fields KK, Geoghegan C, et al. Advancing digital health innovation in oncology: priorities for high-value digital transformation in cancer care. J Med Internet Res 2023;25:e43404 [FREE Full text] [doi: 10.2196/43404] [Medline: 36598811]
8. Davis SW, Oakley-Girvan I. mHealth education applications along the cancer continuum. J Cancer Educ 2015;30(2):388-394. [doi: 10.1007/s13187-014-0761-4] [Medline: 25482319]
9. DiMasi JA, Smith Z, Oakley-Girvan I, Mackinnon A, Costello M, Tenaerts P, et al. Assessing the financial value of decentralized clinical trials. Ther Innov Regul Sci 2023;57(2):209-219 [FREE Full text] [doi: 10.1007/s43441-022-00454-5] [Medline: 36104654]
10. Bruce CR, Harrison P, Nisar T, Giammattei C, Tan NM, Bliven C, et al. Assessing the impact of patient-facing mobile health technology on patient outcomes: retrospective observational cohort study. JMIR Mhealth Uhealth 2020;8(6):e19333 [FREE Full text] [doi: 10.2196/19333] [Medline: 32589161]
11. Tuvesson H, Eriksén S, Fagerström C. mHealth and engagement concerning persons with chronic somatic health conditions: integrative literature review. JMIR Mhealth Uhealth 2020;8(7):e14315 [FREE Full text] [doi: 10.2196/14315] [Medline: 32706686]
12. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
13. Druce KL, Dixon WG, McBeth J. Maximizing engagement in mobile health studies: lessons learned and future directions. Rheum Dis Clin North Am 2019;45(2):159-172 [FREE Full text] [doi: 10.1016/j.rdc.2019.01.004] [Medline: 30952390]
14. Kaveladze BT, Wasil AR, Bunyi JB, Ramirez V, Schueller SM. User experience, engagement, and popularity in mental health apps: secondary analysis of app analytics and expert app reviews. JMIR Hum Factors 2022;9(1):e30766 [FREE Full text] [doi: 10.2196/30766] [Medline: 35099398]
15. Serrano KJ, Coa KI, Yu M, Wolff-Hughes DL, Atienza AA. Characterizing user engagement with health app data: a data mining approach. Transl Behav Med 2017;7(2):277-285 [FREE Full text] [doi: 10.1007/s13142-017-0508-y] [Medline: 28616846]
16. Trifan A, Oliveira M, Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. JMIR Mhealth Uhealth 2019;7(8):e12649 [FREE Full text] [doi: 10.2196/12649] [Medline: 31444874]
17. Kelders SM, van Zyl LE, Ludden GDS. The concept and components of engagement in different domains applied to eHealth: a systematic scoping review. Front Psychol 2020;11:926 [FREE Full text] [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
18. Madujibeya I, Lennie T, Aroh A, Chung ML, Moser D. Measures of engagement with mHealth interventions in patients with heart failure: scoping review. JMIR Mhealth Uhealth 2022;10(8):e35657 [FREE Full text] [doi: 10.2196/35657] [Medline: 35994345]
19. Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, et al. A library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions: scoping review. JMIR Mhealth Uhealth 2019;7(1):e11941 [FREE Full text] [doi: 10.2196/11941] [Medline: 30664463]

XSL•FO
RenderX

20. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022;24(4):e35120 [FREE Full text] [doi: 10.2196/35120] [Medline: 35471414]
21. Mustafa AS, Ali N, Dhillon JS, Alkawsi G, Baashar Y. User engagement and abandonment of mHealth: a cross-sectional survey. Healthcare (Basel) 2022;10(2):221 [FREE Full text] [doi: 10.3390/healthcare10020221] [Medline: 35206837]
22. Oakley-Girvan I, Lavista JM, Miller Y, Davis S, Acle C, Hancock J, et al. Evaluation of a mobile device survey system for behavioral risk factors (SHAPE): app development and usability study. JMIR Form Res 2019;3(1):e10246 [FREE Full text] [doi: 10.2196/10246] [Medline: 30684441]
23. Oakley-Girvan I, Yunis R, Longmire M, Ouillon JS. What works best to engage participants in mobile app interventions and e-health: a scoping review. Telemed J E Health 2022;28(6):768-780 [FREE Full text] [doi: 10.1089/tmj.2021.0176] [Medline: 34637651]
24. Oakley-Girvan I, Docherty JP. A new approach to enhancing engagement in eHealth apps. Interact J Med Res 2022;11(2):e38886 [FREE Full text] [doi: 10.2196/38886] [Medline: 36279587]
25. Lepore SJ, Rincon MA, Buzaglo JS, Golant M, Lieberman MA, Bass SB, et al. Digital literacy linked to engagement and psychological benefits among breast cancer survivors in internet-based peer support groups. Eur J Cancer Care (Engl) 2019;28(4):e13134 [FREE Full text] [doi: 10.1111/ecc.13134] [Medline: 31318132]
26. Haun J, Noland-Dodd V, Varnes J, Graham-Pole J, Rienzo B, Donaldson P. Testing the BRIEF health literacy screening tool. Fed Pract 2009;26(12):24-31 [FREE Full text]
27. Wellness 101 – how to improve your overall health. St. Luke's University Health Network YouTube page. 2019. URL: https://www.youtube.com/watch?v=_nuDp-fded8 [accessed 2024-03-01]
28. Adult BMI calculator. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/english_bmi_calculator/bmi_calculator.html [accessed 2023-07-18]
29. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315 [FREE Full text] [doi: 10.1111/j.1540-5915.2008.00192.x]
30. Rodgers GP. Why should I join a clinical trial? National Institute of Diabetes and Digestive and Kidney Disease YouTube page. 2019. URL: https://www.youtube.com/watch?v=36Sd8WpgR94 [accessed 2024-03-01]
31. 5 ways to make the most of your doctor visit. National Institute on Aging YouTube page. URL: https://www.youtube.com/watch?v=BincCVl-YsI [accessed 2023-07-18]
32. Are clinical trials safe? National Cancer Institute. URL: https://www.cancer.gov/about-cancer/treatment/clinical-trials/patient-safety [accessed 2023-07-19]
33. What are the risks and benefits of participating in clinical research? National Institute of Mental Health. URL: https://www.nimh.nih.gov/news/media/2021/what-are-the-risks-and-benefits-of-participating-in-clinical-research [accessed 2023-07-19]
34. What are clinical trials and studies? National Institute on Aging. URL: https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies#:~:text=Clinical%20trials%20are%20research%20studies,safe%20and%20effective%20in%20people [accessed 2023-07-19]
35. Clinical research: benefits, risks, and safety. National Institute on Aging. URL: https://www.nia.nih.gov/health/clinical-research-benefits-risks-and-safety [accessed 2023-07-19]
36. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022;24(5):e35371 [FREE Full text] [doi: 10.2196/35371] [Medline: 35612886]
37. Sim I. Mobile devices and health. N Engl J Med 2019;381(10):956-968. [doi: 10.1056/NEJMra1806949] [Medline: 31483966]
38. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
39. Moon Z, Zuchowski M, Moss-Morris R, Hunter MS, Norton S, Hughes LD. Disparities in access to mobile devices and e-health literacy among breast cancer survivors. Support Care Cancer 2022;30(1):117-126 [FREE Full text] [doi: 10.1007/s00520-021-06407-2] [Medline: 34236506]
40. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]

Abbreviations

mHealth: mobile health
PCN: Patient Care Network

XSL•FO
RenderX

Edited by G Eysenbach, T de Azevedo Cardoso; submitted 18.08.23; peer-reviewed by S Amagai, B Oh, C Latkin; comments to author 17.10.23; revised version received 14.11.23; accepted 27.02.24; published 26.03.24.

Please cite as:
Monachelli R, Davis SW, Barnard A, Longmire M, Docherty JP, Oakley-Girvan I
Designing mHealth Apps to Incorporate Evidence-Based Techniques for Prolonging User Engagement
Interact J Med Res 2024;13:e51974
URL: https://www.i-jmr.org/2024/1/e51974
doi:10.2196/51974
PMID:38416858

©Rebecca Monachelli, Sharon Watkins Davis, Allison Barnard, Michelle Longmire, John P Docherty, Ingrid Oakley-Girvan. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 26.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Using Routine Data to Improve Lesbian, Gay, Bisexual, and Transgender Health

Catherine L Saunders[1], MBBS, PhD

Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Catherine L Saunders, MBBS, PhD
Department of Psychiatry
University of Cambridge
Herchel Smith Building, Forvie Site
Robinson Way, Cambridge Biomedical Campus
Cambridge, CB2 0SZ
United Kingdom
Phone: 44 1223337106
Email: cs834@medschl.cam.ac.uk

Abstract

The collection of sexual orientation in routine data, generated either from contacts with health services or in infrastructure data resources designed and collected for policy and research, has improved substantially in the United Kingdom in the last decade. Inclusive measures of gender and transgender status are now also beginning to be collected. This viewpoint considers current data collections, and their strengths and limitations, including accessing data, sample size, measures of sexual orientation and gender, measures of health outcomes, and longitudinal follow-up. The available data are considered within both sociopolitical and biomedical models of health for individuals who are lesbian, gay, bisexual, transgender, queer, or of other identities including nonbinary (LGBTQ+). Although most individual data sets have some methodological limitations, when put together, there is now a real depth of routine data for LGBTQ+ health research. This paper aims to provide a framework for how these data can be used to improve health and health care outcomes. Four practical analysis approaches are introduced—descriptive epidemiology, risk prediction, intervention development, and impact evaluation—and are discussed as frameworks for translating data into research with the potential to improve health.

(Interact J Med Res 2024;13:e53311) doi:10.2196/53311

KEYWORDS

lesbian; gay; bisexual; trans; LGBTQ+; routine data; England; United Kingdom; health; viewpoint; sexual orientation; health services; infrastructure data; policy; gender; health outcome; epidemiology; risk prediction; risk

Introduction

Research into health for individuals who are lesbian, gay, bisexual, transgender, queer, or of other identities including nonbinary (LGBTQ+) has consistently found that these populations experience poorer outcomes [1], with particularly strong and consistent evidence around poorer mental health for lesbian, gay, and bisexual adults [2]. LGBTQ+ health research typically uses 2 broad frameworks, sociopolitical or psychosocial models (where experiences of discrimination, victimization, stigma, and harassment are central to our understanding of health), and models of clinical, biomedical, and lifestyle risk factors [3]. Sexual health and HIV research falls mostly within biomedical frameworks; in practice, inequalities are interconnected [4,5]. Experiences of discrimination, harassment, and stigma have a profound impact on clinical health outcomes for LGBTQ+ communities [6]. Poorer health care access and quality among vulnerable groups exacerbate these impacts [7,8].

Data-driven LGBTQ+ health research has historically been based on convenience or purposive, rather than population-based samples [9], even in large-scale studies such as the National LGBT Survey in 2018 [1]. The collection of data on sexual orientation, gender, and transgender status has not been prioritized in routine sources [3]. Challenges in identifying understandable, meaningful, and acceptable measures, and concerns about LGBTQ+ respondents being able to answer questions safely, have been additional barriers to data collection [10].

However, data are improving [11]. The Equality Act in 2010 placed a statutory duty on public bodies in the United Kingdom

XSL•FO
RenderX

to publish equality outcomes and report on progress in addressing disadvantage experiences by (among other characteristics) sexual orientation and gender reassignment [12], which had a strong positive impact. The Office for National Statistics has carried out development work on measures [13], and although the census in 2011 did not include questions about sexual orientation and transgender status, in part because of the concerns about respondent safety [10], by 2021 questions on both were included, with instruction statements that the questions were voluntary. In general, data collection for sexual orientation is more established than collection of data on transgender status and gender, although, for both, collections are improving.

This paper aims to provide a guide for how these improvements in routine data collection can potentially translate to improved health and health care outcomes.

Pathways From Data to Improved Health

LGBTQ+ health research using routine data sits within wider health data science research frameworks which are designed to leverage person-level routine health data (collected either from routine contacts with services or from infrastructure data resources designed and collected for policy and research) to improve outcomes [14]. Analyses fall under four broad translational pathways: (1) descriptive epidemiology, (2) risk prediction, (3) informing innovation and improvement, and (4) impact evaluation.

These pathways have the potential to improve outcomes by (1) providing evidence to inform policy and practice (descriptive epidemiology), (2) better targeting of interventions and understanding of population health needs (risk prediction), (3) more rational health service developments (intervention development), and (4) information on effectiveness informing commissioning or funding decisions (impact evaluation), respectively.

Health data science as a field has struggled with equality, diversity, and inclusion [15] and there are problems across the whole discipline. Algorithmic biases in risk prediction models, including in how models are developed, with differentially poorer functioning for minoritized groups, or inequitable outcomes when the models are implemented, are currently a particular area of concern [16,17]. In addition, missing data contribute not only to poorer risk model development, but to a lack of basic descriptive epidemiology, informed intervention development, or equalities impact evaluation. For LGBTQ+ health research using routine data, the pathways to improved outcomes are the same, as are the challenges of missing data [18].

Data

There are 5 groups or types of routine UK data sources—where information about sexual orientation and gender or transgender status are either well established or now starting to be recorded, beginning to address this lack of data. In the same way that LGBTQ+ health research balances both societal and biomedical models the data sources, which now include a collection of sexual orientation (more likely) or gender and transgender status (beginning to be introduced) reflect a balance of routine data from social and health sources. The five groups are (1) social science or societal data collections (including Understanding Society [19], birth cohort studies [20], educational cohort studies [21], and census data); (2) general and specific health surveys primarily designed to understand population health (including the Health Survey for England [22] and the National Survey of Sexual Attitudes and Lifestyles (NATSAL) [23]); (3) health services or patient surveys primarily designed for health service quality improvement (including the General Practice Patient Survey [24,25] and the Cancer Patient Experience Survey [26]); (4) health cohort studies (UK Biobank [27] and Our Future Health [28]); and (5) health records (including primary care research databases such as the Clinical Practice Research Datalink [29,30], and secondary and community services data sets, including the improving access to psychological therapies and mental health services data sets [31], and registry data, for example, cancer registry data [32]).

Put together, the data are starting to form a comprehensive collection but for each resource, there are strengths and limitations or challenges. For example, the Equality and Human Rights Commission was able to draw on quantitative evidence and data about sexual orientation and gender reassignment in work, education, and health in the State of the Nation report on equality and human rights in Britain published in November 2023 [33]. However, data access, sample size, measures of sexual orientation, gender, and health outcomes, and the ability to carry out longitudinal analysis and data quality vary across sources.

In terms of access to data, social science collections are primarily accessed without cost through the UK Data Archive; for sensitive fields, which often include sexual orientation or gender and transgender status, additional safeguards are in place. The UK Biobank and Our Future Health are 2 large biomedical data research cohorts accessed through trusted research environments (secure data hosting platforms) with relatively low but nonzero costs to researchers [28,34]. For all sources, access to data can require time and perseverance [11]. In terms of longitudinal follow-up, the UK Biobank is a mature cohort study, for which recruitment began in 2006 before sexual orientation and transgender status were routinely collected. Questions are included instead about sexual history, which provides some insight [34]. In contrast, Our Future Health for which recruitment began in 2022 has an inclusive gender question and questions about both sexual history and sexual orientation but only baseline data collection to date (recruitment is ongoing) [28].

Sample size is often a trade-off with detail. Understanding Society is a household panel survey designed to provide estimates about how life in the United Kingdom is changing and what stays the same over many years, with linked health and social data [19,35]. In common with other longitudinal and cohort data collections, the sample size is relatively small (about 40,000 people at baseline), compared, for example, to the General Practice Patient Survey which is a large cross-sectional survey designed to evaluate health care quality, which has a much larger sample size (about 700,000 responses) but with

XSL•FO
RenderX

much less nuanced health and particularly, social measures recorded.

Pooling data across sources is an approach to increasing sample size [36], and again resources are improving, estimates across an in-depth range of health outcomes from the Health Survey for England using data from across 7 years have recently been published [22].

Data from electronic health records (EHRs), or data routinely recorded as part of clinical or health care encounters offer both detailed health outcome data and large sample size. The challenge is often that for EHR data collections, in contrast with research data infrastructure resources, or survey-based health data collections, measurement of sexual orientation and gender or transgender status is less good. In part, this is because these are resources not designed for research but primarily collected as clinical documentation. Pilots have begun to improve recording, to support audit and quality improvement. However, given both historic discrimination experienced by the LGBTQ+ community based on sexual orientation recording in medical records, and the interpersonal interlinking of recording and coming out to a care provider [37,38], this solution to data improvement is not simply a neutral administrative process. The reluctance of health care providers to ask about sexual orientation is a second barrier [39]. Recent research using EHR has provided insight by looking at transgender patients in primary care records based on prescribing and clinical codes, and this is an exciting area of progress [29,30]. These approaches have their own challenges, however, with historic clinical codes including outdated and discriminatory terminology still present in some older coded EHR records [18]. Legal barriers to identifying transgender patients after transition provide a further barrier to research using EHR; legislative changes have been required for recent quantitative analyses [40].

There are notable areas where data are poorer. Data governance and ethical challenges mean that data collections are much less likely to collect information on sexual orientation or gender from children. For example, some research studies have used proxies or less detailed response options where exact measures of ethnicity, gender identity, or disability cannot be asked [41]. Although HIV and sexual health research are well-studied topics in LGBTQ+ health research overall [42], routine data are usually more strongly safeguarded and less available for research, although measured in some collections.

Applied Methodology

The improvement in data collection for LGBTQ+ health research in the last decade mean that the applied methodological research around the use of these data is also developing. Questions about the longitudinal consistency of self-reported sexual orientation and history have been explored; changes are more frequently reported at younger ages [34,43,44]. For sexual orientation, missing data have reduced over time since the question has been routinely introduced in surveys [45]. Secular trends are also being better understood [44,45], meaning that age, period, and cohort effects in LGBTQ+ health research can begin to be untangled [24]. Differential item functioning for new questions among groups for whom English is a second language is a current area of concern for new gender questions, although this is unlikely to be a methodological issue specific to these particular items. The challenges of longitudinal consistency in question wording needing to be balanced against requirements for relevant and up-to-date survey items is again a methodological challenge not specifically limited to questions about sexual orientation and gender. New, nuanced, tools for understanding gender are beginning to be developed [46]; however, space constraints in surveys mean that often only single items are asked. While free text or more in-depth response options (or allowing multiple rather than single responses) are more inclusive [47,48], these nuanced data are often excluded from quantitative reporting. Data for people who identify as asexual are very limited, as are data for people with variations in sexual characteristics.

How Have These Data Translated Into Applied Research?

As LGBTQ+ routine data are improving, the insights that come from descriptive epidemiological LGBTQ+ health research are also developing. For example, historically, studies using routine data have been able to consider cancer risk factors such as, smoking and alcohol consumption [49-51], more easily than rarer cancer outcomes such as incidence. Limited sample size and poorer measurement of outcomes mean that earlier studies looking at cancer were cross-sectional and could only consider cancer prevalence without disaggregation by diagnosis [52]. Larger cross-sectional data sets have allowed disaggregation of diagnoses among lesbian, gay, and bisexual patients with cancer, identifying disparities primarily in HIV and human papillomavirus–associated cancers [26]. More recent work has for the first time in the United Kingdom been able to look at the impact of higher smoking prevalence identified in earlier studies on lung cancer incidence, using the UK Biobank resource [53], connecting both biomedical and sociopolitical frameworks; the LGBTQ+ community has historically been targeted by tobacco marketing.

Inequalities in LGBTQ+ mental health outcomes have also been well established through a series of studies and meta-analyses using routine data from the United Kingdom [54]. In our recent work collaboratively exploring LGBTQ+ research priorities, intersectionality (understanding the interdependent and overlapping systems of discrimination and disadvantage) was identified as an area of research need; and race, ethnicity, and socioeconomic inequalities were particularly highlighted [55]. Larger sample sizes mean that intersectionality can now begin to be explored quantitatively [56]; newer longitudinal collections are providing additional insight [57].

Again, it is not just biomedical models that are important. Routine educational data sets have been important in highlighting the higher levels of bullying experienced by young LGBTQ+ people in schools [21].

However, we also know that on its own research describing inequalities experienced by LGBTQ+ adults will not lead directly to improved outcomes. Process measures of care quality

XSL•FO
RenderX

are often easier to improve than more tangible health outcomes. However, although disparities in primary care access, communication, and satisfaction were measured routinely between 2011 and 2017, inequalities experienced by lesbian, gay, and bisexual adults persisted across the time period [58].

Impact evaluation is a second pathway, therefore, where routine data are beginning to be used to provide insight with the potential to change the care process and improve LGBTQ+ health outcomes. The collection of sexual orientation information in the Improving Access to Psychological Therapies data set has allowed inequalities evaluation of these services for lesbian, gay, and bisexual adults, finding that they were not as effective as for heterosexual service users [31]. In contrast, an inequalities evaluation of the introduction of telephone triage in general practices using the GP Patient Survey found that although there was variation between practices in outcomes, for different groups of patients within the same practice, including lesbian, gay, and bisexual adults, there was no evidence of differential impact on access to primary care [59]. The Millennium Cohort Study has been used to understand the differential impact of the COVID-19 pandemic on sexual minority groups [57], as has Understanding Society [60].

Where to Next—Routine Data Analysis?

Routine data for LGBTQ+ health research are much better today, in 2024, than they were even 10 years ago. Sexual orientation has now been collected in many sources for over a decade and more diverse and inclusive gender measures are being introduced, and are established in some collections. Of course, measurement needs to continue and is continuing to improve, and there are limitations and barriers; no data set alone is perfect. However, across the spectrum of sources, there is a real depth of data now available and in terms of research, the data are good enough now to at least start thinking properly about how we can use these resources to improve LGBTQ+ health and tackle inequalities.

In terms of data development, of course linkage is 1 exciting potential future avenue, with the linked 2011 census and routine health care data in Scotland providing a possible model for future development. But in reality, using routine data for LGBTQ+ health research lies within the wider UK research landscape for using routine data overall. Here the Goldacre review probably shines some light on the direction of travel [61]. Access is becoming more cautious, and data are becoming more securely safeguarded, new frameworks and solutions are needed to ensure that access continues and barriers do not increase [62]. For sensitive fields such as sexual orientation, gender, and transgender status, this is particularly important, but it is likely that time, patience, and perseverance are going to continue to be required when working in this space. As a balance to concerns about the use of person-level data, tools sharing aggregate data such as the census resources from the United Kingdom's Office for National Statistics [63], and the analysis tool for the GP Patient Survey remain important resources and provide real insight.

So, the question remains, how are we going to use these data to improve LGBTQ+ health? Although the data are better, the approaches have not changed and the methodological answers to the pathways from data to improved health remain the same. The four pathways are (1) descriptive epidemiology, (2) risk prediction, (3) informing innovation and improvement, and (4) impact evaluation.

Given the recentness of the data improvements and that data resources are still improving, there remains a real need for basic epidemiological descriptive work using these new data to answer questions and provide insight where simply the data have not been available before. More in-depth analyses, analyses considering longitudinal changes, and better measures of health and health outcomes, as well as sexual orientation and gender and transgender status, are all part of this. Frameworks for addressing health inequalities require researchers to go beyond simply describing known inequities [14], but for LGBTQ+ health there is still an evidence gap where descriptive epidemiology that focuses on areas where research could have an impact on policy has a place.

Maybe the results will be unsurprising, and research may show that inequalities have not disappeared as the data have improved, but the work is still important, and insight is still needed.

Risk prediction as a field has real challenges ahead to get to grips with equality, diversity, and inclusion, and this needs to include LGBTQ+ health. For transgender health specifically, there are some more questions to ask around risk model development; the exclusion of transgender adults from the development of some risk scores [47], and lack of clarity about how to implement scores based on binary gender or sex classifications are some specific issues to add to these [18]. Although methodological work is still needed to understand the best way to develop and implement risk scoring for transgender patients to avoid potentially both under- and overtreatment, the critical first step is to ensure that data used for model development do not exclude transgender populations before the research begins.

In terms of intervention development and audit, the improvement of data is important to ensure that evidence-based interventions are developed and part of wider thinking about how routine data can improve health and LGBTQ+ health in particular. Specific clinical data sets, such as cancer data collections or more in-depth surveys such as NATSAL will be particularly important in this domain. Much local evaluation of LGBTQ+ health interventions remains qualitative [64], and the evidence base for health equity audits to address inequalities remains poor [65].

The importance of including explicit inequality analyses in impact evaluations remains a key analysis strategy for improving health. Even when interventions are not LGBTQ+ specific, there may or may not be an inequitable impact. This kind of routine equalities impact work for LGBTQ+ and other groups is central to the drive the Equality Act has given to the improvement in data that we have seen, and needs to become a routine part of evaluative work.

XSL•FO
RenderX

Where to Next—LGBTQ+ Health Research?

The data are good enough now for routine data to play a substantive part in LGBTQ+ health research, and there are clear and realistic pathways for how this research can potentially improve health. This comes within the wider context of flourishing LGBTQ+ health research overall [66]. Health and health care are complex [67]. It is not a linear pathway from data to improved health outcomes; but good research can play a part.

There are particular challenges for health research with LGBTQ+ children and young people, where data are often less frequently collected and ethical and governance considerations are particularly important, and there is an identified need for more research [68,69].

The co-option of research findings into homophobic or transphobic narratives is a further difficult area, as are avoiding some of the blind spots around equality, diversity, and inclusion in routine data research that are beginning to be identified particularly in risk prediction work [15]. Good communication and cautious interpretations of findings are part of the solution, as are patient and public involvement, and the involvement of LGBTQ+ communities in identifying research priorities and in carrying out research [55]. Best practice guidance for LGBTQ+ health research [70], inclusive public involvement [71], and involvement in LGBTQ+ health research [72] provide some signposts for researchers.

Conclusions

Descriptive epidemiology, risk prediction, informing innovation and improvement, and impact evaluation are 4 practical pathways from data to improved health. Data for LGBTQ+ health research are now good enough and improving. We know that health inequalities exist, within both societal and biomedical frameworks. Research with strong public involvement, good clear communication, and stakeholder involvement is key, as in all research. Overall, this is a positive story for routine data. We are at the stage where the analysis of routine data can contribute to making real practical steps toward informing policy and practice, better targeting of interventions and understanding of population health needs, more rational health service developments, informing commissioning or funding decisions, and improving LGBTQ+ health.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

None declared.

References

1. National LGBT survey: research report. Government Equalities Office. 2018. URL: https://www.gov.uk/government/publications/national-lgbt-survey-summary-report [accessed 2024-04-10]
2. King M, Semlyen J, Tai SS, Killaspy H, Osborn D, Popelyuk D, et al. A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. BMC Psychiatry 2008;8:70 [FREE Full text] [doi: 10.1186/1471-244X-8-70] [Medline: 18706118]
3. McDermott E, Nelson R, Weeks H. The politics of LGBT+ health inequality: conclusions from a UK scoping review. Int J Environ Res Public Health 2021;18(2):826 [FREE Full text] [doi: 10.3390/ijerph18020826] [Medline: 33478019]
4. Mendenhall E, Newfield T, Tsai AC. Syndemic theory, methods, and data. Soc Sci Med 2022;295:114656 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114656] [Medline: 34949486]
5. Tsai AC. Syndemics: a theory in search of data or data in search of a theory? Soc Sci Med 2018;206:117-122 [FREE Full text] [doi: 10.1016/j.socscimed.2018.03.040] [Medline: 29628175]
6. Meyer IH. Minority stress and mental health in gay men. J Health Soc Behav 1995;36(1):38-56. [Medline: 7738327]
7. Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Med Res Methodol 2006;6:35 [FREE Full text] [doi: 10.1186/1471-2288-6-35] [Medline: 16872487]
8. Elliott MN, Kanouse DE, Burkhart Q, Abel GA, Lyratzopoulos G, Beckett MK, et al. Sexual minorities in England have poorer health and worse health care experiences: a national survey. J Gen Intern Med 2015;30(1):9-16 [FREE Full text] [doi: 10.1007/s11606-014-2905-y] [Medline: 25190140]
9. Semlyen J, Hagger-Johnson G. Sampling frame for sexual minorities in public health research. J Public Health (Oxf) 2017;39(3):644 [FREE Full text] [doi: 10.1093/pubmed/fdw078] [Medline: 27613765]
10. McManus S. Sexual Orientation Research Phase 1: A Review of Methodological Approaches. Edinburgh: Scottish Executive Social Research; 2003.
11. Davies M, Moon G, Sabel CE. Navigating the LGB data landscape: a review of appropriate secondary data sources for sexuality and substance use research in the UK. Int J Environ Res Public Health 2022;19(3):1329 [FREE Full text] [doi: 10.3390/ijerph19031329] [Medline: 35162352]

XSL•FO
RenderX

12. Equality Act 2010. legislation.gov.uk. London: HMSO; 2010. URL: https://www.legislation.gov.uk/ukpga/2010/15/contents [accessed 2024-04-17]
13. Question development. Office for National Statistics. 2021. URL: https://www.ons.gov.uk/census/planningforcensus2021/questiondevelopment [accessed 2024-04-10]
14. Ford J, Sowden S, Olivera J, Bambra C, Gimson A, Aldridge R, et al. Transforming health systems to reduce health inequalities. Future Healthc J 2021;8(2):e204-e209 [FREE Full text] [doi: 10.7861/fhj.2021-0018] [Medline: 34286186]
15. Knight HE, Deeny SR, Dreyer K, Engmann J, Mackintosh M, Raza S, et al. Challenging racism in the use of health data. Lancet Digit Health 2021;3(3):e144-e146 [FREE Full text] [doi: 10.1016/S2589-7500(21)00019-4] [Medline: 33549513]
16. Robbins HA, Alcala K, Swerdlow AJ, Schoemaker MJ, Wareham N, Travis RC, et al. Comparative performance of lung cancer risk models to define lung screening eligibility in the United Kingdom. Br J Cancer 2021;124(12):2026-2034 [FREE Full text] [doi: 10.1038/s41416-021-01278-0] [Medline: 33846525]
17. Khor S, Haupt EC, Hahn EE, Lyons LJL, Shankaran V, Bansal A. Racial and ethnic bias in risk prediction models for colorectal cancer recurrence when race and ethnicity are omitted as predictors. JAMA Netw Open 2023;6(6):e2318495 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.18495] [Medline: 37318804]
18. Berner AM, Lund J, Saunders CL. Tackling the complexity of gender bias in primary care. Br J Gen Pract 2021;71(708):296-297 [FREE Full text] [doi: 10.3399/bjgp21X716177] [Medline: 34319880]
19. University of Essex, Institute for Social and Economic Research. Understanding society: 9th release. UK Data Service. 2023. URL: http://doi.org/10.5255/UKDA-Series-2000053 [accessed 2024-04-10]
20. University College London, UCL Institute of Education, Centre for Longitudinal Studies. Millennium cohort study: 15th release. UK Data Service. 2023. URL: http://doi.org/10.5255/UKDA-Series-2000031 [accessed 2024-04-10]
21. Henderson M. Understanding bullying experiences among sexual minority youths in England. Centre for Longitudinal Studies Working paper/8. 2015. URL: https://cls.ucl.ac.uk/wp-content/uploads/2017/04/CLS-WP-20158.pdf [accessed 2024-04-10]
22. Health survey England additional analyses - health and health-related behaviours of lesbian, gay and bisexual adults. NHS Digital. 2021. URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-england-additional-analyses/lesbian-gay-and-bisexual-adults [accessed 2024-04-10]
23. Russell M, NatCen Social Research, Senior M. National survey of sexual attitudes and lifestyles, 2010-2012: secure access. UK Data Service. 2021. URL: https://datacompass.lshtm.ac.uk/id/eprint/633/ [accessed 2024-04-10]
24. Saunders CL, Berner A, Lund J, Mason AM, Oakes-Monger T, Roberts M, et al. Demographic characteristics, long-term health conditions and healthcare experiences of 6333 trans and non-binary adults in England: nationally representative evidence from the 2021 GP patient survey. BMJ Open 2023;13(2):e068099 [FREE Full text] [doi: 10.1136/bmjopen-2022-068099] [Medline: 36731935]
25. Saunders CL, MacCarthy S, Meads C, Massou E, Mant J, Saunders AM, et al. Long-term conditions among sexual minority adults in England: evidence from a cross-sectional analysis of responses to the English GP patient survey. BJGP Open 2021;5(5):BJGPO.2021.0067 [FREE Full text] [doi: 10.3399/BJGPO.2021.0067] [Medline: 34465579]
26. Saunders CL, Meads C, Abel GA, Lyratzopoulos G. Associations between sexual orientation and overall and site-specific diagnosis of cancer: evidence from two national patient surveys in England. J Clin Oncol 2017;35(32):3654-3661 [FREE Full text] [doi: 10.1200/JCO.2017.72.5465] [Medline: 28945501]
27. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK biobank resource with deep phenotyping and genomic data. Nature 2018;562(7726):203-209 [FREE Full text] [doi: 10.1038/s41586-018-0579-z] [Medline: 30305743]
28. Our Future Health. 2023. URL: https://ourfuturehealth.org.uk/about-us/ [accessed 2024-04-10]
29. Brown J, Pfeiffer RM, Shrewsbury D, O'Callaghan S, Berner AM, Gadalla SM, et al. Prevalence of cancer risk factors among transgender and gender diverse individuals: a cross-sectional analysis using UK primary care data. Br J Gen Pract 2023;73(732):e486-e492 [FREE Full text] [doi: 10.3399/BJGP.2023.0023] [Medline: 37365010]
30. Jackson SS, Brown J, Pfeiffer RM, Shrewsbury D, O'Callaghan S, Berner AM, et al. Analysis of mortality among transgender and gender diverse adults in England. JAMA Netw Open 2023;6(1):e2253687 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.53687] [Medline: 36716027]
31. Rimes KA, Ion D, Wingrove J, Carter B. Sexual orientation differences in psychological treatment outcomes for depression and anxiety: national cohort study. J Consult Clin Psychol 2019;87(7):577-589. [doi: 10.1037/ccp0000416] [Medline: 31219292]
32. Henson KE, Elliss-Brookes L, Coupland VH, Payne E, Vernon S, Rous B, et al. Data resource profile: national cancer registration dataset in England. Int J Epidemiol 2020;49(1):16-16h [FREE Full text] [doi: 10.1093/ije/dyz076] [Medline: 31120104]
33. Equality and human rights monitor 2023. Equality and Human Rights Commission. 2023. URL: https://www.gov.uk/government/publications/equality-and-human-rights-monitor-2023 [accessed 2024-04-10]
34. Saunders CL, Massou E. Using UK Biobank for sexual minority health research. J Comp Eff Res 2021;10(13):1001-1010 [FREE Full text] [doi: 10.2217/cer-2021-0052] [Medline: 34231371]

XSL•FO
RenderX

35. Booker CL, Rieger G, Unger JB. Sexual orientation health inequality: evidence from understanding society, the UK longitudinal household study. Prev Med 2017;101:126-132 [FREE Full text] [doi: 10.1016/j.ypmed.2017.06.010] [Medline: 28601620]
36. Semlyen J. Recording sexual orientation in the UK: pooling data for statistical power. Am J Public Health 2017;107(8):1215-1217 [FREE Full text] [doi: 10.2105/AJPH.2017.303910] [Medline: 28657782]
37. Jamieson A, Cross H, Arthur S, Nambiar K, Llewellyn CD. Patient sexual orientation and gender identity disclosure. Clin Teach 2020;17(6):669-673 [FREE Full text] [doi: 10.1111/tct.13182] [Medline: 33217175]
38. Brooks H, Llewellyn CD, Nadarzynski T, Pelloso FC, Guilherme FDS, Pollard A, et al. Sexual orientation disclosure in health care: a systematic review. Br J Gen Pract 2018;68(668):e187-e196 [FREE Full text] [doi: 10.3399/bjgp18X694841] [Medline: 29378698]
39. Pollard A, Bradley J, Cooper M, Sadhwani S, Northam A, Llewellyn C. The NHS England fundamental information standard for monitoring sexual orientation. Br J Gen Pract 2019;69(679):94-95 [FREE Full text] [doi: 10.3399/bjgp19X701213] [Medline: 30705020]
40. Cass review- quantitative research. Cass Review. URL: https://cass.independent-review.uk/cass-review-quantitive-research/ [accessed 2024-04-10]
41. Mansfield KL, Puntis S, Soneson E, Cipriani A, Geulayov G, Fazel M. Study protocol: the OxWell school survey investigating social, emotional and behavioural factors associated with mental health and well-being. BMJ Open 2021;11(12):e052717 [FREE Full text] [doi: 10.1136/bmjopen-2021-052717] [Medline: 34880020]
42. Pakianathan M, Daley N, Hegazi A. Gay, bisexual, and other men who have sex with men: time to end the fixation with HIV. BMJ 2016;354:i4739. [doi: 10.1136/bmj.i4739] [Medline: 27590481]
43. Hu Y, Denier N. Sexual orientation identity mobility in the United Kingdom: a research note. Demography 2023;60(3):659-673 [FREE Full text] [doi: 10.1215/00703370-10769825] [Medline: 37257022]
44. MacCarthy S, Saunders CL, Elliott MN. Increased reporting of sexual minority orientation from 2009 to 2017 in England and implications for measuring sexual minority health disparities. LGBT Health 2020;7(7):393-400 [FREE Full text] [doi: 10.1089/lgbt.2019.0181] [Medline: 32758045]
45. Cross H, Llewellyn CD. A decline in patient disclosure of heterosexuality in the English general practice patient survey: a longitudinal analysis of cross-sectional data. Fam Pract 2020;37(5):661-667 [FREE Full text] [doi: 10.1093/fampra/cmaa033] [Medline: 32270180]
46. Yarrow E, Duschinsky R, Saunders CL. The gender variance scale: developing and piloting a new tool for measuring gender diversity in survey research. J Gend Stud 2023;33(2):244-256 [FREE Full text] [doi: 10.1080/09589236.2023.2277463]
47. Saunders CL, Lund J, Mason AM, Roberts M, Smith J, Duschinsky R. The health and healthcare outcomes of trans and/or non-binary adults in England: protocol for an analysis of responses to the 2021 GP patient survey. Sexes 2022;3(3):325-335 [FREE Full text] [doi: 10.3390/sexes3030025]
48. Elliott MN, Dahlhamer JM, MacCarthy S, Beckett MK, Orr N, Guerino P, et al. Using ancillary sociodemographic data to identify sexual minority adults among those responding "something else" or "don't know" to sexual orientation questions. Med Care 2019;57(12):e87-e95 [FREE Full text] [doi: 10.1097/MLR.0000000000001190] [Medline: 31415342]
49. Hagger-Johnson G, Taibjee R, Semlyen J, Fitchie I, Fish J, Meads C, et al. Sexual orientation identity in relation to smoking history and alcohol use at age 18/19: cross-sectional associations from the Longitudinal Study of Young People in England (LSYPE). BMJ Open 2013;3(8):e002810 [FREE Full text] [doi: 10.1136/bmjopen-2013-002810] [Medline: 23985386]
50. Shahab L, Brown J, Hagger-Johnson G, Michie S, Semlyen J, West R, et al. Sexual orientation identity and tobacco and hazardous alcohol use: findings from a cross-sectional English population survey. BMJ Open 2017;7(10):e015058 [FREE Full text] [doi: 10.1136/bmjopen-2016-015058] [Medline: 29074508]
51. Meads C, Zeeman L, Sherriff N, Aranda K. Prevalence of alcohol use amongst sexual and gender minority (LGBTQ+) communities in the UK: a systematic scoping review. Alcohol Alcohol 2023;58(4):346-356 [FREE Full text] [doi: 10.1093/alcalc/agad029] [Medline: 37114766]
52. Boehmer U, Miao X, Ozonoff A. Cancer survivorship and sexual orientation. Cancer 2011;117(16):3796-3804 [FREE Full text] [doi: 10.1002/cncr.25950] [Medline: 21557209]
53. Underwood S, Lyratzopoulos G, Saunders CL. Breast, prostate, colorectal, and lung cancer incidence and risk factors in women who have sex with women and men who have sex with men: a cross-sectional and longitudinal analysis using UK Biobank. Cancers (Basel) 2023;15(7):2031 [FREE Full text] [doi: 10.3390/cancers15072031] [Medline: 37046692]
54. Pitman A, Marston L, Lewis G, Semlyen J, McManus S, King M. The mental health of lesbian, gay, and bisexual adults compared with heterosexual adults: results of two nationally representative English household probability samples. Psychol Med 2021:1-10 [FREE Full text] [doi: 10.1017/S0033291721000052] [Medline: 33592165]
55. Crowe S, Barker E, Roberts M, Lloyd L, de Barros CM, Rebelo-Harris B, et al. Are we asking the right questions? Working with the LGBTQ+ community to prioritise healthcare research themes. Res Involv Engagem 2021;7(1):64 [FREE Full text] [doi: 10.1186/s40900-021-00298-7] [Medline: 34556178]
56. MacCarthy S, Saunders CL, Elliott MN. Sexual minority adults in England have greater odds of chronic mental health problems: variation by sexual orientation, age, ethnicity, and socioeconomic status. LGBT Health 2022;9(1):54-62 [FREE Full text] [doi: 10.1089/lgbt.2021.0011] [Medline: 34882021]

XSL•FO
RenderX

57. Bécares L, Kneale D. Inequalities in mental health, self-rated health, and social support among sexual minority young adults during the COVID-19 pandemic: analyses from the UK millennium cohort study. Soc Psychiatry Psychiatr Epidemiol 2022;57(10):1979-1986 [FREE Full text] [doi: 10.1007/s00127-022-02291-1] [Medline: 35507033]
58. Saunders CL, Flynn S, Massou E, Lyratzopoulos G, Abel G, Burt J. Sociodemographic inequalities in patients' experiences of primary care: an analysis of the general practice patient survey in England between 2011 and 2017. J Health Serv Res Policy 2021;26(3):198-207 [FREE Full text] [doi: 10.1177/1355819620986814] [Medline: 33517786]
59. Saunders CL, Gkousis E. Impact of telephone triage on access to primary care for people living with multiple long-term health conditions: rapid evaluation. Health Soc Care Deliv Res 2022;10(18):1-48 [FREE Full text] [doi: 10.3310/ucce5549] [Medline: 35793423]
60. Booker CL, Meads C. Sexual orientation and the incidence of COVID-19: evidence from understanding society in the UK longitudinal household study. Healthcare (Basel) 2021;9(8):937 [FREE Full text] [doi: 10.3390/healthcare9080937] [Medline: 34442074]
61. Goldacre B, Morley J. Better, broader, safer: using health data for research and analysis. A Review Commissioned by the Secretary of State for Health and Social Care. 2022. URL: https://assets.publishing.service.gov.uk/media/624ea0ade90e072a014d508a/goldacre-review-using-health-data-for-research-and-analysis.pdf [accessed 2024-04-10]
62. Ford T, Mansfield KL, Markham S, McManus S, John A, O'Reilly D, et al. The challenges and opportunities of mental health data sharing in the UK. Lancet Digit Health 2021;3(6):e333-e336 [FREE Full text] [doi: 10.1016/S2589-7500(21)00078-9] [Medline: 34044999]
63. Romanelli RJ, Picken N, Adams A. Counting LGBTQ+ Lives in England and Wales: A Critical Reflection on Emerging Findings from the 2021 Census. Santa Monica, CA: RAND Corporation; 2023.
64. Braybrook D, Bristowe K, Timmins L, Roach A, Day E, Clift P, et al. Communication about sexual orientation and gender between clinicians, LGBT+ people facing serious illness and their significant others: a qualitative interview study of experiences, preferences and recommendations. BMJ Qual Saf 2023;32(2):109-120 [FREE Full text] [doi: 10.1136/bmjqs-2022-014792] [Medline: 36657773]
65. van Daalen KR, Davey F, Norman C, Ford JA. Health equity audits: a systematic review of the effectiveness. BMJ Open 2021;11(11):e053392 [FREE Full text] [doi: 10.1136/bmjopen-2021-053392] [Medline: 34764176]
66. Participate in research. LGBT Foundation. URL: https://lgbt.foundation/research/participate [accessed 2024-04-10]
67. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. BMJ 2001;323(7313):625-628 [FREE Full text] [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
68. Evidence review: gonadotrophin releasing hormone analogues for children and adolescents with gender dysphoria. National Institute for Health and Care Excellence. 2020. URL: https://cass.independent-review.uk/nice-evidence-reviews/ [accessed 2024-04-10]
69. Evidence review: gender-affirming hormones for children and adolescents with gender dysphoria. National Institute for Health and Care Excellence. 2020. URL: https://cass.independent-review.uk/wp-content/uploads/2022/09/20220726_Evidence-review_Gender-affirming-hormones_For-upload_Final.pdf [accessed 2024-04-10]
70. Ethical research: good practice guide to researching LGBT communities and issues. LGBT Foundation. URL: https://dxfy8lrzbpywr.cloudfront.net/Files/1a884870-453a-429d-a213-399a9502472c/Ethics%2520Guide.pdf [accessed 2024-04-10]
71. Cowan K. INVOLVE: A practical guide to being inclusive in public involvement in health research - lessons from the reaching out programme. National Institute for Health Research. 2020. URL: https://www.invo.org.uk/wp-content/uploads/2020/03/INVOLVE_RO_report_FINAL_180220.pdf [accessed 2024-04-10]
72. de Barros CM, Saunders K, Crowe S. LGBTQ+ research and involvement – what have we learnt? National Institute for Health and Care Research: School for Primary Care Research. 2022. URL: https://www.spcr.nihr.ac.uk/news/blog/lgbtq-research-and-involvement-2013-what-have-we-learnt, [accessed 2024-04-10]

Abbreviations

EHR: electronic health record
LGBTQ+: lesbian, gay, bisexual, transgender, queer, or other identities including nonbinary
NATSAL: National Survey of Sexual Attitudes and Lifestyles

XSL•FO
RenderX

Edited by T de Azevedo Cardoso; submitted 03.10.23; peer-reviewed by C Li, E Fedele; comments to author 16.02.24; revised version received 28.02.24; accepted 26.03.24; published 01.05.24.

Please cite as:
Saunders CL
Using Routine Data to Improve Lesbian, Gay, Bisexual, and Transgender Health
Interact J Med Res 2024;13:e53311
URL: https://www.i-jmr.org/2024/1/e53311
doi:10.2196/53311
PMID:38691398

©Catherine L Saunders. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 01.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Ad Hoc Modifications to a High Dependency Psychiatric Unit for People With Dementia During the COVID-19 Period

Thanita Pilunthanakul[1*], BSc, MA, MD; Giles Ming Yee Tan[2*], MBBS, M Med

[1]MOH Holdings Pte Ltd, Singapore, Singapore
[2]Institute of Mental Health, Singapore, Singapore
[*]all authors contributed equally

Corresponding Author:
Thanita Pilunthanakul, BSc, MA, MD
MOH Holdings Pte Ltd
1 Maritime Square
Singapore, 099253
Singapore
Phone: 65 66220980
Email: thanita.pilunthanakul@mohh.com.sg

Abstract

The COVID-19 pandemic led to behavioral exacerbations in people with dementia. Increased hospitalizations and lack of bed availability in specialized dementia wards at a tertiary psychiatric hospital in Singapore resulted in lodging people with dementia in the High Dependency Psychiatric Unit (HDPCU). Customizations to create a dementia-friendly environment at the HDPCU included: (1) environmental modifications to facilitate orientation and engender familiarity; (2) person-centered care to promote attachment, inclusion, identity, occupation, and comfort; (3) risk management for delirium; and (4) training core competencies. Such practical solutions can also be implemented elsewhere to help overcome resource constraints and repurpose services to accommodate increasing populations of people living with dementia.

(Interact J Med Res 2024;13:e49618) doi:10.2196/49618

KEYWORDS

dementia; COVID-19; high dependency psychiatric unit; psychiatric intensive care unit; caregiver stress; SARS-CoV-2; psychiatric; psychiatry; mental health; health care system; Alzheimer; ward; care facility

Background

The spread of SARS-CoV-2 causing COVID-19 required significant reorganization of the Singapore health care system to support the rising number of COVID-19 cases and associated mental health issues [1,2]. Postponement of nonurgent clinic appointments due to reallocating hospital resources for pandemic control and patients' fear of contracting COVID-19 led to significant delays in treatment and further increased the risk of psychiatric relapses [3]. Reduced physical and cognitive activity from the cessation of dementia daycare programs and limited social engagements due to restricted visits with friends and family further increased social isolation and behavioral exacerbations in people with dementia [4]. On top of this, unemployment, work-from-home policies, quarantine orders, and home-based learning put family members in closer contact with people with dementia. This struggle to cope with the increasing care demands of people with dementia has resulted in greater caregiver burnout and hospitalizations for respite care [5,6].

Due to the shortage of beds in the dementia wards at the Institute of Mental Health (IMH), a tertiary psychiatric care facility in Singapore, people with dementia are occasionally lodged in the High Dependency Psychiatric Care Unit (HDPCU) of the IMH, which was not initially designed for dementia-friendly interventions. In this viewpoint, we provide insight into easily implementable and practical creative solutions that have been used to accommodate people with dementia in the HDPCU during the COVID-19 pandemic, overcoming resource constraints and repurposing services in the face of changing needs using a patient-centered approach (summarized in Table 1).

XSL•FO
RenderX

Table 1. Summarized framework for customizing a care environment for people with dementia.

Principles	Examples

Environmental modifications to facilitate orientation and engender familiarity	• Clear signs and signage
	• Sizable and readable calendars and clocks
	• Frequent reorientation by staff
	• Nurse in areas allowing natural sunlight and views of greenery
	• Simulate home-like surroundings, including
	  • partition areas to create “rooms” for different activities
	  • place photographs of friends/family close to the patient’s bed
	  • arrange regular communication between the patient and their family/friends via video calls

Person-centered care to promote attachment, inclusion, identity, occupation, and comfort	• Obtain a detailed personal history from family/friends regarding the patient’s preferences
	• Surround the patient with items that affirm their personhood, including
	  • playing songs that bring comfort or voice recordings of family/friends in the ward
	  • addressing patients by their usual/preferred nickname
	  • if possible, allow patients to wear/have sentimental items close by
	• Empower patients to exercise choice as much as possible, no matter how small the decisions may be
	• Encourage patients to join together for meals and games
	• Engage patients in meaningful and mentally stimulating activities
	• Express comforting interactions and validate patients’ concerns
	• Patiently answer repeated questions and allow relatively more time to perform tasks

Risk management (delirium)	• Actively take measures to prevent delirium, including
	  • restrict physical restraints to only when necessary, and even then, for the shortest duration required
	  • minimize medications that risk iatrogenic delirium
	• Obtain a corroborative history regarding patients’ expression of discomfort to recognize signs of distress and address agitated behavior promptly

Core competencies (training if required) of staff	• Understand core concepts of and practice person-centered care • Geriatric-specific care, including fall and choking risks, along with activities of daily living support • De-escalation skills for agitated geriatric patients

Environmental Modifications of the HDPCU

The HDPCU is a specialized inpatient unit devised for patients with an acute psychiatric disorder linked to severe agitation or aggression, placing them at significant risk to themselves or others, leading to the requirement of close monitoring. The nursing counter is sandwiched between two locked gender-specific cubicles with 4 and 6 beds, respectively, and single bathrooms. The nurse has a full view of both cubicles and there are various discreetly placed security cameras. Items that could fuel self-harming or suicidal behaviors, such as wires for electronics, plastic bags, detergents, sharp pencils, and utensils, are strictly prohibited in the ward. There is a 2:1 nurse-to-patient ratio. Staff are specially trained in swift de-escalation to ensure safety and prevent violence, including applying physical restraints and administering oral and intramuscular sedation if required. A psychiatrist, a junior doctor, and the nursing and allied health care team are on site to manage the patients.

People with dementia often experience disorientating situations due to separation from familiar settings, people, and routines. Dementia wards have specific modifications to orient people with dementia, such as legible signage; large clocks; brighter lighting; and contrasting-colored walls, furniture, and utensils [7]. Renovating the HDPCU to suit such requirements was not immediately feasible. Hence, modifying the environment to have clear signs indicating the bathrooms; a sizeable hand-drawn daily calendar facing the bed indicating the date, day, month, and year; and verbal reorientation 3 times a day were implemented to facilitate orientation. People with dementia were also nursed opposite a readable digital clock and beside big windows that offered a view of greenery and allowed in natural sunlight.

Dementia wards engender familiarity by creating homey surroundings, including paintings hung along corridors and divided kitchen, bedroom, and living room spaces. Since safety is of utmost priority in the HDPCU, rules are often strict, and cubicles are designed to be relatively smaller than found in other wards along with an open layout for easy monitoring. These restrictions may cause people with dementia, particularly those who like to wander, to feel trapped and anxious. Given the nature of patients admitted to the HDPCU, the noisy and disruptive atmosphere can destabilize and frighten people with dementia. To create a calm environment that minimizes overstimulation and distractions, people with dementia were nursed in a partitioned visitors’ area accessed via a corridor

XSL•FO
RenderX

adjacent to the cubicles and nursing counter. Families of people with dementia were encouraged to bring photographs to place in front of patients' beds and participate in regular video calls from the ward smartphone to lessen the effects of visitor restrictions during the pandemic [8]. The sectioned area also reduced the risk of impaired sleep-wake cycles and sun-downing behaviors that are common among people with dementia and could provoke other patients. The improvised space simulated a private bedroom, while a wheel-in television and movable couches in the cubicles' shared living area imitated a makeshift living room.

Promoting Person-Centered Care

In person-centered care for people with dementia, personhood consists of attachment, inclusion, identity, occupation, and comfort. Emotional distress is usually triggered by unmet needs related to aspects of personhood. Obtaining a detailed personal history from the family regarding the preferences of people with dementia is essential to affirm personhood. For one such patient, playing Chinese songs from his childhood, hearing voice recordings of family, addressing him by his preferred nickname, and wearing a jacket gifted from his daughter in the ward provided a sense of comfort, identity, and continuation of self. Empowering the patient to exercise choice as much as possible, even for tasks as small as choosing a preferred snack, preserves autonomy and dignity. Encouraging people with dementia to join other patients during meals and games instills a sense of inclusion and occupation [9,10]. Engaging people with dementia in meaningful activities mentally stimulates and reduces the restlessness related to the tendency to worry about their situation. Expressing warmth through comforting interactions, patiently answering repeated questions, allowing them more time to perform tasks, and validating concerns can help to settle the wariness and diminished sense of attachment experienced by these patients. These efforts promote the therapeutic relationship and trust between staff and people with dementia, ultimately reducing aggression and distress.

Risk Management

Lastly, because people with dementia are prone to delirium during acute hospitalization, the HDPCU team actively took measures to prevent this risk. Physical restraints were only applied if verbal de-escalation repeatedly failed and the extreme agitation posed a safety risk to themselves or others; when required, patients were restrained for the shortest duration necessary. Wherever possible, medications that risk iatrogenic delirium in people with dementia were avoided, such as short-acting benzodiazepines for tranquilization, anticholinergic drugs, and opioid-containing analgesics. People with dementia often have issues communicating their needs and are likely to only respond to their present state due to verbal difficulties and memory problems. Obtaining a further history regarding the patients' typical behavioral patterns and expression of discomfort from pain, hunger, thirst, or constipation helped the team promptly recognize signs of distress and address agitated behaviors early without escalating to restraints.

Prospects

In conclusion, hospital care conditions can be difficult for people with dementia as they require familiarity, frequent orientation, and a high level of staff trained to handle their needs. The HDPCU adapted to rapid hospital protocol and health care policy changes during the COVID-19 pandemic and the resultant rise in the inpatient dementia population. Although the HDPCU staff were not geriatric-trained, the favorable staffing ratio and expertise in handling agitated and aggressive patients made it easier to implement person-centered care. Such conditions may not be available in nonspecialized wards, posing a challenge for catering to the increasing population of people with dementia admitted to hospitals in Singapore. Nonetheless, creative solutions could be established to customize the environment for such patients aptly. Hospitals could also consider bringing in key "experts" such as psychogeriatricians and geriatric nurses to advise on optimizing nonspecialized wards and provide training to care for people with dementia.

Acknowledgments

The authors thank the nurses and allied health care team from the High Dependency Psychiatric Unit at the Institute of Mental Health, Singapore, for their team effort and support.

Authors' Contributions

TP contributed to conceptualizing and writing the original draft and to reviewing and editing the manuscript. GTMY contributed to conceptualizing, supervising, and reviewing and editing the manuscript.

Conflicts of Interest

None declared.

References

1. Lim WS, Soh KC, Surej J, Goh KH. Examining the role of a psychiatry department in Singapore amidst COVID-19. Asian J Psychiatr 2020 Dec;54:102296 [FREE Full text] [doi: 10.1016/j.ajp.2020.102296] [Medline: 32682304]

XSL•FO
RenderX

2. Poremski D, Subner SH, Lam GF, Dev R, Mok YM, Chua HC, et al. Effective infection prevention and control strategies in a large, accredited, psychiatric facility in Singapore. Infect Control Hosp Epidemiol 2020 Oct 23;41(10):1238-1240 [FREE Full text] [doi: 10.1017/ice.2020.163] [Medline: 32321623]
3. Wang H, Li T, Barbarino P, Gauthier S, Brodaty H, Molinuevo JL, et al. Dementia care during COVID-19. Lancet 2020 Apr 11;395(10231):1190-1191 [FREE Full text] [doi: 10.1016/S0140-6736(20)30755-8] [Medline: 32240625]
4. D'Cruz M, Banerjee D. Caring for persons living with dementia during the COVID-19 pandemic: advocacy perspectives from India. Front Psychiatry 2020 Oct 27;11:603231 [FREE Full text] [doi: 10.3389/fpsyt.2020.603231] [Medline: 33192744]
5. Cohen G, Russo MJ, Campos JA, Allegri RF. Living with dementia: increased level of caregiver stress in times of COVID-19. Int Psychogeriatr 2020 Nov;32(11):1377-1381 [FREE Full text] [doi: 10.1017/S1041610220001593] [Medline: 32729446]
6. Brown EE, Kumar S, Rajji TK, Pollock BG, Mulsant BH. Anticipating and mitigating the impact of the COVID-19 pandemic on Alzheimer's disease and related dementias. Am J Geriatr Psychiatry 2020 Jul;28(7):712-721 [FREE Full text] [doi: 10.1016/j.jagp.2020.04.010] [Medline: 32331845]
7. Waller S, Masterson A. Designing dementia-friendly hospital environments. Future Hosp J 2015 Feb;2(1):63-68 [FREE Full text] [doi: 10.7861/futurehosp.2-1-63] [Medline: 31098081]
8. Garland K, Beer E, Eppingstall B, O'Connor DW. A comparison of two treatments of agitated behavior in nursing home residents with dementia: simulated family presence and preferred music. Am J Geriatr Psychiatry 2007 Jun;15(6):514-521. [doi: 10.1097/01.JGP.0000249388.37080.b4] [Medline: 17293386]
9. Cohen-Mansfield J, Marx MS, Dakheel-Ali M, Regier NG, Thein K, Freedman L. Can agitated behavior of nursing home residents with dementia be prevented with the use of standardized stimuli? J Am Geriatr Soc 2010 Aug 03;58(8):1459-1464 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02951.x] [Medline: 20579167]
10. Hicks-Moore SL. Relaxing music at mealtime in nursing homes: effects on agitated patients with dementia. J Gerontol Nurs 2005 Dec;31(12):26-32. [doi: 10.3928/0098-9134-20051201-07] [Medline: 16375095]

Abbreviations

HDPCU: High Dependency Psychiatric Unit
IMH: Institute of Mental Health

Edited by T de Azevedo Cardoso; submitted 03.06.23; peer-reviewed by HY Goh, V Rashedi; comments to author 12.12.23; revised version received 22.01.24; accepted 13.05.24; published 11.06.24.

Please cite as:
Pilunthanakul T, Tan GMY
Ad Hoc Modifications to a High Dependency Psychiatric Unit for People With Dementia During the COVID-19 Period
Interact J Med Res 2024;13:e49618
URL: https://www.i-jmr.org/2024/1/e49618
doi:10.2196/49618
PMID:38861715

©Thanita Pilunthanakul, Giles Ming Yee Tan. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 11.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Strategies to Alleviate the Burden Experienced by Informal Caregivers of Persons With Severe Mental Disorders in Low- and Middle-Income Countries: Scoping Review

Olindah Silaule[1], MSc; Daleen Casteleijn[2], PhD; Fasloen Adams[3], PhD; Nokuthula Gloria Nkosi[4], PhD

[1]Department of Occupational Therapy, University of the Witwatersrand, Johannesburg, South Africa

[2]Department of Occupational Therapy, University of Pretoria, Pretoria, South Africa

[3]Division of Occupational Therapy, Stellenbosch University, Cape Town, South Africa

[4]Department of Nursing Education, University of the Witwatersrand, Johannesburg, South Africa

Corresponding Author:
Olindah Silaule, MSc
Department of Occupational Therapy
University of the Witwatersrand
7 York Road
Parktown
Johannesburg, 2193
South Africa
Phone: 27 0782035213
Email: osilaule@cartafrica.org

Abstract

Background: There is considerable evidence of the burden of care encountered by informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries. Previous studies have highlighted the need to support these informal caregivers as key players in the care of these patients. To date, limited evidence exists on the extent and types of strategies for supporting these informal caregivers in low- and middle-income countries.

Objective: This scoping review aims to identify and describe the extent and type of evidence on the existing strategies for alleviating the burden of care among informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries.

Methods: A systematic literature search was completed following the Joanna Briggs Institute methodology for scoping reviews. The participants, concept, and context framework was used to guide the search for literature sources across 5 databases: PubMed, MEDLINE, CINAHL, and PsycINFO for published literature and ProQuest for unpublished literature. This review included studies that reported on strategies for alleviating the burden of care among informal caregivers of persons with severe and enduring mental health conditions, with a focus on studies that evaluated or recommended caregiver interventions and support strategies in low- and middle-income countries. The search was limited to studies conducted between 2001 and 2021, and only papers written in English were considered for inclusion. Using the Covidence software (Veritas Health Innovation), 2 reviewers independently screened the papers, applied the inclusion and exclusion criteria, and met biweekly to discuss and resolve conflicts. The relevant studies and reported outcomes were summarized, organized, and analyzed descriptively using numeric summary analysis and deductive content analysis.

Results: Of the 18,342 studies identified, 44 (0.24%) met the inclusion criteria. The included studies were from 16 low- and middle-income countries in Asia, Africa, Europe, and South and North America. Most studies (21/44, 48%) were randomized controlled trials conducted in Asian countries. The identified strategies were grouped into 2 categories: implemented and recommended intervention strategies. Identified strategies included community-based interventions, psychoeducation interventions, support groups, cognitive behavioral therapy, spirituality-based interventions, and smartphone-based interventions. In addition, mindfulness and empowerment, collaborative interventions, standard care, financial and social support, counseling, occupation-based interventions, policy and legislature, and access to mental health care were identified. Psychoeducation and support group interventions were identified as common strategies for alleviating the burden of care among informal caregivers of persons with severe and enduring mental health conditions.

XSL•FO
RenderX

Conclusions: This review provides evidence on the types of implemented and recommended strategies for alleviating the burden of care among informal caregivers in low- and middle-income countries. Although psychoeducational interventions were the most preferred strategy for alleviating burden, their benefits were short-lived when compared with peer-led support groups.

International Registered Report Identifier (IRRID): RR2-10.2196/44268

(Interact J Med Res 2024;13:e48587) doi:10.2196/48587

KEYWORDS

severe mental disorders; informal caregivers; caregiver stress; caregiver support; low- and middle-income country; mobile phone

Introduction

Background Functional Implications of Severe and Enduring Mental Health Conditions

People with severe and enduring mental health conditions are unable to cope with the demands of everyday life and need care after being discharged from a health facility. Serious functional limitations frequently interfere with the ability of people with severe and enduring mental health conditions to perform essential roles such as being a worker, family member, or friend [1]. The lack of independence among people with severe and enduring mental health conditions in daily living activities and, in some cases, behavioral problems result in poor quality of life. Psychosis, bipolar mood, major depressive, anxiety, eating, and personality disorders are often classified as severe and enduring mental health conditions [2]. The amount of care needed varies greatly depending on the diagnosis, severity of symptoms, and level of independent functioning. As such, assistance is needed in areas of basic self-care, healthy eating, following daily routines, medication management and compliance, engagement in meaningful activities, and community integration. Therefore, informal caregivers are essential to fulfill the care needs of people with severe and enduring mental health conditions. This task becomes draining, and the caregiver burden should not be underestimated [3,4].

Informal Caregivers in Mental Health

Informal caregivers are people who deliver care without remuneration to persons with a chronic condition. The caregivers of persons with severe and enduring mental health conditions are most often family members or neighbors who assist with the care of the person with a mental health condition [5]. Dixon et al [6] noted that, often, the caregiver of a person with a severe and enduring mental health condition does not have a choice in being the caregiver, which is an additional burden of care for the carer. Informal caregivers' roles include monitoring medication, being the contact person between the health provider and family, early identification of signs of relapse, taking care of daily tasks such as self-care, providing meals, and ensuring the safety of the person [7,8]. Informal caregivers provide emotional support when needed and deal with the challenging behavior of the person with a severe and enduring mental health condition, which may lead to police involvement [9]. Involving informal caregivers in the routine care and management of persons with severe and enduring mental health conditions has shown a positive influence on the course of the illness but only if the caregivers' needs are addressed and they are supported in one way or another [5,6,10].

Caregiver Burden and Need for Support

The burden of care experienced by caregivers is classified into objective and subjective burden. Objective burden refers to the tangible impact of the demands of caring tasks and encompasses the practical and concrete aspects of caregiving that can be quantified or assessed externally [11,12]. Subjective burden refers to the emotional and psychological experiences, feelings, and perceptions of caregivers related to their caregiving role [8,13]. It focuses on how caregivers perceive the impact of their responsibilities on their well-being and mental health.

In low- and middle-income countries, objective burden can easily overshadow subjective burden, such as a lack of community mental health care services and clinics that are out of stock of medication for people with severe and enduring mental health conditions, causing relapse and often readmission [11]. Other concrete aspects of objective burden, such as long hours of being available to the care recipient, time lost from daily activities, and not being able to earn an income, have serious consequences for caregivers as they are unable to pursue their own goals in life and, accordingly, experience a lower quality of life [9,12]. Cultural beliefs about being cursed by ancestors [14] and stigma from health care professionals [15] further aggravate the caring duties of caregivers, precipitating objective burden. The objective burden of care is less reported in the literature than subjective feelings of burden. Many studies have reported the subjective burden of care, which highlights emotional distress and feelings of anxiety, depression, and sadness related to the challenges of caregiving and witnessing the struggles of their family member. Loss of freedom and autonomy has also been reported owing to the demand of caregiving [16]. Feelings of shame and social isolation stem from the stigma of mental illness and cultural beliefs about being cursed by ancestors. Caregivers can also experience role strain as they have to juggle responsibilities at home, which can cause fatigue and overall dissatisfaction in life [17]. Both dimensions of burden are crucial for understanding the challenges faced by family caregivers of individuals with severe and enduring mental health conditions and play a vital role in informing support strategies and interventions.

There is overwhelming evidence in the literature for the need to support caregivers, and several studies from high-income and low- and middle-income countries have indicated successful strategies for alleviating caregiver burden. Studies on informal caregivers of people with severe and enduring mental health conditions in low- and middle-income countries have also increased over the past 10 years, and the burden of caregivers has been well described [18-21]. Yerriah et al [16] reported on the extent of the burden of caregivers of persons with

XSL•FO
RenderX

schizophrenia in rural South Africa and found that this sample has a higher burden of care. A recommendation from this study was to develop strategies to support caregivers with the aim of improving their quality of life.

Support Strategies for Caregivers

To date, a number of strategies to support caregivers have been investigated with varying results. In Turkey, family-to-family support programs have shown a positive impact on the burden of care [22]. A meta-analysis by Chen et al [23] showed some support for nonpharmacological interventions (mostly psychoeducation) for caregivers of persons with schizophrenia, but the authors reported a potential bias in the results because of the small sample size. In contrast, the systematic review and meta-analysis by Sin et al [24] did not support psychoeducation to improve compliance with treatment and prevent relapse in persons with psychosis, and they reported a lack of available data; thus, no meta-regressions could be conducted. Ewertzon and Hanson [18] conducted a narrative review and identified provision of knowledge, problem-solving, stress management, mutual support groups, and individual-support interventions as successful support interventions for caregivers. Finally, a systematic review by Napa et al [25] revealed that there is insufficient evidence of interventions for psychological distress and expressed emotions in families of persons who experienced first-episode psychosis.

There is limited evidence of web-based health care services and digital health technologies for supporting informal caregivers of individuals with severe and enduring mental health conditions [26]. Sin et al [27] developed an eHealth intervention called Carers for People with Psychosis e-support, but its effectiveness has not yet been investigated. Ploeg et al [28] conducted a rapid review of web-based interventions to improve general caregiver outcomes. More than half of the 17 included studies showed a positive outcome for decreased depressive symptoms, stress, and anxiety among caregivers. If virtual strategies can support caregivers in rural areas, they may be a feasible method to reach people in remote areas who have poor access to health care.

Informal caregivers in rural or remote areas face additional objective burdens such as poor access to services, lack of integration of mental health into community health services [29], and poor intersectoral collaboration. They often have to wait long hours before any support arrives, and in many cases, they have to deal with challenges with the limited resources they have available.

If the types of intervention strategies that could be relevant for low- and middle-income countries, how they were implemented, and the outcome that was achieved were mapped, it could guide health care workers to support informal caregivers on various levels and with various strategies. The availability of virtual support strategies could add another dimension of support to caregivers, which may lead to additional positive outcomes such as immediate support, available information, and contact with support groups. People with severe and enduring mental health conditions benefit as they are most likely to receive optimal care, and their relapse rate may decrease, which means less need for readmission in overextended mental health care wards or hospitals. In addition, people with severe and enduring mental health conditions may also experience a better quality of life if their carers are supported. Thus, it is essential to understand the strategies to alleviate informal caregiver burden and how these strategies should be implemented.

This scoping review aimed to map the strategies to alleviate the objective and subjective burden of informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries. The objectives of this scoping review were to (1) identify the types of existing strategies (virtual and face-to-face) for alleviating the objective or subjective burden of care, (2) describe the characteristics of the identified strategies, and (3) list the positive outcomes that were achieved using the identified strategies.

Methods

Review Methodology

This scoping review followed the Joanna Briggs Institute methodology specific to scoping reviews [30]. An a priori protocol for this review has been published [31].

Review Question

The scoping review addressed two research questions related to the strategies for alleviating the burden of care among informal caregivers of persons with severe and enduring mental health conditions:

1. Which existing strategies are reported in the literature for alleviating the burden of informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries?
2. What are the outcomes reported by the authors of the strategies for alleviating burden among informal caregivers?

Eligibility Criteria

Participants

Studies were included if participants were informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries. Studies that focused on informal caregivers of patients diagnosed with Alzheimer disease and dementia were excluded. This included family, friends, neighbors, and community members who voluntarily provided care without any remuneration. Studies that included both informal caregivers and patients as participants were considered for inclusion, whereas those that reported only on patients were excluded.

Concept

Caregiver burden was defined as the physical, psychological, emotional, social, and financial stresses of providing care to a person with severe and enduring mental health conditions [32]. Severe and enduring mental health conditions in this scoping review included the schizophrenia spectrum and other psychotic disorders, bipolar and related disorders, mood or depressive disorders, and personality disorders. Studies that investigated strategies for alleviating caregiver burden were considered for inclusion. This also included studies that investigated the synonyms used to describe the burden of care among informal caregivers, such as *caregiver strain*, *stress*, and *role fatigue*.

XSL•FO
RenderX

Studies that reported on the burden of care without suggesting strategies for alleviating it were excluded.

Context

Studies from low- and middle-income countries, including countries in Africa, Asia, Latin America, and the Caribbean, were considered for inclusion. In addition, the inclusion criteria comprised studies from lower-middle-income countries using the World Bank classification of the economic status of a country at the time the study was conducted. The term *developing country* was included as it is similar to *lower-middle-income country*, and thus, countries in Africa, Asia, and Latin America (including the Caribbean) were eligible for inclusion.

Types of Sources

This review included published qualitative, quantitative, and mixed methods studies with all types of designs as well as unpublished studies, including dissertations and theses. Over the last decade, there has been a growing body of knowledge from studies outlining strategies for alleviating burden among informal caregivers in low- and middle-income countries. As a result, this review was only restricted to studies conducted between 2011 and 2021. Only papers written in English were included in this review.

Search Strategy

A 3-step search strategy process was used to identify relevant studies. The search strategy commenced with an initial limited search of PubMed, MEDLINE, CINAHL, and PsycINFO for published studies between 2011 and 2021. The keywords were as follows: "[informal caregiver/s OR caregiver/s] AND [caregiver burden OR caregiver stress] AND [support strategy/ies OR intervention/s] AND [severe mental disorder/s OR mental illness] AND [developing country/ies]."

The second search was refined with the assistance of the librarian at the University of the Witwatersrand, and additional terms were added: "[Carers or 'Informal Carers' or 'Male Caregiver' or Women Caregiver' or 'Family Caregivers'] AND [Caregiver* strain' or 'Caregiver *Exhaustion' or 'Caregiver* burnout' or 'Carer* burden'] AND ['underserved' or 'deprived' or 'middle-income' or low income country]." On advice from the librarian, gray literature such as conference presentations and unpublished studies was searched on ProQuest Dissertations and Theses Global and public health and conference databases. Finally, 2 reviewers (OS and DC) screened the reference lists of the included articles for additional studies. Textbox 1 presents the final list of search terms used to search published and unpublished studies across all the selected databases in this review. The identified references were imported into the Mendeley reference manager (Elsevier) for screening.

Textbox 1. Search terms used in the databases.

'Informal caregiver*' or Caregiver* or Carers or 'Informal Carers' or 'Male Caregiver' or 'Women Caregiver' or 'Family Caregivers' or 'Unpaid Caregiver' or 'Spouse Caregiver' or 'Caretaker' or 'Older Caregiver'

AND

'Caregiver* burden' or 'Caregiver* stress' or 'Caregiver* strain' or 'Caregiver *Exhaustion' or 'Caregiver* burnout' or 'Carer* burden' or 'Carer* stress' or 'Caretaker* role fatigue' or 'Burden of Caregiver*' or 'Caretaker* burden' or 'Caretaker* load' or 'Caregiver* Psychology'

AND

'Strategy' or 'Strategies' or 'Intervention' or 'Procedure' or 'Programme*' or 'Management' or 'Protocol* Guidelines' or 'Guide' or 'Policy' or 'Policies'

AND

'Mental disorders' or 'Mental illness' or 'schizophrenia' or 'bipolar mood disorder' or 'Major depressive disorder' or 'Psychotic disorder' or 'Personality disorder' or 'Bipolar affective disorder'

AND

(developing OR (less* N1 developed) OR "under developed" OR underdeveloped OR "under served" OR underserved OR deprived OR poor* OR "middle income" OR (low* N1 income)) N1 (countr* OR nation* OR population* OR world)) OR ((developing OR (less* N1 developed) OR "under developed" OR underdeveloped OR "under served" OR underserved OR deprived OR poor* OR "middle income" OR (low* N1 income)) N1 (countr* OR nation* or population* OR world)) OR ((developing OR (less* N1 developed) OR "under developed" OR underdeveloped

Source of Evidence Selection

Following the search, all identified citations were exported to the Mendeley reference manager and thereafter to a web-based software, Covidence (Veritas Health Innovation), for primary screening and data extraction from the selected articles. Duplicates were removed, and 2 reviewers (OS and DC) independently conducted the screening for titles, abstracts, and full texts using the inclusion and exclusion criteria. To enhance the reliability of the review results, the first 5 abstracts were screened, and the 2 reviewers compared the screening results and clarified conflicts. Following this, the criteria were revised to ensure a detailed description of the various types of burden. The remaining abstracts were screened, and comparisons were made after 10 screenings to manage conflicts. The 2 reviewers (OS and DC) proceeded to screen texts and extract data from the selected articles. Any disagreements regarding the inclusion of the studies were resolved through discussion between the 2 reviewers.

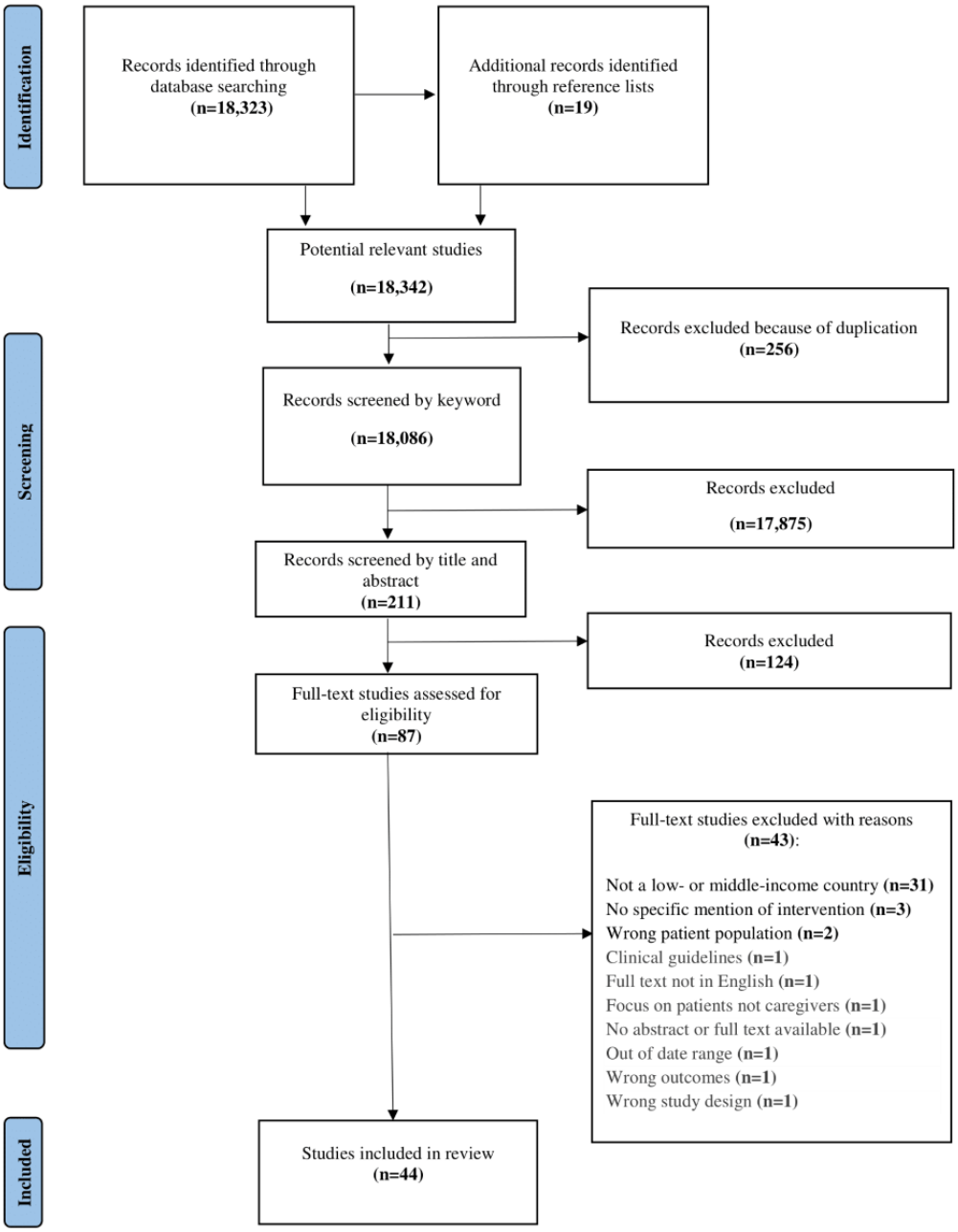
Data Extraction

The Covidence software has templates for data extraction. These templates were modified using the data extraction tables proposed by Peters et al [30]. To ensure the extraction of

XSL•FO
RenderX

relevant data to answer the review question, these templates were piloted by the reviewers before use. In total, 2 separate data extraction templates (Multimedia Appendix 1) were used to extract data from studies that implemented an intervention and from studies that recommended interventions and strategies for alleviating caregiver burden. Data related to the characteristics of the studies were extracted, including study title; study aims; citation details; population of interest; concept of interest; context of the study, including the country and type of setting; type of evidence sources; study approach and designs; and participant characteristics, such as age, gender, and diagnosis of the care recipients. In addition, information on the characteristics of the interventions was extracted. This included the intervention content (ie, type of intervention, intervention developer and deliverer, and type of burden targeted by the intervention). The intervention description included the duration of the intervention, number of sessions, and location for the intervention. The templates for extracting the data are available in Multimedia Appendix 1. The reviewers extracted the data independently. Web-based meetings were held biweekly to discuss and resolve any discrepancies in the extracted data. Regular comparisons were easy and quick to conduct and improved the consistency of the extracted data. After completing the data extraction, the 2 reviewers scanned the references of the included articles to ensure that no articles were missed. Included and excluded studies were reported in a PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flow diagram (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



XSL•FO
RenderX

Data Analysis and Presentation

To increase consistency in the data, this review followed a 3-step analysis process as proposed by Levac et al [33]. This included analyzing the data, reporting the results, and applying meaning to the results [33]. The first step was reviewing the extracted data and identifying the type of data to be analyzed, which were then grouped according to the objectives of the review. This was followed by identifying the type of analysis appropriate for the specific type of data and then analyzing the extracted data. A descriptive quantitative analysis was conducted in a Microsoft Excel (Microsoft Corp) spreadsheet to describe the characteristics of the studies. This included the overall number of studies, types of study design, years of publication, characteristics of the study populations, countries where the studies were conducted, and types of interventions. In addition, descriptive content analysis was conducted deductively using the NVivo software (Lumivero) to code the characteristics of the extracted data into overall categories [34]. Data were categorized into types of strategies—implemented and recommended intervention strategies—and the outcomes of the strategies aimed at alleviating the burden of informal caregivers as reported by the authors. To enhance the clarity of the emerging findings, the data were summarized and presented in graphs and tables.

Results

Study Inclusion

A total of 18,323 studies were identified from the databases using the keywords. In total, 19 additional studies were identified from the reference lists of the included studies, thereby bringing the total number of potentially relevant studies to 18,342. Subsequently, of the 18,342 studies, 256 (1.4%) duplicates were removed, thereby leaving 18,086 (98.6%) studies deemed relevant to the review based on the keywords. A total of 98.83% (17,875/18,086) of these studies were then excluded; this exclusion of numerous studies occurred because of their incongruence with the predefined inclusion criteria, such as instances in which the study did not pertain to a low- or middle-income country context or in which the primary emphasis lay on the patient rather than the caregiver. Therefore, 1.17% (211/18,086) of the studies were screened by title and abstract based on set inclusion and exclusion criteria. This led to the exclusion of 58.8% (124/211) of these studies from the review. A total of 41.2% (87/211) of the studies were then retrieved as full texts based on eligibility by screening their titles and abstracts. Of these 87 studies, 43 (49%) were excluded for various reasons: 31 (72%) were not from low- or middle-income countries; 3 (7%) did not specify the intervention strategy; 2 (5%) focused on the wrong patient population; and another 7 (16%) were excluded for being a clinical guideline, focusing on patients rather than caregivers, having misaligned outcomes and a study design within the exclusion criteria (eg, review articles), being out of the date range of this review, the full text being in a language other than English, and a lack of an abstract and full text. Finally, 44 studies were included in this scoping review. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram was used to present the selection of the studies (Figure 1). It took approximately 8 months, from September 2021 to June 2022, to carry out the study from conception to completion of the project.

Characteristics of the Eligible Studies

Summary of Studies and Evidence Type

The included studies were published between 2011 and 2021. Most studies (31/44, 70%) were published between 2016 and 2021, and 30% (13/44) were published between 2011 and 2016. The types of studies included peer-reviewed journal articles (40/44, 91%), research reports or theses (2/44, 5%), and opinion pieces (2/44, 5%). Of the 44 included studies, 29 (66%) reported on a specific implemented intervention strategy, and 15 (34%) outlined recommended intervention strategies aimed at alleviating the burden of informal caregivers of people with severe and enduring mental health conditions in low- and middle-income countries. The study designs included randomized controlled trials (21/44, 48%), quantitative studies (15/44, 34%), qualitative studies (3/44, 7%), mixed methods studies (3/44, 7%), and gray literature (2/44, 5%; Table 1).

XSL•FO
RenderX

Table 1. Characteristics of the included studies (N=44).

Characteristic	Studies, n (%)
Summary of studies	
Implemented strategy studies	29 (66)
Recommended strategy studies	15 (34)
Type of evidence	
Journal articles	40 (91)
Research reports or theses	2 (5)
Opinion pieces	2 (5)
Year of publication	
2011-2015	13 (30)
2016-2021	31 (70)
Study design	
Randomized controlled trial	21 (48)
Quantitative design	15 (34)
Qualitative design	3 (7)
Mixed methods	3 (7)
Gray literature	2 (5)

Context of the Studies

A total of 44 studies identified from 16 low- and middle-income countries in Asia, Africa, Europe, and South and North America (Mexico) were included. The countries included China (12/44, 27%); Iran (10/44, 23%); India (5/44, 11%); Brazil (3/44, 7%); Greece (2/44, 5%); Ghana (2/44, 5%); and Thailand, Nepal, South Korea, Indonesia, Jordan, Mexico, Ethiopia, Turkey, Botswana, and mixed countries (including Brazil, Iran, Colombia and Mexico 1/44, 2% of the studies each). As shown in Table 2, majority of the studies with implemented strategies (24/29, 83%) and those with recommended strategies (9/15, 60%) were from Asian countries. Overall, 7% (2/29) of the studies each from Europe, South America, and North America implemented the strategies for alleviating burden and 13% (2/15) of the studies from South and North America recommended the strategies. Only 3% (1/29) of the studies conducted in Africa implemented the strategies and 20% (3/15) of the studies recommended the strategies. A study (1/15, 7%) from mixed countries recommended the strategies for alleviating burden.

Table 2. Continents of the recommended and implemented strategies (N=44).

Continent	Implemented strategies, n (%)	Recommended strategies, n (%)
Asia	24 (83)	9 (60)
Africa	1 (3)	3 (20)
Europe	2 (7)	0 (0)
South and North America	2 (7)	2 (13)
Mixed countries	0 (0)	1 (7)

Characteristics of the Study Populations

The targeted populations in the included studies were mainly family caregivers (38/44, 86%), 7% (3/44) of the studies targeted informal carers, and some studies (3/44, 7%) did not specify their targeted population but rather identified them as caregivers. Most studies (33/44, 75%) included both male and female caregivers, and 25% (11/44) did not specify the sex of their sample. Similarly, most studies (24/44, 55%) included both male and female care recipients, only 2% (1/44) of the studies focused only on male care recipients, and 43% (19/44) did not specify the gender of the care recipients included in the studies. The diagnoses of the care recipients in the included studies were grouped into 6 categories: schizophrenia (17/44, 39%), mood disorders (4/44, 9%), multiple diagnoses including schizophrenia and mood disorders (5/44, 11%), mixed severe and enduring mental health conditions (5/44, 11%), other psychotic and chronic conditions (6/44, 14%), and unspecified mental disorders (7/44, 16%; Table 3).

XSL•FO
RenderX

Table 3. Characteristics of the study populations in the included studies (N=44).

Characteristic	Studies, n (%)
Population of interest	
Family caregiver or caregivers	38 (86)
Informal carer or carers	3 (7)
Caregiver or caregivers	2 (5)
Carer or carers	1 (2)
Concept of interest	
Caregiver burden	18 (41)
Family burden	16 (36)
Caregiver stress	9 (20)
Caregiver strain	1 (2)
Sex of caregivers	
Mixed-gender groups (male and female)	33 (75)
Not specified	11 (25)
Gender of care recipients	
Mixed-gender groups (male and female)	24 (55)
Male	1 (2)
Not specified	19 (43)
Diagnosis of care recipients	
Schizophrenia	17 (39)
Mood disorders	4 (9)
Schizophrenia and mood disorders	5 (11)
Mixed diagnoses of severe mental disorders	5 (11)
Unspecified mental disorders	7 (16)
Other (drug dependency, psychotic disorders, and chronic conditions)	6 (14)

Concept of Interest

The concepts of interest reported in the studies included caregiver burden (18/44, 41%), family burden (16/44, 36%), caregiver stress (9/44, 20%), and caregiver strain (1/44, 2%; Table 3).

Review Findings

This section outlines the findings of the scoping review and includes the types of intervention strategies, characteristics of the intervention strategies, and author-reported outcomes from the included studies.

Types of Intervention Strategies

The various strategies for alleviating caregiver burden were grouped into 2 categories: implemented and recommended intervention strategies. The implemented intervention strategies emerged from studies that investigated the effects of a specific intervention to alleviate the burden of informal caregivers. The recommended interventions were identified from studies that investigated caregiver burden and outlined recommended intervention strategies to be used to alleviate the burden of informal caregivers of persons with severe and enduring mental health conditions. Most studies (29/44, 66%) reported on an implemented intervention strategy, and 34% (15/44) of the studies recommended strategies to alleviate caregiver burden.

Most of the studies and articles (33/44, 75%) originated in Asia, with 55% (24/44) implemented and 20% (9/44) recommended intervention strategies reported on, followed by Africa with 2% (1/44) implemented and 7% (3/44) recommended intervention strategies, South and North America with 5% (2/44) implemented and 5% (2/44) recommended intervention strategies, and Europe with 5% (2/44) implemented intervention strategies (Table 2). A study from mixed countries only reported recommended intervention strategies. The identified strategies were grouped into the following categories: community-based interventions, psychoeducation interventions, support groups, cognitive behavioral therapy (CBT), spirituality-based interventions, smartphone-based interventions, mindfulness and empowerment, collaborative interventions, standard care, financial and social support, counseling, occupation-based interventions, policy and legislature, and access to mental health care (Table 4 and Textbox 2). These categories are discussed in the following subsections.

XSL•FO
RenderX

Table 4. Description of the identified evaluated strategies for alleviating informal caregiver burden.

Type of strategy	Evaluated interventions	Author-reported outcomes
Community-based interventions	• CoMHIP[a] • RESHAPE[b]	• Positive outcomes (positive effects on family caregivers)
Psychoeducation interventions	• Psychoeducation program • Group psychoeducation • FLEP[c] • Peer-assisted education	• Improved coping skills • Enhanced recovery • Improved QoL[d] • Knowledge and skill acquisition • Reduced burden of care • Reduced anxiety and stress • Reduced psychological strain • Enhanced family functioning • Relapse prevention • Reduced prolonged admissions • Negative outcomes (no efficacy on burden of care, QoL, or self-esteem among BMD[e] caregivers)
Support groups	• FPGP[f] • Family-led mutual support program	• Knowledge and skill acquisition • Enhance help seeking • Alleviation of guilt • Improved family and patient functioning • Decreased demand on mental health services • Improved psychosocial health • Improved QoL
Cognitive behavioral therapy	• Emotional regulation training	• Reduced burden of care • Reduced anxiety and stress • Increased resilience
Spirituality-based intervention	• Spirituality-based program	• Reduced anxiety and stress
Guided self-help interventions	• Manual-guided PBSP[g] • The Good Mood Guide	• Strengthened positive caregiving experience • Reduced negative caregiving experience • Increased access to support
Smartphone-based interventions	• MHapps[h]	• None
Mindfulness and empowerment interventions	• IEP[i] • MBSR[j] program	• Caregiver empowerment • Improved QoL • Decreased depressive symptoms • Increased self-efficacy
Collaborative interventions	• Participatory care model	• Reduced burden of care • Increased resilience
Standard care	• Multimodal intervention (general medicine, psychiatry, psychology, family therapy, neuropsychological rehabilitation, and occupational therapy)	• Reduced burden of care • Increased social support

[a]CoMHIP: community mental health early intervention project.
[b]RESHAPE: Reducing Stigma Among Healthcare Providers.
[c]FLEP: Family Link Education Programme.
[d]QoL: quality of life.
[e]BMD: bipolar mood disorder.
[f]FPGP: Family-led Peer Support Group Program.
[g]PBSP: problem-solving–based self-learning program.
[h]MHapps: mental health apps.
[i]IEP: integrated empowerment program.
[j]MBSR: mindfulness-based stress reduction.

XSL•FO
RenderX

Textbox 2. Description of the identified recommended strategies for alleviating informal caregiver burden.

Community-based interventions

- Strengthening existing primary health care system
- Establishment of day nursing or care
- Respite care
- Vocational training for patients
- Single-family caregiver organization
- Supported employment for patients
- Home visits

Psychoeducation interventions

- Group psychoeducation
- Family education
- Continuing education programs for effective practice
- Educational lectures
- Psychoeducation program

Support groups

- Nongovernmental mental health–related support groups
- Family caregiver assistance programs

Cognitive behavioral therapy

- Family therapy groups

Spirituality-based intervention

- Turning to traditional healers and spiritual leaders

Smartphone-based interventions

- Mental health apps

Mindfulness and empowerment interventions

- Stress management training

Collaborative interventions

- Caregiver involvement in all program elements
- Pluralistic and ecological approach to service delivery

Standard care

- Telepsychiatry
- Availability of emergency teams

Financial support

- Disability grant
- Fee-free mental health services
- Medical insurance and free medication

Social support

- Medical care social security
- Social resources

Counseling

XSL•FO
RenderX

- Basic counseling
- Supportive psychotherapy

Occupation-based interventions

- Physical and leisure activities

Policy and legislature

- National health insurance schemes
- Advocacy for a strong mental health policy
- Integration of caregiver actions and interventions into national mental health care plans
- Fee-free mental health services

Access to mental health care

- Caregiver care inclusion in daily treatment facilities
- Access to therapeutic tools
- Availability of mental health units or departments
- Periodic health checks for caregivers (every 6 months on average)
- Improvement in institutional mental health care

Psychoeducational Interventions

The first strategy was psychoeducational interventions, which was identified as the most common strategy implemented and recommended to alleviate caregiver burden. A total of 27% (12/44) of the studies implemented 4 different psychoeducational interventions, including 58% (7/12) of these studies using a psychoeducation program [35-40], 17% (2/12) using group psychoeducation [41,42], 17% (2/12) using the Family Link Education Programme [43,44], and 8% (1/12) of the studies implementing peer-assisted education [45] (Table 4). Similarly to the implemented intervention strategies, the recommended psychoeducational interventions identified included psychoeducational programs, group psychoeducation, family psychoeducation, continuing education programs for effective practice, and educational lectures [7,46] (Textbox 2).

Support Group Interventions

The second strategy was support groups. In total, 7% (3/44) of the studies implemented support group strategies, including the Family-led Peer Support Group Program, family-led mutual support programs, and nongovernmental mental health–related support groups [47-49] (Table 4). The recommended support group strategies identified in this review included nongovernmental mental health–related support groups and family caregiver support programs [50-54] (Textbox 2).

Community-Based Interventions

Community-based interventions were identified as a third strategy to alleviate caregiver burden, and 5% (2/44) of the included studies reported on implementing a community mental health early intervention project and Reducing Stigma Among Healthcare Providers [55,56] as community-based interventions (Table 4). The recommended community-based interventions included strengthening existing primary health care systems, establishing day nursing or care for the care recipients, respite care, vocational rehabilitation training for patients, single-family caregiver organizations, supported employment for patients, and home visits [46,52,53,57] (Textbox 2).

Guided Self-Help Interventions

A total of 5% (2/44) of the studies implemented guided self-help interventions, namely, a manual-guided problem-solving–based self-learning program [58] and the Good Mood Guide [59] (Table 4). No recommended guided self-help interventions were identified in this review.

Mindfulness and Empowerment Interventions

Mindfulness and empowerment interventions were implemented in 5% (2/44) of the studies, and the programs implemented were a mindfulness-based stress reduction program and an integrated empowerment program [60,61] (Table 4). Stress management training was identified as a recommended mindfulness and empowerment intervention strategy in this review [7] (Textbox 2).

CBT Interventions

Emotional regulation training was the only implemented CBT strategy identified in 5% (2/44) of the included studies [62,63] (Table 4). The recommended cognitive behavioral strategy identified in this review was family therapy groups [64] (Textbox 2).

Spirituality-Based Interventions

In total, 2% (1/44) of the studies implemented a spirituality-based program for informal caregivers in which the comparison intervention was 2 standard group training sessions related to general mental disorders [65] (Table 4). Turning to traditional healers and spiritual leaders was identified as a recommended spirituality-based intervention for alleviating informal caregiver burden [52] (Textbox 2).

XSL•FO
RenderX

Participatory Care Model

In total, 2% (1/44) of the studies implemented a participatory care model as a collaborative intervention strategy [66] (Table 4). The recommended collaborative interventions were caregiver involvement in all program elements and a pluralistic and ecological approach to service delivery [56,67] (Textbox 2).

Standard Care

The implemented standard care interventions were multimodal interventions that encompassed general medicine, psychiatry, psychology, family therapy, neuropsychological rehabilitation, and occupational therapy with no comparison intervention [68] (Table 4). The recommended standard care interventions were telepsychiatry and ensuring the availability of emergency teams [53] (Textbox 2).

Smartphone-Based Interventions

Only 2% (1/44) of the studies implemented smartphone-based interventions using mental health apps to alleviate the burden of informal caregivers with no comparison intervention [67] (Table 4). Although the study indicated the use of the mental health apps with care recipients and their caregivers, the specific intervention offered was not described. No recommended smartphone-based interventions were identified in this review.

Additional Strategies Recommended for Alleviating Caregiver Burden

The additional recommended intervention strategies were financial support, including the provision of a disability grant, fee-free mental health services, medical insurance, and free medication for care recipients [7,57,68]. The recommended social support comprised medical care social security and social resources [52,54,56,57,69]. Recommended counseling interventions comprised basic counseling and supportive psychotherapy [64,69]. Occupation-based interventions included participation in physical and leisure activities [64]. Policy and legislature recommended interventions comprised the implementation of national health insurance schemes, advocacy for a strong mental health policy, integration of caregiver actions and interventions into national mental health care plans, and fee-free mental health services [53,62,68]. Finally, other recommended interventions were access to mental health care, which considers caregiver care inclusion in daily treatment facilities; access to therapeutic tools; availability of mental health units or departments; periodic health checks for caregivers (every 6 months on average); and improved institutional mental health care [35,52,54,67] (Textbox 2).

Characteristics of the Implemented Intervention Strategies

Most of the implemented intervention strategies (18/29, 62%) were targeted at alleviating both objective and subjective burden, and 38% (11/29) were exclusively aimed at alleviating subjective burden. The intervention developers were reported as being researchers based on evidence (13/29, 45%) and researchers with expert input (5/29, 17%). In total, 28% (8/29) were existing interventions implemented without any adaptations, and 10% (3/29) of the interventions did not specify the intervention developer. The intervention deliverers identified in the studies included trained peer facilitators (7/29, 24%), researchers (7/29, 24%), psychiatrists and nurses (5/29, 17%), multiple health care professionals (4/29, 14%), and psychologists (3/29, 10%), and 10% (3/29) of the interventions did not specify who delivered them.

The number of sessions offered in the implemented intervention strategies was between 6 and 8 sessions (13/29, 45%), ≤5 sessions (8/29, 28%), and ≥11 sessions (5/29, 17%), and 10% (3/44) of the interventions did not specify the number of sessions. The duration of the implemented interventions in the studies was 6 to 15 weeks (10/29, 34%), ≤5 weeks (7/29, 24%), and ≥16 weeks (4/29, 14%), and 28% (8/29) of the interventions did not specify the duration. Most interventions (9/29, 31%) were delivered once per week, some were offered twice a week (8/29, 28%), and some of the interventions (12/29, 41%) did not specify the frequency of their sessions. The length of the sessions in the implemented interventions ranged between 2 and 2.5 hours (8/29, 28%), 1 to 1.5 hours (6/29, 21%), and ≤1 hour (1/29, 3%), and 48% (14/29) of the interventions did not specify the length of their sessions (Table 5).

XSL•FO
RenderX

Table 5. Characteristics of the implemented strategies for alleviating informal caregiver burden (N=29).

Study	Year of publication	Type of intervention	Type of burden targeted by the intervention	Intervention developer	Intervention deliverer	Number of sessions	Duration of intervention	Frequency of sessions per week	Length of the sessions
Ng et al [56]	2018	CoMHIP[a]	Subjective and objective	Not specified	The Integrated Community Centre for Mental Wellness	Not specified	Not specified	Not specified	Not specified
Rai et al [55]	2018	RESHAPE[b]	Subjective and objective burden	Developed as part of WHO[c] mhGAP[d] and PRIME[e]	TPO[f] Nepal, a Nepali non-governmental organization	5 sessions	4 days	Not specified	Not specified
Rajai et al [45]	2021	Peer- assisted education	Subjective and objective burden	Researchers and approved by 3 faculty members of the Army College of Medical Sciences	Trained peer facilitators	6 sessions	3 weeks	Twice weekly	1 hour
Zhou et al [44]	2020	FLEP[g]	Subjective and objective burden	FLEP: peer-led psychoeducation program developed from the stress and coping model by Pearlin et al [70]	Group facilitators who were peer specialists of experienced caregivers	8 sessions	8 weeks	Once weekly	2 hours
Dewi et al [37]	2019	Family psychoeducation and care decision without pasung	Subjective and objective burden	Researchers	Researchers	3 sessions	3 weeks	Not specified	35-40 minutes
Tabeleão et al [36]	2018	Psychoeducation	Subjective and objective burden	Not specified	10 psychologists	6 sessions	Not specified	Not specified	Not specified
Ntsayagae [40]	2017	Psychoeducation	Subjective and objective burden	Researcher based on existing literature	Researcher	2 sessions	Not specified	Not specified	Not specified
De Souza et al [39]	2016	Psychoeducation	Subjective and objective burden	Psychoeducational intervention for patients with BD[h] by Colom et al [71] translated and adapted by Dell-Aglio et al [72]	Psychiatrist responsible for specific case	6 sessions	6 to 8 weeks	Twice weekly	Not specified
Kolostoumpis et al [38]	2015	Psychoeducation	Subjective burden	Adapted from the treatment protocol developed by Reinares et al [73] in the Barcelona Bipolar Disorders Program in Spain	Psychiatrist and psychologist	7 sessions	Not specified	Not specified	Not specified
Jessy and Kalaimathy [35]	2014	Psychoeducation	Subjective and objective burden	Researcher based on WHO guidelines for educational interventions and existing literature. Manual translated into Tamil.	Researcher	5 modules	10 weeks	Twice weekly	1 hour

XSL•FO
RenderX

Study	Year of publication	Type of intervention	Type of burden targeted by the intervention	Intervention developer	Intervention deliverer	Number of sessions	Duration of intervention	Frequency of sessions per week	Length of the sessions
Fallahi Khoshknab et al [41]	2014	Group psychoeducation	Subjective and objective burden	Researchers based on educational program of psychiatric nursing and psychiatric textbooks	Organized based on educational program of psychiatric nursing and psychiatric textbooks (Campbell [74]) and converted to understandable text for patients and families	4 sessions	4 weeks	Once weekly	2 hours
Navidian et al [42]	2012	Psychoeducation	Subjective burden	Researchers based on the families' needs and the existing literature	Psychiatrist and mental health nurse	4 sessions	4 weeks	Once weekly	2 hours
Sharif et al [75]	2012	Psychoeducation	Subjective and objective burden	Psychiatrist and psychiatric nurse based on the literature and needs of the families	Psychiatrist, nurse, and guest speakers	10 sessions	5 weeks	Not specified	1.5 hours
Chiu et al [43]	2011	FLEP	Subjective burden	Task force consisting of a mental health social worker, a recovered patient with editorial experience, and a caregiver based on available related local educational materials and the NAMI[i] Family-to-Family program	Trainers who were themselves family members of people with SMI[j]. They received training and a trainer manual.	8 sessions	Not specified	Once weekly	Not specified
Chien et al [48]	2018	Family-led support program	Subjective and objective burden	Family-led mutual support group—contents were based on similar program protocols and the researcher-developed family mutual support groups for psychotic disorders	Family-led mutual support group—co-led by 2 peer family caregivers along with a researcher and rehabilitation nurse	16 sessions	36 weeks	Twice weekly	2 hours
Mentis et al [49]	2015	Support group intervention	Subjective burden	Nongovernmental mental health organization (NGOMH[k])	Not mentioned	Not specified	Not specified	Not specified	Not specified
Chien et al [48]	2018	Family-led mutual support program	Subjective burden	6 experts on psychiatric rehabilitation (including psychiatrists, clinical psychologists, and nurse specialists)	Peer leader who received training from the researchers and worked closely with a group leader who was a trained psychiatric nurse	14 sessions	39 weeks	Twice weekly	2 hours

XSL•FO
RenderX

Study	Year of publication	Type of intervention	Type of burden targeted by the intervention	Intervention developer	Intervention deliverer	Number of sessions	Duration of intervention	Frequency of sessions per week	Length of the sessions
Chien and Chan [47]	2013	FPGP[l]	Subjective and objective burden	Content and format were based on previous programs conducted by Li and Arthur [76] in mainland China. Appropriateness of the content was rated by 7 experts, including psychiatrists, psychologists, and nursing specialists.	Peer support—peer leader supported by principal researcher	14 group sessions	39 weeks	Twice weekly	2 hours
Behrouian et al [62]	2021	Emotional regulation	Subjective burden	On the basis of the Dialectical Behavior Therapy Skills Workbook and CBT[m] principles	Clinical psychologist	8 sessions	8 weeks	Once weekly	Not specified
Behrouian et al [63]	2020	Emotion regulation training	Subjective and objective burden	Sessions were based on previous studies (Gratz and Gunderson [77]). Trainings were based on the Dialectical Behavior Therapy Skills Workbook (McKay et al [78]).	Clinical psychologist	8 sessions	Not specified	Not specified	Not specified
Faghih and Pahlavanzadeh [79]	2019	CBT	Subjective burden	Researchers based on existing literature	Researcher	16 sessions	8 weeks	Twice weekly	1.5 hours
Khosravi et al [65]	2021	Spirituality-based program	Subjective burden	Researcher	Researcher	6 sessions	8 weeks	Twice weekly	1 hour
Chien et al [58]	2020	Manual-guided PBSP[n]	Subjective and objective burden	Developed by McCann et al [80] in Australia, and its Chinese translated version was validated and refined by the research team.	Psychiatric nurse	5 modules	21 weeks	Not specified	Not specified
McCann et al [59]	2015	Guided self-help manual	Subjective and objective burden	The Good Mood Guide developed by Lifeline South Coast NSW[o], Australia	Researcher via telephone	8 modules	8 weeks	Once weekly	Not specified
Deb et al [67]	2018	Smartphone-based interventions (MHapps[p])	Subjective and objective burden	Researcher	Psychiatrist	Not specified	Not specified	Not specified	Not specified

XSL•FO
RenderX

Study	Year of publication	Type of intervention	Type of burden targeted by the intervention	Intervention developer	Intervention deliverer	Number of sessions	Duration of intervention	Frequency of sessions per week	Length of the sessions
Hyun et al [61]	2018	IEP[q]	Subjective burden	Developed by Hyun et al [61] for community-living PMIs[r] based on the empowerment theories of Kanter [81] and McLean [82]	Mental health professionals who received training with a written structured intervention protocol from the research team	4 sessions	4 weeks	Once weekly	2 hours
Hou et al [60]	2014	MBSR[s] program	Subjective burden	Not stated but assumed that the researchers modeled the program on the original MBSR by Kabat-Zinn [83]	3 trained instructors with >3 years of experience in MBSR	8 sessions	8 weeks	Once weekly	2 hours, with a CD of 30-45 minutes for home practice
Zoladl et al [66]	2020	Participatory care model	Subjective burden	Researchers	Researchers and staff at the research site	8 sessions	12 weeks	Once weekly	1.5 hours
Ramirez et al [68]	2017	Standard care—included care from general medicine, psychiatry, psychology, family therapy, neuropsychological rehabilitation, and occupational therapy	Subjective and objective burden	Not stated	Medical officer, psychiatrist, psychologist, and occupational therapist	12-18 sessions	10 weeks	Not specified	Not specified

[a]CoMHIP: community mental health early intervention project.
[b]RESHAPE: Reducing Stigma Among Health Care Providers.
[c]WHO: World Health Organization.
[d]mhGAP: Mental Health Gap Action Programme.
[e]PRIME: Programme for Improving Mental Health Care.
[f]TPO: Transcultural Psychosocial Organization.
[g]FLEP: Family Link Education Programme.
[h]BD: bipolar disorder.
[i]NAMI: National Alliance on Mental Illness.
[j]SMI: severe mental illness.
[k]NGOMH: nongovernmental mental health.
[l]FPGP: Family-led Peer Support Group Program.
[m]CBT: cognitive behavioral therapy.
[n]PBSP: problem-solving–based self-learning program.
[o]NSW: New South Wales.
[p]MHapps: mental health apps.
[q]IEP: integrated empowerment program.
[r]PMI: person with mental illness.
[s]MBSR: mindfulness-based stress reduction.

XSL•FO
RenderX

Perceived Effectiveness of the Implemented Intervention Strategies

This review outlines the effectiveness of the implemented intervention strategies as reported by the authors who conducted the studies (Table 4). It should be noted that this scoping review included both published and gray literature, and therefore, no critical appraisal of the studies or meta-analyses were conducted. Therefore, care should be taken as these interpretations may be clouded by author bias [84].

The authors expressed the effectiveness of the implemented intervention strategies by highlighting whether the strategy resulted in a positive or negative outcome for the informal caregivers or their care recipients. Overall, the implemented strategies were reported to be effective in reducing the burden of care and improving the quality of life of informal caregivers. Psychoeducation intervention strategy outcomes included improved coping skills, improved quality of life, reduced anxiety and stress, reduced burden of care, reduced psychological strain, improved knowledge and skills in caregiving, enhanced family functioning, enhanced recovery, relapse prevention, and reduced prolonged admissions. Support group outcomes included alleviation of guilt, enhanced help-seeking behavior, improved quality of life, improved knowledge and skills in caregiving, improved psychosocial health, improved family and patient functioning, and decreased demand for mental health services. CBT outcomes included reduction in burden of care, anxiety, and stress and increased resilience of informal caregivers. Guided self-help outcomes included strengthened positive caregiving experiences, reduced negative caregiving experiences, and increased access to support. The reported outcomes for mindfulness and empowerment interventions included empowerment of caregivers, increased self-efficacy, improved quality of life, and decreased depressive symptoms. Collaborative intervention outcomes included reduced burden of care and increased resilience. The standard care outcomes reported included positive outcomes in caregiving, reduction in the burden of care, and increase in social support for caregivers.

Negative outcomes were only reported for the psychoeducation and mindfulness and empowerment intervention strategies. The reported negative outcomes of psychoeducation interventions were no improvement in the burden of care, quality of life, or self-esteem. The reported negative outcomes of mindfulness and empowerment interventions included short-lived improvement in anxiety, and the authors highlighted no improvement after 3 months of follow-up with the participants and no effect on perceived stress, quality of life, and self-compassion, indicating that the effects of the interventions were not sustainable.

Discussion

Principal Findings

This scoping review set out to map the literature on strategies for alleviating the burden of informal caregivers of persons with severe and enduring mental health conditions in low- and middle-income countries. The review identified types of strategies, strategy characteristics, and outcomes of the strategies as reported by the authors. The types of strategies identified were categorized as implemented and recommended intervention strategies and included community-based interventions, psychoeducational interventions, support groups, CBT, spirituality-based interventions, guided self-help, smartphone-based interventions, mindfulness and empowerment, collaborative interventions, standard care, financial and social support, counseling, occupation-based interventions, policy and legislature, and access to mental health.

Most of the implemented and recommended intervention studies (33/44, 75%) were conducted in Asian countries and targeted both subjective and objective burdens, and some studies (18/29, 62%) focused exclusively on subjective burden. This shows an increasing research interest in caregiver interventions in mental health research in Asian countries over the last decade. In contrast, the small number of studies from other low- and middle-income countries possibly confirms the limited research on support strategies for alleviating caregiver burden in mental health. Despite consensus on the high levels of burden of informal caregivers of persons with severe and enduring mental health conditions in these countries [56,85-87] and the importance of supporting caregivers, few studies (44/18,086, 0.24%) implemented and recommended intervention strategies for alleviating caregiver burden. This evidence gap raises significant concerns considering the scarcity of mental health care professionals and limited access to quality mental health services in many low- and middle-income countries. As a result, informal caregivers assume a vital role in the support and care of persons with severe and enduring mental health conditions in these regions [88,89]. Having limited strategies for alleviating the burden of informal caregivers may have dire consequences for the management of persons with severe and enduring mental health conditions in these countries.

The implemented and recommended intervention strategies were mainly focused on both male and female family caregivers. These findings are consistent with the literature as families have long been acknowledged as key stakeholders in the care and management of mental disorders [90-93]. Therefore, it is important to ensure that the strategies for alleviating caregiver burden are targeted specifically to this population, particularly in low- and middle-income countries, where occupying this role is often obligatory [94]. The fact that most implemented intervention strategies were aimed at both male and female caregivers is important to note. Although caregiving is identified as a female-oriented role, there is evidence that men also occupy the primary caregiver role and are also likely to experience high levels of caregiver burden [20,94]. This emphasizes the need to ensure that, where strategies for alleviating caregiver burden are implemented, they focus on all caregivers irrespective of their gender. The care recipients in these studies were mostly diagnosed with schizophrenia and mood disorders. This is consistent with the literature highlighting that depressive disorders, schizophrenia, and bipolar disorders are among the top 10 leading causes of disability in low- and middle-income countries [95].

The findings of this scoping review emphasize the need for evidence-based intervention strategies aimed at alleviating the

XSL•FO
RenderX

burden of informal caregivers in low- and middle-income countries. Most implemented intervention strategies were informed by evidence, in which researchers consulted existing literature and sought expert input to develop their interventions. Trained peer facilitators delivered most of the implemented interventions and were informal caregivers themselves. This is important to note as it aligns with task shifting, which is focused on transferring skills and responsibilities to local people with the aim of increasing access to mental health services in low- and middle-income countries where there is a shortage of human resources [96,97]. In addition, the findings of this review reveal that researchers, nurses, and psychiatrists offered some of the interventions, which may not be sustainable in low- and middle-income countries given the shortage of mental health professionals [95]. Most studies (13/29, 45%) offered 6 to 8 sessions over 6 to 8 weeks, and the sessions were facilitated once or twice a week for 1 to 2 and a half hours. Given the number and frequency of sessions, it may be useful to use peer facilitators to offer these interventions as most are already in the community, which will ensure sustainability in resource-constrained contexts in low- and middle-income countries.

Overall, the implemented intervention strategies were reported to have a positive effect on alleviating the burden of informal caregivers of people with severe and enduring mental health conditions in low- and middle-income countries. In some studies, these effects were reported for both informal caregivers and their care recipients. The authors reported the effectiveness of the strategies in their studies in terms of the help, benefit, and effect that a specific implemented intervention strategy had on the informal caregivers. Some compared their interventions mainly with standard psychiatric care and psychoeducation, which focused on providing caregivers with information on specific or general mental disorders. Psychoeducational interventions were frequently implemented intervention strategies and were identified as helpful, beneficial, and effective in reducing caregiver burden [35,38,42,75]. This type of intervention was reported to improve caregiver knowledge and skills to enable them to cope with the demands of caregiving. In addition, psychoeducational interventions were linked with positive patient outcomes as they were reported to enhance patient recovery, which subsequently reduced relapses and prolonged admissions. These interventions were commended for being simple, feasible, and cost-effective, making them the most preferred form of intervention strategy to address the burden of informal caregivers.

Similarly, implemented intervention strategies such as support groups, community-based interventions, guided self-help, mindfulness and empowerment, and CBT were reported to have a positive effect on the burden of informal caregivers in low- and middle-income countries. It is important to note that the family-led mutual support program, a peer-facilitated intervention, was reported to have long-term desirable benefits on the psychosocial health of both caregivers and their care recipients compared with the psychoeducational program and standard care offered by psychiatrists, clinical psychologists, and nursing specialists. Similarly, the manual-guided problem-solving–based self-learning program, a guided self-help intervention, was reported to have a superior treatment effect on caregiver burden, care recipients' symptom severity, and the duration of rehospitalizations at the 6-month follow-up compared with a well-accepted family psychoeducation group program. This implies that support groups and guided self-help, although not frequently implemented, should be considered beneficial strategies for alleviating caregiver burden. In addition, as self-directed and peer-evaluated interventions, these intervention strategies may be considered cost-effective in extending the services to informal caregivers, especially in low- and middle-income countries where a shortage of human resources affects the delivery and quality of mental health services.

Although not frequently implemented, intervention strategies such as spirituality-based, collaborative, and standard care interventions were also reported to reduce caregiver burden. Spirituality-based interventions were identified as an inexpensive and readily available resource for alleviating the burden of informal caregivers. Similarly, the participatory care model was reported to be an efficient and low-cost method for reducing caregiver burden and increasing caregiver resilience. Therefore, it is necessary to further explore the use of these interventions with informal caregivers of people with severe and enduring mental health conditions in low- and middle-income countries.

Despite their potential, the use of virtual interventions or telehealth strategies for informal caregivers remains underrepresented in the literature. This may be attributed to the relatively slow adoption of telehealth among informal caregivers. Notably, only 1 smartphone app was identified in this scoping review. However, telehealth holds promise as a viable approach for disseminating psychoeducational information, providing real-time support, and facilitating participation in virtual caregiver support groups, particularly in remote areas. The increasing proliferation of smartphones and internet access in Africa further emphasizes opportunities for the development of telehealth programs tailored to caregivers of individuals with severe and enduring mental health conditions. Although most of the implemented intervention strategies reported positive effects on caregiver burden, there were negative outcomes reported for some interventions. Hou et al [60] reported that the mindfulness-based stress reduction program led to a short-lived improvement in the stress and anxiety of the informal caregivers but that no improvement was noted after the 3-month follow-up. In addition, this intervention did not demonstrate a sustained effect on health-related quality of life, perceived stress, or self-compassion. These outcomes were attributed to the loss to follow-up and the fact that the measurement instruments could have been less sensitive to changes in quality of life and stress. Similarly, the 6-session individual psychoeducation intervention for informal caregivers of persons with bipolar mood disorder showed no effect on burden, quality of life, or self-esteem. This was attributed to the reduction in the number of sessions from 21 in the original instruction [71] to 6 in this study [35]. Another reason was that the sessions were conducted individually, whereas the literature highlights that multifamily intervention groups are effective [71,73,98].

XSL•FO
RenderX

Implications for Practice and Policy

The findings of this review provide evidence for the reported evaluated and recommended interventions having benefits in reducing the burden of care among informal caregivers in low- and middle-income countries. Previous studies conducted in low- and middle-income countries have since emphasized the urgent need to strengthen informal caregiver support regarding mental health. The need for information and skills in handling mental health care users, as well as the need for emotional and tangible support, has been highlighted in previous studies [42,99]. The findings of this review revealed that psychoeducation and support groups were highly used strategies for improving knowledge and skills as well as building support for informal caregivers in low- and middle-income countries. Furthermore, these strategies were reported to be beneficial and cost-effective, thereby making them a viable option for implementation in low- and middle-income countries, where limited access to mental health resources prevails [100]. The need to ensure that these strategies are offered on a continuous basis was highlighted in many studies, indicating the need for 6 to 10 sessions over a period of 6 to 10 weeks [7]. To ensure sustainability, training of peer facilitators to deliver these interventions may be realistic as this aligns with a task-shifting approach, which has long been advocated for as a cost-effective strategy for increasing access to mental health services in low- and middle-income countries [101]. It is interesting to note that peer support groups and guided self-help were reported to have long-term benefits on caregiver burden compared with psychoeducation; this highlights that using self-directed interventions is considered a practical option for alleviating informal caregiver burden. In low- and middle-income countries where a shortage of human resources prevails, it is important to consider such interventions as they empower informal caregivers to build support and take ownership of their health and well-being [35]. Spiritual-based interventions, although not frequently implemented, are important to note in the context of low- and middle-income countries as most informal caregivers have been identified as relying on spiritual and religious coping [102,103]. Recommended strategies such as the provision of financial and social support as well as policy and legislature strategies call for an urgent need for policies in mental health to shift focus toward integrating caregiver-oriented services into practice. In addition, these findings highlight the need to adopt an intersectoral approach [104] in which various sectors such as religious and spiritual organizations provide mental health services to extend their accessibility to informal caregivers.

Limitations

In alignment with Arksey and O'Malley, this review did not include a critical appraisal to ascertain the quality of the studies as the purpose of this scoping review was to map existing strategies for alleviating caregiver burden and report on the outcomes as stated by the authors. In their study, Woo et al [105] cautioned that the exclusion of critical appraisals in scoping reviews means that the review cannot ascertain the research gaps that it aims to address if the included studies are of poor methodological quality. Although this scoping review provides evidence on the existing strategies for alleviating caregiver burden, it is important that the suggested strategies be evaluated in the specific context to ensure their effectiveness before implementation in clinical practice. This step was omitted as the purpose of this review was to map the available intervention strategies and the outcomes reported by the authors in alleviating burden among informal caregivers in low- and middle-income countries. This review only included studies conducted in low- and middle-income countries, and only papers written in English were considered, thus reducing the extent and scope of the evidence on the strategies for alleviating burden among informal caregivers regarding mental health.

Conclusions

The findings of this scoping review provided the authors with categories they can use to develop semistructured interview guides to use when exploring the existing formal and informal community mental health services to alleviate the burden of informal caregivers in rural South Africa. The categories outline the different types of strategies that can be used to alleviate caregiver burden, in particular the types of strategies offered to informal caregivers and the intervention developers and deliverers in and outside the mental health care system. Although most of the included studies (29/44, 66%) implemented these strategies, a few studies conducted in other low- and middle-income countries (15/44, 34%) recommended the use of these interventions to alleviate caregiver burden. Future studies from low- and middle-income countries in other continents, including Africa and South and North America, should address this gap in the research by evaluating these intervention strategies to alleviate the burden of informal caregivers.

Critical appraisal, which is used to ascertain the quality of the studies, was omitted as the purpose of this scoping review was to map existing strategies for alleviating caregiver burden and report on the outcomes as stated by the authors. Future studies should conduct quality appraisals to establish the effectiveness of these strategies in alleviating the burden of informal caregivers. Peer-facilitated support group interventions, although not frequently implemented, were identified as having long-term benefits compared with frequently implemented interventions such as psychoeducation. It is recommended that future research be directed at implementing and evaluating these interventions to alleviate burden in low- and middle-income countries.

Acknowledgments

The authors wish to acknowledge Mr Devind Peter, a librarian at the Wits Health Sciences Library, for assisting with the development of the search strategy and retrieval of the literature sources for this scoping review. This study was supported by the Consortium for Advanced Research Training in Africa. The Consortium for Advanced Research Training in Africa is jointly led by the African Population and Health Research Center and the University of the Witwatersrand and funded by the Carnegie

XSL•FO
RenderX

Corporation of New York (grant G-19-57145); Swedish International Development Cooperation Agency (grant 54100113); Uppsala Monitoring Center; Norwegian Agency for Development Cooperation; and Wellcome Trust (reference 107768/Z/15/Z) and UK Foreign, Commonwealth, and Development Office with support from the Developing Excellence in Leadership, Training, and Science in Africa program. The statements made and views expressed are solely the responsibility of the fellow.

Data Availability

All data generated or analyzed during this study are included in this published paper.

Authors' Contributions

OS completed the literature search. OS and DC completed the screening of titles, abstracts, and full texts of the identified studies. In addition, they extracted data from all the included studies and completed the draft of the manuscript. FA and NGN reviewed the manuscript draft. OS incorporated the feedback from the authors for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data extraction templates.
[DOCX File , 23 KB - ijmr_v13i1e48587_app1.docx]

Multimedia Appendix 2
PRISMA-ScR Checklist.
[PDF File (Adobe PDF File), 156 KB - ijmr_v13i1e48587_app2.pdf]

References

1. Pillay AL. Is deinstitutionalisation a cheap alternative to chronic mental health care? S Afr J Psychol 2017 May 10;47(2):141-147 [FREE Full text] [doi: 10.1177/0081246317709959]
2. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
3. Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: a reappraisal from population-based studies. Gerontologist 2015 Apr;55(2):309-319 [FREE Full text] [doi: 10.1093/geront/gnu177] [Medline: 26035608]
4. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016 Feb;3(2):171-178 [FREE Full text] [doi: 10.1016/s2215-0366(15)00505-2]
5. Lynch SH, Shuster G, Lobo ML. The family caregiver experience - examining the positive and negative aspects of compassion satisfaction and compassion fatigue as caregiving outcomes. Aging Ment Health 2018 Nov;22(11):1424-1431 [FREE Full text] [doi: 10.1080/13607863.2017.1364344] [Medline: 28812375]
6. Dixon LB, Glynn SM, Cohen AN, Drapalski AL, Medoff D, Fang LJ, et al. Outcomes of a brief program, REORDER, to promote consumer recovery and family involvement in care. Psychiatr Serv 2014 Jan 01;65(1):116-120 [FREE Full text] [doi: 10.1176/appi.ps.201300074] [Medline: 24177229]
7. Chadda RK. Caring for the family caregivers of persons with mental illness. Indian J Psychiatry 2014 Jul;56(3):221-227 [FREE Full text] [doi: 10.4103/0019-5545.140616] [Medline: 25316932]
8. Paruk S, Ramdhial M. Prevalence of caregiver burden, depressive and anxiety symptoms in caregivers of children with psychiatric disorders in Durban, South Africa. S Afr J Psychiatr 2018 Sep 20;24:a1314 [FREE Full text] [doi: 10.4102/sajpsychiatry.v24i0.1314]
9. Naughton-Doe R, Moran N, Wakeman E, Wilberforce M, Bennett L, Webber M. Interventions that support unpaid carers of adult mental health inpatients: a scoping review. J Ment Health 2022 May 09:1-17 [FREE Full text] [doi: 10.1080/09638237.2022.2069702] [Medline: 35532039]
10. Yesufu-Udechuku A, Harrison B, Mayo-Wilson E, Young N, Woodhams P, Shiers D, et al. Interventions to improve the experience of caring for people with severe mental illness: systematic review and meta-analysis. Br J Psychiatry 2015 Apr;206(4):268-274 [FREE Full text] [doi: 10.1192/bjp.bp.114.147561] [Medline: 25833867]
11. Chisholm D, Docrat S, Abdulmalik J, Alem A, Gureje O, Gurung D, et al. Mental health financing challenges, opportunities and strategies in low- and middle-income countries: findings from the Emerald project. BJPsych Open 2019 Aug 06;5(5):e68 [FREE Full text] [doi: 10.1192/bjo.2019.24] [Medline: 31530327]
12. Alexander G, Bebee CE, Chen KM, Vignes RM, Dixon B, Escoffery R, et al. Burden of caregivers of adult patients with schizophrenia in a predominantly African ancestry population. Qual Life Res 2016 Feb;25(2):393-400 [FREE Full text] [doi: 10.1007/s11136-015-1077-5] [Medline: 26205769]

XSL•FO
RenderX

13. Mokgothu MC, Du Plessis E, Koen MP. The strengths of families in supporting mentally-ill family members. Curationis 2015 Apr 10;38(1):1258 [FREE Full text] [doi: 10.4102/curationis.v38i1.1258] [Medline: 26018195]
14. Kimotho SG. Understanding the nature of stigma communication associated with mental illness in Africa: a focus on cultural beliefs and stereotypes. In: Deconstructing Stigma in Mental Health. Hershey, PA: IGI Global; 2018.
15. Egbe CO, Brooke-Sumner C, Kathree T, Selohilwe O, Thornicroft G, Petersen I. Psychiatric stigma and discrimination in South Africa: perspectives from key stakeholders. BMC Psychiatry 2014 Jul 4;14(1):191 [FREE Full text] [doi: 10.1186/1471-244x-14-191]
16. Yerriah J, Tomita A, Paruk S. Surviving but not thriving: burden of care and quality of life for caregivers of patients with schizophrenia spectrum disorders and comorbid substance use in South Africa. Early Interv Psychiatry 2022 Feb;16(2):153-161. [doi: 10.1111/eip.13141] [Medline: 33733599]
17. Kayaalp A, Page KJ, Rospenda KM. Caregiver burden, work-family conflict, family-work conflict, and mental health of caregivers: a mediational longitudinal study. Work Stress 2021;35(3):217-240 [FREE Full text] [doi: 10.1080/02678373.2020.1832609] [Medline: 34483432]
18. Ewertzon M, Hanson E. Support interventions for family members of adults with mental illness: a narrative literature review. Issues Ment Health Nurs 2019 Sep;40(9):768-780. [doi: 10.1080/01612840.2019.1591547] [Medline: 31145025]
19. Soo SA, Zhang ZW, Khong SJ, Low JE, Thambyrajah VS, Alhabsyi SH, et al. Randomized controlled trials of psychoeducation modalities in the management of bipolar disorder: a systematic review. J Clin Psychiatry 2018;79(3):17r11750. [doi: 10.4088/JCP.17r11750] [Medline: 29727072]
20. Day K, Petrakis M. Family interventions in early psychosis service delivery: a systematized review. Soc Work Ment Health 2017;15(6):632-650 [FREE Full text] [doi: 10.1080/15332985.2016.1271381]
21. Claxton M, Onwumere J, Fornells-Ambrojo M. Do family interventions improve outcomes in early psychosis? A systematic review and meta-analysis. Front Psychol 2017 Mar 27;8:371 [FREE Full text] [doi: 10.3389/fpsyg.2017.00371] [Medline: 28396643]
22. Bademli K, Cetinkaya Duman Z. Family to family support programs for the caregivers of schizophrenia patients: a systematic review. Turk Psikiyatri Derg 2011;22(4):255-265 [FREE Full text] [Medline: 22143951]
23. Chen LF, Liu J, Zhang J, Lu XQ. Non-pharmacological interventions for caregivers of patients with schizophrenia: a meta-analysis. Psychiatry Res 2016 Jan 30;235:123-127 [FREE Full text] [doi: 10.1016/j.psychres.2015.11.037] [Medline: 26639649]
24. Sin J, Gillard S, Spain D, Cornelius V, Chen T, Henderson C. Effectiveness of psychoeducational interventions for family carers of people with psychosis: a systematic review and meta-analysis. Clin Psychol Rev 2017 Aug;56:13-24 [FREE Full text] [doi: 10.1016/j.cpr.2017.05.002] [Medline: 28578249]
25. Napa W, Tungpunkom P, Pothimas N. Effectiveness of family interventions on psychological distress and expressed emotion in family members of individuals diagnosed with first-episode psychosis: a systematic review. JBI Database Syst Rev Implement Rep 2017 Apr;15(4):1057-1079. [doi: 10.11124/JBISRIR-2017-003361] [Medline: 28398985]
26. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
27. Sin J, Henderson C, Woodham LA, Sesé Hernández A, Gillard S. A multicomponent eHealth intervention for family carers for people affected by psychosis: a coproduced design and build study. J Med Internet Res 2019 Aug 06;21(8):e14374 [FREE Full text] [doi: 10.2196/14374] [Medline: 31389333]
28. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-based interventions to improve mental health, general caregiving outcomes, and general health for informal caregivers of adults with chronic conditions living in the community: rapid evidence review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]
29. Silaule O, Nkosi NG, Adams F. Extent of caregiver burden among informal caregivers of persons with severe mental disorders in rural South Africa. Rural Remote Health 2023 Jun;23(2):7509 [FREE Full text] [doi: 10.22605/RRH7509] [Medline: 37264594]
30. Peters MD, Godfrey C, McInerney P, Munn Z, Trico AC, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, editors. Joanna Briggs Institute Reviewer's Manual. Adelaide, Australia: The Joanna Briggs Institute; 2020.
31. Silaule O, Adams F, Nkosi NG. Strategies for alleviating the burden experienced by informal caregivers of persons with severe mental disorders in transitional countries: protocol for a scoping review. JMIR Res Protoc 2023 Jul 24;12:e44268 [FREE Full text] [doi: 10.2196/44268] [Medline: 37486756]
32. Liu Z, Heffernan C, Tan J. Caregiver burden: a concept analysis. Int J Nurs Sci 2020 Jul 25;7(4):438-445 [FREE Full text] [doi: 10.1016/j.ijnss.2020.07.012] [Medline: 33195757]
33. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
34. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. BMC Med Res Methodol 2016 Feb 09;16:15 [FREE Full text] [doi: 10.1186/s12874-016-0116-4] [Medline: 26857112]

XSL•FO
RenderX

35. Jessy MF. Empowerment of caregivers of persons with schizophrenia: a group psychoeducational intervention. Mother Teresa Womens University. 2014. URL: https://shodhganga.inflibnet.ac.in:8443/jspui/handle/10603/289368 [accessed 2023-11-27]
36. Tabeleão V, Tomasi E, de Avila Quevedo L. A randomized, controlled trial of the effectiveness of a psychoeducational intervention on family caregivers of patients with mental disorders. Community Ment Health J 2018 Feb 25;54(2):211-217. [doi: 10.1007/s10597-017-0126-7] [Medline: 28444574]
37. Dewi AR, Daulima NH, Wardani IY. Managing family burden through combined family psychoeducation and care decision without pasung therapies. Enfermería Clínica 2019 Sep;29:76-80. [doi: 10.1016/j.enfcli.2019.04.012]
38. Kolostoumpis D, Bergiannaki JD, Peppou LE, Louki E, Fousketaki S, Patelakis A, et al. Effectiveness of relatives' psychoeducation on family outcomes in bipolar disorder. Int J Ment Health 2015 Sep 15;44(4):290-302. [doi: 10.1080/00207411.2015.1076292]
39. de Souza MS, da Silva RA, Molina MA, Jansen K, de Lima Ferreira L, Kelbert EF, et al. Six-session caregiver psychoeducation on bipolar disorder: does it bring benefits to caregivers? Int J Soc Psychiatry 2016 Jun 14;62(4):377-385. [doi: 10.1177/0020764016636909] [Medline: 26975694]
40. Ntsayagae EI. A psycho-educational programme for caregivers of mentally ill persons to facilitate their mental health. University of Johannesburg. 2017 Oct. URL: https://www.proquest.com/openview/cfa667c9ca773bc0b59f3df3d6d863f5/1?pq-origsite=gscholar&cbl=2026366&diss=y [accessed 2023-11-28]
41. Fallahi Khoshknab M, Sheikhona M, Rahgouy A, Rahgozar M, Sodagari F. The effects of group psychoeducational programme on family burden in caregivers of Iranian patients with schizophrenia. J Psychiatr Ment Health Nurs 2014 Jun 26;21(5):438-446. [doi: 10.1111/jpm.12107] [Medline: 23980535]
42. Navidian A, Kermansaravi F, Rigi SN. The effectiveness of a group psycho-educational program on family caregiver burden of patients with mental disorders. BMC Res Notes 2012 Aug 01;5:399 [FREE Full text] [doi: 10.1186/1756-0500-5-399] [Medline: 22853873]
43. Chiu MY, Wei GF, Lee S, Choovanichvong S, Wong FH. Empowering caregivers: impact analysis of FamilyLink Education Programme (FLEP) in Hong Kong, Taipei and Bangkok. Int J Soc Psychiatry 2013 Feb;59(1):28-39 [FREE Full text] [doi: 10.1177/0020764011423171] [Medline: 21971981]
44. Zhou DH, Chiu YL, Lo TL, Lo WF, Wong SS, Luk KL, et al. Outside-in or inside-out? A randomized controlled trial of two empowerment approaches for family caregivers of people with schizophrenia. Issues Ment Health Nurs 2020 Sep 04;41(9):761-772. [doi: 10.1080/01612840.2020.1734992] [Medline: 32497453]
45. Rajai N, Lami B, Pishgooie AH, Habibi H, Alavizerang F. Evaluating the effect of peer-assisted education on the functioning in family caregivers of patients with schizophrenia: a clinical trial study. Korean J Fam Med 2021 Sep;42(5):356-362 [FREE Full text] [doi: 10.4082/kjfm.20.0098] [Medline: 34607410]
46. Stanley S, Balakrishnan S, Ilangovan S. Correlates of caregiving burden in schizophrenia: a cross-sectional, comparative analysis from India. Soc Work Ment Health 2017;15(3):284-307 [FREE Full text] [doi: 10.1080/15332985.2016.1220440]
47. Chien WT, Chan SW. The effectiveness of mutual support group intervention for Chinese families of people with schizophrenia: a randomised controlled trial with 24-month follow-up. Int J Nurs Stud 2013 Oct;50(10):1326-1340. [doi: 10.1016/j.ijnurstu.2013.01.004] [Medline: 23433723]
48. Chien WT, Bressington D, Chan SW. A randomized controlled trial on mutual support group intervention for families of people with recent-onset psychosis: a four-year follow-up. Front Psychiatry 2018 Dec 18;9:710 [FREE Full text] [doi: 10.3389/fpsyt.2018.00710] [Medline: 30618880]
49. Mentis M, Messinis L, Kotrotsiou E, Angelopoulos NV, Marneras C, Papathanasopoulos P, et al. Efficacy of a support group intervention on psychopathological characteristics among caregivers of psychotic patients. Int J Soc Psychiatry 2015 Jun;61(4):373-378. [doi: 10.1177/0020764014547075] [Medline: 25149101]
50. Seshadri K, Sivakumar T, Jagannathan A. The family support movement and schizophrenia in India. Curr Psychiatry Rep 2019 Sep 14;21(10):95. [doi: 10.1007/s11920-019-1081-5] [Medline: 31522258]
51. Sagut P, Çetinkaya Duman Z. Comparison of caregiver burden in first episode versus chronic psychosis. Arch Psychiatr Nurs 2016 Dec;30(6):768-773 [FREE Full text] [doi: 10.1016/j.apnu.2016.07.011] [Medline: 27888973]
52. AlMakhamreh SS. Exploring experiences of informal carers of mental health: developing community intervention in social work in Jordan. Int Soc Work 2017 Mar 16;61(6):1042-1053 [FREE Full text] [doi: 10.1177/0020872817695391]
53. Bonsu AS, Salifu Yendork J, Teye-Kwadjo E. The influence of caregiver stress and affiliate stigma in community-based mental health care on family caregiver wellbeing. Int J Ment Health 2020;49(1):35-55 [FREE Full text] [doi: 10.1080/00207411.2020.1717053]
54. Wang M, Jin G, Wei Y, Wang F, Pan Z, Chen L, et al. Awareness, utilization and influencing factors of social supports for main informal caregivers of schizophrenia patients: a cross-sectional study in primary care settings in Beijing, China. BMC Fam Pract 2020 Sep 17;21(1):192 [FREE Full text] [doi: 10.1186/s12875-020-01257-z] [Medline: 32943006]
55. Rai S, Gurung D, Kaiser BN, Sikkema KJ, Dhakal M, Bhardwaj A, et al. A service user co-facilitated intervention to reduce mental illness stigma among primary healthcare workers: utilizing perspectives of family members and caregivers. Fam Syst Health 2018 Jun;36(2):198-209 [FREE Full text] [doi: 10.1037/fsh0000338] [Medline: 29902036]

XSL•FO
RenderX

56. Ng P, Young DK, Pan J, Law KK. The effects of a community mental health intervention project upon family caregivers of persons with suspected mental illness in the Chinese cultural context. Int Soc Work 2017 Mar 20;61(6):1067-1082. [doi: 10.1177/0020872817695642]
57. Chen L, Zhao Y, Tang J, Jin G, Liu Y, Zhao X, et al. The burden, support and needs of primary family caregivers of people experiencing schizophrenia in Beijing communities: a qualitative study. BMC Psychiatry 2019 Feb 20;19(1):75 [FREE Full text] [doi: 10.1186/s12888-019-2052-4] [Medline: 30786852]
58. Chien WT, Bressington D, Lubman DI, Karatzias T. A randomised controlled trial of a caregiver-facilitated problem-solving based self-learning program for family carers of people with early psychosis. Int J Environ Res Public Health 2020 Dec 14;17(24):9343 [FREE Full text] [doi: 10.3390/ijerph17249343] [Medline: 33327452]
59. McCann TV, Songprakun W, Stephenson J. A randomized controlled trial of guided self-help for improving the experience of caring for carers of clients with depression. J Adv Nurs 2015 Jul;71(7):1600-1610. [doi: 10.1111/jan.12624] [Medline: 25656334]
60. Hou RJ, Wong SY, Yip BH, Hung AT, Lo HH, Chan PH, et al. The effects of mindfulness-based stress reduction program on the mental health of family caregivers: a randomized controlled trial. Psychother Psychosom 2014;83(1):45-53. [doi: 10.1159/000353278] [Medline: 24281411]
61. Hyun MS, Nam KA, Kim H. Effects of a brief empowerment program for families of persons with mental illness in South Korea: a pilot study. Issues Ment Health Nurs 2019 Jan;40(1):8-14 [FREE Full text] [doi: 10.1080/01612840.2018.1453000] [Medline: 29847184]
62. Behrouian M, Ramezani T, Dehghan M, Sabahi A, Ebrahimnejad Zarandi B. The effect of the emotion regulation training on the resilience of caregivers of patients with schizophrenia: a parallel randomized controlled trial. BMC Psychol 2021 Mar 02;9(1):39 [FREE Full text] [doi: 10.1186/s40359-021-00542-5] [Medline: 33653410]
63. Behrouian M, Ramezani T, Dehghan M, Sabahi A, Ebrahimnejad Zarandi B. The effect of emotion regulation training on stress, anxiety, and depression in family caregivers of patients with schizophrenia: a randomized controlled trial. Community Ment Health J 2020 Aug;56(6):1095-1102 [FREE Full text] [doi: 10.1007/s10597-020-00574-y] [Medline: 32056061]
64. Souza AL, Guimarães RA, de Araújo Vilela D, de Assis RM, de Almeida Cavalcante Oliveira LM, Souza MR, et al. Factors associated with the burden of family caregivers of patients with mental disorders: a cross-sectional study. BMC Psychiatry 2017 Oct 25;17(1):353 [FREE Full text] [doi: 10.1186/s12888-017-1501-1] [Medline: 29070012]
65. Khosravi F, Fereidooni-Moghadam M, Mehrabi T, Moosavizade SR. The effect of a spirituality-based program on stress, anxiety, and depression of caregivers of patients with mental disorders in Iran. J Relig Health 2022 Feb;61(1):93-108 [FREE Full text] [doi: 10.1007/s10943-021-01372-w] [Medline: 34417681]
66. Zoladl M, Afroughi S, Nooryan K, Kharamin S, Haghgoo A, Parandvar Y. Applying collaborative care model on intensive caregiver burden and resilient family caregivers of patients with mental disorders: a randomized controlled trial. Iran J Psychiatry 2020 Feb 25. [doi: 10.18502/ijps.v15i1.2436]
67. Sinha Deb K, Tuli A, Sood M, Chadda R, Verma R, Kumar S, et al. Is India ready for mental health apps (MHApps)? A quantitative-qualitative exploration of caregivers' perspective on smartphone-based solutions for managing severe mental illnesses in low resource settings. PLoS One 2018 Sep 19;13(9):e0203353 [FREE Full text] [doi: 10.1371/journal.pone.0203353] [Medline: 30231056]
68. Ramírez A, Palacio JD, Vargas C, Díaz-Zuluaga AM, Duica K, Agudelo Berruecos Y, et al. Expressed emotions, burden and family functioning in schizophrenic and bipolar I patients of a multimodal intervention programme: PRISMA. Revista Colombiana de Psiquiatría (English ed.) 2017 Jan;46(1):2-11 [FREE Full text] [doi: 10.1016/j.rcpeng.2017.03.008]
69. Zeng Y, Zhou Y, Lin J. Perceived burden and quality of life in Chinese caregivers of people with serious mental illness: a comparison cross-sectional survey. Perspect Psychiatr Care 2017 Jul 23;53(3):183-189. [doi: 10.1111/ppc.12151] [Medline: 26909775]
70. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
71. Colom F, Vieta E, Scott J. Psychoeducation Manual for Bipolar Disorder. Cambridge, MA: Cambridge University Press; 2010.
72. Dell'Aglio JCJ, Figueoredo AL, Souza LD, Argimon IL. Cognitive behavioral model of bipolar disorder. In: Andretta I, Oliveira MS, editors. Pratical Manual of Cognitive Behavioral Therapy. São Paulo, Brazil: Casa do Psicólogo; 2011.
73. Reinares M, Colom F, Rosa AR, Bonnín CM, Franco C, Solé B, et al. The impact of staging bipolar disorder on treatment outcome of family psychoeducation. J Affect Disord 2010 Jun;123(1-3):81-86. [doi: 10.1016/j.jad.2009.09.009] [Medline: 19853922]
74. Campbell RJ. Campbell's Psychiatric Dictionary, 9th edition. New York, NY: Oxford University Press; 2009.
75. Sharif F, Shaygan M, Mani A. Effect of a psycho-educational intervention for family members on caregiver burdens and psychiatric symptoms in patients with schizophrenia in Shiraz, Iran. BMC Psychiatry 2012 May 25;12(1):48 [FREE Full text] [doi: 10.1186/1471-244x-12-48]
76. Li Z, Arthur D. Family education for people with schizophrenia in Beijing, China: randomised controlled trial. Br J Psychiatry 2005 Oct;187:339-345. [doi: 10.1192/bjp.187.4.339] [Medline: 16199793]

XSL•FO
RenderX

77. Gratz KL, Gunderson JG. Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. Behav Ther 2006 Mar;37(1):25-35. [doi: 10.1016/j.beth.2005.03.002] [Medline: 16942958]
78. McKay M, Wood JC. The Dialectical Behavior Therapy Skills Workbook: Practical Dbt Exercises for Learning Mindfulness, Interpersonal Effectiveness, Emotion Regulation, Emotion Regulation, and Distress Tolerance. Oakland, CA: New Harbinger Publications; 2007.
79. Faghih M, Pahlavanzadeh S. The effect of cognitive behavioral therapy on the burden in drug dependent persons' caregivers: a randomized controlled clinical trial. Iran J Nurs Midwifery Res 2019;24(2):131-136 [FREE Full text] [doi: 10.4103/ijnmr.IJNMR_86_18] [Medline: 30820225]
80. McCann TV, Cotton SM, Lubman DI. Social problem solving in carers of young people with a first episode of psychosis: a randomized controlled trial. Early Interv Psychiatry 2017 Aug;11(4):346-350 [FREE Full text] [doi: 10.1111/eip.12301] [Medline: 26592195]
81. Kanter RM. Men and Women of the Corporation. New York, NY: Basic Books; 1993.
82. McLean A. Empowerment and the psychiatric consumer/ex-patient movement in the United States: contradictions, crisis and change. Soc Sci Med 1995 Apr;40(8):1053-1071. [doi: 10.1016/0277-9536(94)00179-w] [Medline: 7597459]
83. Kabat-Zinn J, Hanh TN. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. New York, NY: Delacorte Press; 1990.
84. Olivier B, Verdonck M, Caseleijn D. Digital technologies in undergraduate and postgraduate education in occupational therapy and physiotherapy: a scoping review. JBI Evid Synth 2020 May;18(5):863-892. [doi: 10.11124/JBISRIR-D-19-00210] [Medline: 32813350]
85. Ayalew M, Workicho A, Tesfaye E, Hailesilasie H, Abera M. Burden among caregivers of people with mental illness at Jimma University Medical Center, Southwest Ethiopia: a cross-sectional study. Ann Gen Psychiatry 2019 Jun 24;18:10 [FREE Full text] [doi: 10.1186/s12991-019-0233-7] [Medline: 31285750]
86. Udoh EE, Omorere DE, Sunday O, Osasu OS, Amoo BA. Psychological distress and burden of care among family caregivers of patients with mental illness in a neuropsychiatric outpatient clinic in Nigeria. PLoS One 2021 May 6;16(5):e0250309 [FREE Full text] [doi: 10.1371/journal.pone.0250309] [Medline: 33956799]
87. Marimbe BD, Cowan F, Kajawu L, Muchirahondo F, Lund C. Perceived burden of care and reported coping strategies and needs for family caregivers of people with mental disorders in Zimbabwe. Afr J Disabil 2016 Aug 24;5(1):209 [FREE Full text] [doi: 10.4102/ajod.v5i1.209] [Medline: 28730046]
88. National mental health strategy 2012/13 - 2015/16. Federal Democratic Republic of Ethiopia Ministry of Health. URL: http://repository.iifphc.org/bitstream/handle/123456789/264/104.12%20National%20Mental%20Health%20Strategy%20--%20ETHIOPIA.pdf?sequence=1&isAllowed=y [accessed 2023-11-28]
89. Caregivers and mental illness: living with Schizophrenia. World Federation for Mental Health (WFMH). 2014. URL: http://discovery.ucl.ac.uk/137713/ [accessed 2023-11-28]
90. The world health report, 2001: mental health; new understanding, new hope. World Health Organization. 2001. URL: https://unesdoc.unesco.org/ark:/48223/pf0000123325 [accessed 2023-11-28]
91. Reupert A, Ward B, McCormick F, Ward C, Waller S, Kidd S. Developing a model of family focused practice with consumers, families, practitioners and managers: a community based participatory research approach. BMC Health Serv Res 2018 Jan 30;18(1):66 [FREE Full text] [doi: 10.1186/s12913-018-2844-0] [Medline: 29382350]
92. Maybery D, Jaffe IC, Cuff R, Duncan Z, Grant A, Kennelly M, et al. Mental health service engagement with family and carers: what practices are fundamental? BMC Health Serv Res 2021 Oct 09;21(1):1073 [FREE Full text] [doi: 10.1186/s12913-021-07104-w] [Medline: 34627245]
93. Cameron SL, Tchernegovski P, Maybery D. Mental health service users' experiences and perspectives of family involvement in their care: a systematic literature review. J Ment Health 2023 Jun;32(3):699-715 [FREE Full text] [doi: 10.1080/09638237.2022.2091760] [Medline: 35808821]
94. Olawale KO, Mosaku KS, Fatoye FO, Mapayi BM, Oginni OA. Caregiver burden in families of patients with depression attending Obafemi Awolowo University teaching hospitals complex Ile-Ife Nigeria. Gen Hosp Psychiatry 2014;36(6):743-747 [FREE Full text] [doi: 10.1016/j.genhosppsych.2014.08.008] [Medline: 25217492]
95. Rathod S, Pinninti N, Irfan M, Gorczynski P, Rathod P, Gega L, et al. Mental health service provision in low- and middle-income countries. Health Serv Insights 2017 Mar 28;10:1178632917694350 [FREE Full text] [doi: 10.1177/1178632917694350] [Medline: 28469456]
96. Ojha SP, Pant SB. Community based rehabilitation for task shifting in mental health. J Psychosoc Rehabil Ment Health 2018 Nov 29;5(2):109-110 [FREE Full text] [doi: 10.1007/s40737-018-0128-y]
97. Deimling Johns L, Power J, MacLachlan M. Community-based mental health intervention skills: task shifting in low- and middle-income settings. Int Perspect Psychol 2018 Oct;7(4):205-230. [doi: 10.1037/ipp0000097]
98. Madigan K, Egan P, Brennan D, Hill S, Maguire B, Horgan F, et al. A randomised controlled trial of carer-focussed multi-family group psychoeducation in bipolar disorder. Eur Psychiatry 2012 May;27(4):281-284 [FREE Full text] [doi: 10.1016/j.eurpsy.2010.12.008] [Medline: 21334858]

XSL•FO
RenderX

99. Ebrahimi H, Seyedfatemi N, Namdar Areshtanab H, Ranjbar F, Thornicroft G, Whitehead B, et al. Barriers to family caregivers' coping with patients with severe mental illness in Iran. Qual Health Res 2018 May 24;28(6):987-1001. [doi: 10.1177/1049732318758644] [Medline: 29478404]
100. Braathen SH, Vergunst R, Mji G, Mannan H, Swartz L. Understanding the local context for the application of global mental health: a rural South African experience. Int Health 2013 Mar;5(1):38-42. [doi: 10.1093/inthealth/ihs016] [Medline: 24029844]
101. Javadi D, Feldhaus I, Mancuso A, Ghaffar A. Applying systems thinking to task shifting for mental health using lay providers: a review of the evidence. Glob Ment Health (Camb) 2017 Jul 31;4:e14 [FREE Full text] [doi: 10.1017/gmh.2017.15] [Medline: 29230310]
102. Ae-Ngibise KA, Doku VC, Asante KP, Owusu-Agyei S. The experience of caregivers of people living with serious mental disorders: a study from rural Ghana. Glob Health Action 2015 May 11;8(1):26957 [FREE Full text] [doi: 10.3402/gha.v8.26957] [Medline: 25967587]
103. Modise TP, Mokgaola IO, Sehularo LA. Coping mechanisms used by the families of mental health care users in Mahikeng sub-district, North West province. Health SA 2021 Aug 16;26:1586 [FREE Full text] [doi: 10.4102/hsag.v26i0.1586] [Medline: 34522392]
104. Adeleye OA, Ofili AN. Strengthening intersectoral collaboration for primary health care in developing countries: can the health sector play broader roles? J Environ Public Health 2010;2010:272896 [FREE Full text] [doi: 10.1155/2010/272896] [Medline: 20454703]
105. Woo BF, Tam WW, Williams MY, Ow Yong JQ, Cheong ZY, Ong YC, et al. Characteristics, methodological, and reporting quality of scoping reviews published in nursing journals: a systematic review. J Nurs Scholarsh 2023 Jul 09;55(4):874-885. [doi: 10.1111/jnu.12861] [Medline: 36494752]

Abbreviations

CBT: cognitive behavioral therapy
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by T de Azevedo Cardoso; submitted 29.04.23; peer-reviewed by DM Ramafikeng, Y Goh; comments to author 12.07.23; revised version received 20.08.23; accepted 04.10.23; published 18.01.24.

Please cite as:
Silaule O, Casteleijn D, Adams F, Nkosi NG
Strategies to Alleviate the Burden Experienced by Informal Caregivers of Persons With Severe Mental Disorders in Low- and Middle-Income Countries: Scoping Review
Interact J Med Res 2024;13:e48587
URL: https://www.i-jmr.org/2024/1/e48587
doi:10.2196/48587
PMID:38236636

©Olindah Silaule, Daleen Casteleijn, Fasloen Adams, Nokuthula Gloria Nkosi. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 18.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Digital Methods for the Spiritual and Mental Health of Generation Z: Scoping Review

Susanna Y Park[1*], PhD; Bridgette Do[2*], MPH, PhD; Jacqlyn Yourell[2], MS, PhD; Janice Hermer[3], MA; Jennifer Huberty[2], PhD

[1]Skylight, Radiant Foundation, Salt Lake City, UT, United States
[2]Fit Minded Inc, Phoenix, AZ, United States
[3]Arizona State University, Tempe, AZ, United States
[*]these authors contributed equally

Corresponding Author:
Susanna Y Park, PhD
Skylight
Radiant Foundation
55 N 300 W Suite 800
Salt Lake City, UT, 84101
United States
Phone: 1 (720) 675 7228
Email: susanna@skylight.org

Abstract

Background: Generation Z (Gen Z) includes individuals born between 1995 and 2012. These individuals experience high rates of anxiety and depression. Most Gen Z individuals identify with being spiritual, and aspects from religion and spirituality can be integrated into mental health treatment and care as both are related to lower levels of depression. However, research on the spiritual and mental health of Gen Z is sparse. To date, there are no systematic or scoping reviews on digital methods to address the spiritual and mental health of Gen Z.

Objective: This scoping review aimed to describe the current state of digital methods to address spiritual and mental health among Gen Z, identify the knowledge gaps, and make suggestions for how to leverage digital spiritual and mental health interventions for Gen Z.

Methods: A comprehensive literature search was conducted in PubMed, Scopus, PsycInfo, CINAHL, Education Full Text, Google Scholar, SocIndex, and Sociological Abstracts. The inclusion criteria were as follows: (1) study population born between 1995 and 2012 (ie, Gen Z); (2) reporting on spiritual health or well-being, spirituality or religion, and mental health or well-being; (3) reporting on using digital methods; (4) publication in 1996 or beyond; (5) human subject research; (6) full text availability in English; (7) primary research study design; and (8) peer-reviewed article. Two authors screened articles and subsequently extracted data from the included articles to describe the available evidence.

Results: A total of 413 articles were screened at the title and abstract levels, of which 27 were further assessed with full text for eligibility. Five studies met the inclusion criteria, and data were extracted to summarize study characteristics and findings. The studies were performed across 4 different countries. There were 2 mixed-methods studies (South Africa and Canada), 2 cross-sectional studies (China and United States), and 1 randomized controlled trial (United States). Of these studies, only 2 discussed digital interventions (a text messaging–based intervention to improve spiritual and mental health, and a feasibility study for a mental health app). Other studies had a digital component with minor or unclear spiritual and mental health measures. Overall, there was a lack of consistency in how spiritual and mental health were measured.

Conclusions: Few studies have focused on assessing the spiritual and mental health of Gen Z in the digital context, and no research to date has examined a digital spiritual and mental health application among Gen Z. Research is needed to inform the development and evaluation of approaches to address the spiritual and mental health of Gen Z via digital means (eg, mobile apps).

(Interact J Med Res 2024;13:e48929) doi:10.2196/48929

XSL•FO
RenderX

KEYWORDS

Generation Z; Gen Z; spiritual health; digital mental health; spirituality

Introduction

The digital age is marked by widespread internet use and the ability to quickly communicate and find information online. Individuals born from 1995 to 2012 are considered "digital natives," as they are the first generation to live in an age where technology and the internet are accessible at all times [1]. The Pew Research Center considers these individuals as part of Generation Z ("Gen Z") and identifies the beginning of this generation to be in 1997 [1]. For the purpose of our research, we used the definition of Gen Z by Katz et al [2], who defined Gen Z as individuals born from 1995 to 2012, to account for those born when the World Wide Web made its public debut in 1995.

As a generation that grew up with technology, digital devices are familiar and seamlessly woven into the daily routines of Gen Z. It has been reported that 95% of Gen Z individuals have access to a smartphone [3]. In a study of 1000 Gen Z individuals aged 13 to 25 years, more than half spent 4 or more hours online compared with just 28% of all US adults who spent 4 or more hours online [4]. In a world where it is nearly impossible to socialize, work, and get an education without technology, Gen Z individuals are "always on," and this is associated with higher rates of depression, attention deficit disorder or attention deficit hyperactivity disorder, and technology addiction [5]. Compared with other generations, Gen Z individuals spend more time alone or on digital communication platforms than engaging in in-person interactions. Between 2000 and 2015, Gen Z high school seniors spent an hour less on in-person social interactions compared with early millennials [6]. Adolescents who spend more time on social media than in in-person interactions are the loneliest compared with those who spend less time on social media. Moreover, between 2007 and 2018, there were great increases in the relative percentage rates for self-injury (47%), seriously considering suicide (76%), and suicide attempts (58%) among Gen Z [7].

The American Psychological Association (APA) states that Gen Z individuals are more likely to report mental health concerns (eg, depression and anxiety) than previous generations [8]. In a 2022 survey of 1055 Gen Z adults, 1 out of 4 reported having more bad days than good within a 1-month time frame. More than 2 out of 5 (42%) had a diagnosed mental health condition, with more than a quarter of those being diagnosed during the COVID-19 pandemic (March 2020) or later. Anxiety and depression are the 2 largest mental health issues among Gen Z, with 9 out of 10 individuals diagnosed with a mental health condition having anxiety and 8 out of 10 having depression [3]. Notably, Gen Z individuals are the most comfortable discussing their mental health [3,8]. One third of Gen Z individuals report posting about their mental health on social media. They also attend therapy and are willing to pay out of pocket for mental health care and services [3]. Despite the comfort of Gen Z in talking about their mental health, there is a crucial need to address the high rates of anxiety, depression, and other health issues that they experience [8].

Spirituality may be an untapped resource to address the mental health crisis experienced by Gen Z today. While spirituality can serve as a component within organized religion, the 2 aspects are distinct. Religion is an organized belief or specific set of practices focusing on a higher power (ie, Christian, Muslim, Buddhist, etc) [9]. Spirituality is a broader concept in which individuals seek connection to self, others, nature, and a sacred or higher being [10]. Individuals may identify with being either religious or spiritual, or both. Gen Z individuals do not necessarily identify with a particular religion or belief but instead practice spirituality. Only half of Gen Z individuals report turning to their faith for support in times of uncertainty [11], and they are more likely to engage in spiritual practices than religious practices [12]. In a study of 10,000 Gen Z individuals aged 13 to 25 years, 68% considered themselves religious and 77% considered themselves spiritual [13]. Gen Z individuals define spirituality as autonomous and faith unbundled, and it is inclusive of all faiths and practices [11,13].

Spirituality is related to several positive health and psychosocial outcomes, namely greater mental health [14]. A recently updated review of the literature on the relationship between spirituality and mental health found that greater spirituality was associated with lower depressive symptoms, lower suicidality, and lower substance abuse [14]. Gen Z individuals face some of the highest rates of mental health conditions (eg, depression) [15,16]; thus, spirituality should be considered in addressing youth mental health today. In the aforementioned report by Singer [13], the majority of Gen Z individuals attributed their spiritual connection to their positive mental health state. Another aspect of mental health that is influenced by spirituality is quality of life among chronically and terminally ill patients. Palliative care patients who struggle with spirituality report poorer quality of life compared with those who feel stable with their spirituality [17]. Additionally, teens and young adults with cancer mention searching for meaning, hope, and life perspectives, even though they may not consider themselves as spiritual [18]. Interventions that promote spiritual well-being (one's sense of purpose, meaning in life, and connection to something greater [19]) may be a powerful resource for improving mental health in Gen Z.

Research on digital mental health interventions and spirituality exist separately. Little is known about digital methods (eg, mobile apps, text messaging, etc) that incorporate both spiritual and mental health among Gen Z. Scoping or systematic reviews on this topic are nonexistent, and research on this topic is very limited. Given that technology is woven into the daily lives of Gen Z, digital mobile apps that promote spirituality may offer a novel approach to supporting the mental health of Gen Z adolescents and young adults. Therefore, the purpose of this scoping review was to describe the current state of digital methods to address spiritual and mental health among Gen Z, identify the knowledge gaps, and make suggestions for how to leverage digital spiritual and mental health interventions for Gen Z.

XSL•FO
RenderX

Methods

Eligibility Criteria

The inclusion criteria for targeted articles were as follows: (1) study population born between 1995 and 2012 (ie, Gen Z); (2) reporting on spiritual health or well-being, spirituality or religion, and mental health or well-being; (3) reporting on using digital methods; (4) publication in 1996 or beyond; (5) human subject research; (6) full text availability in English; (7) primary research study design; and (8) peer-reviewed article.

Information Sources

Guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for scoping reviews (PRISMA-ScR) [20], searches were conducted in PubMed, Scopus, PsycInfo (EBSCO), CINAHL Plus with Full Text (ProQuest), Education Full Text (H.W. Wilson), Google Scholar, SocIndex with Full Text, and Sociological Abstracts (ProQuest) on May 2 and 3, 2023, by a health sciences librarian (J Hermer) at Arizona State University. The same information sources were searched on September 21-26, 2023, and again on November 7-8, 2023, and were also included in the results. ERIC and Atla Religion Database were also searched, but they returned no results.

Search Strategy

The searches were optimized for each individual database but included a combination of keywords and subject headings for the following 4 categories: Generation Z or Gen Z; spirituality or religion; mental well-being; and mobile health, mHealth, and eHealth (Multimedia Appendix 1). Owing to the very limited nature of the results, only language filters were applied to ensure that all relevant literature was available to be screened. All records were imported into Zotero to check if any articles were retracted and then into Covidence systematic review software for deduplication and screening [21]. Apart from the initial search, 2 additional searches were conducted to ensure that all relevant literature was found and to add any recent literature from the following 6 months. A search of all databases was performed on September 21-26, 2023, using the additional keywords of “faith,” “transcenden*,” and “life purpose” or “existential needs,” and a final search was performed on November 7, 2023, using the additional keywords of “spiritu*” and “relig*.” The scoping review protocol was registered online in Open Science Framework (OSF) on May 16, 2023.

Selection Process

The selection process was completed entirely on Covidence. Prior to reviewing titles and abstracts, duplicate titles were eliminated by Covidence. Two authors (SYP and BD) screened all titles and abstracts independently and were blinded to each other’s decisions. Any disagreements were discussed between the authors SYP and BD, and agreed upon before full-text review. Agreement scores for article selection between the 2 authors were not logged in Covidence for the initial search; however, the agreement scores were 85% (81/95) and 90% (80/87) for reviewing titles and abstracts between the 2 authors for the September and November searches, respectively. For full-text article review, the author SYP independently reviewed half of the articles, and the author BD independently reviewed the remaining articles. The authors SYP and BD deliberated with each other, and with J Huberty and JY if there were any questions regarding inclusion based on the article eligibility criteria. Articles were excluded if the study population included Gen Z but did not explicitly distinguish Gen Z in the population sample and the results were not disaggregated by age. The final articles included in the review were agreed upon by all 4 authors.

Data Collection Process and Synthesis Methods

Prior to the search, all authors agreed on the following characteristics for data extraction and synthesis: title; authors; study country; study objectives; study design; data collection timeframe; recruitment methods; sample size; participant characteristics; description of digital methods; constructs related to religion, spirituality, or spiritual well-being; assessment or measure for religion, spirituality, or spiritual well-being; constructs related to mental health; assessment or measure for mental health; main findings; and study limitations. These characteristics were selected to ensure a detailed understanding of available literature as it relates to the goal of the scoping review. For articles included in the scoping review, the author SYP independently extracted data based on the a priori characteristics for half of the articles and the author BD independently extracted data from the remaining articles. Given that each article had data extracted by a single author, there were no agreement scores. After evaluating the included articles, study characteristics and main findings were summarized in a descriptive manner.

Some studies that initially appeared eligible for this review were ultimately excluded because they did not meet specific criteria. For instance, a cross-sectional study on the perceptions of 475 Gen Z individuals and young millennials and their use of a spiritual self-care app [22] was originally included when reviewing titles and abstracts. However, upon full-text review, we found that the study results did not distinguish between Gen Z individuals and young millennials, thus failing to meet our review’s criteria (ie, Gen Z only).

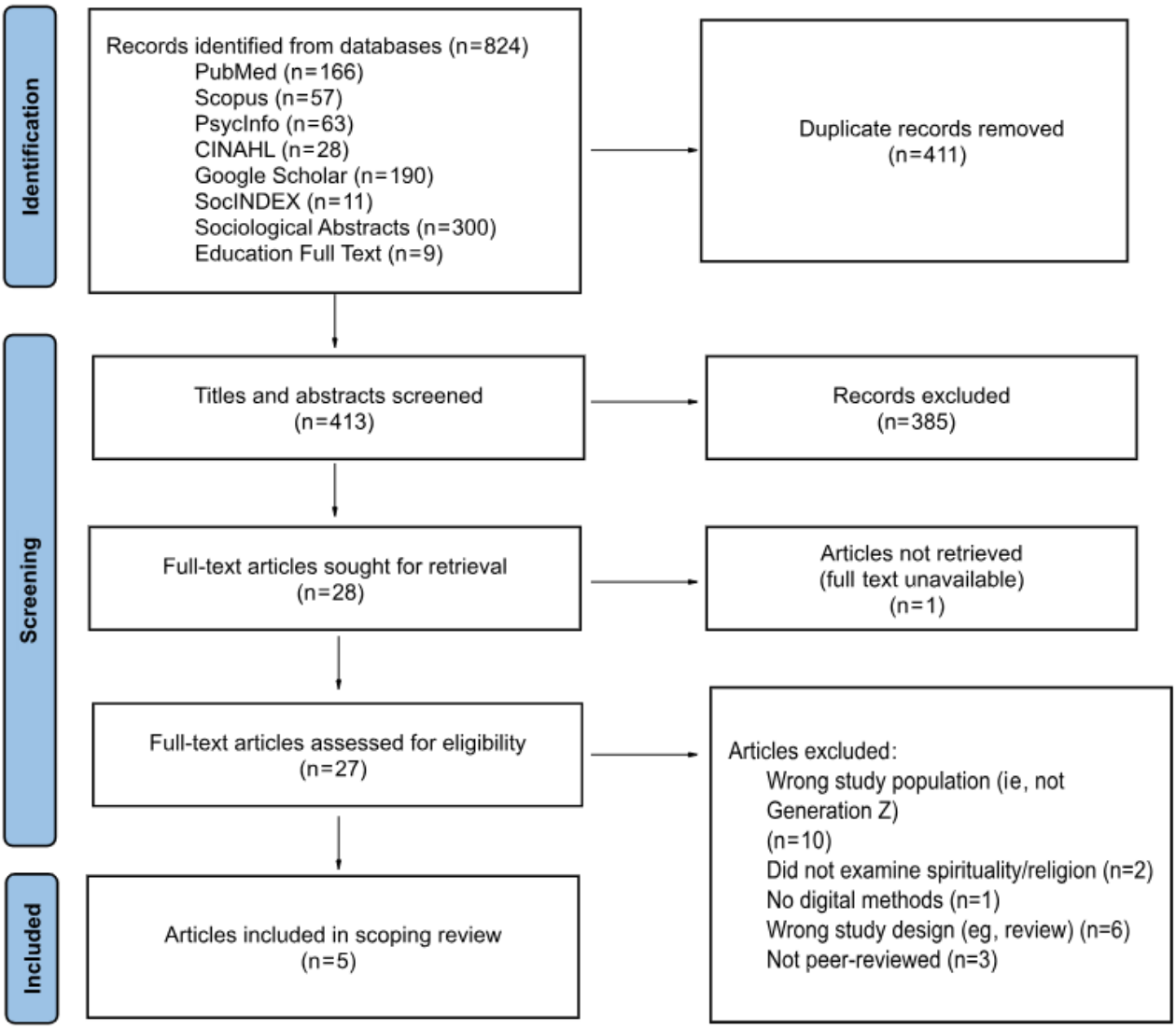
Results

Study Characteristics

Details on article selection are illustrated in a PRISMA diagram (Figure 1). Our search identified 824 articles from 8 databases based on the search terms. After removing 411 duplicates in Covidence, 413 articles were screened by title and abstract. After the initial screening based on the inclusion criteria, the full texts of the remaining 27 articles were screened. Review articles were not included in the final review, but references were screened to see if any additional literature was admissible. Ultimately, the scoping review included 5 articles. Of these 5 articles, 1 [23] was included in the review from the updated search conducted in September 2023 and 1 [24] was included from the updated search in November 2023. Characteristics and results of the studies are summarized in Table 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of article selection.



XSL•FO
RenderX

Table 1. Characteristics and main findings of the included studies.

Reference and study country	Study design	Sample characteristics	Digital aspect[a]	Religion/spirituality measures[b]	Mental health measures[b]	Main findings
Mindu et al [25], 2023 (South Africa)	Mixed-methods study	93 youth and young adults aged 16-24 years; 44% female	Assessed participants' knowledge and preferences for a digital mental health app	Where do youth seek treatment or assistance when they have mental health problems (Response options: Visit a spiritual healer; Go to a church for prayers; Traditional medicine; Clinic/hospital; Visit a health care worker)	Mental health awareness (eg, taught about mental health and no prior education on mental health) Mental health conditions affecting youth in the community (eg, substance abuse or misuse, posttraumatic stress disorder, depression, and anger)	• No participants had experience using a mental health app, but 99% indicated mental apps are important and can benefit youth. • Religious and cultural beliefs were a barrier to using digital platforms (eg, social media).
Gao et al [26], 2021 (China)	Cross-sectional study	1017 first-year college students (mean age 19 years); 77.8% female	Participants were recruited through an eHealth application to complete a survey	Religion (Response options: No religion; Buddhism; Christian; Others)	Generalized Anxiety Disorder-7 (GAD-7) and depression; Patient Health Questionnaire-9 (PHQ-9) [27,28]	• 95.3% indicated having no religion. • Belief in Christianity and in Buddhism were associated with greater anxiety.
Craig Rushing et al [29], 2021 (United States)	Randomized controlled trial	833 American Indian or Alaska Native teenagers and young adults aged 15-24 years; 66.3% female	The 2 intervention arms included 3 SMS text messages per week for 8 weeks with information, role model videos, images, and engagement opportunities (eg, reply for more information, resource links, etc)	Rate your spiritual health (Response options: Excellent; Very good; Good; Fair) [30]	Rate your mental health (Response options: Excellent; Very good; Good; Fair) [30]	• No significant differences between the 2 intervention arms within subjects. • Mean scores of perceived health (physical, mental, and spiritual) significantly increased over time for both intervention arms. • Those who reported better health also reported greater cultural resilience, identity, and cultural pride.
Reed et al [23], 2022 (United States)	Cross-sectional study	349 American Indian or Alaska Native youth aged 15-24 years; 71.1% female	Assessed participants' use of media technologies (ie, media types, frequency, and duration) and how they use media technologies (ie, online behaviors and activities)	Select the top 3 health topics from a list of 15, including spiritual health	Self-reported mental health (How good is your mental health? Response options: 4-point Likert-type scale; 4=Excellent and 1=Poor) Select the top 3 health topics from a list of 15, including mental health	• 53.5% of participants relied heavily on the internet to access health information. • Nonsexual and gender minority participants reported better mental health than sexual and gender minority participants. • The top 3 most important health topics were Native identity, mental health, and social justice and equality. • Spiritual health was selected as the most important health topic 14% of the time.

XSL•FO
RenderX

Reference and study country	Study design	Sample characteristics	Digital aspect[a]	Religion/spirituality measures[b]	Mental health measures[b]	Main findings
Au-Yeung et al [24], 2023 (Canada)	Mixed-methods study	5 Indigenous youth aged 11-16 years; 60.0% female	Assessed participants' opinions on the JoyPop app, and the accessibility and feasibility of the app	Qualitative question: "What does a 'Good Mind' (Haudenosaunee concept) mean to you?" "The Good Mind is a physical, psychological, and spiritual journey that includes a reflective awareness of thoughts and intentions, and a way of being that is expressed through self-compassion and compassion for other beings."	Perceived mental wellness (5-point Likert scale) Qualitative question: "Would you please describe what mental wellness means to you?"	• Participants had mixed ratings for their mental health from "fair" to "very good." • Participants reported nature to be important for a "Good Mind." • Participants reported enjoying the app's features (eg, games and breathing techniques) and made suggestions to make it relevant to their culture.

[a]Refers to the digital component of the study, which may include the mode through which an intervention was given (eg, mobile text messages) or the constructs measured (eg, social media use and preferences for a mobile app).

[b]Measures refer to assessments, tools, or specific survey items that were used in the study to assess the construct of interest (eg, religion or spirituality, and mental health).

Study Design and Study and Sample Characteristics

All 5 studies examined Gen Z (adolescents and young adults; aged 11-24 years; born between 1995 and 2012). Of the 5 articles, 2 were published in 2023, 1 was published in 2022, and 2 were published in 2021. There were 2 mixed-methods studies (1 in South Africa [25] and 1 in Canada [24]), 2 cross-sectional studies (1 in China [26] and 1 in the United States [23]), and 1 randomized controlled trial in the United States [29]. Sample sizes varied across the 5 studies, ranging from 5 to 1017.

Study by Mindu et al, 2023

Mindu et al [25] conducted a mixed-methods study in South Africa to inform the implementation of a mobile phone–based mental health intervention [25]. The study did not use any specified measures for mental health but asked questions via quantitative and qualitative methods to understand participants' understandings or perceptions of mental health. Questions addressed (1) mental health awareness; (2) digital interventions for mental health; (3) access to digital devices and the internet; (4) preferences for digital mental health interventions; and (5) barriers to the use of digital mental health innovations. Quantitative results showed that almost half (49%) of the participants had heard about mental health apps, none had experience using one, and 99% indicated that mental health apps are important and can benefit youth. These results underscore the severe lack of mobile mental health apps designed for individuals in South Africa. Qualitative results revealed that religious and cultural beliefs were a barrier to using digital platforms (eg, social media), and the authors highlighted the crucial need to develop culturally appropriate and relevant digital apps that represent those who they serve. Overall, participants expressed high interest in using a digital mental health app (eg, social media) to learn about mental health and seek resources. The study authors acknowledged that while no digital methods or interventions were implemented in the study, the results provided a unique Gen Z perspective on the usability of digital methods.

Study by Gao et al, 2021

Gao et al [26] conducted a cross-sectional study in China among first-year Chinese college students. Using a health management app (Residents e-Health), a questionnaire was distributed to examine depression and anxiety, and their associations with other health-related constructs, such as stress and nutrition. The authors examined participants' religious beliefs (ie, Christian, Buddhism, other, and no religion) as a correlate of anxiety and depression. However, most of the sample (95%) did not have a religion. Belief in Christianity and belief in Buddhism were associated with greater anxiety in the sample. The study authors noted that the large number of participants ascribing to no religion limited the understanding of the relationship between religion and mental health. The overall conclusion of the study indicated that early lifestyle interventions assessing religion, as well as other demographic and behavioral factors, are important for understanding the factors contributing to mental health in Gen Z.

Study by Craig Rushing et al, 2021

Craig Rushing et al [29] conducted a randomized controlled trial in the United States to examine the efficacy of an mHealth intervention (BRAVE) for physical, mental, and spiritual health; resilience; self-esteem; and coping and help-seeking skills among American Indian or Alaska Native teenagers and young adults. Participants were randomized to participate in 1 of 2 groups: (1) An 8-week intervention arm involving 3 SMS text messages per week highlighting common coping strategies, preferred wellness strategies, help-seeking skills, and related protective factors such as cultural resilience, identity, and cultural pride; or (2) An 8-week control arm involving 3 SMS text messages per week designed to elevate and reaffirm Native voices in STEM (science, technology, engineering, and mathematics) and medicine. Participants were in both arms, and both had messaging that included a combination of information,

XSL•FO
RenderX

role model videos, images, and opportunities of engagement (eg, reply for more information and links to access resources). The findings indicated that there were no significant differences between the 2 study arms within subjects, such that participants in the intervention arm did not report better outcomes than those in the control arm. However, mean scores of perceived health (ie, physical, mental, and spiritual) significantly increased over time in both arms. In addition, participants who reported better health also reported greater cultural resilience, identity, and cultural pride. The percentage of participants who used the resources and information in the intervention arm text messages also increased over 5 months. The study authors acknowledged some limitations, including high or favorable survey measure outcomes at baseline, only a 1-week break between receiving interventions in the arms, and control messaging likely being novel and helpful. Another limitation to note is the assessment of perceived health, which combined 3 separate survey measures assessing perceived physical, mental, and spiritual health. The combination of 3 different aspects of health makes it challenging to derive valid inferences regarding mental or spiritual health on their own. Overall, the intervention demonstrated improved health outcomes and underscores the acceptability of text messaging to promote and support well-being.

Study by Reed et al, 2022

Reed et al [23] conducted a cross-sectional study in the United States, in which 349 American Indian or Alaska Native youth (aged 15-24 years) were asked about the extent to which they use media technologies, how they use technologies, and their health priorities. Several trends were revealed. The majority of participants (64.7%) reported sending 1 to 50 text messages a day. Instagram was the most popular daily technology used, and 65.3% of participants reported using social media 3-7 hours per day. Participants also self-rated their mental health. The findings indicated that nonsexual and gender minority youth (56.7%) reported better mental health than sexual and gender minority youth (36.4%). To better understand important health topics, participants were asked to select their top 3 health topics from a list. The most popular topic selected was Native identity or cultural pride (73%), followed by mental health (57%) and social justice and equality (31%). Spiritual health was selected by 14% of youth in their top 3 health topics. While spiritual health was not among the top health topics selected, it is important to note that the range of options offered spanned across several categories (eg, social justice or inequality, alcohol and drug use or abuse, and the environment). Overall, the study authors concluded that building resources that foster cultural pride and positive identity must be included in any programs or technologies for addressing mental health among American Indian or Alaska Native youth.

Study by Au-Yeung et al, 2023

Au-Yeung et al [24] conducted a Haudenosaunee (Canada's largest First Nations reserve, Six Nations of the Grand River) community-based study in which 5 Haudenosaunee youth (aged 11-16 years) tested the JoyPop mobile app (available in English and French) that is designed to promote resilience among youth. The app offers breathing exercises, mood tracking, journaling, personalizable social support, a 24-hour helpline, and games. The results indicated that participants had mixed ratings on their self-reported mental wellness, ranging from "fair" to "very good," and that they used the app 1 to 3 times a day. Of the 5 participants, 4 were interviewed about their experiences using the app, perspectives of mental wellness, and characteristics of a "Good Mind" (an Indigenous concept on the physical, psychological, and spiritual journey that maintains balance and harmony in a person). Interview participants reported that the app was easy to use and esthetically pleasing. They also enjoyed all of the app's features, with the exception of the Circle of Trust feature (ie, personalizable social support). Interview participants identified positivity and happiness, understanding emotions, acts of kindness, personal hobbies, and positive body language as important to their mental wellness. Participants also discussed important characteristics of a Good Mind, such as positivity, kindness, and connecting with nature. Overall, the app was favorable to the participants, but they suggested incorporating specific features like words in their own language and Indigenous visuals (eg, feathers and clan animals). While the app lacked explicit content on a Good Mind, the authors recommended incorporating concepts of a Good Mind to enhance its relatability to Haudenosaunee youth, given its cultural significance to the Haudenosaunee people. Although the study had a small sample size, the authors concluded that mobile health interventions can be beneficial to Indigenous youth, as mental health apps continue to be of interest and Indigenous cultures value the promotion of health and resilience. Further, Indigenous tribes across North America have unique perspectives, and pan-Indigenous resilience apps like JoyPop will need to be tailored to specific cultural contexts.

Discussion

Overview

The purpose of this scoping review was to describe the current state of digital methods to address spiritual and mental health among Gen Z, identify the knowledge gaps, and make suggestions for how to leverage digital spiritual and mental health interventions for Gen Z. A comprehensive literature search across 8 databases identified only 5 relevant studies, emphasizing the significant lack of published research on digital methods to address the spiritual and mental health of Gen Z. Among the 5 studies, only 2 discussed digital interventions, and of these studies, 1 examined a text messaging–based intervention to improve spiritual and mental health and reported improvements in spiritual and mental health over time [27], and 1 examined the feasibility of a mental health app [24]. The sparse available literature limits conclusions on the current state of digital methods to address spiritual and mental health, and warrants future research to address these gaps.

Current State of Digital Methods for Spiritual and Mental Health

Gen Z individuals are facing a mental health crisis as they experience high rates of depression and anxiety [8,16]. For example, one of the included studies reported that the prevalences of anxiety and depression among college freshmen were 40.3% and 45.3%, respectively [26]. Gen Z individuals also report feeling lonely, having low self-confidence, and being

XSL•FO
RenderX

distressed about the future [3,5,12]. All 5 studies included in this review discussed the overall well-being of Gen Z and, to some extent, aspects of spirituality. The findings suggest that spirituality may play a role in the mental health of Gen Z and should be considered in the development and implementation of future digital applications that address mental health. The study by Mindu et al [25] particularly underscores the need for customizability in mental health applications considering that participants expressed that using digital platforms (ie, social media) conflicted with their religious and cultural beliefs. Thus, the introduction of applications that allow users to engage in practices and view content that aligns with their values and beliefs is a potential avenue for combating this barrier and could in turn strengthen one's spirituality. However, research is needed to determine the acceptability of digital mental health applications that involve spiritual content.

Nearly every facet of the lives of Gen Z involves technology, for example, using computers for school or work, using mobile apps to order food, and using social media and texting to communicate with friends. Despite the widespread use of technology among Gen Z and the increasing number of research studies employing digital methods to test and deliver mental health programs and interventions [31], only 1 of the studies included in this review examined the effects of a digital intervention on spiritual and mental health [27] and 1 assessed the feasibility of a mental health mobile app [24]. The lack of digital methods or interventions to address the spiritual and mental health of Gen Z warrants the development of accessible ways for Gen Z to practice spiritual and mental self-care. Gen Z individuals spend more time (4 or more hours daily) on social media compared with other generations [3], and poor mental health is often attributed to social media use [32]. However, social media is not necessarily harmful and is instead dependent upon what Gen Z individuals do and see online, their pre-existing strengths or vulnerabilities, and the environment in which they are raised (eg, parental monitoring) [32,33]. Digital methods for spiritual wellness that are specifically targeted for Gen Z and built to empower Gen Z to practice self-care and build healthy coping mechanisms may benefit the mental health of Gen Z. Results from this review illustrate the acceptability of a mobile intervention promoting spiritual and mental health [24,29], and the interest Gen Z individuals have toward using digital methods to address their spiritual and mental health [25]. However, given that little is known about this topic, more research is needed to truly grasp the feasibility and efficacy of digital approaches to address spiritual and mental health with regard to the well-being of Gen Z.

Gaps in the Literature

Published literature or research about the spiritual and mental health of Gen Z is limited, reporting mostly mental health statistics rather than examining determinants of mental health or interventions for improving mental health (eg, spirituality) [1,33]. Research to date on spirituality in Gen Z has only been performed by faith-based organizations or by nonacademic research institutions and has mainly focused on comparing the views of Gen Z on religion and spirituality to the views of other generations [11,12,34]. For example, the majority of Gen Z individuals (77%) identify as spiritual, preferring to ascribe to a set of values from various beliefs [13]. While we know that the majority of Gen Z individuals identify as spiritual, empirical research that specifically examines their spiritual practices and preferences, overall spiritual well-being, and associations with mental and physical health outcomes is warranted. Among the 5 studies included in this review, 3 examined spiritual health. The randomized controlled trial [29] asked participants to rate their spiritual health using a combined measure for overall health that included physical, mental, and spiritual health, limiting the ability to assess spiritual well-being explicitly among these participants. The cross-sectional survey [23] asked participants to select the top 3 health topics important to them (eg, mental health, spiritual health, and Native identity), but the study did not examine any associations between spiritual health and mental health. Spiritual health was not defined for participants in either study. The mixed-methods study in Canada [24] interviewed participants who used an app designed to promote resilience in youth with regard to a Good Mind, an Indigenous concept on the physical, psychological, and spiritual journey that maintains balance and harmony in a person. However, this concept is specific to the Haudenosaunee people; therefore, the findings may not be applicable to the perspectives on spirituality of other young communities. In addition, only 1 article administered a valid measure to assess mental health, spirituality, or spiritual well-being (ie, Generalized Anxiety Disorder-7 [GAD-7]) [26]. The validity and reliability of instruments used to measure mental health and spirituality or spiritual well-being in the study samples were not reported in any of the included articles. Identified gaps offer research opportunities to comprehensively examine spiritual and mental health among Gen Z.

Suggestions for Leveraging Digital Spiritual and Mental Health Interventions for Gen Z

Gen Z individuals are remarkably familiar with navigating digital spaces and integrating spirituality into their lives, and are the most comfortable talking about their mental health compared with other generations [3]. Gen Z individuals who have a spiritual connection have better perceptions about their mental health and believe their spiritual health contributes to their mental health, compared with those who do not have a spiritual connection [11]. Since Gen Z individuals primarily consume information through technology, digital interventions are a promising method to teach and facilitate various practices in spiritual and mental self-care. For example, the digital intervention content of Craig Rushing et al [29] included wellness strategies, such as self-care and goal setting, which resulted in improved perceived health (eg, physical, mental, and spiritual) for participants. In the study by Mindu et al [25], digital mobile interventions through social media were concluded to be potentially useful to increase mental health literacy and knowledge of resources. Additionally, the study by Gao et al [26] suggested that early interventions that target the lifestyle behaviors of Gen Z (ie, smoking) can improve their depression and anxiety. A spiritual self-care app tailored for Gen Z, for example, may allow Gen Z to engage with full autonomy and convenience. Rather than requiring Gen Z to seek out places to practice religious or spiritual beliefs, a spiritual self-care app can be available wherever they are and

XSL•FO
RenderX

whenever they want. Digital spiritual self-care interventions should also consider incorporating topics that are deemed important to Gen Z, such as cultural relevance, inclusivity, social justice, and nature [23,24].

Mobile apps grant users the ability to customize their experience, which vastly differs from traditional means of practicing religion that ascribe to a predetermined set of beliefs, values, and practices that are often fixed [13]. Access to technology enables Gen Z individuals to autonomously decide which aspects of different spiritual beliefs and practices they resonate with. Most digital apps include features that allow users to customize their experience and the content they engage with. Thus, a mental health app that includes various components of spiritual practices that users can choose from may provide a new way for young people to tend to their spiritual and mental health. Based on this review, research examining the feasibility, acceptability, and efficacy of digital health tools specifically targeting spiritual and mental health among Gen Z is absent, indicating the need for research in this area. Researchers, companies, and nonprofit organizations can leverage existing digital spaces that Gen Z individuals frequently use (eg, Instagram and TikTok) to garner feedback on what they might desire in a digital spiritual and mental health intervention. For example, 2 of the studies [24,29] incorporated feedback from Indigenous youth on existing digital methods (eg, SMS text messaging and JoyPop) to inform the development of their digital content. Spiritual well-being apps, for example, can be used to target Gen Z and deliver evidence-based content that integrates spirituality and mental health. These studies also emphasize the importance of culturally relevant interventions that can speak to diverse cultural backgrounds and beliefs among Gen Z.

Strengths and Limitations of This Review

There are multiple strengths of this review. First, this scoping review indicates that there is a limited knowledge base surrounding digital methods for addressing the spiritual and mental health of Gen Z and summarizes the current state of the literature on this topic. This is the first scoping review to address this topic and highlights a crucial gap in supporting young people's mental health. Additionally, a librarian was consulted and involved in the search process to bolster the rigor and accuracy of the review. Covidence was used to minimize human error in screening eligible articles. Along with these strengths, there are some limitations in this review. First, a scoping review limits the objective understanding of a topic such that quantitative results cannot be compiled to determine effect sizes across studies. Second, database searches are not uniform and require nuanced search methodologies, which can result in relevant studies being missed. Third, manuscripts during the full-text screening stage were split between the authors SYP and BD, and thus, manuscripts were assessed by a single rater. Finally, research on this topic is severely lacking, which limits the number of articles included in this review and impacts the ability to construct a cohesive narrative or draw definitive conclusions about the state of the field. This underscores the necessity of this review to highlight the gaps and urge further investigation.

Scientific Contribution

The goal of this scoping review was to assess the extent to which research on digital methods to address the mental and spiritual health of Gen Z has been conducted. The scoping review revealed a lack of available research on spirituality and mental health. Specifically, there is a dearth of studies on the use of digital methods to deliver spiritual well-being for mental health in Gen Z. Despite an increase in mental health concerns among Gen Z and the growing body of evidence on the beneficial effects of spiritual self-care on mental health, few published articles touch on this topic. The findings from this review highlight the opportunity for addressing the mental and spiritual health of Gen Z through digital methods (eg, mobile apps). The use of digital methods to address mental health is a growing area of research; however, spirituality and spiritual self-care have received little attention. There is potential for researchers to examine spiritual self-care, which is delivered through digital methods, and its impact on populations experiencing significant mental health problems. Overall, the scoping review underscores the need for future research to examine the acceptability and feasibility of digital approaches to address spiritual and mental health among Gen Z.

Conclusion

This scoping review underscores the dearth of research surrounding digital methods to address spiritual and mental health among Gen Z. Considering that digital methods to address aspects of mental health are increasingly popular and effective [35,36], research is needed to examine digital platforms that address spiritual and mental health. This is especially pertinent for Gen Z individuals as they have some of the greatest rates of mental health issues and are the most digitally savvy generation to date, and most indicate that they are spiritual. Leveraging spirituality as a way to address mental health among Gen Z via digital means offers a novel and relevant approach for addressing the mental health crisis impacting young people today.

Acknowledgments

The authors would like to thank Clare Beatty for the organization of this project and Justin Jones for inspiring us to write this paper.

Data Availability

The full review process, including the search terms and databases used, is available as a supplementary file.

XSL•FO
RenderX

Authors' Contributions

SYP and BD wrote and reviewed this paper. J Hermer searched and curated articles for review. JY reviewed and edited the manuscript. J Huberty supervised the review.

Conflicts of Interest

SYP is currently the Behavioral Research Scientist for the Radiant Foundation and does not receive incentives for the outcomes of the research. J Huberty is the Chief Science Officer for the Radiant Foundation's Skylight app. J Huberty was hired by the Radiant Foundation to lead scientific strategy and mentor the scientists for Skylight, but does not receive incentives for the outcomes of the research. BD and JY are employees of J Huberty and are independent from the Radiant Foundation leadership. The role of SYP and J Huberty is to ensure the quality of science regarding the Skylight app. J Hermer is a librarian at Arizona State University and is not associated with the Radiant Foundation or with J Huberty. The authors do not have stock in Skylight and receive no financial incentives from the sales of Skylight. The authors do not receive any financial benefit from the outcomes of this study.

Multimedia Appendix 1
Databases and search terms used for review.
[DOCX File , 23 KB - ijmr_v13i1e48929_app1.docx]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 239 KB - ijmr_v13i1e48929_app2.pdf]

References

1. Defining generations: Where Millennials end and Generation Z begins. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/ [accessed 2024-01-25]
2. Katz R, Ogilvie S, Shaw J, Woodhead L. Gen Z, Explained: The Art of Living in a Digital Age. Chicago, IL: The University of Chicago Press; 2021.
3. State of Gen Z Mental Health. Harmony Healthcare IT. 2022. URL: https://www.harmonyhit.com/state-of-gen-z-mental-health/ [accessed 2024-01-25]
4. Gen Z Is Extremely Online. Morning Consult. 2022. URL: https://pro.morningconsult.com/instant-intel/gen-z-social-media-usage [accessed 2023-05-10]
5. Twenge JM. iGen: Why Today's Super-Connected Kids Are Growing Up Less Rebellious, More Tolerant, Less Happy–and Completely Unprepared for Adulthood–and What That Means for the Rest of Us. New York, NY: Atria Books; 2017.
6. Pichler S, Kohli C, Granitz N. DITTO for Gen Z: A framework for leveraging the uniqueness of the new generation. Business Horizons 2021 Sep;64(5):599-610. [doi: 10.1016/j.bushor.2021.02.021]
7. Duffy ME, Twenge JM, Joiner TE. Trends in mood and anxiety symptoms and suicide-related outcomes among U.S. undergraduates, 2007-2018: Evidence from two national surveys. J Adolesc Health 2019 Nov;65(5):590-598. [doi: 10.1016/j.jadohealth.2019.04.033] [Medline: 31279724]
8. Gen Z more likely to report mental health concerns. APA. URL: https://www.apa.org/monitor/2019/01/gen-z [accessed 2023-04-11]
9. Arrey AE, Bilsen J, Lacor P, Deschepper R. Spirituality/religiosity: A cultural and psychological resource among Sub-Saharan African migrant women with HIV/AIDS in Belgium. PLoS One 2016 Jul 22;11(7):e0159488 [FREE Full text] [doi: 10.1371/journal.pone.0159488] [Medline: 27447487]
10. Jaysawal N, Saha S. COVID-19 and spiritual well-being: implications for social work. Journal of Religion & Spirituality in Social Work: Social Thought 2022 Nov 18;42(2):135-151. [doi: 10.1080/15426432.2022.2127391]
11. Springtide Research Institute. The State of Religion & Young People 2022: Mental Health–What Faith Leaders Need to Know. Winona, MN: Springtide Research Institute; 2022.
12. Faith Like a Roundabout: Religion and Gen Z in an Uncertain World. Patheos. 2022. URL: https://www.patheos.com/blogs/religionnow/2022/01/faith-like-a-roundabout-religion-and-gen-z-in-an-uncertain-world/ [accessed 2023-10-09]
13. Singer K. Gen Zers Are Religious, And Their Mental Health Depends On It. Religion Unplugged. 2023. URL: https://religionunplugged.com/news/2023/1/6/gen-z-is-religious-and-their-mental-health-depends-on-it [accessed 2024-01-25]
14. Lucchetti G, Koenig HG, Lucchetti ALG. Spirituality, religiousness, and mental health: A review of the current scientific evidence. World J Clin Cases 2021 Sep 16;9(26):7620-7631 [FREE Full text] [doi: 10.12998/wjcc.v9.i26.7620] [Medline: 34621814]
15. Generation Z and Mental Health. The Annie E. Casey Foundation. URL: https://www.aecf.org/blog/generation-z-and-mental-health [accessed 2024-01-25]
16. Most U.S. Teens See Anxiety and Depression as a Major Problem Among Their Peers. Pew Research Center. URL: http://tinyurl.com/2rnxseh8 [accessed 2024-01-25]

XSL•FO
RenderX

17. Damen A, Exline J, Pargament K, Yao Y, Chochinov H, Emanuel L, et al. Prevalence, predictors and correlates of religious and spiritual struggles in palliative cancer patients. J Pain Symptom Manage 2021 Sep;62(3):e139-e147 [FREE Full text] [doi: 10.1016/j.jpainsymman.2021.04.024] [Medline: 33984462]
18. Barton KS, Tate T, Lau N, Taliesin KB, Waldman ED, Rosenberg AR. "I'm not a spiritual person." How hope might facilitate conversations about spirituality among teens and young adults with cancer. J Pain Symptom Manage 2018 Jun;55(6):1599-1608 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.02.001] [Medline: 29428188]
19. Spiritual Wellbeing. Australian National University. URL: https://www.anu.edu.au/spiritual-wellbeing [accessed 2024-01-25]
20. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
21. Covidence. URL: https://www.covidence.org/ [accessed 2024-01-25]
22. Park SY, Huberty J, Yourell J, McAlister KL, Beatty CC. A spiritual self-care mobile app (Skylight) for mental health, sleep, and spiritual well-being among generation Z and young millennials: Cross-sectional survey. JMIR Form Res 2023 Sep 27;7:e50239 [FREE Full text] [doi: 10.2196/50239] [Medline: 37597179]
23. Reed N, Peterson R, Ghost Dog T, Kaufman C, Kelley A, Craig Rushing S. Centering native youths' needs and priorities: Findings from the 2020 Native Youth Health Tech Survey. Am Indian Alsk Native Ment Health Res 2022 Sep;29(3):1-17. [doi: 10.5820/aian.2903.2022.1] [Medline: 36178745]
24. Au-Yeung A, Marfatia D, Beers K, General DA, McQueen KCD, Martin-Hill D, Six Nations Youth Mental Wellness Committee. Exploring the feasibility of a mental health application (JoyPop) for Indigenous youth. Front Psychiatry 2023 Oct 6;14:1269347 [FREE Full text] [doi: 10.3389/fpsyt.2023.1269347] [Medline: 37867769]
25. Mindu T, Mutero IT, Ngcobo WB, Musesengwa R, Chimbari MJ. Digital mental health interventions for young people in rural South Africa: Prospects and challenges for implementation. Int J Environ Res Public Health 2023 Jan 13;20(2):1453 [FREE Full text] [doi: 10.3390/ijerph20021453] [Medline: 36674209]
26. Gao C, Sun Y, Zhang F, Zhou F, Dong C, Ke Z, et al. Prevalence and correlates of lifestyle behavior, anxiety and depression in Chinese college freshman: A cross-sectional survey. Int J Nurs Sci 2021 Jul 10;8(3):347-353 [FREE Full text] [doi: 10.1016/j.ijnss.2021.05.013] [Medline: 34307785]
27. Spitzer R, Kroenke K, Williams J, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
28. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
29. Craig Rushing S, Kelley A, Bull S, Stephens D, Wrobel J, Silvasstar J, et al. Efficacy of an mHealth Intervention (BRAVE) to promote mental wellness for American Indian and Alaska native teenagers and young adults: Randomized controlled trial. JMIR Ment Health 2021 Sep 15;8(9):e26158 [FREE Full text] [doi: 10.2196/26158] [Medline: 34524092]
30. Hernandez BC, Vigna JF, Kelley ML. The Youth Coping Responses Inventory: development and initial validation. J Clin Psychol 2010 Sep;66(9):1008-1025. [doi: 10.1002/jclp.20697] [Medline: 20694963]
31. Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: Systematic overview. JMIR Ment Health 2021 Apr 29;8(4):e25847 [FREE Full text] [doi: 10.2196/25847] [Medline: 33913817]
32. Gen Z mental health: The impact of tech and social media. McKinsey Health Institute. URL: https://www.mckinsey.com/mhi/our-insights/gen-z-mental-health-the-impact-of-tech-and-social-media [accessed 2024-01-25]
33. Health advisory on social media use in adolescence. APA. URL: https://www.apa.org/topics/social-media-internet/health-advisory-adolescent-social-media-use [accessed 2024-01-25]
34. When You Say You Believe In God, What Do You Mean? The Pew Charitable Trusts. URL: https://www.pewtrusts.org/en/trust/archive/fall-2018/when-you-say-you-believe-in-god-what-do-you-mean [accessed 2024-01-25]
35. Laird B, Puzia M, Larkey L, Ehlers D, Huberty J. A mobile app for stress management in middle-aged men and women (Calm): Feasibility randomized controlled trial. JMIR Form Res 2022 May 24;6(5):e30294 [FREE Full text] [doi: 10.2196/30294] [Medline: 34989677]
36. Neher T, Green J, Puzia M, Huberty J. Describing the use of a mindfulness-based app for sleep and mental well-being, across age, in children. Child Youth Care Forum 2021 Sep 24;51(4):749-768. [doi: 10.1007/s10566-021-09651-z]

Abbreviations

Gen Z: Generation Z
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

Edited by G Eysenbach, T de Azevedo Cardoso; submitted 11.05.23; peer-reviewed by F Dos Santos; comments to author 14.09.23; revised version received 11.10.23; accepted 15.01.24; published 06.02.24.

Please cite as:
Park SY, Do B, Yourell J, Hermer J, Huberty J
Digital Methods for the Spiritual and Mental Health of Generation Z: Scoping Review
Interact J Med Res 2024;13:e48929
URL: https://www.i-jmr.org/2024/1/e48929
doi:10.2196/48929
PMID:38261532

©Susanna Y Park, Bridgette Do, Jacqlyn Yourell, Janice Hermer, Jennifer Huberty. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 06.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Improvements in Neoplasm Classification in the International Classification of Diseases, Eleventh Revision: Systematic Comparative Study With the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision

Yicong Xu[1*], MPH; Jingya Zhou[2,3*], MPH; Hongxia Li[4], MD; Dong Cai[1], BA; Huanbing Zhu[5*], MD, PhD; Shengdong Pan[6*], MPA, LLM

[1]Medical Records Room, Department of Medical Administration, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

[2]Department of Medical Records, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[3]Collaborating Center for the WHO Family of International Classifications in China, Beijing, China

[4]Department of Gastroenterology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

[5]Department of Hepatobiliary and Pancreatic Surgery, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

[6]Department of Medical Administration, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Shengdong Pan, MPA, LLM
Department of Medical Administration
The Second Affiliated Hospital, Zhejiang University School of Medicine
88 Jiefang Road
Hangzhou, 310009
China
Phone: 86 1 366 666 9123
Email: psd618@zju.edu.cn

Abstract

Background: The International Classification of Diseases, Eleventh Revision (ICD-11) improved neoplasm classification.

Objective: We aimed to study the alterations in the ICD-11 compared to the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision (ICD-10-CCM) for neoplasm classification and to provide evidence supporting the transition to the ICD-11.

Methods: We downloaded public data files from the World Health Organization and the National Health Commission of the People's Republic of China. The ICD-10-CCM neoplasm codes were manually recoded with the ICD-11 coding tool, and an ICD-10-CCM/ICD-11 mapping table was generated. The existing files and the ICD-10-CCM/ICD-11 mapping table were used to compare the coding, classification, and expression features of neoplasms between the ICD-10-CCM and ICD-11.

Results: The ICD-11 coding structure for neoplasms has dramatically changed. It provides advantages in coding granularity, coding capacity, and expression flexibility. In total, 27.4% (207/755) of ICD-10 codes and 38% (1359/3576) of ICD-10-CCM codes underwent grouping changes, which was a significantly different change (χ^2_1=30.3; P<.001). Notably, 67.8% (2424/3576) of ICD-10-CCM codes could be fully represented by ICD-11 codes. Another 7% (252/3576) could be fully described by uniform resource identifiers. The ICD-11 had a significant difference in expression ability among the 4 ICD-10-CCM groups (χ^2_3=93.7; P<.001), as well as a considerable difference between the changed and unchanged groups (χ^2_1=74.7; P<.001). Expression ability negatively correlated with grouping changes (r=–.144; P<.001). In the ICD-10-CCM/ICD-11 mapping table, 60.5% (2164/3576) of codes were postcoordinated. The top 3 postcoordinated results were specific anatomy (1907/3576, 53.3%), histopathology (201/3576, 5.6%), and alternative severity 2 (70/3576, 2%). The expression ability of postcoordination was not fully reflected.

XSL•FO
RenderX

Conclusions: The ICD-11 includes many improvements in neoplasm classification, especially the new coding system, improved expression ability, and good semantic interoperability. The transition to the ICD-11 will inevitably bring challenges for clinicians, coders, policy makers and IT technicians, and many preparations will be necessary.

(Interact J Med Res 2024;13:e52296) doi:10.2196/52296

KEYWORDS

Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision; ICD-10; ICD-10-CCM; ICD-11; improvement; International Classification of Diseases, Eleventh Revision; International Classification of Diseases, Tenth Revision; International Classification of Diseases; neoplasm; transition

Introduction

The World Health Organization (WHO) adopted the International Classification of Diseases, Tenth Revision (ICD-10) in May 1990 [1]. The ICD-10 has been widely used in over 120 countries over the past 30 years. In multiple countries, expansions of the ICD-10, such as the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision (ICD-10-CCM), based on the second edition [2], have been developed to meet specific requirements. The ICD-10-CCM has been used for national performance assessment and medical insurance payment in public hospitals. Given the advances in medical knowledge and health information, the WHO started working on the International Classification of Diseases, Eleventh Revision (ICD-11) in 2007 [3], and it was expected to be implemented by WHO member countries starting in January 2022.

The most critical changes in the ICD-11 were the establishment of a semantic knowledge base and the reconstruction of the coding system. It introduced the Foundation Component, the Common Ontology, and linearization, as well as the new concepts of precoordination and postcoordination [3-5]. The chapter on neoplasms has also changed. It adds crucial morphology to precoordination, having the highest proportion (98.9%) and the most dimensions (3.5) of postcoordination. Cancer has been one of the top 3 causes of death in China since 2005 [6]; thus, the coding system of the ICD-11 will significantly impact cause of death reporting, cancer registration, and disease diagnosis records. Stakeholders need to have a good understanding of the classification of neoplasms in the ICD-11.

This study analyzes the changes between the ICD-10-CCM and ICD-11 in terms of coding features, classification features, and expression features in neoplasm classification and hopes to provide evidence supporting the transition in China.

Methods

Ethical Considerations

According to the Measures for Ethical Review of Human Life Science and Medical Research issued by the National Health Commission of the People's Republic of China, this study utilized public data and did not involve human subjects, and thus, the requirement of ethical permission was waived. All examples were constructed and neither correspond to real clinical cases nor to any datasets.

Materials

The following 3 existing, publicly available files were used in this paper: (1) ICD-10/ICD-11 mapping tables [7], (2) ICD-11 simple tabulation [8], and (3) the second revision of the ICD-10-CCM [9].

The first 2 files were downloaded from the WHO website, and the ICD-10-CCM was released by the National Health Commission of the People’s Republic of China.

Research Methods

The ICD-10-CCM neoplasm codes were manually recoded by the ICD-11 coding tool [10] based on the International Classification of Diseases, Eleventh Revision, Mortality and Morbidity Statistics (ICD-11 MMS) codes to generate an ICD-10-CCM/ICD-11 mapping table for neoplasms. The 3 existing files and the ICD-10-CCM/ICD-11 mapping table were used to analyze ICD-11 features, namely, the coding structure, coding capacity, grouping changes, expression ability, expression flexibility, and the expression of postcoordination in the neoplasm classification.

Mapping ICD-10-CCM to ICD-11 Codes

Due to the homology of the morphology section between the ICD-10 and ICD-11, mapping was performed for only the topography codes. To ensure the accuracy of the results, manual recoding was independently implemented by 2 authors who both had more than 10 years of coding experience and had received ICD-11 training. Inconsistent results were resolved by consulting a senior coder from the Collaborating Center for the WHO Family of International Classifications in China.

Statistical Standards

Leaf Codes

The ICD-10, ICD-10-CCM, and ICD-11 MMS codes that can be used at the lowest level are called leaf codes. All statistical analyses were based on the leaf codes.

Coding Capacity

This capacity involves the number of leaf codes that can be used in actual coding.

Grouping Changes

Grouping was based on the block structure of the ICD. The equivalent groups included ICD-10 group 1 (malignant neoplasms) and ICD-11 group 3 (malignant neoplasms, except

XSL•FO
RenderX

for lymphoid, hematopoietic, central nervous system, or related tissues), ICD-10 group 2 (in situ neoplasms) and ICD-11 group 4 (in situ neoplasms, except for lymphoid, hematopoietic, central nervous system, or related tissues), and ICD-10 group 3 (benign neoplasms) and ICD-11 group 5 (benign neoplasms, except for lymphoid, hematopoietic, central nervous system, or related tissues). If an ICD-10/ICD-10-CCM leaf code was not classified into the equivalent ICD-11 group, it was considered to have undergone a grouping change.

Expression Ability

For each ICD-10-CCM code, we identified the best-matching ICD-11 MMS leaf code. When all the clinical details in the diagnosis were expressed without redundant information, the code was considered fully represented. This study also defined the synonyms contained in the Foundation Component as a full representation.

Statistical Analysis

All data were analyzed with SPSS (version 25.0, IBM). The changes in coding capacity, groups, and expression ability between the ICD-10 and ICD-11 were described as rates and percentages. The chi-square test was used to determine the difference in grouping changes when mapped to the ICD-11 between the ICD-10 and ICD-10-CCM. The chi-square test was also used to analyze the difference in full expression ability among the 4 ICD-10-CCM groups, and the Bonferroni method was used for pairwise comparisons. The chi-square test was used to analyze the difference in full expression ability between the changed and unchanged groups mapped to the ICD-11. The ϕ correlation coefficient was used to analyze the correlation between expression ability and grouping changes. Statistical significance was set at $P<.05$.

Results

Coding Features

Coding Structure

Chapter 2 in both the ICD-10 and ICD-11 addresses neoplasms, and the codes in both versions are alphanumeric but have different structures. The ICD-10 codes for neoplasms consist of topography codes and morphology codes. The topography codes range from C00 to D48. Except for leaf codes, the coding range, categories, and subcategories of the ICD-10-CCM are the same as those of the ICD-10. The ICD-10 morphology codes consist of 5 digits. The first 4 digits identify the histological type of the neoplasm, and the fifth digit, following a slash (/), indicates its behavior. In the ICD-10-CCM, the morphology codes consist of 6 digits, with a fifth number before the slash.

In contrast, the coding structure of the ICD-11 MMS codes has dramatically changed. It is composed of stem codes and extension codes that are connected by an ampersand (&). The precoordinated stem codes consist of sites and essential morphology types. The newly added chapter on extension codes addresses other morphology types and greater site specificity [11], as well as stage, grading, laterality, and the diagnostic method. These codes can be used for postcoordination. The stem codes range from 2A00 to 2F9Z. The extension codes for morphology are 6-digit codes composed of letters and numbers, starting with the letter *X*. Examples of the ICD-10 and ICD-11 complete neoplasm codes are shown in Table 1.

Table 1. Examples of the International Classification of Diseases, Tenth Revision (ICD-10), Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision (ICD-10-CCM), and International Classification of Diseases, Eleventh Revision (ICD-11) complete neoplasm codes.

Neoplasm	ICD-10		ICD-10-CCM		ICD-11	
	Topography code	Morphology code	Topography code	Morphology code	Stem code	Extension codes
Adenocarcinoma of the common bile duct, stage III, diagnosis confirmed by histology	C24.0	8140/3	C24.003	81400/3	2C15.0	• &XS6H • &XY9Q

In the ICD-10, C24.0 represents a malignant neoplasm of the extrahepatic bile duct, and 8140/3 represents adenocarcinoma. In the ICD-10-CCM, C24.003 represents a malignant neoplasm of the common bile duct, and 81400/3 represents adenocarcinoma. In the ICD-11, 2C15.0 represents adenocarcinoma of the distal bile duct, XS6H represents stage III, and XY9Q represents a diagnosis confirmed by histology. We used the the ICD-11 coding tool website [10].

Coding Capacity

Categories, subcategories, and leaf codes can be used for statistics, but only leaf codes can be used for actual coding. In the ICD-10, chapter 2 includes 759 leaf codes [7]. In the ICD-10-CCM, the topography codes are basically expanded by refining the sites, with 3634 leaf codes for neoplasms. Taking malignant neoplasm of the nasal cavity (C30.0) as an example, the ICD-10-CCM contains 5 additional leaf codes, such as a malignant neoplasm of nasal cartilages (C30.001) and a malignant neoplasm of the nasal concha (C30.002).

In the ICD-11, chapter 2 includes 1037 leaf codes [8]. Compared with the ICD-10, the number of ICD-11 leaf codes was expanded by 36.6%, which is conducive to better granularity of statistics and classification.

Classification Feature

The ICD-11 has readjusted the neoplasm groups. In the ICD-10, chapter 2 was divided into 4 groups: malignant neoplasms, in situ neoplasms, benign neoplasms, and neoplasms of uncertain or unknown behavior. In the ICD-11, chapter 2 was increased to 7 groups (Table 2). The newly added ICD-11 group 1 includes all neoplasms of the brain and central nervous system, regardless of behavior, and the newly added ICD-11 group 2 includes all hematopoietic and lymphoid tissues. The ICD-10 group 4 was split into 2 separate groups: ICD-11 group 6 (neoplasms of

XSL•FO
RenderX

uncertain behavior) and ICD-11 group 7 (neoplasms of unknown behavior).

Table 2. Grouping changes between the International Classification of Diseases, Tenth Revision (ICD-10) and the International Classification of Diseases, Eleventh Revision (ICD-11).

Classification system	ICD-11								
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Other chapters	No mapping
ICD-10 codes, n									
Group 1	23[a]	91[a]	352	N/A[b]	N/A	N/A	N/A	8[a]	2
Group 2	N/A	N/A	N/A	57	N/A	N/A	N/A	N/A	N/A
Group 3	11[a]	N/A	1[a]	N/A	146	N/A	N/A	1[a]	N/A
Group 4	11[a]	18[a]	N/A	N/A	N/A	48[a]	51[a]	2[a]	2
ICD-10-CCM codes, n									
Group 1	93[a]	499[a]	1252	N/A	N/A	N/A	N/A	17[a]	3
Group 2	N/A	N/A	N/A	169	N/A	N/A	N/A	N/A	N/A
Group 3	118[a]	N/A	1[a]	N/A	796	N/A	N/A	3[a]	N/A
Group 4	85[a]	60[a]	N/A	N/A	N/A	316[a]	149[a]	18[a]	55

[a]Indicates grouping changes. The ICD-10 and ICD-10-CCM no mapping codes were not included in the statistical analysis of this study. Hence, 755 ICD-10 codes and 3576 ICD-10-CCM codes were used for percentage and chi-square analyses.

[b]N/A: not applicable.

In total, 27.4% (207/755) of ICD-10 codes underwent grouping changes. Among them, 150 codes were migrated by 1 group, 56 by 2 groups, and one by 3 groups. In the ICD-10-CCM, 38% (1359/3576) leaf codes underwent grouping changes (Table 2). A chi-square test revealed significant differences in grouping changes between the ICD-10 and ICD-10-CCM (χ^2_1=30.3; P<.001).

Expression Features

Expression Ability

Because of the classification changes, 58 of the 3634 ICD-10-CCM codes for neoplasms could not be recoded. In total, 3576 codes were included in the manual recoding study. The results of 2 separate recodings showed that 6% (213/3576) codes were mapped inconsistently. A total of 32 stem codes were inconsistent, and 181 extension codes differed. The 213 codes were all identified after consultation with the senior coder from the Collaborating Center for the WHO Family of International Classifications in China.

The final results showed that 16.6% (594/3576) codes were fully represented without postcoordination, 51.2% (1830/3576) codes were fully characterized with postcoordination, and the remaining 32.2% (1152/3576) codes were only partially described (Table 3). In addition, 7% (252/3576) codes were fully represented when using uniform resource identifiers (URIs).

Table 3. Comparison of the International Classification of Diseases, Eleventh Revision (ICD-11) expression ability among the 4 Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision (ICD-10-CCM) groups.

ICD-10-CCM group	Total codes, N	Full representation				Partial representation	
		With PC[a] (a), n	Without PC (b), n	Codes (n1=a+b), n	Codes (n1/N), %	Codes (n2), n	Codes (n2/N), %
Group 1	1861	935	406	1341	72.1	520	27.9
Group 2	169	106	44	150	88.8	19	11.2
Group 3	918	501	73	574	62.5	344	37.5
Group 4	628	288	71	359	57.2	269	42.8
Total	3576	1830	594	2424	67.8	1152	32.2

[a]PC: postcoordination.

Based on the 4 ICD-10-CCM groups, ICD-11 significantly differed in expression ability (χ^2_3=93.7; P<.001). The Bonferroni method showed that the ICD-11 had the most robust expression ability in the ICD-10-CCM group 2, followed by the ICD-10-CCM group 1, and there was no significant difference between the remaining 2 groups.

Based on whether grouping changes occurred during mapping, 73.1% (1620/2217) ICD-10-CCM codes in the unchanged group

XSL•FO
RenderX

and 59.2% (804/1359) ICD-10-CCM codes in the changed group were fully expressed, and the difference was significant (χ^2_1=74.7; *P*<.001). Expression ability had a negative correlation with grouping changes (*r*=–.144; *P*<.001).

Expression Flexibility

The expression flexibility of the ICD-11 is reflected in many aspects. For instance, in the ICD-10, the subcategory ".8" generally describes overlapping neoplasm sites, while in the ICD-11, several methods are used. Specifically, the ICD-11 uses multiple extension codes. Sometimes, only 1 extension code is used, such as XA4YW8 (overlapping sites of the esophagus). Occasionally, the ICD-11 describes this condition through stem codes, such as 2B71.0 (adenocarcinoma of the esophagogastric junction). ICD-11 classifications can also be represented through URIs, for instance, http://id.who.int/icd/entity/419755630 (Kaposi sarcoma of multiple organs).

Expression flexibility is also reflected in the additional option of postcoordination, which can meet the different requirements of most hospitals for clinical phenotype mining. For example, a patient with hepatocellular carcinoma in the left lobe of the liver that causes chronic intermittent cancer pain and tumor anemia would be coded as 2C12.02&XA5766/MG30.10&XT5G/3A71.0. Postcoordination fully expresses the clinical phenotypes and demonstrates the relationship between diseases and clinical phenotypes.

Expression of Postcoordination

According to the ICD-10-CCM/ICD-11 mapping table, 60.5% (2164/3576) of codes had postcoordination, and the average amount of postcoordination per code was approximately 0.7. Regarding the type of postcoordination, the proportion of specific anatomy (1907/3576, 53.3%) was the highest, followed by histopathology (201/3576, 5.6%), and the remaining dimensions were basically less than 2% (Table 4).

Table 4. Expression of postcoordination in the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision/International Classification of Diseases, Eleventh Revision (ICD-10-CCM/ICD-11) mapping table.

Number and dimensions of postcoordination	Codes, n (%)
0	1412 (39.5)
No	
1	
Specific anatomy	1750 (48.9)
Histopathology	176 (4.9)
Alternative severity 2	60 (1.7)
Laterality	5 (0.1)
Course of the condition	3 (0.1)
Stem codes	3 (0.1)
2	
Specific anatomy/specific anatomy	118 (3.3)
Specific anatomy/histopathology	17 (0.5)
Specific anatomy/laterality	8 (0.2)
Histopathology/alternative severity 2	8 (0.2)
Alternative severity 2/stem codes	2 (0.1)
3	
Specific anatomy/specific anatomy/specific anatomy	11 (0.3)
Specific anatomy/specific anatomy/laterality	3 (0.1)

Discussion

Principal Findings

The different coding structure of the ICD-11 provides advantages in terms of coding granularity, coding capacity, and expression flexibility. According to the mapping tables, the grouping changes between the ICD-10 and ICD-10-CCM differed (χ^2_1=30.3; *P*<.001). Meanwhile, neither the ICD-10 group 2 nor the ICD-10-CCM group 2 exhibited group migration when mapped to the ICD-11. In this study, ICD-11 expression ability (67.8%) was slightly higher than that in other studies (60%) [12,13]. Among the 4 ICD-10-CCM groups, there were significant differences in expression ability (χ^2_3=93.7; *P*<.001), with ICD-10-CCM group 2 having the highest expression ability. The expression ability negatively correlated with grouping changes (*r*=–0.144; *P*<.001).

The expression ability of the ICD-11 was still underestimated. There were many reasons why 1152 ICD-10-CCM codes did not have full representation, some of which can be avoided in actual cases. First, of the 296 ICD-10-CCM codes mapped to the ICD-11 group 1, only 26 (8.8%) codes were fully expressed,

XSL•FO
RenderX

which is far below the overall level. Among them, 217 codes cannot be fully expressed because most of the stem codes in ICD-11 group 1 do not include behavior. However, in coding actual cases with morphological types, some stem codes, including behavior, would be used instead of the residual category. Therefore, in actual coding, the expression ability of this group would be higher than that determined in this study.

Second, 714 codes could not be fully expressed due to site classification. Fortunately, 109 codes detailing specific sites may be resolved by refining the value set of extension codes. Examples include the frenulum of the upper lip, the ileocecal valve, and the rectouterine recess. However, there were also some codes for which the classification was different, such as peripheral nervous system neoplasms. This condition could be solved by coding the actual cases. In addition, some words, such as canceration of the gastric stump and cervical stump, need to be addressed.

Third, URIs can supplement the function of the ICD-11 MMS codes, especially for some diseases that cannot be identified by the ICD-11 codes. For instance, URIs (http://id.who.int/icd/entity/1595913346) make classic Kaposi sarcoma classifiable. The new coding system can also make other diseases identifiable [14-16], such as chronic pain and rare diseases. In summary, the ICD-11 has advantages in terms of actual coding and can address the expression needs of neoplasms.

The ICD-11 Vs the ICD-10

Compared to the ICD-10, the ICD-11 has undergone significant changes from design to use. First, the design purpose of the ICD-11 has changed. For more than a century, the ICD has been the basis for comparable statistics on causes of mortality and morbidity between places and over time. As a statistically friendly classification system, when the ICD-10 is used for clinical term records and diagnosis-related grouping, it cannot meet practical needs, resulting in various expanded versions of the WHO ICD-10 in multiple countries, which often leads to inconsistent statistical standards for data. In contrast, the ICD-11 is a clinically friendly classification system that meets diverse goals beyond mere health statistics, including clinical term records, patient safety and quality, reimbursement, decision support, and more. The ICD-11 is entirely digital, terminology is coded with the coding tool and application programming interface, and it has a semantic knowledge base; these features are beneficial for standardized data collection. In short, the ICD-11 has benefits in terms of obtaining statistics as well as multiaxial coding, coding granularity [4,17,18] and standardization, achieving the integration of terminology and classification. As shown in Table 1, although the ICD-10-CCM codes are expanded and refined at the leaf code level, the capture of clinical details is still weaker compared to the ICD-11 codes.

Second, the design concept of the ICD-11 is different. The ICD-10 is an independent classification system, and the standard terminology set for the ICD-10 has been developed independently. Differently, the ICD-11 has good semantic interoperability with other classification systems through harmonized methods. The Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) is considered one of the most comprehensive clinical terminologies in the world [19]. The ICD-11 Foundation Component, which includes semantic network concepts and their relationships, is organized around the Common Ontology from a subset of the SNOMED CT [20,21]. The Common Ontology has been harmonized with ICD text definitions, primarily from the SNOMED CT clinical findings hierarchy (findings, disorders, and disease) and secondarily from other hierarchies (situations, events, social context, and so on) [22]. The rich Foundation Component has approximately 80,000 entries and 40,000 synonyms [5]. The ICD-11 linearizations, including the ICD-11 MMS and the International Classification of Diseases for Oncology (ICD-O), are subsets derived from the Foundation Component. The ICD-11 integrates the morphology section of the ICD-O, ICD-O linearization, and tumor node metastasis classification, and the histopathology codes of the ICD-11 are also compatible with the ICD-O [4]. Additionally, the ICD-11 integrates numerous clinical terminologies from some expanded versions of the WHO ICD-10. Compared to other classifications, the ICD-11 has advantages in terms of concept coverage and compatibility.

Third, the ICD-11 has stronger logical links between codes. If multiple codes are needed for disease expression, there is a lack of practical connections between ICD-10 codes. In the ICD-11, precoordination contains the site and morphology, and the postcoordination of morphology and the clinical phenotype is linked to stem codes through an ampersand (&) and slash (/).

In addition, the ICD-11 contains rich dimensions of postcoordination. Compared to the ICD-10, which provides information only about topography and morphology, the ICD-11 can also include other dimensions of postcoordination, such as stage, grading, laterality, and the diagnostic method. The ICD-11 coding system is more conducive to unifying national cancer registration in these dimensions. However, due to the limitations of the ICD-10 coding structure, these dimensions of postcoordination were missing when mapped to the ICD-11, with only 0.7 postcoordination per code, and its expression ability cannot be truly reflected.

Moreover, the ICD-11 has more expression flexibility. The ICD-10 coding system is fixed and single. The ICD-11 provides 2 sets of codes: ICD-11 MMS codes and URIs. In ICD-11 MMS codes, a stem code can be used alone or with optional extension codes. A URI is a string of characters that uniquely identifies a particular entity. Therefore, a coding system with good expression flexibility can meet the coding granularity and clinical phenotype mining requirements of hospitals at different levels.

Finally, the ICD-11 has different update mechanisms. The ICD-10 updates relatively slowly. In 1999, the WHO established the Update Advisory Committee as the only authoritative body revising the ICD-10. In 2005, the WHO officially published the second edition of the ICD-10 and continued to issue revisions to the relevant content of the ICD-10 through official channels every year. The ICD-11 provides a web-based coding tool instead of paper environments, which is highly beneficial for timely updates. It has a real-time updated orange version and an annually updated blue version.

XSL•FO
RenderX

Overall, the eleventh revision is more extensive than any other revision since the sixth in 1948. These design-level features make the ICD-11 competitive with other classification systems.

Challenges of Replacement

The improvements described above also pose challenges for replacing the ICD-10 with the ICD-11. Compared to the ICD-10, the ICD-11 leaf codes for neoplasms have expanded by 36.6%, and there are also approximately 16,000 extension codes. Hence, the expression ability of the ICD-11 for clinical details far exceeds that of the ICD-10 and ICD-10-CCM. However, codes with better granularity require clinicians to record diagnoses and treatments in greater detail. It is essential to carry out the necessary interventions to enhance medical record documentation according to ICD-11 before or simultaneously with country-wide implementation [23]. Clinicians will be required to have a good understanding of the ICD-11. Otherwise, the best classification system is just a decoration.

Second, coders are accustomed to using the old classification system and need time to familiarize themselves with the new coding system and tool. Studies have shown that some coders have difficulties coding on the web due to a lack of familiarity with software [24,25]. In this study, 2 skilled coders who received ICD-11 training still had 213 inconsistent codes. Web-based coding requires high professional ability. Studies have also shown that the reliability and accuracy of ICD-11 coding are lower than those of ICD-10 coding [23-25]. In the early stages of ICD-11 promotion, countries involved in the transition need to conduct ICD-11 training and transform coders' coding concepts. To maintain coding accuracy, developing high-quality ICD-11 training materials, training processes, and local ICD-11 guidelines is effective [23,26].

Moreover, many countries have applied the ICD-10 in different fields, such as mortality and morbidity statistics, diagnosis-related grouping, and cancer registries. There is also a potential challenge with IT systems being able to adopt a new classification that allows for unlimited width in fields. Stakeholders must upgrade their information systems to meet the needs of the ICD-11, which requires a significant amount of time and money to hire medical, IT, and management personnel to complete this transformation. The experience of the United States in replacing ICD versions can provide a better reference. In the United States, because of significant opposition and reservations expressed by stakeholders, it took 6 years from the adoption of the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) to complete the transition [27], and the costs associated with the transition were estimated to be between US $475 million and US $1.5 billion dollars, including training, productivity losses, and system changes [28].

Importantly, the ICD-10 and ICD-11 will coexist for a long time before replacement, which contributes to solving some problems. The parallel implementation of the 2 classification systems can ensure government policy continuity. In January 2022, the pilot application program to promote ICD-11 was launched in China, with 59 large general public hospitals from all provinces participating. However, ICD-11 is not fully understood by other hospitals. The impacts of the new classification system implementation included coding accuracy, ICD version mapping, and more [29]. The ICD version mapping can be used for interoperability between coded data sets [30]. Hospitals do not require dual coding and use ICD-11 to report data. If a hospital needs dual coding, a mapping table can significantly reduce the time, at least by half, spent on duplicate coding. A map can also maintain the same accuracy of ICD-11 coding for all hospitals and alert coders if there are any coding errors. Usually, a high-quality map requires substantial manual curation, and some studies use algorithmic mapping approaches, such as sequential mapping [30]. Some scholars have developed hybrid methods, such as automatic mapping and manual review [31], semiautomatic mapping, and manual evaluation [32]. Although automatic coding has many benefits, supervised and manual mapping are still necessary [33]. Undeniably, there may be some issues with the use of mapping tables. A study has shown that 14.5% of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes used by internists, when mapped to the ICD-10-CM, resulted in potential clinical inaccuracies [34]. In general, a high-quality map would contribute to the smooth transition of ICD-11, and multiple mapping methods can be explored. This study confirms the feasibility of using manual mapping tables for neoplasms between the ICD-10-CCM and ICD-11. To date, the results of some studies on the transition to the ICD-11 have also been optimistic [17,18].

In addition, although the WHO provides an ICD-11 coding tool, countries must continue optimizing this tool. A mapping table can improve the intelligence of the coding tool. If the terms are the same as those in the mapping table, the coding tool can automatically load the coding cluster without selecting postcoordination one by one. The accuracy and efficiency of ICD-11 coding largely depend on how well the coding tool is optimized. It is necessary to continuously optimize the coding tool before fully promoting the ICD-11.

Limitations

This study had some limitations that must be considered. First, this study focused only on the clinical condition of neoplasms. However, the chapter on neoplasms is quite complex and one of the most varied chapters on other conditions, with significant changes. Moreover, patients with cancer may experience different manifestations, including chronic, surgical, and emergency conditions. For example, a patient with long-term chronic hepatitis B causing cirrhosis and liver cancer was admitted to the hospital for emergency surgery due to a tumor rupture. Furthermore, the burden of cancer in China continues to grow with the aging population. According to statistics, the crude cancer mortality rate increased from 108.3 per 100,000 individuals in the 1990-1992 period to 170.1 per 100,000 individuals in 2015 [35]. Significantly, the research method is well thought out. Manual mapping can provide a good understanding of the new ICD-11 features and help individuals familiarize themselves with the characteristics of the new classification systems. As a method, manual mapping can be extended to other conditions. More conditions can be gradually incorporated, especially emergency and surgical conditions, in the future.

XSL•FO
RenderX

Second, this study used only manual mapping, which is time-consuming and laborious. Other methods combined with manual mapping can be continuously explored, especially automatic mapping between the ICD-10-CCM and ICD-11, including algorithmic mapping and machine learning approaches.

Conclusion

Neoplasm classification has undergone many improvements in the ICD-11, especially the new coding system, improved expression ability, and good semantic interoperability. The new coding system provides advantages in coding granularity, coding capacity, and expression flexibility. Moreover, 67.8% of ICD-10-CCM neoplasm codes can be fully represented by the ICD-11, and expression ability negatively correlates with grouping changes (r=–0.144, P<.001). The more significant the changes in a new classification system are, the less information can be expressed when mapped to other databases. The use of URIs and maintenance mechanisms can increase the expression ability for coding in actual cases. In addition, the good semantic interoperability of the ICD-11, integrating numerous clinical terminologies from the SNOMED CT, the ICD-O, the tumor node metastasis classification, and expanded versions of the WHO ICD-10, makes the ICD-11 competitive with other classification systems. The transition to the ICD-11 will inevitably bring numerous challenges for clinicians, coders, policy makers and IT technicians, and many preparations will be necessary.

Acknowledgments

The authors would like to thank Dr Xiaoding Liu from the Department of Pathology at Peking Union Medical College Hospital for providing information about pathology and answering our questions. This work was supported by the Fundamental Research Funds for the Central Universities (grant 2022FZZX06-09), and the Zhejiang Kangenbei Hospital-managed soft science research project (grant 2023ZHA-KEB305).

Conflicts of Interest

None declared.

References

1. ICD history. World Health Organization. 2022. URL: https://www.who.int/standards/classifications/classification-of-diseases [accessed 2022-11-16]
2. Peking Union Medical College Hospital, World Health Organization, International Classification Family Cooperation Center. The International Statistical Classification of Diseases and Related Health Problems tenth Revision (ICD-10)—volume 1-Tabular List. Beijing, China: People's Medical Publishing House; 2010.
3. Chute CG, Çelik C. Overview of ICD-11 architecture and structure. BMC Med Inform Decis Mak 2022;21(Suppl 6):378 [FREE Full text] [doi: 10.1186/s12911-021-01539-1] [Medline: 35578335]
4. Xu Y, Zhou J, Wang Y. The novel and powerful ICD-11 classification system for neoplasm coding: a comparative study with the ICD-O. BMC Med Inform Decis Mak 2022;22(1):333 [FREE Full text] [doi: 10.1186/s12911-022-02077-0] [Medline: 36527027]
5. Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. BMC Med Inform Decis Mak 2021;21(Suppl 6):206 [FREE Full text] [doi: 10.1186/s12911-021-01534-6] [Medline: 34753471]
6. National Health Commission of the People's Republic of China. China Health Statistical Yearbook. Beijing, China: NHC; 2023.
7. ICD-10/ICD-11 mapping tables. World Health Organization. 2022. URL: https://icd.who.int/browse11/l-m/en [accessed 2022-11-30]
8. Simple tabulation. World Health Organization. 2022. URL: https://icd.who.int/browse11/l-m/en [accessed 2022-11-30]
9. The second revision of the Chinese clinical modification of ICD-10 (ICD-10-CCM). National Health Commission of the People's Republic of China. 2019. URL: http://www.nhc.gov.cn/yzygj/s3593g/201904/b8323261bb8a4175a2046d2fffa93936.shtml [accessed 2022-04-15]
10. ICD-11 coding tool (version: 02/2022). World Health Organization. 2022. URL: https://icd.who.int/ct11/icd11_mms/zh/2022-02 [accessed 2024-01-13]
11. ICD-11 reference guide. World Health Organization. 2022. URL: https://icdcdn.who.int/icd11referenceguide/en/html/index.html [accessed 2022-06-26]
12. Fung KW, Xu J, Bodenreider O. The new international classification of diseases 11th edition: a comparative analysis with ICD-10 and ICD-10-CM. J Am Med Inform Assoc 2020;27(5):738-746 [FREE Full text] [doi: 10.1093/jamia/ocaa030] [Medline: 32364236]
13. Fung KW, Xu J, McConnell-Lamptey S, Pickett D, Bodenreider O. Feasibility of replacing the ICD-10-CM with the ICD-11 for morbidity coding: a content analysis. J Am Med Inform Assoc 2021;28(11):2404-2411 [FREE Full text] [doi: 10.1093/jamia/ocab156] [Medline: 34383897]

XSL•FO
RenderX

14. 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 [FREE Full text] [doi: 10.1186/s13023-015-0251-8] [Medline: 25887186]
15. Chute CG. The rendering of human phenotype and rare diseases in ICD-11. J Inherit Metab Dis 2018;41(3):563-569 [FREE Full text] [doi: 10.1007/s10545-018-0172-5] [Medline: 29600497]
16. Zinboonyahgoon N, Luansritisakul C, Eiamtanasate S, Duangburong S, Sanansilp V, Korwisi B, et al. Comparing the ICD-11 chronic pain classification with ICD-10: how can the new coding system make chronic pain visible? A study in a tertiary care pain clinic setting. Pain 2021;162(7):1995-2001. [doi: 10.1097/j.pain.0000000000002196] [Medline: 33449507]
17. Mabon K, Steinum O, Chute CG. Postcoordination of codes in ICD-11. BMC Med Inform Decis Mak 2022;21(Suppl 6):379 [FREE Full text] [doi: 10.1186/s12911-022-01876-9] [Medline: 35581649]
18. Drösler SE, Weber S, Chute CG. ICD-11 extension codes support detailed clinical abstraction and comprehensive classification. BMC Med Inform Decis Mak 2021;21(Suppl 6):278 [FREE Full text] [doi: 10.1186/s12911-021-01635-2] [Medline: 34753461]
19. Tavakoli K, Kalaw FGP, Bhanvadia S, Hogarth M, Baxter SL. Concept coverage analysis of ophthalmic infections and trauma among the standardized medical terminologies SNOMED-CT, ICD-10-CM, and ICD-11. Ophthalmol Sci 2023;3(4):100337 [FREE Full text] [doi: 10.1016/j.xops.2023.100337] [Medline: 37449050]
20. Rodrigues JM, Schulz S, Rector A, Spackman K, Millar J, Campbell J, et al. ICD-11 and SNOMED CT common ontology: circulatory system. Stud Health Technol Inform 2014;205:1043-1047. [Medline: 25160347]
21. Rodrigues JM, Schulz S, Rector A, Spackman K, Üstün B, Chute CG, et al. Sharing ontology between ICD 11 and SNOMED CT will enable seamless re-use and semantic interoperability. Stud Health Technol Inform 2013;192:343-346. [Medline: 23920573]
22. Rodrigues JM, Robinson D, Della Mea V, Campbell J, Rector A, Schulz S, et al. Semantic alignment between ICD-11 and SNOMED CT. Stud Health Technol Inform 2015;216:790-794. [Medline: 26262160]
23. Zarei J, Golpira R, Hashemi N, Azadmanjir Z, Meidani Z, Vahedi A, et al. Comparison of the accuracy of inpatient morbidity coding with ICD-11 and ICD-10. Health Inf Manag 2023;0(0):18333583231185355. [doi: 10.1177/18333583231185355] [Medline: 37491819]
24. Lee H, Kim S. Impact of the ICD-11 on the accuracy of clinical coding in Korea. Health Inf Manag 2023;52(3):221-228. [doi: 10.1177/18333583221095147] [Medline: 35614871]
25. Eisele A, Dereskewitz C, Oberhauser C, Kus S, Coenen M. Reliability, usability and applicability of the ICD-11 beta draft focusing on hand injuries and diseases: results from German field testing. Int J Qual Health Care 2019;31(10):G174-G179 [FREE Full text] [doi: 10.1093/intqhc/mzz121] [Medline: 31838492]
26. Eastwood CA, Southern DA, Doktorchik C, Khair S, Cullen D, Boxill A, et al. Training and experience of coding with the World Health Organization's international classification of diseases, eleventh revision. Health Inf Manag 2023;52(2):92-100 [FREE Full text] [doi: 10.1177/18333583211038633] [Medline: 34555947]
27. Monestime JP, Mayer RW, Blackwood A. Analyzing the ICD-10-CM transition and post-implementation stages: a public health institution case study. Perspect Health Inf Manag 2019;16(Spring):1a [FREE Full text] [Medline: 31019430]
28. Libicki M, Brahmakulam I. The costs and benefits of moving to the ICD-10 code sets. RAND Corporation. 2004. URL: https://www.rand.org/pubs/technical_reports/TR132.html [accessed 2024-01-13]
29. Kusnoor SV, Blasingame MN, Williams AM, DesAutels SJ, Su J, Giuse NB. A narrative review of the impact of the transition to ICD-10 and ICD-10-CM/PCS. JAMIA Open 2020;3(1):126-131 [FREE Full text] [doi: 10.1093/jamiaopen/ooz066] [Medline: 32607494]
30. Xu J, Fung KW, Bodenreider O. Sequential mapping—a novel approach to map from ICD-10-CM to ICD-11. Stud Health Technol Inform 2022;290:96-100 [FREE Full text] [doi: 10.3233/SHTI220039] [Medline: 35672978]
31. Hong N, Zhang Y, Ren Y, Hou L, Wang C, Li J, et al. Cross-language terminology mapping between ICD-10-CN and SNOMED-CT. Stud Health Technol Inform 2022;290:42-46. [doi: 10.3233/SHTI220028] [Medline: 35672967]
32. Ohlsen T, Kruse V, Krupar R, Banach A, Ingenerf J, Drenkhahn C. Mapping of ICD-O tuples to OncoTree codes using SNOMED CT post-coordination. Stud Health Technol Inform 2022;294:307-311. [doi: 10.3233/SHTI220464] [Medline: 35612082]
33. Grosjean J, Billey K, Charlet J, Darmoni SJ. Manual evaluation of the automatic mapping of International Classification of Diseases (ICD)-11 (in French). Stud Health Technol Inform 2020;270:1335-1336. [doi: 10.3233/SHTI200429] [Medline: 32570646]
34. Caskey RN, Abutahoun A, Polick A, Barnes M, Srivastava P, Boyd AD. Transition to international classification of disease version 10, clinical modification: the impact on internal medicine and internal medicine subspecialties. BMC Health Serv Res 2018;18(1):328 [FREE Full text] [doi: 10.1186/s12913-018-3110-1] [Medline: 29728145]
35. Sun D, Li H, Cao M, He S, Lei L, Peng J, et al. Cancer burden in China: trends, risk factors and prevention. Cancer Biol Med 2020;17(4):879-895 [FREE Full text] [doi: 10.20892/j.issn.2095-3941.2020.0387] [Medline: 33299641]

XSL•FO
RenderX

Abbreviations

ICD-10: International Classification of Diseases, Tenth Revision
ICD-10-CCM: Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision
ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification
ICD-11 MMS: International Classification of Diseases, Eleventh Revision, Mortality and Morbidity Statistics
ICD-11: International Classification of Diseases, Eleventh Revision
ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification
ICD-O: International Classification of Diseases for Oncology
SNOMED CT: Systematized Nomenclature of Medicine-Clinical Terms
URI: uniform resource identifier
WHO: World Health Organization

Edited by M Sheth; submitted 29.08.23; peer-reviewed by T Karen, D Gartner; comments to author 28.11.23; revised version received 13.01.24; accepted 31.01.24; published 08.03.24.

Please cite as:
Xu Y, Zhou J, Li H, Cai D, Zhu H, Pan S
Improvements in Neoplasm Classification in the International Classification of Diseases, Eleventh Revision: Systematic Comparative Study With the Chinese Clinical Modification of the International Classification of Diseases, Tenth Revision
Interact J Med Res 2024;13:e52296
URL: https://www.i-jmr.org/2024/1/e52296
doi:10.2196/52296
PMID:38457228

©Yicong Xu, Jingya Zhou, Hongxia Li, Dong Cai, Huanbing Zhu, Shengdong Pan. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 08.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original

Application of AI in Sepsis: Citation Network Analysis and Evidence Synthesis

MeiJung Wu[1,2], MSc; Md Mohaimenul Islam[3], PhD; Tahmina Nasrin Poly[1], MSc; Ming-Chin Lin[1,4,5], PhD

[1]Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan

[2]Department of Nursing, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan

[3]Department of Outcomes and Translational Sciences, College of Pharmacy, The Ohio State University, Columbus, OH, United States

[4]Taipei Neuroscience Institute, Taipei Medical University, Taipei, Taiwan

[5]Department of Neurosurgery, Shuang Ho Hospital, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Ming-Chin Lin, PhD
Graduate Institute of Biomedical Informatics
College of Medical Science and Technology
Taipei Medical University
250 Wuxing St
Xinyi District
Taipei, 110
Taiwan
Phone: 886 66202589
Email: arbiter@tmu.edu.tw

Abstract

Background: Artificial intelligence (AI) has garnered considerable attention in the context of sepsis research, particularly in personalized diagnosis and treatment. Conducting a bibliometric analysis of existing publications can offer a broad overview of the field and identify current research trends and future research directions.

Objective: The objective of this study is to leverage bibliometric data to provide a comprehensive overview of the application of AI in sepsis.

Methods: We conducted a search in the Web of Science Core Collection database to identify relevant articles published in English until August 31, 2023. A predefined search strategy was used, evaluating titles, abstracts, and full texts as needed. We used the Bibliometrix and VOSviewer tools to visualize networks showcasing the co-occurrence of authors, research institutions, countries, citations, and keywords.

Results: A total of 259 relevant articles published between 2014 and 2023 (until August) were identified. Over the past decade, the annual publication count has consistently risen. Leading journals in this domain include *Critical Care Medicine* (17/259, 6.6%), *Frontiers in Medicine* (17/259, 6.6%), and *Scientific Reports* (11/259, 4.2%). The United States (103/259, 39.8%), China (83/259, 32%), United Kingdom (14/259, 5.4%), and Taiwan (12/259, 4.6%) emerged as the most prolific countries in terms of publications. Notable institutions in this field include the University of California System, Emory University, and Harvard University. The key researchers working in this area include Ritankar Das, Chris Barton, and Rishikesan Kamaleswaran. Although the initial period witnessed a relatively low number of articles focused on AI applications for sepsis, there has been a significant surge in research within this area in recent years (2014-2023).

Conclusions: This comprehensive analysis provides valuable insights into AI-related research conducted in the field of sepsis, aiding health care policy makers and researchers in understanding the potential of AI and formulating effective research plans. Such analysis serves as a valuable resource for determining the advantages, sustainability, scope, and potential impact of AI models in sepsis.

(Interact J Med Res 2024;13:e54490) doi:10.2196/54490

KEYWORDS

AI; artificial intelligence; bibliometric analysis; bibliometric; citation; deep learning; machine learning; network analysis; publication; sepsis; trend; visualization; VOSviewer; Web of Science; WoS

XSL•FO
RenderX

Introduction

Sepsis is a life-threatening medical emergency [1] affecting approximately 48.9 million individuals globally each year and potentially contributing to over 11 million deaths [2]. Previous studies indicated that sepsis-related hospitalization can result in fatal outcomes in 30%-50% of cases [3,4]. However, prompt stratification and the timely administration of specific treatments have the potential to lower sepsis-related mortality. Identifying sepsis at an early stage can be challenging due to the complex pattern of the disease [5,6] and the diversity of the septic population [7].

Artificial intelligence (AI) has piqued interest in its excellent potential to stratify patients with a high risk of sepsis [8]. In recent times, AI models have seen widespread application in the prediction of sepsis and have shown superior performance compared with conventional statistical methods [9,10]. Yet, no study has shed light on the variety of AI applications and their potential and limitations in sepsis through a scientific consolidation of knowledge. Bibliometric analysis aids researchers in comprehending specific research fields, a crucial aspect for guiding both future research endeavors (eg, what else should we know) and practical implementation (eg, what should we do) [11]. This research aims to address the following questions, with the intent of advancing the previous research on the application of AI in sepsis: (1) What countries, institutions, sources, and documents have demonstrated the highest productivity within the realm of AI applied to sepsis? (2) What are the hot research topics and themes of research in the application of AI in sepsis? (3) What methods are mainly applied in the existing body of literature? (4) What types of limitations appeared in the existing literature regarding the application of AI in sepsis? and (5) What are the literature gaps and future research agendas?

In this study, we could systematically investigate shifts in publication growth, offering more valuable insights to fellow researchers and policy makers engaged in priority setting and assessment.

Methods

Data Source

We leveraged extracted data from the Web of Science Core Collection as of August 31, 2023. We used Web of Science for its comprehensive coverage across multiple databases, comprising a wide range of bibliometric indicators and literature from various disciplines. Using a predefined search strategy, we intended to include all relevant literature for bibliometric analysis. We used the following key words: *artificial intelligence* OR *computational intelligence* OR *deep learning* OR *computer aided* OR *machine learning* OR *support vector machine* OR *data learning* OR *artificial neural network* OR *digital image* OR *convolutional neural network* OR *evolutionary algorithms* OR *feature learning* OR *reinforcement learning* OR *big data* OR *image segmentation* OR *hybrid intelligent system* OR *hybrid intelligent system* OR *recurrent neural network* OR *natural language processing* OR *Bayesian network* OR *Bayesian learning* OR *random forest* OR *evolutionary algorithms* OR *multiagent system* AND *sepsis.* The collected records contained essential attributes, including publication date, authorship, institutional affiliation, geographic origin, and cited references. This data set served as the foundation for our subsequent analytical investigations.

Inclusion and Exclusion Criteria

The titles and abstracts underwent initial screening by 2 independent authors (MW and TNP). If there was uncertainty from one reviewer regarding whether the article met the inclusion criteria, it was included for a thorough full-text review. Following this, both authors independently assessed the full text, and any differences in opinion were resolved through consensus with the research team. We considered studies for inclusion if they met the following criteria: (1) they were written in English, and (2) they applied AI models in the context of sepsis. In this study's screening process, we included research or review articles published in peer-reviewed journals, conference proceedings, reviews, and early access articles. We excluded studies if they were published as letters, editorials, book chapters, or books.

Data Collection and Preprocessing

To ensure compatibility with Bibliometrix and VOSviewer [12], we saved the data in the "**.txt" format, a format recognized by both tools for conducting analyses. Our data set encompasses a comprehensive range of information, including titles, list of authors, name of countries, list of institutions, abstracts, keywords, name of journals, and publication dates.

Statistical Analysis

Bibliometrix and VOSviewer tools were used to uncover the knowledge structure, most influential countries, research hot spots, and productive authors, along with various bibliometric insights. The processed data were uploaded into these bibliometric tools, and analysis was conducted based on the information included within the data documents [13]. Afterward, we generated network maps among journals, authors, countries, and institutions, where individual points symbolized authors, countries, or institutions. Moreover, connected lines in the network maps depicted the relationships between these entities. Larger points and more robust lines indicated a higher number of articles and more substantial collaborative relationships, respectively [14,15].

We computed the annual growth rate of publications. The annual publication count, annual growth, and average growth rate of publications were determined through the following methods:

Where N is the total number of articles in the current year, and N_{k-1} is the total number of articles in the previous year.

Furthermore, we conducted an analysis of publication trends based on the following criteria: the top 10 most prolific countries, institutions, journals, authors, and studies in this area. The rankings of countries, institutions, journals, and authors were determined based on the number of articles.

XSL•FO
RenderX

Results

Distribution of Articles by Publication Year

The initial search yielded 327 articles focused on the application of AI in sepsis. After applying predefined inclusion criteria, 68 articles were deleted, leaving 259 articles for the final analysis (Figure 1).

Over time, there has been a substantial rise in the number of publications in this field. Notably, the yearly publication number increased from just 2 articles in 2014 to 72 articles in 2022. It is important to note that before 2018, the yearly publication count did not cross 10 articles. The calculated annual growth rate was found to be 44.81% (Figure 2).

Figure 1. A diagram for the detailed selection criteria and bibliometric analysis steps of applying artificial intelligence to sepsis in the Web of Science Core Collection database.

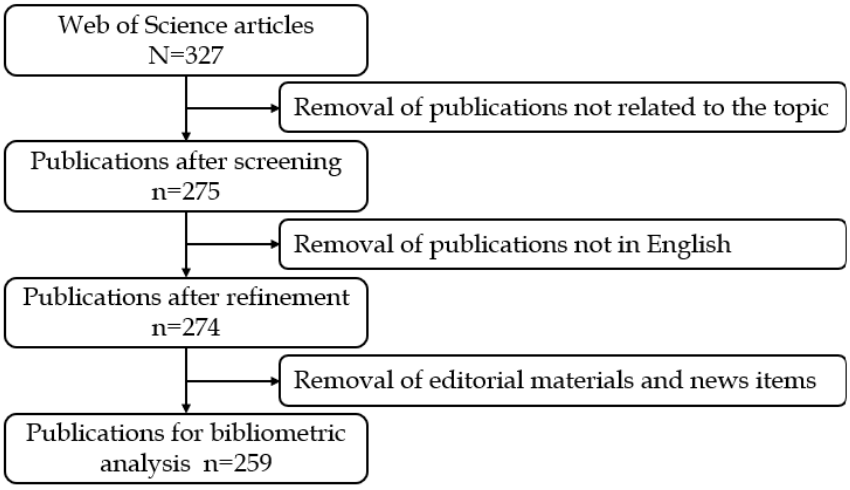
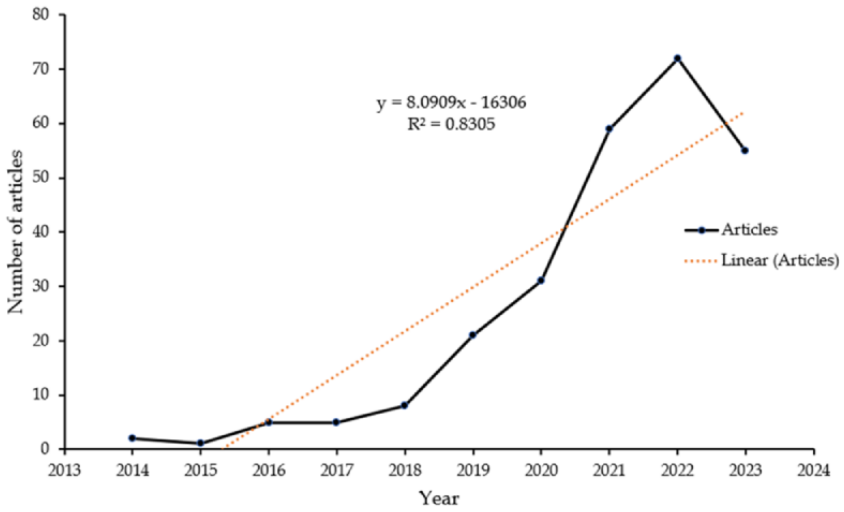


Figure 2. Trends in the number of publications on the application of artificial intelligence to the study of sepsis from 2014 to 2023 (August).



Distribution of Source Journals

A total of 122 journals published articles on the application of AI in sepsis. Among them, the *Critical Care Medicine* journal was the most productive, having published 6.6% (17/122) of articles in this domain (Table 1). *Frontiers in Medicine*, *Scientific Reports*, and the *American Journal of Respiratory and Critical Care Medicine* were in the second, third, and fourth positions, publishing 17, 11, and 7 articles, respectively, on this topic. However, the top 10 journals published 86 articles, accounting for 33.2% (86/259) of all publications in this area.

XSL•FO
RenderX

Table 1. The top 10 journals with publications on the application of artificial intelligence in sepsis from 2014 to August 2023.

Rank	Journal	Country	Category	Publication frequency, n (%)	Impact factor in 2022	5-year impact factor
1	*Critical Care Medicine*	United States	Engineering, electrical, and electronics	17 (6.6)	8.8	8.4
2	*Frontiers in Medicine*	Switzerland	Multidisciplinary science	17 (6.5)	3.9	4.2
3	*Scientific Reports*	United Kingdom	Multidisciplinary science	11 (4.2)	4.6	4.9
4	*American Journal of Respiratory and Critical Care Medicine*	United States	Multidisciplinary science	7 (2.7)	24.7	21.9
5	*Frontiers in Immunology*	Switzerland	Clinical neurology	6 (2.3)	7.3	8.0
6	*Intensive Care Medicine*	United States	Neurosciences	6 (2.3)	38.9	27
7	*Journal of the American Medical Informatics Association*	United States	Computer science and interdisciplinary applications	6 (2.3)	6.4	6.3
8	*PLoS One*	United States	Neurosciences	6 (2.3)	3.7	3.8
9	*BMC Medical Informatics and Decision Making*	United Kingdom	Engineering and multidisciplinary	5 (1.9)	3.5	3.9
10	*Computers in Biology and Medicine*	United States	Engineering and biomedical	5 (1.9)	7.7	6.9

Distribution of Countries and Regions

This study revealed that researchers from 73 countries and regions engaged in research on these subjects and published their work in various international peer-reviewed journals. Out of the total 259 articles, the United States made the most substantial contribution with 103 publications (39.8%), followed by China with 83 publications (32%), United Kingdom with 14 publications (5.4%), and Taiwan with 12 publications (4.6%) (Table 2).

Table 2. The top 10 countries and regions with publications on the application of artificial intelligence in sepsis from 2014 to August 2023.

Rank	Country	Articles, n (%)
1	United States	103 (39.8)
2	China	83 (32)
3	United Kingdom	14 (5.4)
4	Taiwan	12 (4.6)
5	India	11 (4.2)
6	Netherlands	10 (3.9)
7	Australia	8 (3.1)
8	Canada	8 (3.1)
9	Spain	7 (2.7)
10	Germany	7 (2.7)

Distribution of Institutions

Table 3 shows the top 10 most productive institutes that used AI applications in sepsis. The University of California system (22/259 articles, 8.5%) ranked first among all research institutions, followed by Emory University (10/259 articles, 3.9%), Harvard University (10/259 articles, 3.9%), and Central South University (8/259 articles, 3.1%).

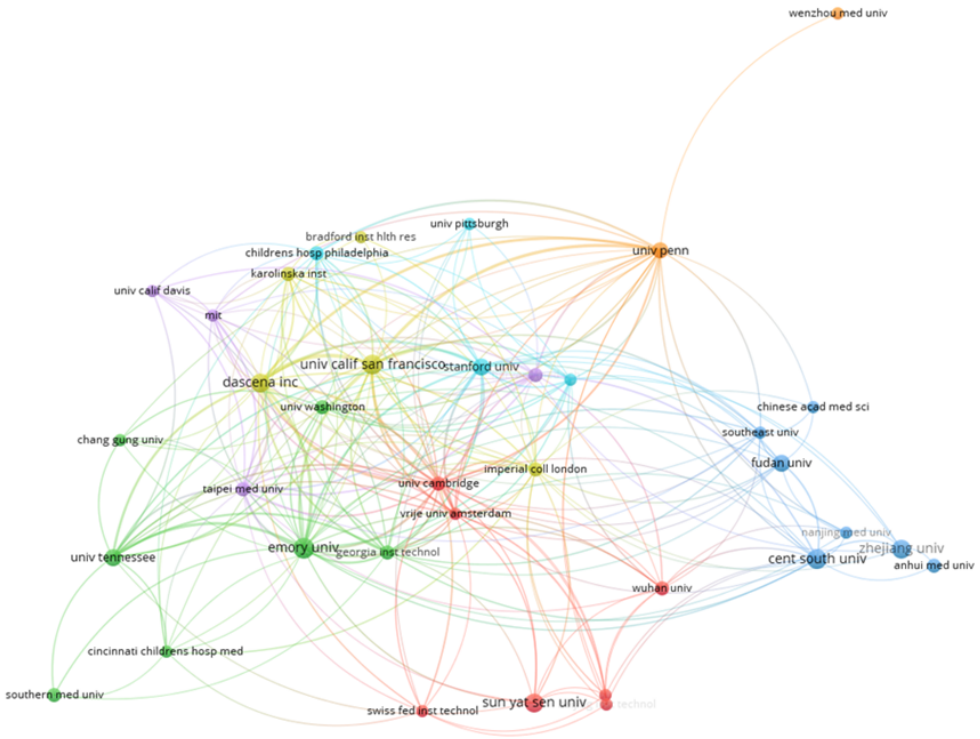
Figure 3 shows the institution cooperation network of 117 institutions that published at least 1 article.

XSL•FO
RenderX

Table 3. The top 10 institutes with publications on the application of artificial intelligence in sepsis from 2014 to August 2023.

Rank	Institutions	Country	Publications, n (%)
1	University of California system	United States	22 (8.5)
2	Emory University	United States	10 (3.9)
3	Harvard University	United States	10 (3.9)
4	Central South University	China	8 (3.1)
5	Dascena Inc	United States	8 (3.1)
6	University of Pennsylvania	United States	8 (3.1)
7	Zhejiang University	China	8 (3.1)
8	Stanford University	United States	7 (2.7)
9	Sun Yat-sen University	China	7 (2.7)
10	Fudan University	China	6 (2.3)

Figure 3. Institution co-operation network.



Distribution of Authors

A total of 259 articles were authored by 1444 individuals, each of whom had at least 1 article to their name. In Table 4, we present the top 10 most prolific authors who conducted and published research in these fields. Ritankar Das secured the top position with 8 articles, closely followed by Chris Barton (6 articles), Rishikesan Kamaleswaran (6 articles), and Suchi Saria (6 articles).

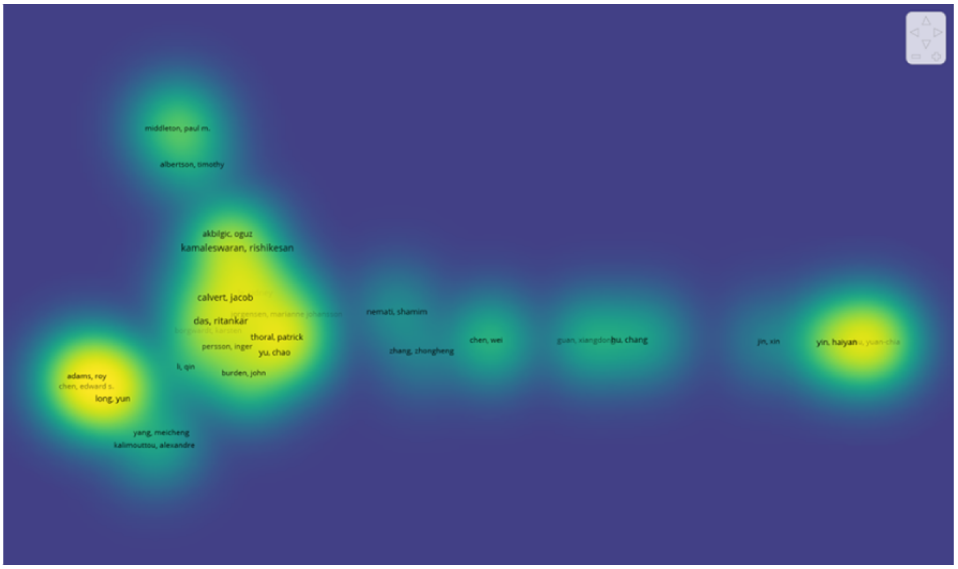
Our analysis shows that 1444 authors have published at least 1 article. The largest set of associated authors consisted of 20 authors in 3 clusters (Figure 4).

XSL•FO
RenderX

Table 4. The top 10 authors with publications on the application of artificial intelligence in sepsis from 2014 to August 2023.

Rank	Author	Articles, n	Citations, n	h-index	Affiliation
1	Das	8	1417	18	Dascena Inc
2	Barton	6	2277	25	University of California San Francisco
3	Kamaleswaran	6	861	14	Emory University
4	Saria	6	2635	23	Johns Hopkins University
5	Calvert	5	1152	17	University of California Berkeley
6	Hoffman	5	957	14	Dascena Inc
7	Li	5	2	1	Sun Yat-sen University
8	Nemati	5	2402	23	University of California San Diego
9	Adams	4	118	6	Johns Hopkins University
10	Davis	4	1253	17	University of Tennessee Health Science Center

Figure 4. The co-authorship network of authors who contributed research on the application of artificial intelligence to sepsis from 2014 to 2023 (August).



Articles Cocitation Analysis

Table 5 shows the top 10 most frequently cited publications. The publication that received the most citations was by Komorowski et al [16], titled "The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care," published in *Nature Medicine* in 2018 and received a total of 408 citations as of August 31, 2023.

XSL•FO
RenderX

Table 5. Top 10 cited articles in the application of artificial intelligence on sepsis research from 2014 to August 2023.

Rank	Author	Journal	Title	Citation, n
1	Komorowski et al [16]	*Nature Medicine*	The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care	408
2	Nemati et al 2018 [17]	*Critical Care Medicine*	An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU	329
3	Taylor et al 2016 [18]	*Academic Emergency Medicine*	Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach	257
4	Desautels et al 2016 [19]	*JMIR Medical Informatics*	Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach	226
5	Fleuren et al [10]	*Intensive Care Medicine*	Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy	184
6	Horng et al 2017 [20]	*PlosOne*	Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning	148
7	Gultepe et al 2014 [21]	*Journal of the American Medical Informatics Association*	From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system	101
8	Giannini et al 2019 [22]	*Critical Care Medicine*	A Machine Learning Algorithm to Predict Severe Sepsis and Septic Shock: Development, Implementation, and Impact on Clinical Practice	100
9	Hou et al 2020 [23]	*Journal of Translational Medicine*	Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost	98
10	Mani et al 2014 [24]	*Journal of the American Medical Informatics Association*	Medical decision support using machine learning for early detection of late-onset neonatal sepsis	97

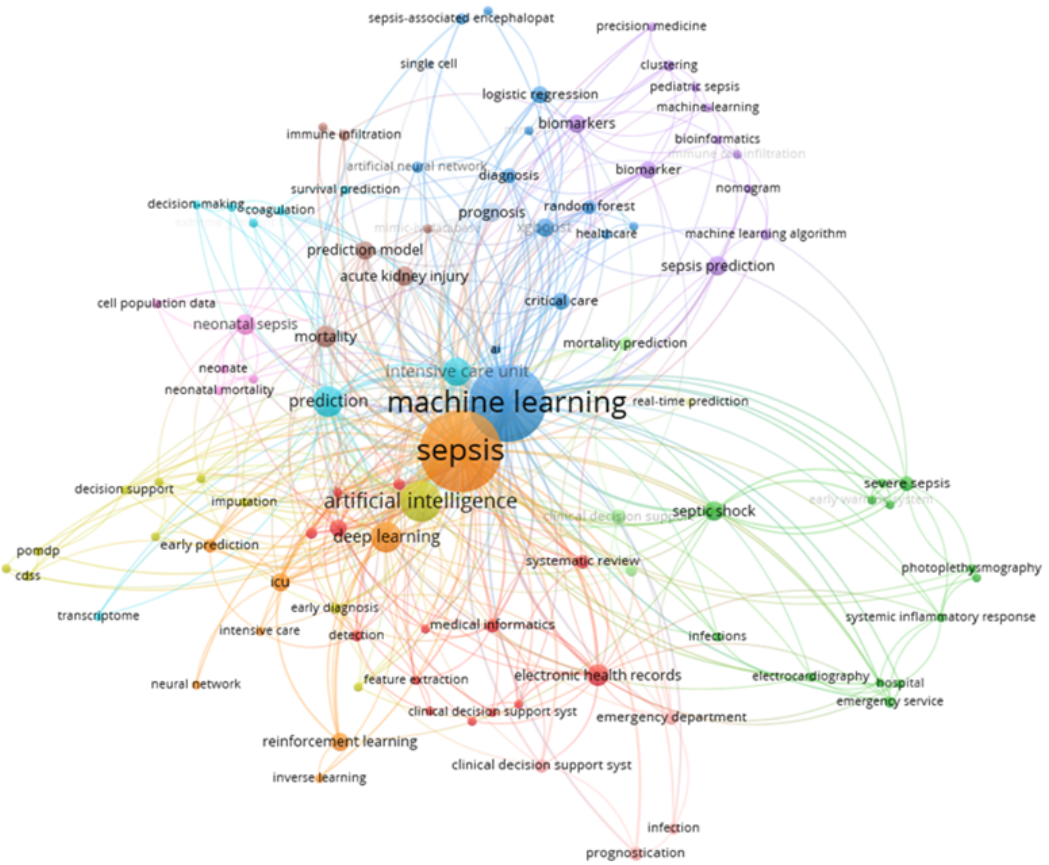
Co-Occurrence Analysis of Top 100 Keywords

Keywords encapsulate the central themes within a publication and are ideal for examining interconnected areas of research. In this study, we performed co-occurrence analysis to pinpoint the prominent research focal points in the field of AI application in sepsis research, using the top 100 keywords. The extraction and clustering of these top 100 keywords were performed using VOSviewer.

Figure 5 illustrates our use of VOSviewer to create a visual network map, consisting of 6 clusters based on the co-occurrence of the top 100 keywords. The core of this visualization network map is occupied by the following keywords: sepsis (n=138), machine learning (n=122), artificial intelligence (n=35), and deep learning (n=20).

XSL•FO
RenderX

Figure 5. The co-occurrence network of the top 100 keywords in artificial intelligence research for sepsis from 2014 to 2023 (August).



Discussion

Main Findings

This study investigated the citation analyses of AI models in sepsis research, scrutinizing the publication patterns related to their application. This study reports a noteworthy upswing in interest in this subject over the past decade, particularly increasing from 2018 to 2022.

The substantial increase in the use of AI in health care research is noteworthy. For instance, there was an 88.88% increase in the use of AI in health care research after 2012, with only 10 countries contributing to over 96% of the studies [25]. Our findings also indicate an annual growth rate of 44.81% in the use of AI in sepsis-related research. If this growth rate continues in the future, we can anticipate the publication volume in the domain of AI in sepsis-related research doubling approximately every 5 years. This increased number of publications indicates advancements, improved functionality, and the overall progress of AI in sepsis, especially when compared to other application areas.

The number of publications exhibits geographical disparities [26-28]. Recent research indicates a rise in the application of AI in health care research, particularly in developed countries [12,29,30]. These nations are investing increased funds into AI research and the development of AI tools to improve health outcomes. This study aligns with these findings, highlighting an increased number of publications originating from developed countries. However, it underscores the importance of researchers from developing nations stepping forward to contribute toward achieving ultimate health goals. Effective collaboration among clinical experts, AI model developers, and health care providers is essential to addressing the challenges at hand [31].

Selecting the appropriate journals for publication can be a complex decision [32]. Authors take various factors into account when submitting, such as the accessibility of open-access journals and the higher impact factor associated with certain subscription journals [33]. Some authors prefer open-access journals for their widespread availability after publication and the potential for a higher number of citations [34]. Conversely, subscription journals from reputable publishers attract attention due to their high impact factors [35,36]. New researchers often struggle with the dilemma of where to submit their work. This study shows that authors frequently weigh both open-access and higher-impact–factor journals when publishing research related to the application of AI in sepsis.

This study aimed primarily to highlight the most commonly used data sets and algorithms in the current literature. The majority of studies used the Multiparameter Intelligent Monitoring in Intensive Care III (MIMIC-III) data set and supervised machine learning models [37-39]. Notably, Komorowski et al [16] introduced a reinforcement learning model to predict sepsis in patients, demonstrating average reliability levels higher than those of human clinicians. While many studies used extensive training and testing data sets, the

XSL•FO
RenderX

majority focused on single-centered data [39]. To apply these models in real-world clinical settings, external validation becomes necessary.

Strengths and Limitations

This study has several strengths. First, it is the first comprehensive bibliometric analysis that sheds light on the research trends of the application of AI in sepsis, illustrating how this field has evolved. Second, this study gauges productivity in terms of sources, authors, institutions, and countries, while also visualizing word trends. This provides novel and in-depth insights for both researchers and practitioners. This study also has some limitations to address. First, we only collected relevant publications from the Web of Science, a widely used academic resource, for bibliometric analyses [13,40-42]. Nevertheless, using other databases, such as PubMed or Scopus, might have provided slightly varied findings. Second, our inclusion criteria comprised articles published solely in English. However, inclusion of other languages, gray literature, and books might have influenced outcomes, particularly considering diverse cultural perspectives among scholars on the application of AI in sepsis. Finally, relying solely on article titles for the search may pose limitations. However, our aim was to focus on publications specifically addressing the application of AI in sepsis. Therefore, a title screening was deemed more suitable than a broader topic search.

Conclusion

This study aimed to present a comprehensive overview of the use of AI in sepsis through a systematic analysis of existing literature. The findings of this study reveal a noticeable increase in the number of publications over the last 10 years. Until now, developed countries have been the primary contributors in this field. Researchers from developing countries should step forward, leveraging population advantages and core technologies in different regions to foster collaboration.

Leading multidisciplinary science journals, including *Frontiers in Medicine*, *Scientific Reports*, and the *American Journal of Respiratory and Critical Care Medicine*, emerge as key contributors to this topic based on the volume of published articles. As the application of AI in sepsis research continues to rise, this study serves as a valuable resource for researchers seeking direction and opportunities for collaboration.

Authors' Contributions

All authors contributed to the conception and design of the study; data collection, analysis, and interpretation; drafting and revising the article critically for content; and approval of the final version to be submitted.

Conflicts of Interest

None declared.

References

1. Islam MM, Nasrin T, Walther BA, Wu CC, Yang HC, Li YC. Prediction of sepsis patients using machine learning approach: a meta-analysis. Comput Methods Programs Biomed 2019;170:1-9. [doi: 10.1016/j.cmpb.2018.12.027] [Medline: 30712598]
2. Wang W, Liu CF. Sepsis heterogeneity. World J Pediatr 2023;19(10):919-927. [doi: 10.1007/s12519-023-00689-8] [Medline: 36735197]
3. Wardi G, Tainter CR, Ramnath VR, Brennan JJ, Tolia V, Castillo EM, et al. Age-related incidence and outcomes of sepsis in California, 2008-2015. J Crit Care 2021;62:212-217 [FREE Full text] [doi: 10.1016/j.jcrc.2020.12.015] [Medline: 33429114]
4. Rhee C, Jones TM, Hamad Y, Pande A, Varon J, O'Brien C, et al. Prevalence, underlying causes, and preventability of sepsis-associated mortality in US acute care hospitals. JAMA Netw Open 2019;2(2):e187571 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.7571] [Medline: 30768188]
5. Vincent JL. The clinical challenge of sepsis identification and monitoring. PLoS Med 2016;13(5):e1002022 [FREE Full text] [doi: 10.1371/journal.pmed.1002022] [Medline: 27187803]
6. Talisa VB, Yende S, Seymour CW, Angus DC. Arguing for adaptive clinical trials in sepsis. Front Immunol 2018;9:1502 [FREE Full text] [doi: 10.3389/fimmu.2018.01502] [Medline: 30002660]
7. de Grooth HJ, Postema J, Loer SA, Parienti JJ, Oudemans-van Straaten HM, Girbes AR. Unexplained mortality differences between septic shock trials: a systematic analysis of population characteristics and control-group mortality rates. Intensive Care Med 2018;44(3):311-322 [FREE Full text] [doi: 10.1007/s00134-018-5134-8] [Medline: 29546535]
8. Goh KH, Wang L, Yeow AYK, Poh H, Li K, Yeow JJL, et al. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nat Commun 2021;12(1):711 [FREE Full text] [doi: 10.1038/s41467-021-20910-4] [Medline: 33514699]
9. Luz CF, Vollmer M, Decruyenaere J, Nijsten MW, Glasner C, Sinha B. Machine learning in infection management using routine electronic health records: tools, techniques, and reporting of future technologies. Clin Microbiol Infect 2020;26(10):1291-1299 [FREE Full text] [doi: 10.1016/j.cmi.2020.02.003] [Medline: 32061798]
10. 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 [FREE Full text] [doi: 10.1007/s00134-019-05872-y] [Medline: 31965266]

XSL•FO
RenderX

11. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: bibliometric analysis. J Med Internet Res 2020;22(7):e18228 [FREE Full text] [doi: 10.2196/18228] [Medline: 32723713]
12. Chen X, Xie H, Li Z, Cheng G, Leng M, Wang FL. Information fusion and artificial intelligence for smart healthcare: a bibliometric study. Inf Process Manage 2023;60(1):103113 [FREE Full text] [doi: 10.1016/j.ipm.2022.103113]
13. Wu CC, Huang CW, Wang YC, Islam MM, Kung WM, Weng YC, et al. mHealth research for weight loss, physical activity, and sedentary behavior: bibliometric analysis. J Med Internet Res 2022;24(6):e35747 [FREE Full text] [doi: 10.2196/35747] [Medline: 35675126]
14. Poly TN, Islam MM, Walther BA, Lin MC, Jack Li YC. Artificial intelligence in diabetic retinopathy: bibliometric analysis. Comput Methods Programs Biomed 2023;231:107358. [doi: 10.1016/j.cmpb.2023.107358] [Medline: 36731310]
15. Yeung AWK, Parvanov ED, Hribersek M, Eibensteiner F, Klager E, Kletecka-Pulker M, et al. Digital teaching in medical education: scientific literature landscape review. JMIR Med Educ 2022;8(1):e32747 [FREE Full text] [doi: 10.2196/32747] [Medline: 35138260]
16. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018;24(11):1716-1720 [FREE Full text] [doi: 10.1038/s41591-018-0213-5] [Medline: 30349085]
17. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. Crit Care Med 2018;46(4):547-553 [FREE Full text] [doi: 10.1097/CCM.0000000000002936] [Medline: 29286945]
18. Taylor RA, Pare JR, Venkatesh AK, Mowafi H, Melnick ER, Fleischman W, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. Acad Emerg Med 2016;23(3):269-278 [FREE Full text] [doi: 10.1111/acem.12876] [Medline: 26679719]
19. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. JMIR Med Inform 2016;4(3):e28 [FREE Full text] [doi: 10.2196/medinform.5909] [Medline: 27694098]
20. Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. PLoS One 2017;12(4):e0174708 [FREE Full text] [doi: 10.1371/journal.pone.0174708] [Medline: 28384212]
21. Gultepe E, Green JP, Nguyen H, Adams J, Albertson T, Tagkopoulos I. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. J Am Med Inform Assoc 2014;21(2):315-325 [FREE Full text] [doi: 10.1136/amiajnl-2013-001815] [Medline: 23959843]
22. Giannini HM, Ginestra JC, Chivers C, Draugelis M, Hanish A, Schweickert WD, et al. A machine learning algorithm to predict severe sepsis and septic shock: development, implementation, and impact on clinical practice. Crit Care Med 2019;47(11):1485-1492 [FREE Full text] [doi: 10.1097/CCM.0000000000003891] [Medline: 31389839]
23. Hou N, Li M, He L, Xie B, Wang L, Zhang R, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. J Transl Med 2020;18(1):462 [FREE Full text] [doi: 10.1186/s12967-020-02620-5] [Medline: 33287854]
24. Mani S, Ozdas A, Aliferis C, Varol HA, Chen Q, Carnevale R, et al. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. J Am Med Inform Assoc 2014;21(2):326-336 [FREE Full text] [doi: 10.1136/amiajnl-2013-001854] [Medline: 24043317]
25. Jimma BL. Artificial intelligence in healthcare: a bibliometric analysis. Telemat Inform Rep 2023;9:100041 [FREE Full text] [doi: 10.1016/j.teler.2023.100041]
26. Hong C, Wang L. Virtual reality technology in nursing professional skills training: bibliometric analysis. JMIR Serious Games 2023;11:e44766 [FREE Full text] [doi: 10.2196/44766] [Medline: 37603389]
27. Tang R, Zhang S, Ding C, Zhu M, Gao Y. Artificial intelligence in intensive care medicine: bibliometric analysis. J Med Internet Res 2022;24(11):e42185 [FREE Full text] [doi: 10.2196/42185] [Medline: 36449345]
28. Wang J, Liang Y, Cao S, Cai P, Fan Y. Application of artificial intelligence in geriatric care: bibliometric analysis. J Med Internet Res 2023;25:e46014 [FREE Full text] [doi: 10.2196/46014] [Medline: 37351923]
29. Suhail F, Adel M, Al-Emran M, Shaalan K. A bibliometric analysis on the role of artificial intelligence in healthcare. In: Mishra S, Tripathy HK, Mallick P, Shaalan K, editors. Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis. Singapore: Springer; 2022:1-14.
30. Javidan AP, Li A, Lee MH, Forbes TL, Naji F. A systematic review and bibliometric analysis of applications of artificial intelligence and machine learning in vascular surgery. Ann Vasc Surg 2022;85:395-405. [doi: 10.1016/j.avsg.2022.03.019] [Medline: 35339595]
31. Gundersen T, Bærøe K. The future ethics of artificial intelligence in medicine: making sense of collaborative models. Sci Eng Ethics 2022;28(2):17 [FREE Full text] [doi: 10.1007/s11948-022-00369-2] [Medline: 35362822]
32. Rowley J, Sbaffi L, Sugden M, Gilbert A. Factors influencing researchers' journal selection decisions. J Inf Sci 2022;48(3):321-335 [FREE Full text] [doi: 10.1177/0165551520958591]
33. Triggle CR, MacDonald R, Triggle DJ, Grierson D. Requiem for impact factors and high publication charges. Account Res 2022;29(3):133-164 [FREE Full text] [doi: 10.1080/08989621.2021.1909481] [Medline: 33787413]

XSL•FO
RenderX

34. Hardwicke TE, Thibault RT, Kosie JE, Tzavella L, Bendixen T, Handcock SA, et al. Post-publication critique at top-ranked journals across scientific disciplines: a cross-sectional assessment of policies and practice. R Soc Open Sci 2022;9(8):220139 [FREE Full text] [doi: 10.1098/rsos.220139] [Medline: 36039285]
35. Bartocci L, Grossi G, Mauro SG, Ebdon C. The journey of participatory budgeting: a systematic literature review and future research directions. Int Rev Adm Sci 2023;89(3):757-774 [FREE Full text] [doi: 10.1177/00208523221078938]
36. Maddi A, Sapinho D. Article processing charges, altmetrics and citation impact: is there an economic rationale? Scientometrics 2022;127(12):7351-7368. [doi: 10.1007/s11192-022-04284-y]
37. Yan MY, Gustad LT, Nytrø Ø. Sepsis prediction, early detection, and identification using clinical text for machine learning: a systematic review. J Am Med Inform Assoc 2022;29(3):559-575 [FREE Full text] [doi: 10.1093/jamia/ocab236] [Medline: 34897469]
38. Deng HF, Sun MW, Wang Y, Zeng J, Yuan T, Li T, et al. Evaluating machine learning models for sepsis prediction: a systematic review of methodologies. iScience 2022;25(1):103651 [FREE Full text] [doi: 10.1016/j.isci.2021.103651] [Medline: 35028534]
39. Kijpaisalratana N, Sanglertsinlapachai D, Techaratsami S, Musikatavorn K, Saoraya J. Machine learning algorithms for early sepsis detection in the emergency department: a retrospective study. Int J Med Inform 2022;160:104689. [doi: 10.1016/j.ijmedinf.2022.104689] [Medline: 35078027]
40. Yeung AWK, Kulnik ST, Parvanov ED, Fassl A, Eibensteiner F, Völkl-Kernstock S, et al. Research on digital technology use in cardiology: bibliometric analysis. J Med Internet Res 2022;24(5):e36086 [FREE Full text] [doi: 10.2196/36086] [Medline: 35544307]
41. Yang K, Hu Y, Qi H. Digital health literacy: bibliometric analysis. J Med Internet Res 2022;24(7):e35816 [FREE Full text] [doi: 10.2196/35816] [Medline: 35793141]
42. Yeung AWK, Tosevska A, Klager E, Eibensteiner F, Tsagkaris C, Parvanov ED, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022;24(1):e28152 [FREE Full text] [doi: 10.2196/28152] [Medline: 34951864]

Abbreviations

AI: artificial intelligence
MIMIC-III: Multiparameter Intelligent Monitoring in Intensive Care III

Edited by T de Azevedo Cardoso; submitted 11.11.23; peer-reviewed by E Kawamoto, D Amancio; comments to author 19.01.24; revised version received 27.01.24; accepted 04.03.24; published 15.04.24.

Please cite as:
Wu M, Islam MM, Poly TN, Lin MC
Application of AI in Sepsis: Citation Network Analysis and Evidence Synthesis
Interact J Med Res 2024;13:e54490
URL: https://www.i-jmr.org/2024/1/e54490
doi:10.2196/54490
PMID:38621231

©MeiJung Wu, Md Mohaimenul Islam, Tahmina Nasrin Poly, Ming-Chin Lin. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 15.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Theoretical Perspectives Underpinning Research on the Physician-Patient Relationship in a Digital Health Practice: Scoping Review

Damoun Nassehi[1], MD, PhD; Birgitta Haga Gripsrud[1], Prof Dr; Ellen Ramvi[1], Prof Dr

Department of Caring and Ethics, Faculty of Health Sciences, University of Stavanger, Stavanger, Norway

Corresponding Author:
Damoun Nassehi, MD, PhD
Department of Caring and Ethics
Faculty of Health Sciences
University of Stavanger
Kjell Arholms gate 41
Stavanger, 4036
Norway
Phone: 47 45426096
Email: damoun.nassehi@uis.no

Abstract

Background: The advent of digital health technologies has transformed the landscape of health care, influencing the dynamics of the physician-patient relationship. Although these technologies offer potential benefits, they also introduce challenges and complexities that require ethical consideration.

Objective: This scoping review aims to investigate the effects of digital health technologies, such as digital messaging, telemedicine, and electronic health records, on the physician-patient relationship. To understand the complex consequences of these tools within health care, it contrasts the findings of studies that use various theoretical frameworks and concepts with studies grounded in relational ethics.

Methods: Using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines, we conducted a scoping review. Data were retrieved through keyword searches on MEDLINE/PubMed, Embase, IEEE Xplore, and Cochrane. We screened 427 original peer-reviewed research papers published in English-language journals between 2010 and 2021. A total of 73 papers were assessed for eligibility, and 10 of these were included in the review. The data were summarized through a narrative synthesis of the findings.

Results: Digital health technologies enhance communication, improve health care delivery efficiency, and empower patients, leading to shifts in power dynamics in the physician-patient relationship. They also potentially reinforce inequities in health care access due to variations in technology literacy among patients and lead to decreases in patient satisfaction due to the impersonal nature of digital interactions. Studies applying a relational ethics framework have revealed the nuanced impacts of digital health technologies on the physician-patient relationship, highlighting shifts toward more collaborative and reciprocal care. These studies have also explored transitions from traditional hierarchical relationships to mutual engagement, capturing the complexities of power dynamics and vulnerabilities. Other theoretical frameworks, such as patient-centered care, and concepts, such as patient empowerment, were also valuable for understanding these interactions in the context of digital health.

Conclusions: The shift from hierarchical to collaborative models in the physician-patient relationship not only underscores the empowering potential of digital tools but also presents new challenges and reinforces existing ones. Along with applications for various theoretical frameworks and concepts, this review highlights the unique comprehensiveness of a relational ethics perspective, which could provide a more nuanced understanding of trust, empathy, and power dynamics in the context of digital health. The adoption of relational ethics in empirical research may offer richer insights into the real-life complexities of the physician-patient relationship, as mediated by digital technologies.

(Interact J Med Res 2024;13:e47280) doi:10.2196/47280

XSL•FO
RenderX

KEYWORDS

digital health; mobile health; telemedicine; physician-patient relations; relational ethics; primary care; patient-provider; physician-patient; telehealth; relationship; eHealth; scoping review; review method; mobile phone

Introduction

Rationale

Digital health technologies are revolutionizing the practice of health care [1]. Driven by societal, political, and technological advancements, these technology-mediated changes are occurring at an unprecedented pace. Envisioning a future in which medicine can be personalized to individuals and diseases can be foreseen and prevented, some experts have high hopes for the potential of technology to improve human life [1,2]. However, skeptics argue that achieving personalization through technology is an elusive paradox, as it fails to account for the fundamental uncertainties and complexities of medicine and the value of human interactions [3].

In the realm of medical practice, a physician's primary goal is to diagnose and treat patients' diseases by drawing upon their biomedical expertise and years of clinical experience. Nevertheless, a physician's practice must also be guided and adapted based on their relationship with the patient. A sustained physician-patient relationship facilitates the tailoring of therapeutic interventions to best suit individual patients' needs [4]. Interpersonal skills honed over time through relational interactions with patients cannot be directly replaced by technology. Future technologies, including artificial intelligence, can only interpret what is explicitly documented as text or images in the electronic health record (EHR), which could lead to potentially missing valuable information generated within the physician-patient relationship, such as the patient's values, preferences, trust, and rapport [5,6]. Moreover, a doctor's personality and nonreflective actions, such as active listening or disregarding patient preferences, can have a significant impact not only on curing the patient's illness but also on maintaining or exacerbating it [4]. Continuity of care with the same physician has been shown to reduce hospital admissions and mortality rates [7]. The rapid integration of technology introduces complex dimensions to an already intricate relationship, necessitating an exploration of technology's impact on the physician-patient dynamic.

Digital technology offers potential for both improvements and risk, as it can efficiently reach a larger number of patients with less effort. At best, this technology-mediated efficiency may offer improved accessibility, real-time monitoring, and personalized treatments. It can also bridge geographical divides, connecting remote patients to specialists. At worst, it may be harmful on a scale far beyond what a single physician could achieve in a lifetime, potentially introducing systematic biases on a massive, automatic scale.

A potential problem with research on digital health technologies is the risk of bias, for example, when research funding comes from for-profit organizations and technology providers. As there may also be a dearth of independent research being published in this domain, we risk developing a skewed view of the evidence [8]. Additionally, while technological advancements surge ahead, ethical and regulatory frameworks have struggled to keep pace with these developments [9], making the need for new knowledge even more pressing.

Undeniably, digital technologies are here to stay and are already reshaping medical practice and the physician-patient relationship. Yet, more research is necessary to truly understand the nuanced implications of this shift. While technology promises efficiency and precision, the physician-patient relationship is rooted in intersubjective trust, empathy, and a deep understanding that transcends quantifiable data. As we further integrate technology into health care, it is therefore essential to explore its impact on this foundational relationship. Does it have the potential to augment the bond, creating new avenues for connection and understanding, or does it carry a more detrimental potential to create distance, becoming a screen that separates rather than connects human beings? By delving deeper into these questions, we can ensure that, as we advance technically, we do not lose sight of the human touch that remains at the heart of healing.

Objectives

This review explores the existing knowledge gap concerning the impact of digital health on the physician-patient dynamic. Additionally, we analyze various theoretical frameworks and concepts used in empirical studies concerning this relationship and contrast the findings of these studies with the results of research grounded in relational ethics frameworks [10]. This research focus has enabled us to identify the consequences of using digital health technologies (such as digital messaging, telemedicine, health-related websites, and EHRs) in physician-patient relationships.

Methods

Overview

Due to the rapid advancements in digital health care, there is a noticeable knowledge gap in the existing literature. A scoping review was deemed a suitable review approach, as it is well suited for providing an updated understanding of the current state of knowledge [11]. This scoping review was conducted according to the guidelines of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [12]. The study protocol was preregistered with the Center for Open Science [13].

Eligibility Criteria

The combined emergence of smartphones and 4G networks from around 2010 has played a crucial role in driving the development and use of digital health services [14]. This convergence has provided individuals with convenient access to health care services through mobile health (mHealth) apps, remote monitoring, and telemedicine, while also enabling health care providers to leverage real-time health data for personalized and data-driven care approaches. Only original peer-reviewed research published in English-language journals from January

XSL•FO
RenderX

1, 2010, to December 31, 2021, was included in this review. The decision to conclude the literature review on December 31, 2021, was driven by the aim to provide a comprehensive and current overview of the field up to our initial submission at the beginning of 2023. We carefully evaluated the scope and depth of available literature within this timeframe, determining that extending the search would likely not alter our core findings.

Research papers on patients and general practitioners or physicians involving health technology were included. Studies that did not specifically mention the physician-patient relationship were excluded.

Information Sources

Studies for inclusion were identified through searches in the following databases: MEDLINE/PubMed, Embase, IEEE Xplore, and Cochrane. Since our review specifically targeted empirical research on the physician-patient relationship, we primarily focused on searching medical databases. However, we also included IEEE Xplore to ensure that we did not miss any significant research in the fields of computer science and electrical engineering. To ensure trustworthy findings and direct insights, our search prioritized original papers and solely focused on high-quality research papers published in peer-reviewed medical journals. Reviews, opinion pieces, and other nonempirical texts were not included. Duplicates were removed.

Search

In collaboration with a librarian, we devised a comprehensive search string to explore the application of digital health technology within the physician-patient relationship. To ensure clarity in our scope, we used the mnemonic strategy PCC (population, concept, and context) to determine which papers should be included. Specifically, the population consisted of physicians, patients, or both; the concept involved digital health technologies used in medical care; and the context revolved around their implementation within the physician-patient relationship. The search string encompassed relevant terms, such as “digital health,” “mHealth,” “eHealth,” “telemedicine,” or “telehealth,” combined with the key phrases “physician-patient relations” or “relational ethics.” Detailed explanations of these terms are provided in Table 1. The search string used was (“digital health” OR “mHealth” OR “eHealth” OR “telemedicine” OR “telehealth”) AND (“physician-patient relations” OR “relational ethics”).

Our particular focus was on “relational ethics” due to its potential to inform a more ethical analysis of the physician-patient relationship, thus addressing the intricacies and challenges unique to this context.

We searched the 4 databases applying the following limitations: “scholarly (peer-reviewed) journals,” “date of publication from January 1, 2010 to December 31, 2021,” and “English language.” The search resulted in the identification of 429 journal papers. Of these papers, 427 were discovered in MEDLINE/PubMed, while 1 paper was identified as a duplicate appearing in both the Cochrane and MEDLINE/PubMed libraries. Additionally, 1 paper obtained from Embase was identified as a scoping review and was therefore excluded from our screening analysis. It is worth noting that no papers were found in IEEE Xplore during our search process. In short, all 427 papers selected for screening were found in the MEDLINE/PubMed database.

Table 1. Search terms.

Search terms	Explanation
digital health	Refers to technology related to health care services and solutions
mHealth[a]	Short for “mobile health,” involves the use of mobile devices
eHealth	Refers to “electronic health,” including digital health services
telemedicine	Involves remote diagnosis and treatment using technology
telehealth	Broad term covering health care services through telecommunication
Physician-patient relations	Refers to the interactions and dynamics between doctors and patients
relational ethics	Refers to an ethical framework focused on understanding and navigating the intricacies of interpersonal dynamics in the doctor-patient relationship

[a]mHealth: mobile health.

Selection of Sources of Evidence

The screening process was performed by 2 researchers. It involved reviewing the abstracts initially found for inclusion, and in cases where abstracts lacked sufficient information, the entire research paper was examined. In instances where 2 researchers could not reach a consensus on whether to include a research paper, a third researcher was consulted as an arbiter to resolve any disagreements. This process led to the inclusion of 73 research papers. Subsequently, all 73 included papers underwent a detailed and thorough examination by 2 researchers. In the event of any disagreement during this phase, a third researcher was consulted to ensure accuracy. Studies that did not specifically mention the physician-patient relationship (n=63) were excluded, resulting in 10 studies to be reviewed.

Data Charting Process

A standardized data charting form was created for the data extraction process, and data charting was performed by 2 reviewers.

Data Items

Our objective was to collect a range of data items relevant to the research. These included the year of the study, the country

XSL•FO
RenderX

where it was conducted, the study type (qualitative or quantitative), the theoretical or ethical framework and concepts used, the type of technology studied, the research objectives, and the participants involved (patients, physicians, and other health care personnel [HCP]; Table 2). Additionally, we extracted text samples that described the impact of digital health technologies on the physician-patient relationship. Our approach involved presenting the main findings in a textual manner.

Table 2. Overview of the data from the included studies.

Authors	Year	Journal	Country	Study type	Type of technology	Aims	Participants
Audrain-Pontevia and Menvielle [15]	2018	*Health Services Management Research*	Canada	Quantitative	Online health community	Examine how online health communities impact the physician-patient relationship	Patients
Balato et al [16]	2013	*British Journal of Dermatology*	Italy	Quantitative	Mobile phone messages	Evaluate the use of telemedicine in improving treatment adherence, patient outcomes, and the physician-patient relationship	Patients and physicians
Grünloh et al [17]	2018	*Journal of Medical Internet Research*	Sweden	Qualitative	Web-based patient portal	Investigate how physicians view the idea of patient participation	Physicians
Győrffy et al [18]	2020	*PLOS ONE*	Hungary	Qualitative	Social media	Explore physicians' knowledge and attitudes toward digital health technologies and the transformation of the doctor-patient relationship	Physicians
Jiang [19]	2019	*Health Communication*	China	Quantitative	Digital messaging	Examine how the quality of face-to-face communication with providers is associated with their subsequent internet use for patient-provider communication	Patients
Kludacz-Alessandri et al [20]	2021	*PLOS ONE*	Poland	Qualitative	Teleconsultation (phone)	Study patients' satisfaction with teleconsultation in primary care and the impact of teleconsultations on GP[a]-patient communication	Patients and physicians
Macdonald et al [21]	2018	*Journal of Medical Internet Research*	Canada	Qualitative	Digital messaging and EHR[b]	Examine HCP's[c] perspectives on how eHealth affects their relationships with patients, as well as its ethical ramifications	Patients, physicians, and other HCPs
Tasneem et al [22]	2019	*American Journal of Hospice and Palliative Medicine*	United States	Qualitative	Video consultation	Investigate the need for web-based videoconferences for oncology patients	Patients and physicians
Townsend et al [23]	2015	*Journal of Medical Internet Research*	Canada	Qualitative	Health-related websites	Focus on patients' and HCP's use of health-related internet information and how it influences the patient-HCP relationship	Patients, physicians, and other HCPs
Yan et al [24]	2020	*International Journal of Environmental Research and Public Health*	China	Qualitative	Digital messaging	Understand the underlying reasons for poor doctor-patient relationships	Patients and physicians

[a]GP: general practitioner.

[b]EHR: electronic health record.

[c]HCP: health care personnel.

Synthesis of Results

To synthesize the results, we first created a summary based on the extracted data items to capture the key findings and statements from the included research papers concerning the physician-patient relationship. This summary was then compared with each research paper in its entirety to ensure accuracy.

Next, we conducted a narrative synthesis of the findings from the included studies, incorporating both the extracted text samples and the summaries of the papers. This approach allowed us to provide a comprehensive overview of the results from both quantitative and qualitative studies, enabling a comprehensive summary of the physician-patient relationship in the context of digital health technologies (Table 3). In addition, we assessed whether the papers applied any ethical frameworks or simple concepts and analyzed how their use contributed to the research (Table 4).

XSL•FO
RenderX

Table 3. Summary of evidence.

Authors	Summary: What does the paper state about the physician-patient relationship?
Audrain-Pontevia and Menvielle [15]	• The paper discusses the impact of online health communities on the patient-physician relationship. The authors explore how online health communities, which provide users with computer-mediated social support and empowerment, impact this relationship. • The authors acknowledge that, traditionally, doctors were the main source of medical information and therefore benefited from authority and power over their patients. However, with the advent of online health communities, patients now have access to social support, resources, and aid, which can make them feel more empowered and influence their relationships with their physicians. • The authors propose that online health communities offer patients the opportunity to gain the power to handle their illnesses and their health, presumably leading to increased participation during the consultation and improving their commitment to their relationship with their physician.
Balato et al [16]	• The study found that patient-physician communication improved in a group receiving SMS text message interventions, whereas it remained unchanged in the control group. This suggests that the use of digital interventions, such as SMS text messages, could potentially enhance the patient-physician relationship.
Grünloh et al [17]	• The models of the doctor-patient relationship presented in the paper describe patients as being static and unchanged, but the authors note that patients with chronic conditions often encounter new situations and need to engage in a sensemaking and learning process. • The paper suggests that the use of patient-accessible EHRs[a] can contribute to the development of the doctor-patient relationship by allowing patients to play an active role. This increased patient participation makes it more difficult for physicians to maintain a strategy that potentially excludes patients. • The authors state that eHealth does not have to be a "power struggle" in the doctor-patient relationship but can potentially help both partners improve their relationship collectively and grow individually. • The authors mention the importance of patient participation for patient safety.
Győrffy et al [18]	• The role of the doctor is in transition, with doctors expected to perform more complex tasks including health information technology and aiding in the digital orientation of patients. They see themselves transforming into mediators based on efficient communication with their patients. • Digitally engaged physicians consider themselves guides, undertaking a guardian and information managing function in the description, collection, and sharing of credible content in the online space. • For a successful leap from hierarchical patterns to the 21st-century doctor-patient relationship, the future generation of physicians should be trained differently and prepared for all the above-described changes. • Medical school curricula should emphasize health and prevention rather than only diseases and pathology via the newest digital technological solutions. • Medical students need to prepare for predictive and proactive working environments, including their new role as a guide or mediator for digitally empowered patients, in contrast with the paternalistic physicians of previous generations.
Jiang [19]	• The results of the study emphasize the important roles of patient-centered communication and the physician-patient relationship in the eHealth and mHealth[b] movement, particularly in the Chinese health care system. • The interplay of physician-patient communication in face-to-face environments and relationship factors (eg, patient trust and patient satisfaction) could exert significant effects in promoting eHealth adoption. • To encourage patients to adopt eHealth technologies, health care providers should first build a patient-centered environment (eg, responding to patients' informational and emotional needs and engaging patients in medical decision-making).
Kludacz-Alessandri et al [20]	• The study concerns patients' satisfaction with teleconsultation in primary care and the impact of teleconsultations on GP[c]-patient communication during the COVID-19 pandemic in Poland. The paper suggests that the quality of GP-patient communication is an essential factor that can improve the results of treatment and patient satisfaction. Only 55% (n=99) of the patients found teleconsultation to be as good as in-person visits with their physician.
Macdonald et al [21]	• The study discusses the concept of a "two-way conversation" that is evolving in the health care provider-patient dynamic. This shift toward more collaborative interactions with patients is, in part, facilitated by eHealth technologies. The authors examine the impact of eHealth on the current state of collaborative consultation, highlighting how it aids in having, using, and supporting conversations with patients. • Some health care professionals in the study embraced the idea of patients as "partners," as they see a partner as someone who helps in improving an outcome by educating themselves and conscientiously monitoring their condition and behavior. • One of the health care professionals stated that patients who are engaged through eHealth and informed about their condition are more useful clinically. • The paper also mentions a pedagogical approach to interacting with the eHealth users among patients.

XSL•FO
RenderX

Authors	Summary: What does the paper state about the physician-patient relationship?
Tasneem et al [22]	• The study assessed the needs of patients receiving palliative care and their perception of how telemedicine video visits might influence their care. Despite concerns about truncated physical examinations and prescription limits, the majority of patients favored having the opportunity for telemedicine video visits. They felt that the physician-patient relationship would not diminish and had few cost concerns. They believed that a video alternative to an in-person visit might increase access, save time, and increase comfort and safety by avoiding a trip to the hospital.
Townsend et al [23]	• In this study, patients reported using personal websites, blogs, chat rooms, and online links to medical test results as part of their eHealth resources. This suggests that patients are actively engaging with digital resources in managing their health, which can have implications for the physician-patient relationship. The paper discusses how the rapid explosion in online digital health resources is seen as transformational, accelerating the shift from traditionally passive patients to patients as partners. This is altering the current patient-health care professional relationship. • The proliferation of eHealth strategies is accelerating a shift in health care from a traditional and paternalistic delivery model to a more mutual patient-health care professional relationship in which informed patients are actively involved in their care and treatment decisions. • The authors mention that eHealth resources provide patients with extensive and up-to-date information, access to medical research, connections to people with similar conditions, immediacy, and convenience in patient-health care professional communications. These factors can significantly impact the physician-patient relationship.
Yan et al [24]	• The paper investigates the underlying reasons for poor doctor-patient relationships in mobile consultation from the perspective of computer-mediated communication. This suggests that the physician-patient relationship may be influenced by the mode of communication, particularly in a digital context. • The paper emphasizes the emerging use of mobile medical consultation in China, which has propelled the establishment of doctor-patient relationships in the mobile context. This again underscores the impact of digital technologies on the physician-patient relationship. • The authors also mention different models or concepts used to assess the doctor-patient relationship, suggesting that the nature of this relationship can be complex and multifaceted.

[a]EHR: electronic health record.

[b]mHealth: mobile health.

[c]GP: general practitioner.

Table 4. Theoretical frameworks and concepts used in the studies.

Authors	Framework or concepts	Relevance of using specific theoretical frameworks and concepts
Audrain-Pontevia and Menvielle [15]	Patient empowerment	Evaluates the impact of online health communities on the physician-patient relationship.
Balato et al [16]	None mentioned	N/A[a]
Grünloh et al [17]	Shared decision-making, patient-centered care, and paternalism	Assesses the impact of a web-based patient portal on patient participation.
Győrffy et al [18]	Patient empowerment and shared decision-making	Studies the attitudes of digitally engaged physicians toward transforming the physician-patient relationship.
Jiang [19]	Patient-centered care, mutual trust, and patient satisfaction	Examines the relationship between face-to-face and online patient-provider communication.
Kludacz-Alessandri et al [20]	Mutual trust	Investigates the satisfaction of patients with teleconsultations.
Macdonald et al [21]	Relational ethics	Relational ethics addresses the ethical content and decisions implicit in everyday relationships and conversations.
Tasneem et al [22]	None mentioned	N/A
Townsend et al [23]	Relational ethics	Core elements of relational ethics are applicable to everyday experiences, practice, and interactions. Applying relational ethics helps with focusing on what is valued in interactions and relationships and what is at risk rather than specific aspects of eHealth such as the nature of self-monitoring devices.
Yan et al [24]	None mentioned	N/A

[a]N/A: not applicable.

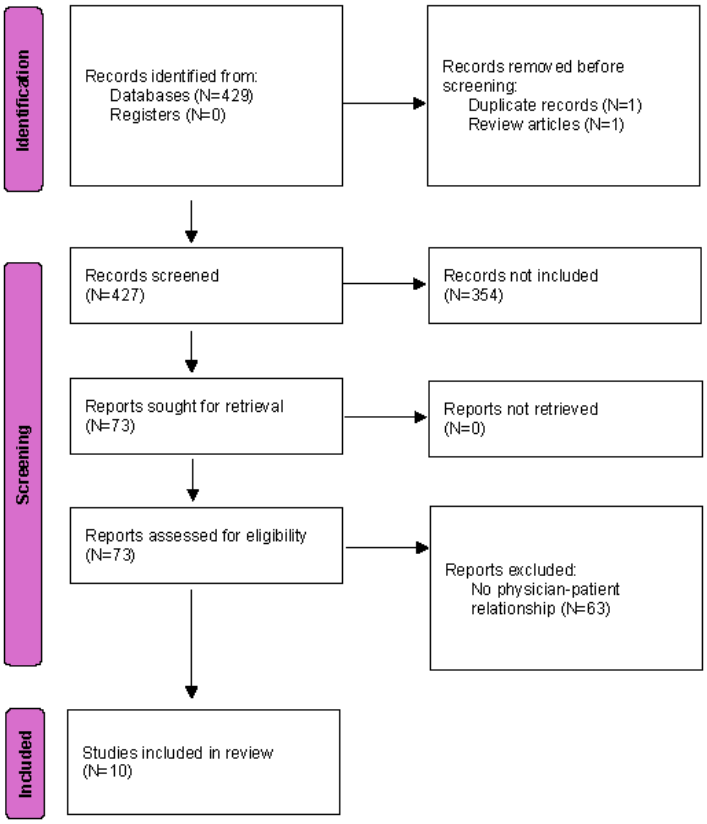
XSL•FO
RenderX

Results

Selection of Sources of Evidence

Out of the 427 studies screened, 73 studies were sought out for retrieval, and after excluding 63 studies that did not focus on the physician-patient relationship, 10 studies were included in this review. Figure 1 illustrates the search process.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) diagram: papers included in the review.



Characteristics of Sources of Evidence

The research papers included in this study [15-24] were conducted in various countries, with 3 in Canada, 2 in China, 1 in Italy, 1 in Sweden, 1 in Hungary, 1 in Poland, and 1 in the United States. These papers were published between 2013 and 2021, with half of them published in 2019 or more recently. The majority of the papers (n=7) were qualitative studies, with the remaining (n=3) papers being quantitative studies. Only 2 research papers, authored by Macdonald et al [21] and Townsend et al [23], referred to a specific ethical framework. Both used relational ethics. A total of 6 other papers adopted different theoretical frameworks and concepts, while the remaining 2 did not use any theoretical frameworks at all.

Regarding the technologies investigated, 3 papers focused on online communities, 3 papers focused on messaging, 2 papers focused on telephone or video consultation, and 2 papers focused on patient portals or EHR. The majority of the papers (n=6) included a mix of patients and physicians or other health care professionals, while 2 papers focused solely on patients and 2 others exclusively on physicians.

For further details about the characteristics of the included research papers, please refer to Table 2.

The Impact of Digital Health Technologies on the Physician-Patient Relationship

Kludacz-Alessandri et al [20] noted that, while teleconsultations offer convenience and safety, especially in situations like the COVID-19 pandemic, they may not always provide the same level of patient satisfaction as in-person visits. According to the authors, this could be due to several reasons such as the inability to read nonverbal cues, technical difficulties, or the impersonal nature of digital communication. The authors suggested that, while digital technologies offer new avenues for communication, they may not fully mirror the richness and depth of face-to-face interactions.

Yan et al [24] discussed the use of mobile medical consultation in China and noted the potential challenges in establishing effective physician-patient relationships in a digital context. According to the authors, 1 key challenge could be the varying degrees of technology literacy among patients. Not all patients may be comfortable using digital platforms for health care, and some may lack access to the necessary technology. This study indicates that this could lead to inequities in access to health care services and affect the quality of the physician-patient relationship.

When patients become more empowered through access to digital health resources and online communities, changes could occur in the power dynamics of the physician-patient

XSL•FO
RenderX

relationship. Some physicians may struggle to adapt to these changes, particularly if they are trained in a more traditional, paternalistic model of health care. This was suggested in the study by Audrain-Pontevia and Menvielle [15].

The advent of online health communities and access to a wealth of health-related information on the internet has empowered patients to take an active role in managing their health. As discussed by Audrain-Pontevia and Menvielle [15], this empowerment can lead to increased participation during consultations and improve commitment to the relationship with the physician.

Digital technologies, such as SMS text messaging, teleconsultations, and mHealth apps, can enhance communication between patients and physicians, making it more frequent, timely, and convenient. As pointed out by Balato et al [16], such digital interventions can potentially enhance the patient-physician relationship.

The use of eHealth and mHealth technologies facilitates a more patient-centered approach to care. As Jiang [19] noted, health care providers who build a patient-centered environment—responding to patients' informational and emotional needs and engaging them in medical decision-making—can promote eHealth adoption and improve health care outcomes.

Digital health technologies can improve the efficiency of health care delivery and make health services more accessible. For example, Tasneem et al [22] found that patients receiving palliative care favored the opportunity for video consultations, as they could save time while increasing access, comfort, and safety by avoiding a trip to the hospital.

eHealth resources provide patients with extensive and up-to-date information, access to medical research, and connections to people with similar conditions. Townsend et al [23] mentioned that these factors can significantly impact the physician-patient relationship, transforming patients from passive recipients of care into active partners in their health care journey.

The availability of health information online and through digital technologies enables patients to be more informed about their health conditions and treatment options. This can lead to better-shared decision-making between patients and physicians [17,18,21].

The Application of Theoretical Frameworks and Concepts

In the study by Grünloh et al [17], the use of 2 theoretical frameworks (shared decision-making and patient-centered care) and 1 concept of the physician-patient relationship (paternalism) helped highlight the roles of the medical professional and the patient, as well as the ways in which patients can contribute to the relationship.

The concept of patient empowerment (used in the study by Audrain-Pontevia and Menvielle [15]) helped with understanding how online health communities influence the patient-physician relationship, but it did not capture the complexities of the power dynamics and relational aspects inherent in these interactions.

Győrffy et al [18] focused on the attitudes of digitally engaged physicians toward transforming the physician-patient relationship. The researchers used the theoretical frameworks of patient empowerment and physician-patient collaboration to assess current and ideal physician-patient relationships. Through this, they recognized that the digital age requires physicians to transition from a role of authority to one of guidance.

Jiang [19] examined the relationship between face-to-face and online patient-provider communication through the concepts of patient-centered care, trust, and satisfaction.

Kludacz-Alessandri et al [20] investigated patients' satisfaction with teleconsultations by considering the concepts of respect and dialogue. These concepts are indeed valuable in understanding the impacts of teleconsultations on the physician-patient relationship.

The 3 studies that did not apply any theoretical frameworks had either superficial findings regarding the physician-patient relationship or it was not clear how they reached their conclusions. For example, Balato et al [16] simply quantified the physician-patient relationship through a 10-point scale questionnaire. Tasneem et al [22] asked patients only about how the technology would affect their relationship with their physician in 2 out of 15 questions. Finally, Yan et al [24] seemed to apply the concepts of mutual respect and patient satisfaction in a deep and meaningful way, but we have no way of knowing since they do not specifically state what concepts they used.

The Application of Relational Ethics

In contrast with the above-applied frameworks, the use of a relational ethics framework can help reveal the subtle and complex impacts of digital health technologies on the physician-patient relationship. For example, in the study by Macdonald et al [21], a relational ethics lens was used to examine how eHealth technology contributes to changes in relations between HCPs and patients, evolving toward more collaborative care. By focusing on every day relationships and conversations, the study was able to understand how these technologies incorporated the relational ethics of patient-centered care into practice.

Relational ethics can help address the power dynamics and vulnerabilities that come into play in the physician-patient relationship, especially with the use of digital technologies. In the study by Townsend et al [23], relational ethics was used to understand how technology impacts relational shifts in ethical patient-HCP relationships. They found that technology use could lead to a transition from a traditional hierarchical relationship to a more reciprocal relationship, which could reveal mutual vulnerabilities.

Discussion

Principal Results

The studies included in this review were diverse, with research conducted across various countries and contexts and exploring a range of technologies, such as online communities, messaging, teleconsultations, and patient portals or EHRs. In terms of user groups, they examined the experiences of patients, physicians,

XSL•FO
RenderX

and other HCPs, offering a comprehensive view of the phenomenon from different user perspectives.

Through the review, we identified several consequences of using digital health technologies in health care. The convenience and accessibility offered by these technologies have the potential to transform the health care landscape by enhancing communication, improving efficiency, and empowering patients. However, they also provide challenges and complexities such as increasing inequities in access to health care services due to variations in technology literacy among patients and decreases in patient satisfaction due to limitations in nonverbal communication and the impersonal nature of digital interactions.

The review further highlighted the shift in power dynamics in the physician-patient relationship from a traditional hierarchical model toward a more reciprocal and collaborative model. This shift is facilitated by the empowering potential of digital health technologies and online health communities but may also present challenges for physicians trained in a more paternalistic model of approach to medicine.

In addition to mapping key findings, we analyzed the theoretical frameworks used in the studies to contrast the use of relational ethics with the application of other theoretical frameworks. Of the 10 studies in this review, 2 did not apply any theoretical frameworks to their research. Others used frameworks or concepts such as patient-centered care, patient empowerment, or shared decision-making. Only 2 studies used relational ethics as their framework.

Relational ethics, by its very nature, emphasizes the value of intersubjective qualities such as empathy, trust, respect, and mutual responsibility. These qualities are critical to the practice of effective and compassionate health care, and their importance is highlighted in the context of digital health technologies interjected between the physician and the patient. While theoretical frameworks and concepts, such as patient-centered care and patient empowerment, provided valuable insights into the impacts of digital health technologies, a relational ethics perspective could provide a more comprehensive understanding. For instance, it could explore how trust is built and maintained in online versus face-to-face interactions, how the power dynamics between physicians and patients might shift in an online context, and how empathy and understanding are conveyed through digital mediums. A relational ethics framework could further enrich this understanding by examining how respect and dialogue contribute to a sense of mutual trust and understanding, how they shape the power dynamics in a teleconsultation, and how they can foster a sense of connection and empathy in a digital environment.

Comparison to Prior Work

This review underscores the potential value of using a relational ethics framework in research on the physician-patient relationship in the complex real-life context of digital health technologies. While previous works have acknowledged the importance of relational aspects in health care delivery [3,5], few have explicitly used relational ethics as a framework to examine the nuanced ethical implications of digital health technologies on the physician-patient relationship.

Limitations and Strengths of This Study

Because this review amassed studies from a variety of countries, we acknowledge that the specific cultural, societal, and health care contexts of these regions can have a significant effect on the physician-patient relationship and the use and adoption of digital health technologies. As such, our findings might not be applicable everywhere.

The review only included studies published in English, which could introduce language bias. If there were relevant studies published in other languages, they would have been left out, potentially narrowing the range of perspectives and findings we were able to consider.

While our review touched on a selection of digital health technologies—including online communities, messaging services, teleconsultations, and patient portals or EHRs—many more kinds of technology are being used in health care. As such, our findings are limited to the specific digital health technologies covered in this review.

Scoping reviews, like this one, are primarily intended to provide a broad overview of the existing literature rather than evaluate the strength of evidence or perform a meta-analysis. This means that it is difficult to draw definitive conclusions or provide firm recommendations based on our findings.

While our review encapsulates literature up to the end of 2021, reflecting the state of research at the time of our initial submission, it may not include the latest developments or studies published post-2021.

The quality of the studies included in a review can have a large effect on its findings. Since we did not conduct a critical appraisal of the quality of the studies we included, this could be seen as a limitation.

Conclusions

Overall, this review offers a comprehensive overview of the current state of evidence regarding the impacts of digital health technologies on the physician-patient relationship. It underscores the potential of these technologies to transform health care delivery, while also highlighting the challenges and complexities they introduce. The review emphasizes the need for further research using a relevant ethics framework to provide a deeper understanding of the impact of digital health technologies on the physician-patient relationship. This will be particularly crucial as digital health technologies continue to expand, evolve, and become more integrated into health care delivery.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank librarian Grete Mortensen at the University of Stavanger for her support in developing the search strategy. The study was conducted as part of the research project "Caring Futures: Developing Care Ethics for Technology-Mediated Care Practices" (QUALITECH). QUALITECH was funded by the Research Council of Norway (grant NFR 301827).

Authors' Contributions

DN was the principal author and was involved in all aspects of the scoping review. BHG was a coauthor and screened the abstracts along with DN. ER also coauthored this paper, as well as assessed the research papers for eligibility and reviewed the included research papers along with DN.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 547 KB - ijmr_v13i1e47280_app1.pdf]

References

1. Mesko B. Digital health technologies and well-being in the future. IT Prof 2020 Jan 1;22(1):20-23 [FREE Full text] [doi: 10.1109/mitp.2019.2963121]
2. Wadhwa V. Exponential health technology bringing personal 'check engine lights'. Forbes. 2013. URL: https://www.forbes.com/sites/singularity/ [accessed 2021-03-26]
3. Vogt H, Ulvestad E, Eriksen TE, Getz L. Getting personal: can systems medicine integrate scientific and humanistic conceptions of the patient? J Eval Clin Pract 2014;20(6):942-952 [FREE Full text] [doi: 10.1111/jep.12251] [Medline: 25312489]
4. Rudebeck CE. Relationship based care - how general practice developed and why it is undermined within contemporary healthcare systems. Scand J Prim Health Care 2019;37(3):335-344 [FREE Full text] [doi: 10.1080/02813432.2019.1639909] [Medline: 31299870]
5. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44-56 [FREE Full text] [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
6. Boers SN, Jongsma KR, Lucivero F, Aardoom J, Büchner FL, de Vries M, et al. SERIES: eHealth in primary care. part 2: exploring the ethical implications of its application in primary care practice. Eur J Gen Pract 2020;26(1):26-32 [FREE Full text] [doi: 10.1080/13814788.2019.1678958] [Medline: 31663394]
7. Sandvik H, Hetlevik Ø, Blinkenberg J, Hunskaar S. Continuity in general practice as predictor of mortality, acute hospitalisation, and use of out-of-hours care: a registry-based observational study in Norway. Br J Gen Pract 2022;72(715):e84-e90 [FREE Full text] [doi: 10.3399/BJGP.2021.0340] [Medline: 34607797]
8. Björndell C, Premberg Å. Physicians' experiences of video consultation with patients at a public virtual primary care clinic: a qualitative interview study. Scand J Prim Health Care 2021;39(1):67-76 [FREE Full text] [doi: 10.1080/02813432.2021.1882082] [Medline: 33650941]
9. The Lancet Digital Health. A digital (r)evolution: introducing the lancet digital health. Lancet Digit Health 2019;1(1):e1 [FREE Full text] [doi: 10.1016/S2589-7500(19)30010-X] [Medline: 33323234]
10. Moore J, Engel J, Prentice D. Relational ethics in everyday practice. Can Oncol Nurs J 2014;24(1):31-39. [doi: 10.5737/1181912x2413134] [Medline: 24707705]
11. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
12. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
13. Nassehi D. Physicians' use of digital technologies and its impact on the physician-patient relationship - a scoping review. OSF Registries. 2022. URL: https://osf.io/jkq2f [accessed 2023-03-12]
14. Bock W, Candelon F, Chai S, Choi E, Corwin J, DiGrande S, et al. The mobile revolution: how mobile technologies drive a trillion-dollar impact. The Boston Consulting Group. 2015. URL: https://www.bcg.com/publications/2015/telecommunications-technology-industries-the-mobile-revolution [accessed 2023-11-25]
15. Audrain-Pontevia AF, Menvielle L. Do online health communities enhance patient-physician relationship? an assessment of the impact of social support and patient empowerment. Health Serv Manage Res 2018;31(3):154-162 [FREE Full text] [doi: 10.1177/0951484817748462] [Medline: 29280679]

XSL•FO
RenderX

16. Balato N, Megna M, Di Costanzo L, Balato A, Ayala F. Educational and motivational support service: a pilot study for mobile-phone-based interventions in patients with psoriasis. Br J Dermatol 2013;168(1):201-205 [FREE Full text] [doi: 10.1111/j.1365-2133.2012.11205.x] [Medline: 23240729]
17. Grünloh C, Myreteg G, Cajander Å, Rexhepi H. "Why do they need to check me?" patient participation through ehealth and the doctor-patient relationship: qualitative study. J Med Internet Res 2018;20(1):e11 [FREE Full text] [doi: 10.2196/jmir.8444] [Medline: 29335237]
18. Győrffy Z, Radó N, Mesko B. Digitally engaged physicians about the digital health transition. PLoS One 2020;15(9):e0238658 [FREE Full text] [doi: 10.1371/journal.pone.0238658] [Medline: 32986733]
19. Jiang S. The relationship between face-to-face and online patient-provider communication: examining the moderating roles of patient trust and patient satisfaction. Health Commun 2020;35(3):341-349 [FREE Full text] [doi: 10.1080/10410236.2018.1563030] [Medline: 30821506]
20. Kludacz-Alessandri M, Hawrysz L, Korneta P, Gierszewska G, Pomaranik W, Walczak R. The impact of medical teleconsultations on general practitioner-patient communication during COVID- 19: a case study from Poland. PLoS One 2021;16(7):e0254960 [FREE Full text] [doi: 10.1371/journal.pone.0254960] [Medline: 34270587]
21. Macdonald GG, Townsend AF, Adam P, Li LC, Kerr S, McDonald M, et al. eHealth technologies, multimorbidity, and the office visit: qualitative interview study on the perspectives of physicians and nurses. J Med Internet Res 2018;20(1):e31 [FREE Full text] [doi: 10.2196/jmir.8983] [Medline: 29374004]
22. Tasneem S, Kim A, Bagheri A, Lebret J. Telemedicine video visits for patients receiving palliative care: a qualitative study. Am J Hosp Palliat Care 2019;36(9):789-794 [FREE Full text] [doi: 10.1177/1049909119846843] [Medline: 31064195]
23. Townsend A, Leese J, Adam P, McDonald M, Li LC, Kerr S, et al. eHealth, participatory medicine, and ethical care: a focus group study of patients' and health care providers' use of health-related internet information. J Med Internet Res 2015;17(6):e155 [FREE Full text] [doi: 10.2196/jmir.3792] [Medline: 26099267]
24. Yan M, Tan H, Jia L, Akram U. The antecedents of poor doctor-patient relationship in mobile consultation: a perspective from computer-mediated communication. Int J Environ Res Public Health 2020;17(7):2579 [FREE Full text] [doi: 10.3390/ijerph17072579] [Medline: 32283741]

Abbreviations

EHR: electronic health record
HCP: health care personnel
mHealth: mobile health
PCC: population, concept, and context
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Review

Edited by T de Azevedo Cardoso; submitted 14.03.23; peer-reviewed by Y Wang, E Assing Hvidt, C Beauvais; comments to author 31.05.23; revised version received 26.08.23; accepted 27.02.24; published 15.05.24.

Please cite as:
Nassehi D, Gripsrud BH, Ramvi E
Theoretical Perspectives Underpinning Research on the Physician-Patient Relationship in a Digital Health Practice: Scoping Review
Interact J Med Res 2024;13:e47280
URL: https://www.i-jmr.org/2024/1/e47280
doi:10.2196/47280
PMID:38748465

©Damoun Nassehi, Birgitta Haga Gripsrud, Ellen Ramvi. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 15.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of an mHealth App on Self-Management of Osteoporosis: Prospective Survey Study

Magnus Grønlund Bendtsen[1], MSc; Bodil Marie Thuesen Schönwandt[1], MSc; Mette Rubæk[1,2], MSc; Mette Friberg Hitz[1,3], PhD

[1]Research Unit, Medical Department, Zealand University Hospital, Koege, Denmark

[2]University of Southern Denmark, Odense, Denmark

[3]Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Magnus Grønlund Bendtsen, MSc
Research Unit, Medical Department, Zealand University Hospital
Lykkebækvej 1
Koege, 4600
Denmark
Phone: 45 22893352
Email: mag.bendtsen@gmail.com

Abstract

Background: Mobile health (mHealth) technologies can be used for disease-specific self-management, and these technologies are experiencing rapid growth in the health care industry. They use mobile devices, specifically smartphone apps, to enhance and support medical and public health practices. In chronic disease management, the use of apps in the realm of mHealth holds the potential to improve health outcomes. This is also true for mHealth apps on osteoporosis, but the usage and patients' experiences with these apps are underexplored.

Objective: This prospective survey study aimed to investigate the eHealth literacy of Danish patients with osteoporosis, as well as the usability and acceptability of the app "My Bones."

Methods: Data on patient characteristics, disease knowledge, eHealth literacy, usability, and acceptability were collected using self-administered questionnaires at baseline, 2 months, and 6 months. The following validated questionnaires were used: eHealth Literacy Questionnaire, System Usability Scale, and Service User Technology Acceptability Questionnaire.

Results: Mean scores for eHealth literacy ranged from 2.6 to 3.1, with SD ranging from 0.5 to 0.6 across the 7 domains. The mean (SD) System Usability Scale score was 74.7 (14.4), and the mean (SD) scores for domains 1, 2, and 6 of the Service User Technology Acceptability Questionnaire were 3.4 (1.2), 4.5 (1.1), 4.1 (1.2), respectively.

Conclusions: Danish patients with osteoporosis are both motivated and capable of using digital health services. The app's usability was acceptable, and it has the potential to reduce visits to general practitioner clinics, enhance health outcomes, and serve as a valuable addition to regular health or social care services.

(Interact J Med Res 2024;13:e53995) doi:10.2196/53995

KEYWORDS

eHealth literacy; health literacy; mHealth; mobile health; eHealth; mobile health apps; self-management; osteoporosis; usability; acceptability

Introduction

Background

Osteoporosis ranks as the fourth most prevalent chronic disease globally, carrying substantial negative personal and economic consequences [1]. Osteoporosis affects approximately 40% of women and 17% of men aged 50 years or older. Despite its high prevalence, osteoporosis is significantly underdiagnosed [2]. Although osteoporosis may remain asymptomatic for many individuals, it presents significant risks such as fractures, chronic pain, reduced daily activity, compromised quality of life, and increased mortality [1]. Lack of disease-specific knowledge among patients with osteoporosis is a global issue [3]. It has been shown that improved support for patients to understand

XSL•FO
RenderX

osteoporosis upon diagnosis is required, along with support in self-management of the disease [4,5].

Mobile health (mHealth) technologies can be used for disease-specific self-management, and the technologies are experiencing rapid growth in the health care industry. These technologies use mobile devices, specifically smartphone apps, to enhance and support medical and public health practices [6]. Studies have demonstrated that, particularly in chronic disease management, the use of apps in the realm of mHealth holds the potential to enhance health outcomes [7]. Despite Slomian et al [8] addressing the potential of using mHealth apps to aid patients in self-managing osteoporosis in 2014, the implementation of such apps in the field of osteoporosis remains inadequate. A recently published systematic review and meta-analysis of digital health technologies for long-term self-management of osteoporosis identified 23 relevant apps for osteoporosis self-management and concluded that osteoporosis apps have the potential to support and improve the management of the disease. Furthermore, mHealth osteoporosis apps also appear to be valuable tools for patients and health care professionals. However, most of the identified apps that are currently available in the field of osteoporosis lack clinically validated evidence of their efficacy [9].

From 2015 to 2018, a team of researchers and health care professionals from Odense University Hospital, Denmark, developed and tested an mHealth app for women recently diagnosed with osteoporosis [10]. The development of the app was based on identified needs among patients and health care professionals [11,12]. The app evaluation revealed that patients perceived the app as providing confidence and reassurance, fostering equitable dialog during consultations, and offering readily accessible assistance for self-managing osteoporosis [13]. After the test period, the app was implemented at Odense University Hospital, Denmark. In 2019, the Danish Health Authority decided to support a nationwide rollout of the app, named “My Bones.” A group of health care professionals with expertise in osteoporosis were engaged in the app’s further development and testing. In September 2021, the app “My Bones” was launched as a freely available app on both the App Store and Google Play platforms. Visuals of the app are available in Multimedia Appendix 1.

After the app’s launch, the Research Unit of the Medical Department at Zealand University Hospital, Denmark, initiated an evaluation of the app focusing on usability, acceptability, eHealth literacy, and self-perceived knowledge of osteoporosis. The Consumer Health Information System Adoption Model was used as a theoretical framework, as a basis for the choice of questionnaires used in this study. In this model, eHealth literacy of the user, usefulness, and usability are all determining factors for the adoption of consumer health information systems like the app “My Bones.”

Objective

The aim of this study was to investigate the eHealth literacy of Danish patients with osteoporosis and to assess the usability and acceptability of the app “My Bones.” Additionally, the aim was to assess the current level of disease knowledge within the patient group.

Methods

Study Design

This prospective survey study was conducted at Zealand University Hospital, Denmark, from January 2020 to February 2023. A total of 100 patients were recruited to test the app “My Bones.” At the time of the study, the app was owned and operated by OSAIA Health. The app provided patients with comprehensive information about osteoporosis, its risk factors, available treatment options, the diagnostic procedure for osteoporosis, and the process of a dual-energy x-ray absorptiometry (DXA) scan. Alongside this general information, the app provided guidance on maintaining a bone-healthy lifestyle, dietary requirements including calcium and vitamin D supplements, and recommendations for physical activity. App users also gained access to basic and safe training exercises tailored to their individual fracture risk and functional level, enabling them to engage in a foundational level of physical activity. The app includes 2 interactive modules. One of these was a calcium calculator that assists patients in measuring their daily calcium intake and determining whether they require calcium supplements. The second module was an interactive DXA graph that gives patients the option to plot their *T* score values for the spine and hip from each DXA scan.

Three self-administered questionnaires were developed and internally pretested before distribution at baseline (Q0), at 2 months (Q2), and at 6 months (Q6). The questionnaires are described under the Measures section. The results from the baseline (Q0) and 2-month (Q2) questionnaires are presented here. However, the results from the 6-month (Q6) questionnaire are excluded because of technical issues preventing data analysis.

Sampling and Recruitment

Patients were recruited from the Outpatient Clinic of Endocrinology at Zealand University Hospital and through advertisements on the Danish Osteoporosis Association’s homepage and a Facebook page.

The inclusion criteria include patients diagnosed with osteoporosis based on either the *T* score criterion ($T \leq -2.5$ at the lumbar spine, total hip, or hip neck) or a diagnostic osteoporotic fracture (fragility fracture of a vertebra with >20% compression or a hip fracture). Postmenopausal women or men over 45 years of age who can read and understand Danish and have access to, as well as the ability to use, a smartphone, tablet, or computer.

The exclusion criteria include patients with mental and cognitive conditions impairing their ability to use an app and read and understand questionnaire questions.

Procedures

Patients with osteoporosis interested in participating were provided with both written patient-oriented materials and oral information about the study. Informed consent was obtained from the patients by the data controller. The signed consent forms were stored in a secure location, only accessible by the data controller. After signing the informed consent, participants were asked to indicate their preferred method of receiving the

XSL•FO
RenderX

study questionnaires—either by email, through e-Boks (a digital postbox for communication between companies, public authorities, and private citizens), or in a printed copy. Detailed instructions on how to download the app from the App Store or Google Play were provided. If patients required assistance, the study staff aided in the process.

SurveyXact, a tool for creating electronic questionnaire-based surveys, was used to distribute the questionnaires to the patients and to establish a database. The data analyst gained access to the database only after all data had been collected and anonymized by way of a respondent key in the survey tool and transferred to the database. The survey tool is compliant with the European Union's General Data Protection Regulation.

In cases of nonresponse, electronic reminders were dispatched, followed by telephone follow-ups. If participants still did not complete the questionnaire despite these attempts, they were classified as dropouts.

Measures

Baseline Questionnaire (Q0)

Sociodemographic data include age, sex, education, occupation, time since diagnosis, and knowledge of the disease (self-reported *T* score of the spine and hip, medical treatment, and fracture history).

Data regarding education were initially reported on 6 levels and were subsequently consolidated into two levels for statistical analysis: (1) shorter education and (2) longer education. Longer education was defined as a bachelor's degree or higher. Data concerning occupation were reported across 8 levels and were later condensed into two levels for statistical analysis: (1) currently working and (2) not currently working. The various education and occupation levels can be found in Table 1.

The eHealth Literacy Questionnaire (eHLQ) consists of 35 items representing 7 scales covering the eHealth Literacy Framework dimensions. Each scale has 4 to 6 items with a 4-point response option. Mean scores are calculated from each scale with equal weighting [14].

The International Physical Activity Questionnaire measures the amount of time an individual spends physically active during a normal 7-day week [15].

Table 1. Participant characteristics (N=95).

Characteristics	Participants
Age (years), mean (SD)	66.3 (8.4)
Sex, n (%)	
Male	5 (5)
Female	90 (95)
Shorter education, n (%)	42 (44)
Primary school	8 (8)
Vocational training, high school	21 (22)
Higher education—short	12 (14)
Longer education, n (%)	53 (56)
Higher education—intermediate	41 (43)
Higher education—long	12 (13)
Currently working, n (%)	33 (35)
Self-employed (professional)	3 (3)
Civil servant	23 (24)
Vocational	3 (3)
Unskilled or semiskilled worker	4 (4)
Not currently working, n (%)	62 (65)
Unemployed	3 (3)
Early retirement	17 (18)
Retirement pension	41 (43)
Leave of absence	1 (1)

XSL•FO
RenderX

Two-Month Questionnaire (Q2)

The System Usability Scale (SUS) is a 10-item Likert scale questionnaire that offers a global perspective on subjective assessments of usability. The SUS generates a single number that represents a composite measure of the overall usability of the system under study, in this case, an app [16-18]. The questionnaire is validated in the Danish population.

The Service User Technology Acceptability Questionnaire (SUTAQ) is a questionnaire that consists of 22 items divided into 6 different subscales. For the purpose of this study, 3 of the 6 domains were included in the Q2 questionnaire. The included domains are (1) "Enhanced Care," (2) "Increased Accessibility," and (3) "Satisfaction." The sum of each subscale indicates the degree of average internal agreement with it [19]. The domains "Privacy and Discomfort," "Care Personnel Concerns," and "Kit as a Substitution" were excluded because of their irrelevance resulting from the lack of health monitoring and interaction between patients and health care professionals via the app. The questionnaire is validated in the Danish population.

Statistical Analysis

Statistical analysis was conducted using the SPSS statistical software (version 21; IBM Corp). Descriptive statistics were generated for participants' characteristics and other dependent variables, with calculation of mean values and SD for normally distributed data and median with range for data not normally distributed. Frequency data were calculated for categorical data.

The relation between each of the 7 eHLQ domains and the covariates age, education, and occupation was investigated using a backward stepwise linear regression analysis.

The correlation between the SUS questionnaire score and age was investigated using the Pearson 2-tailed correlation analysis.

All statistical outcomes were examined against a P value of .05 to determine statistical significance. The study was not an intervention study, and no power calculation was performed.

Ethical Considerations

The study was conducted in accordance with the principles of the Declaration of Helsinki. It received approval from the local Data Protection Authority (Reg-152-2020). Although the study did not meet the criteria that necessitate approval from the ethics committee, guidelines for obtaining informed consent were adhered to.

Results

Participant Characteristics

A total of 100 patients signed the informed consent. Of these, 95 responded to the entire or part of the baseline questionnaire (Q0), as shown in the flow diagram in Figure 1.

Of the 95 respondents, 90 were women and 5 were men. The mean (SD) age of the study population was 66.3 (8.4) years, and education levels were reported as shorter by 44% (42/95) of the participants and as longer by 56% (53/95) of the respondents. Regarding occupational status, 35% (33/95) stated a current connection to the labor market, while 65% (62/95) stated that they were either retired, unemployed, or on leave of absence (Table 1).

The mean time since the diagnosis of osteoporosis was 6 years (0-34 years). Among the 95 respondents, 52% (49/95) did not know the T score of the lumbar spine, and 68% (64/95) did not know the T score of the hip. Regarding self-reported osteoporosis status, 22% (21/95) and 4% (4/95) reported severe osteoporosis (T score<–3.0) in the spine region and the hip region, respectively (Table 2). Previous major osteoporotic fractures were reported by 49% (46/95) of the respondents (Table 2), while 67% (64/95) stated that they were currently undergoing medical treatment, the majority in treatment with bisphosphonate (Table 3).

Too much missing data hindered a meaningful analysis of the data from the International Physical Activity Questionnaire, and as a consequence, the results cannot be presented.

XSL•FO
RenderX

Figure 1. Participant flow diagram.

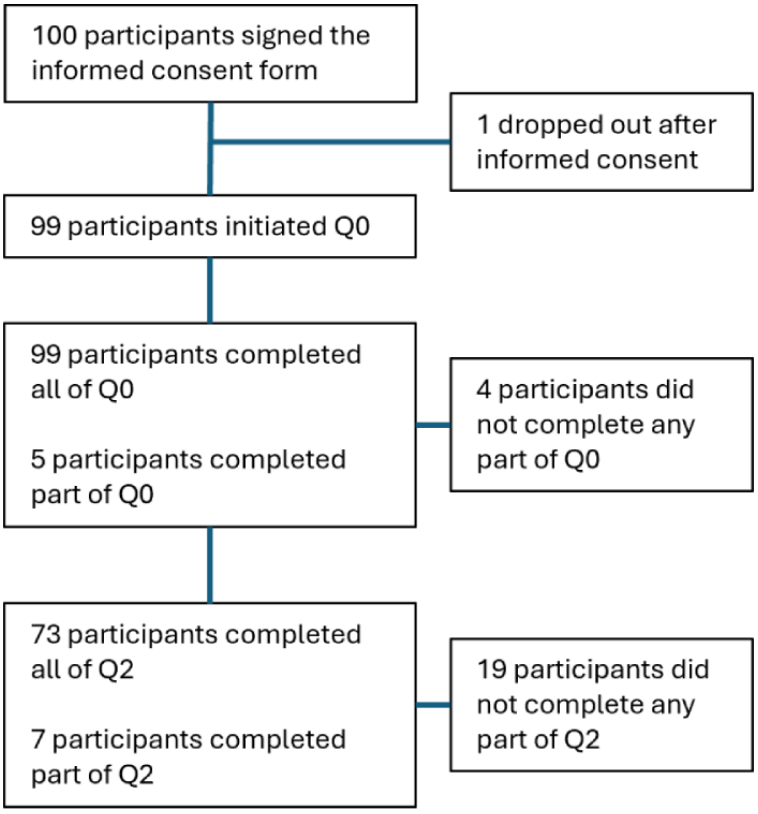


Table 2. *T* scores and fractures (N=95).

Characteristics	Participants
Time since diagnosis(years), mean (range)	6 (0-34)
***T* score lumbar spine, n (%)**	
–2.5 to –3.0	24 (26)
–3.0 to –4.0	14 (15)
≥–4.0	7 (7)
Do not know my *T* score	49 (52)
***T* score hip, n (%)**	
–2.5 to –3.0	26 (28)
–3.0 to –4.0	3 (3)
≥–4.0	1 (1)
Do not know my *T* score	64 (68)
Previous fractures, n (%)	46 (49)
Wrist	21 (22)
Upper arm	3 (3)
Vertebrae	15 (16)
Hip	6 (6)
Other[a]	19 (20)

[a]Other fractures reported: fracture of foot, ankle, heel, toe, finger, elbow, and tibia.

XSL•FO
RenderX

Table 3. Medication (N=93).

Characteristics	Participants
Current treatment plan, n (%)	64 (67)
Oral BP[a]	16 (25)
IV[b] BP	26 (41)
Denosumab	11 (17)
PTH[c] analog	1 (2)
Romosozumab	2 (3)
Unknown to me	6 (9)
Other	2 (3)

[a]BP: bisphosphonate.
[b]IV: intravenous.
[c]PTH: parathyroid hormone.

eHLQ, SUS, and SUTAQ

A total of 90 participants completed the eHLQ section of the Q0 questionnaire. Mean (SD) scores reported ranged from 2.6 (0.5) to 3.1 (0.6) across the 7 domains (Table 4). A definition of each domain can be found in Multimedia Appendix 1.

A total of 79 participants responded to the SUS questionnaire with a mean (SD) score of 74.7 (14.4), while 74 participants responded to the SUTAQ, with mean (SD) scores for domains 1, 2, and 6 being 3.4 (1.2), 4.5 (1.1), and 4.1 (1.2), respectively (Table 5). Answers to individual questions of the SUTAQ are presented in Table 6.

The covariate "occupation" had an effect on eHLQ domain 4 (P=.04) with an adjusted R^2 of 0.036 (Tables 7 and 8). No effects of the other covariates were found for the remaining domains of the eHLQ. No correlation was found between the SUS score and age (P=.37).

Table 4. eHealth Literacy Questionnaire (eHLQ, N=90).

Domain	Value, mean (SD)
eHLQ1: using technology to process health information	2.9 (0.6)
eHLQ2: understanding of health concepts and language	3.1 (0.5)
eHLQ3: ability to actively engage with digital services	3.0 (0.6)
eHLQ4: feel safe and in control	3.1 (0.5)
eHLQ5: motivated to engage with digital services	2.9 (0.6)
eHLQ6: access to digital services that work	2.9 (0.5)
eHLQ7: digital services that suit individual needs	2.6 (0.6)

Table 5. System Usability Score (SUS, n=80) and Service User Technology Acceptability Questionnaire (SUTAQ, n=73) scores at 2 months (Q2).

Questionnaire	Value, mean (SD)
SUS	74.6 (15.3)
SUTAQ (scores from 1 to 6)	
Domain 1	4.1 (1.2)
Domain 2	3.4 (1.1)
Domain 3	4.5 (1.1)

XSL•FO
RenderX

Table 6. Service User Technology Acceptability Questionnaire individual statements (n=73).

Statement	Agreement with statement, n (%)					
	1[a]	2[a]	3[a]	4[a]	5[a]	6[a]
Domain 1						
The kit has allowed me to be less concerned about my health status.	12 (16)	3 (4)	19 (26)	13 (18)	16 (22)	10 (14)
The kit has made me more actively involved in my health.	5 (7)	2 (3)	15 (21)	21 (29)	16 (22)	14 (19)
The kit allows the people looking after me, to better monitor me and my condition.	15 (21)	14 (19)	13 (18)	11 (15)	13 (18)	7 (10)
The kit can be or should be recommended to people in a similar condition to mine.	2 (3)	1 (1)	6 (8)	7 (10)	24 (33)	33 (45)
The kit can certainly be a good addition to my regular health or social care.	4 (5)	1 (1)	7 (10)	17 (23)	20 (27)	24 (33)
Domain 2						
The kit I received has saved me time in that I did not have to visit my GP clinic or other health or social care professional as often.	12 (16)	6 (8)	14 (19)	18 (25)	15 (21)	8 (11)
The kit I received has increased my access to care (health or social care professionals)	12 (16)	7 (10)	19 (26)	22 (30)	10 (14)	3 (4)
The kit I received has helped me to improve my health	6 (8)	7 (10)	13 (18)	25 (34)	16 (22)	6 (8)
The kit has made it easier to get in touch with health and social care professionals.	15 (21)	8 (11)	20 (27)	16 (22)	12 (16)	2 (3)
Domain 6						
The kit has been explained to me sufficiently.	5 (7)	10 (14)	12 (16)	12 (16)	15 (21)	19 (26)
The kit can be trusted to work appropriately.	1 (1)	1 (1)	5 (7)	11 (15)	32 (44)	23 (32)
I am satisfied with the kit I received.	2 (3)	3 (4)	12 (16)	7 (10)	28 (38)	21 (29)

[a]1=strongly disagree; 2=moderately disagree; 3=mildly disagree; 4=mildly agree; 5=moderately agree; 6=strongly agree.

Table 7. Summary of regression analysis showing the relationship between occupation and eHealth Literacy Questionnaire domain 4.

Model	*R*	R^2	Adjusted R^2	SE of the estimate
1	0.216[a]	0.047	0.036	0.5046

[a]Predictors: (Constant). Occupation.

Table 8. ANOVA[a] for occupation's effect on eHealth Literacy Questionnaire domain 4.

Model 1	Sum of squares	*df*	Mean square	*F* (*df*)	*P* value
Regression	1.099	1	1.099	4.315 (1,88)	.04[b]
Residual	22.405	88	0.255	__[c]	—
Total	23.504	89	—	—	—

[a]Dependent variable: eHealth Literacy Questionnaire domain 4.

[b]Predictors: (Constant). Occupation.

[c]Not available.

Discussion

Principal Results

Our analysis revealed eHLQ scores ranging between 2.6 and 3.1 across the 7 domains, with domain 7 "Digital services that suit individual needs" being the lowest-scoring domain and domains 2 "Understanding of health concepts and language" and 4 "Feel safe and in control" being the highest-scoring domains. These findings are comparable to those of another Danish study by Holt et al [20] on eHealth literacy conducted on 246 patients diagnosed with diabetes, other endocrine conditions, or gastrointestinal diseases. Holt et al [20] demonstrated eHLQ scores ranging between 2.6 and 3.1, with

XSL•FO
RenderX

slightly lower scores across the 7 domains. This suggests that our findings are representative of Danish patients with chronic diseases. A Spanish study investigating electronic health literacy in 166 primary care patients, with a median age of 65 (52-78) years, revealed somewhat lower eHLQ scores ranging from 1.7 to 2.8 across the 7 domains [21]. An Australian study of 525 patients at 3 primary care clinics, with a mean age of 56.7, showed eHLQ scores ranging from 2.4 to 3.0 [22]. In comparison with participants from other countries, Danish patients seem to score higher on the eHLQ. These elevated scores among Danish individuals could be attributed to the extensive digitization of public services in Denmark. According to the latest report on digitalization in Denmark (2022), 95% of citizens reported receiving messages from public services through the digital mailbox known as "e-Boks." Additionally, 74% of citizens aged 15 to 89 years use digital public services at least once a week, and 66% of citizens aged 16 to 74 years have booked a doctor's appointment via the web and accessed health information through online sources [23]. On examining the eHLQ scores in this study, high scores are observed in domains 1-6, suggesting that Danish patients with osteoporosis are both motivated and able to use digital services related to health. The lowest-scoring domain, domain 7, indicates the current state of technology regarding digital services and points to a potential for developing more individualized services designed to meet the needs of patients.

The statistical analysis revealed a significant but relatively small effect of the covariate "occupation" on the scores of eHealth literacy domain 4 "Feel safe and in control." This suggests that among the study sample, individuals who are currently working tend to feel safer and more in control when using electronic health services. However, given the small effect size within a limited sample, no definitive conclusions can be drawn. In the field of eHealth literacy, conflicting results have been reported regarding determinants of eHealth literacy scores. A recent systematic review demonstrated an association between eHealth literacy and age, sex, educational level, and family income [24]. Another study conducted by Arcury et al [25] failed to find a connection between eHealth literacy and sociodemographic factors. However, they did discover associations between eHealth literacy and the number of e-devices owned, as well as computer stress [25]. A related finding was reported by Richtering et al [26], who concluded that more time spent on the internet was associated with higher eHealth literacy among 453 patients with moderate to high cardiovascular risk. These findings suggest that factors other than sociodemographic variables play a role in eHealth literacy.

The app "My Bones" aims to assist patients in managing osteoporosis. To achieve this objective, it is crucial for the app to be both usable and well received by its users, prompting an investigation into its usability and acceptability. The analysis of the SUS questionnaire at the 2-month mark resulted in a mean (SD) score of 74.6 (15.3). According to Bangor et al [17], a score of approximately 70 indicates a user-friendliness rating of "Good." Conversely, scores around 35 and 50 correspond to ratings of "Poor" and "OK," respectively. On the other end of the spectrum, scores around 85 and 90 are associated with ratings of "Excellent" and "Best imaginable," respectively [17]. Based on the SUS score, the usability of the "My Bones" app is considered acceptable.

The statistical analysis showed no correlation with age, contradicting the finding by Bangor et al that indicated a slightly negative correlation between the SUS score and age [27]. This discrepancy could be attributed to both the small sample size and the higher average age of this study population.

Analyzing the data from the 3 SUTAQ domains revealed a mean (SD) score of 4.1 (1.2) for domain 1 "Enhanced care," which indicates a general agreement that the app "My Bones" can improve the care patients receive. The mean (SD) score of 4.5 (1.1) for domain 6 "Satisfaction" suggests acceptance and satisfaction with the app. However, the mean (SD) score of 3.4 (1.1) for domain 2 "Increased accessibility" indicates a lower level of agreement with beliefs that the app has facilitated the receipt of care from health care professionals.

In a Danish study involving 68 patients recruited from primary care and an outpatient clinic, the acceptability of a telehealth service was assessed using the SUTAQ. The scores obtained for domains 1, 2, and 6 were 5.0, 4.2, and 5.5, respectively [28]. Another Danish study focused on telehealth services for patients with chronic obstructive pulmonary disease, diabetes, and inflammatory bowel disease, and pregnant women with either diabetes or a need for enhanced care, demonstrated similar results. Scores for domain 1 ranged from 4.4 to 5.1, domain 2 from 3.8 to 4.7, and domain 6 from 5.2 to 5.6 [29]. Both studies exhibited higher scores across all 3 domains, indicating greater satisfaction and acceptance with the telehealth services compared with the "My Bones" app. One possible explanation for the lower scores in this study is that the SUTAQ was originally designed to assess the acceptability of telehealth services involving remote monitoring by health care professionals [19]. However, the app evaluated in our study does not function as a telehealth system, as it lacks monitoring, data collection by patients, and direct interaction with care personnel. Therefore, the applicability of SUTAQ results in describing the acceptability of the app among our study population may be limited. Certain statements within SUTAQ domains 1 and 2 are rendered irrelevant by the nature of the app. For instance, one statement in domain 2 referred to increased access to care through the received kit, which is not applicable because the app does not connect users with health care professionals. Similar issues were observed in other statements within these domains.

Nevertheless, it is worth noting that specific statements regarding the app's usefulness showed promising results. Among the 74 participants who answered the SUTAQ questions, 56% agreed that the app saved them a visit to the general practitioner clinic. In total 65% agreed that the app helped them improve their health, 54% agreed that the app made them less concerned about their health status, and 70% agreed that the app made them more involved in their own health. Furthermore, 84% of participants agreed that the app can be a valuable addition to regular health or social care.

In summary, our findings from both the SUS and SUTAQ questionnaires underscore a high level of satisfaction with the "My Bones" app. While the usefulness of SUTAQ scores may

XSL•FO
RenderX

be questionable, the responses to individual statements indicate that the app holds value for patients with osteoporosis.

Other Findings

We discovered that over half of the 95 patients with osteoporosis had a significant gap in their knowledge about their condition, as they were not aware of their *T* scores. The *T* score reflects the severity of bone loss and risk of fractures. Knowing the *T* score can empower patients to take preventive measures in their daily lives to avoid injuries. Osteoporosis is a silent disease that often exhibits no symptoms; thus, the *T* score becomes a crucial indicator for patients to manage their condition. The fact that half of the patients lack awareness of the *T* score suggests a communication issue within the Danish health care system. Lack of disease-specific knowledge among patients with osteoporosis is a global issue [3], which might be addressed by apps like "My Bones." However, because we did not inquire about the *T* score in our follow-up questionnaires, further research is needed to confirm whether the app can improve *T* score knowledge among patients with osteoporosis.

Perspectives

Based on our findings, it appears that the app "My Bones" has the potential to effectively support the self-management of osteoporosis in this population as a supplement to current health care services. Further investigation should be undertaken to fully assess the app's ability to enhance self-management among patients.

Limitations

This study is subject to several limitations. First, the small population size restricts the generalizability of our findings. Additionally, individuals who are already accustomed to using smartphones and health management apps on a daily basis may have been more likely to participate, potentially introducing a bias toward higher eHealth literacy, acceptability, and satisfaction with the app. Recruitment primarily occurred at a single outpatient clinic and through online advertisements on the patient organization's homepage and a Facebook page, introducing a sampling bias as most patients with osteoporosis in Denmark are typically treated at general practitioner clinics.

The use of questionnaires also presents limitations. Self-reported data on fractures, *T* scores, and medication are susceptible to response bias and may be less reliable than data obtained from registries. Conversely, the questions on *T* score and medication were specifically designed to provide insight into the patients' understanding of their disease.

Despite these limitations, our study provides valuable insights into the eHealth literacy of patients with osteoporosis and the acceptability and usability of the first publicly available osteoporosis management app in Denmark.

Conclusions

We uncovered a high level of eHealth literacy, indicating that Danish patients with osteoporosis possess both the motivation and ability to use mHealth services. The app demonstrated acceptable usability and garnered general satisfaction among users. These findings bolster the viability of an app for self-management of osteoporosis to support Danish patients. The app holds the potential to reduce visits to the general practitioner clinic, enhance health outcomes, and serve as a valuable addition to regular health or social care. However, further investigation is necessary to thoroughly evaluate its effectiveness in improving self-management of osteoporosis.

Acknowledgments

The authors would like to thank Swinburne University for granting permission to use the eHealth Literacy Questionnaire (eHLQ) and Lars Kayser (National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark) for providing assistance in accessing the questionnaire and offering guidance during the research period. Thanks are also extended to Kristian Kidholm (Department of Clinical Research, University of Southern Denmark) and Professor Stanton Newman (University of London) for granting access to the original Service User Technology Acceptability Questionnaire (SUTAQ), and to Julie Hvidt (Health Innovation Centre of Southern Denmark) for providing access to the Danish version of System Usability Scale (SUS). Additionally, the authors would like to express their thanks to OSAIA Health for granting access to the app "My Bones," owned and operated by OSAIA Health. This research was supported by Union Chimique Belge Nordic (20141114).

Conflicts of Interest

MGB has received speaker honorariums from Union Chimique Belge (UCB), Gedeon Richter, and Novo Nordisk. MFH has received grants from Amgen, UCB, and the Ellab-Foundation; received speaker honorariums from UCB and Amgen; and participated in advisory boards with UCB and Novo Nordic. BMTS and MR declare that they have no conflict of interest.

Multimedia Appendix 1
Scale names and construct definitions of the eHealth Literacy Questionnaire (eHLQ) and visuals of the app "My Bones".
[DOCX File , 743 KB - ijmr_v13i1e53995_app1.docx]

References

1. Kanis JA, Norton N, Harvey NC, Jacobson T, Johansson H, Lorentzon M, et al. SCOPE 2021: a new scorecard for osteoporosis in Europe. Arch Osteoporos 2021;16(1):82 [FREE Full text] [doi: 10.1007/s11657-020-00871-9] [Medline: 34080059]

XSL•FO
RenderX

2. Vestergaard P, Rejnmark L, Mosekilde L. Osteoporosis is markedly underdiagnosed: a nationwide study from Denmark. Osteoporos Int 2005;16(2):134-141 [FREE Full text] [doi: 10.1007/s00198-004-1680-8] [Medline: 15197546]
3. des Bordes J, Prasad S, Pratt G, Suarez-Almazor ME, Lopez-Olivo MA. Knowledge, beliefs, and concerns about bone health from a systematic review and metasynthesis of qualitative studies. PLoS One 2020;15(1):e0227765 [FREE Full text] [doi: 10.1371/journal.pone.0227765] [Medline: 31940409]
4. Hansen C, Konradsen H, Abrahamsen B, Pedersen BD. Women's experiences of their osteoporosis diagnosis at the time of diagnosis and 6 months later: a phenomenological hermeneutic study. Int J Qual Stud Health Well-being 2014;9:22438 [FREE Full text] [doi: 10.3402/qhw.v9.22438] [Medline: 24559545]
5. Hansen CA, Abrahamsen B, Konradsen H, Pedersen BD. Women's lived experiences of learning to live with osteoporosis: a longitudinal qualitative study. BMC Womens Health 2017;17(1):17 [FREE Full text] [doi: 10.1186/s12905-017-0377-z] [Medline: 28279157]
6. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016;39(6):981-994 [FREE Full text] [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
7. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
8. Slomian J, Appelboom G, Ethgen O, Reginster JY, Bruyère O. Can new information and communication technologies help in the management of osteoporosis? Womens Health (Lond) 2014;10(3):229-232 [FREE Full text] [doi: 10.2217/whe.14.15] [Medline: 24956289]
9. Alhussein G, Hadjileontiadis L. Digital health technologies for long-term self-management of osteoporosis: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022;10(4):e32557 [FREE Full text] [doi: 10.2196/32557] [Medline: 35451968]
10. Jakobsen PR, Hermann AP, Søndergaard J, Wiil UK, Clemensen J. Development of an mHealth application for women newly diagnosed with osteoporosis without preceding fractures: a participatory design approach. Int J Environ Res Public Health 2018;15(2):330 [FREE Full text] [doi: 10.3390/ijerph15020330] [Medline: 29438343]
11. Jakobsen PR, Hermann AP, Søndergaard J, Wiil UK, Dixon RF, Clemensen J. Left in limbo—experiences and needs among postmenopausal women newly diagnosed with osteoporosis without preceding osteoporotic fractures: a qualitative study. Post Reprod Health 2018;24(1):26-33. [doi: 10.1177/2053369118755189] [Medline: 29392982]
12. Jakobsen PR, Hermann AP, Søendergaard J, Wiil UK, Jensen CM, Clemensen J. The gap between women's needs when diagnosed with asymptomatic osteoporosis and what is provided by the healthcare system: a qualitative study. Chronic Illn 2021;17(1):3-16 [FREE Full text] [doi: 10.1177/1742395318815958] [Medline: 30525980]
13. Jakobsen PR, Hermann AP, Søndergaard J, Wiil UK, Clemensen J. Help at hand: women's experiences of using a mobile health application upon diagnosis of asymptomatic osteoporosis. SAGE Open Med 2018;6:2050312118807617 [FREE Full text] [doi: 10.1177/2050312118807617] [Medline: 30397471]
14. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
15. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395 [FREE Full text] [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
16. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry. London: Taylor & Francis; 1995:189-194.
17. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123.
18. Hvidt JCS, Christensen LF, Sibbersen C, Helweg-Jørgensen S, Hansen JP, Lichtenstein MB. Translation and validation of the system usability scale in a Danish mental health setting using digital technologies in treatment interventions. Int J Hum Comput Interact 2019;36(8):709-716. [doi: 10.1080/10447318.2019.1680922]
19. Hirani SP, Rixon L, Beynon M, Cartwright M, Cleanthous S, Selva A, et al. Quantifying beliefs regarding telehealth: development of the whole systems demonstrator service user technology acceptability questionnaire. J Telemed Telecare 2017;23(4):460-469. [doi: 10.1177/1357633X16649531] [Medline: 27224997]
20. Holt KA, Karnoe A, Overgaard D, Nielsen SE, Kayser L, Røder ME, et al. Differences in the level of electronic health literacy between users and nonusers of digital health services: an exploratory survey of a group of medical outpatients. Interact J Med Res 2019;8(2):e8423 [FREE Full text] [doi: 10.2196/ijmr.8423] [Medline: 30950809]
21. García-García D, Bazán MJA, Pérez-Rivas FJ. Factors influencing eHealth literacy among Spanish primary healthcare users: cross-sectional study. Int J Environ Res Public Health 2022;19(23):15497 [FREE Full text] [doi: 10.3390/ijerph192315497] [Medline: 36497572]

XSL•FO
RenderX

22. Cheng C, Elsworth G, Osborne RH. Validity evidence based on relations to other variables of the eHealth Literacy Questionnaire (eHLQ): Bayesian approach to test for known-groups validity. J Med Internet Res 2021;23(10):e30243 [FREE Full text] [doi: 10.2196/30243] [Medline: 34647897]
23. Tassy A, Törnfelt C. It-anvendelse i befolkningen 2022. Danmarks Statistik. 2023. URL: https://www.dst.dk/da/Statistik/nyheder-analyser-publ/Publikationer/VisPub?cid=44692 [accessed 2023-08-25]
24. Chang SJ, Jang SJ, Lee H, Kim H. Building on evidence to improve eHealth literacy in older adults: a systematic review. Comput Inform Nurs 2021;39(5):241-247. [doi: 10.1097/CIN.0000000000000674] [Medline: 32925185]
25. Arcury TA, Sandberg JC, Melius KP, Quandt SA, Leng X, Latulipe C, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020;39(2):141-150 [FREE Full text] [doi: 10.1177/0733464818807468] [Medline: 30353776]
26. Richtering SS, Hyun K, Neubeck L, Coorey G, Chalmers J, Usherwood T, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017;4(1):e4 [FREE Full text] [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
27. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
28. Sandvei M, Kynde I, Lee A. Fronløber projektet: evaluering af tilfredshed hos borgere og organisatoriske erfaringer. Syddansk Universitet. 2013. URL: https://www.sdu.dk/~/media/Files/Om_SDU/Centre/CAST/PDF_filer/Frontloeber.pdf [accessed 2023-08-25]
29. Lee A, Sandvei M, Hosbond T, Petersen J, Christiansen KR. Klinisk Integreret Hjemmemonitorering (KIH) Slutrapportering til Fonden for Velfaerdsteknologi. Digitaliseringsstyrelsen. 2015. URL: https://findresearcher.sdu.dk/ws/files/115999662/KIH_Slutrapport.pdf [accessed 2024-02-24] [doi: 10.13140/RG.2.1.5081.1363]

Abbreviations

DXA: dual-energy x-ray absorptiometry
eHLF: eHealth Literacy Framework
eHLQ: eHealth Literacy Questionnaire
IPAQ: International Physical Activity Questionnaire
mHealth: mobile health
SUS: System Usability Scale
SUTAQ: Service User Technology Acceptability Questionnaire

Edited by M Yousufuddin; submitted 26.10.23; peer-reviewed by P Nohl-Deryk, C Jacob; comments to author 24.01.24; revised version received 05.02.24; accepted 10.02.24; published 01.04.24.

Please cite as:
Bendtsen MG, Schönwandt BMT, Rubæk M, Hitz MF
Evaluation of an mHealth App on Self-Management of Osteoporosis: Prospective Survey Study
Interact J Med Res 2024;13:e53995
URL: https://www.i-jmr.org/2024/1/e53995
doi:10.2196/53995
PMID:38557362

©Magnus Grønlund Bendtsen, Bodil Marie Thuesen Schönwandt, Mette Rubæk, Mette Friberg Hitz. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 01.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Intramural Health Care Through Video Consultations and the Need for Referrals and Hospital Admissions: Retrospective Quantitative Subanalysis of an Evaluation Study

Katharina Schmalstieg-Bahr[1,2], MD; Miriam Giovanna Colombo[3], MPH; Roland Koch[3], MD; Joachim Szecsenyi[4], MD; Friedrich Völker[2]; Eva Elisabeth Blozik[2,5], MD; Martin Scherer[1,2], MD

[1]Department of General Practice and Primary Care, University Medical Center Eppendorf, Hamburg, Germany
[2]A+ Videoclinic GmbH, Gräfelfing, Germany
[3]Institute for General Practice and Interprofessional Care, University Hospital Tübingen, Tübingen, Germany
[4]Department of General Practice and Health Services Research, University Hospital Heidelberg, Heidelberg, Germany
[5]Institute of Primary Care, University Hospital Zurich, Zurich, Switzerland

Corresponding Author:
Katharina Schmalstieg-Bahr, MD
Department of General Practice and Primary Care
University Medical Center Eppendorf
Martinistrasse 52
Bldg. W37, 5th Fl
Hamburg, 20246
Germany
Phone: 49 40 7410 52400
Fax: 49 40 7410 4022
Email: k.schmalstieg-bahr@uke.de

Abstract

Background: In comparison to the general population, prison inmates are at a higher risk for drug abuse and psychiatric, as well as infectious, diseases. Although intramural health care has to be equivalent to extramural services, prison inmates have less access to primary and secondary care. Furthermore, not every prison is constantly staffed with a physician. Since transportation to the nearest extramural medical facility is often resource-intensive, video consultations may offer cost-effective health care for prison inmates.

Objective: This study aims to quantify the need for referrals to secondary care services and hospital admissions when video consultations with family physicians and psychiatrists are offered in prison.

Methods: In 5 German prisons, a mixed methods evaluation study was conducted to assess feasibility, acceptance, and reasons for conducting video consultations with family physicians and psychiatrists. This analysis uses quantitative data from these consultations (June 2018 to February 2019) in addition to data from a sixth prison added in January 2019 focusing on referral and admission rates, as well as reasons for encounters.

Results: At the initiation of the project, 2499 prisoners were detained in the 6 prisons. A total of 435 video consultations were conducted by 12 physicians (3 female and 7 male family physicians, and 2 male psychiatrists during the study period). The majority were scheduled consultations (341/435, 78%). In 68% (n=294) of all encounters, the patient was asked to consult a physician again if symptoms persisted or got worse. In 26% (n=115), a follow-up appointment with either the video consultant or prison physician was scheduled. A referral to other specialties, most often psychiatry, was necessary in 4% (n=17) of the cases. Only in 2% (n=8) of the consultations, a hospital admission was needed. Usually, hospital admissions were the result of unscheduled consultations, and the videoconferencing system was the method of communication in 88% (n=7) of these cases, while 12% (n=1) were carried out over the phone. Reasons for admissions were severe abdominal pain, hypotension, unstable angina or suspected myocardial infarction, or a suspected schizophrenic episode.

Conclusions: Most scheduled and unscheduled consultations did not require subsequent patient transport to external health care providers. Using telemedicine services allowed a prompt patient-physician encounter with the possibility to refer patients to other specialties or to admit them to a hospital if necessary.

XSL•FO
RenderX

(Interact J Med Res 2024;13:e44906) doi:10.2196/44906

KEYWORDS

intramural health care; prison; telemedicine; primary care; family medicine; referral; hospital admission; admission rate; intramural; penal; video consult; e-consult; remote care; virtual care; health care delivery; service delivery; health care system

Introduction

In 2022 approximately 56,000 people were incarcerated in a total of 172 German prisons. While 19,339 (34%) out of 56,557 inmates face imprisonment of over 1 year to 4 years, 5340 (9%) out of 56,557 inmates serve prison sentences of 5 years or longer [1,2]. By law, these prisoners have the same right to access health care as patients with statutory health insurance outside the correctional system. However, caring for incarcerated patients is challenging. They are more likely to experience alcohol or drug abuse, mental illnesses, or communicable diseases, such as hepatitis C or HIV infections [3,4], and not every prison is constantly staffed with a physician on-site. Specialized (secondary care) services are less available to inmates since a resource-intensive transport to the next extramural facility has to be organized. Telemedicine offers the potential to close this gap and improve intramural health care [5]. In this context, the collective term telemedicine describes heterogeneous concepts that aim at providing medical diagnostics, therapy, and rehabilitation despite physical distance or time lag [6]. In many countries, it has been used to facilitate or enhance intramural care [7-10], but it has not been implemented on a broad scale in the German correctional system yet.

This study aims to assess the need and the reasons for referrals to secondary care services, as well as hospital admissions, when scheduled or unscheduled telemedicine consultations with family physicians and psychiatrists are offered in prison.

Methods

Overview

The pilot project to establish video consultations in German prisons was initiated by the Ministry of Justice Baden-Württemberg in cooperation with A+ Videoclinic (VC), a provider of telemedical services, and initially, involved 5 German prisons. Between June and December 2018, the pilot project was evaluated in terms of feasibility and acceptance of the video consultations, as well as consultation reasons by conducting a mixed methods evaluation study. Quantitative and qualitative data were collected through site visits in the prisons, questionnaires, semistructured interviews, and consultation documentation. Further details are reported elsewhere [11]. This analysis is a retrospective subanalysis using a quantitative VC data set that was generated during the evaluation study period containing the information depicted in Table 1, as well as additional data from a sixth prison, which was not part of the initial pilot project.

All 6 prisons were located in the federal state of Baden-Württemberg, Germany. Inmates were male and female (adults and adolescents) and 18 years of age and older. Participation in the pilot study was voluntary. Patients could choose either a video consultation or regular medical care. If the patient opted for a video consultation, he or she had to sign an informed consent form.

Consultations that were conducted between June 2018 and February 2019 were analyzed in this study. The videoconferencing system (VCS) was the preferred method of communication. The phone was used in case of any technical problems with the VCS. Consultations were carried out by a team of 12 physicians (3 female and 7 male family physicians, and 2 male psychiatrists) employed by the telemedicine provider. Scheduled encounters were conducted during fixed weekly timeslots—either with VC-family physicians or a VC psychiatrist. Outside of these consultation hours, prison nursing staff could reach the on-call VC-family physician 24 hours 7 days per week. These patient-physician-contacts outside of consultation hours were counted as unscheduled consultations. Depending on the time of contact, there was not always a trained nurse present in prison. If the on-call family physician required help regarding a psychiatric problem, he or she could contact a VC psychiatrist.

The VC provided the telemedical infrastructure for VC physicians and prisons. Physicians documented the consultations electronically with a VC laptop using a virtual private network to access the VC software called Videoclinic Portal (Videoclinic) developed by the Videoclinic. The participating prisons were also equipped with VC laptops to receive the documentation. No remote medical devices, such as stethoscopes, that would have allowed the physician to directly auscultate a patient were used during the pilot study. Either prison nursing staff or correctional officers were present during each encounter. Prison nursing staff could obtain the patient's vital signs (pulse, blood pressure, temperature, and oxygen saturation) and write an electrocardiogram if necessary and available. A total of 4 out of 6 prisons had an electrocardiogram on-site. Further details on the medical equipment available in the 5 prisons that were part of the initial pilot project can be found elsewhere [11].

For this study, an anonymized data set was exported from the VC software containing the variables depicted in Table 1. The data analysis for this study comprised descriptive statistical methods and was performed using Microsoft Excel 2016.

XSL•FO
RenderX

Table 1. Data provided by A+ Videoclinic.

Variable	Description
Date of consultation	• Date and time of the encounter
Prison	• Name of the prison
Physician	• Treating physician (pseudonymized number D1-12)
Medical specialty	• Family medicine • Psychiatry
Assessment	• Current assessment (free text)
Diagnosis	• Current diagnosis (free text and International Classification of Diseases, 10th revision code)
Plan	• Recommended treatment (free text)
Medication	• Current medication if any prescribed (free text)
Type of consultation	• Scheduled • Unscheduled
Method of communication	• Video • Phone
Interpreter	• Foreign language if interpreter was used
Follow-up	• No further treatment • Follow-up if symptoms persist or worsen • Planned follow-up appointment • Referral • Hospital admission

Ethics Approval

Ethics approval for the evaluation of the pilot project was obtained from the ethics committee of the Eberhard-Karls-University Tübingen (728/2018BO1) and the ethics committee of the State Medical Association Baden-Württemberg (F-2018-054) [11].

Results

Baseline Characteristics of the Participating Prisons

The ratio of prisoners to medical staff differed between the 6 prisons. Table 2 shows the sex and number of inmates of each prison at the initiation of the pilot study, as well as the number of physicians and nurses. Except for P4, prison staff comprised at least 1 physician. In P4, 2 external physicians offered scheduled consultation hours twice a week. Scheduled consultations with external physicians from other specialties (not family medicine) varied greatly. In most prisons, regular visits from a dentist were established. P2 offered appointments with a gynecologist and psychiatrist. In addition to dental care and the care provided by the prison physician, P6 offered consultations with a dermatologist and surgeon. Furthermore, nursing staff from P6 was able to contact an external psychiatrist if needed.

Table 2. Occupancy and staff per prison.

Prison	Characteristics	Number of prisoners (n=2499), n (%)	Number of physicians (n=12), n (%)	Number of nurses (n=69), n (%)
P1	Male adults	350 (14)	2 (17)	9 (13)
P2	Female adults or adolescents	350 (14)	1 (8)	6 (9)
P3	Male adolescents	395 (16)	4 (33)	9 (13)
P4	Male adults	52 (2)	0 (0)	3 (4)
P5	Male adults	772 (31)	4 (33)	22 (32)
P6	Male adults	580 (23)	1 (8)	20 (29)

XSL•FO
RenderX

Referrals Following Video Consultations

From June 2018 to February 2019, VC physicians conducted 435 consultations. Out of that, 78% (n=341) of all consultations were scheduled, and the remainder were unscheduled (94/435, 22%). In 68% (n=294), patients were asked to consult a physician again if symptoms persisted or got worse. In 26% (n=115) a follow-up appointment with the video consultant or prison physician was scheduled. A referral to other specialties was necessary in 4% (n=17). The VCS was the method of communication in all 17 cases during which a referral was necessary. A total of 3 of these encounters were unscheduled. Patient transport to an extramural facility was indicated in none of these cases since the secondary care physician could either be contacted via video consultation, was able to come to the prison, or was employed by the prison but off duty during the initial encounter (Table 3).

Due to the small number of necessary referrals, it is hard to rank specialties based on the frequency of necessary referrals. The data showed mixed results with psychiatry as the most common specialty referred to (3/17, 18%), followed by urology (2/17, 12%) and proctology (2/17, 12%). In 2 cases, patients were referred to an external psychiatrist; and 1 patient had a follow-up appointment with a prison psychiatrist. Referral diagnoses also varied between the cases (Table 3). A total of 5 out of the 12 VC physicians (4 family physicians and 1 psychiatrist) conducted the encounters resulting in a referral, however, the number of overall encounters per physician varied greatly ranging from 2 to 143. An interpreter was not needed in any of the 17 consultations that resulted in a referral.

Table 3. Overview of the referrals to other specialties (n=17).

Type of consultation	Medical specialty (physician number)	Diagnosis	Medical specialty referred to
Unscheduled	FM[a] (D7)	Drug abuse, opiate withdrawal	Psychiatry (coming to site)
Scheduled	PSY[b] (D8)	Adjustment disorder	Child and adolescent psychiatry
Unscheduled	FM (D9)	Acneiform dermatitis	Dermatology (video consultation)
Scheduled	FM (D5)	Nausea, suspected adverse reaction, suspected melena	Rheumatology
Scheduled	FM (D10)	(missing data)	(missing data)
Scheduled	FM (D10)	Presbyopia	Ophthalmology
Scheduled	FM (D10)	(missing data)	(missing data)
Scheduled	FM (D5)	Ankle sprain R	Physician licensed to treat work accidents (German: *Durchgangsarzt*)
Scheduled	FM (D10)	(missing data)	(missing data)
Scheduled	FM (D5)	Adverse drug reaction	Urology
Scheduled	FM (D10)	Toothache	Dentistry
Unscheduled	FM (D9)	Prison admission exam	Psychiatry or drug counseling (prison psychiatrist)
Scheduled	PSY (D8)	Chronic arm pain after car accident, insomnia[c]	Neurology
Scheduled	PSY (D8)	Liver cirrhosis, suspected ascites[c]	Gastroenterology
Scheduled	FM (D5)	Priapism	Urology
Scheduled	FM (D5)	First degree hemorrhoids	Proctology[d]
Scheduled	FM (D5)	First degree hemorrhoids	Proctology[d]

[a]FM: family medicine.

[b]PSY: psychiatry.

[c]No diagnosis coded—information taken from medical history.

[d]Referral only necessary if symptoms persist despite ordered treatment.

Hospital Admissions Following Video Consultations

In 2% (n=8) of the cases, a hospital admission was required. These cases were independent of the 17 cases that required a referral to another specialty in an ambulatory care setting. Hospital admission was usually the result of an unscheduled consultation (7/8, 88%), and the VCS was used in 88% (n=7). Gastrointestinal problems or pain were the most common reason for admission (4/8, 50%) and 6 VC physicians (5 family physicians and 1 psychiatrist) conducted the encounters (Table 4). An interpreter was not needed in any of these consultations.

XSL•FO
RenderX

Table 4. Overview of hospital admissions (n=8).

Type of consultation	Method of communication	Medical specialty (physician number)	Diagnosis
Unscheduled	Video	FM[a] (D1)	Suspected tuberculosis, hypotension
Unscheduled	Video	FM (D7)	Unstable angina
Unscheduled	Video	FM (D1)	Severe hypotension suspected due to antipsychotics
Unscheduled	Video	FM (D11)	Gastrointestinal hemorrhage
Unscheduled	Phone	FM (D2)	Abdominal pain, kidney stones
Unscheduled	Video	FM (D2)	Gallbladder disease
Unscheduled	Video	FM (D4)	Abdominal pain (differential diagnosis: myocardial infarction)
Scheduled	Video	PSY[b] (D8)	Suspected schizophrenia

[a]FM: family medicine.
[b]PSY: psychiatry.

Hospital Admission and Referrals per Prison and Physician

Almost half of the video consultations (194/435, 45%) were conducted in prison P2 (Table 5). It was the only prison with female inmates (adults and adolescents). At the time of the initial visit to the participating prisons in the framework of the pilot project (September 2018), the number of inmates was similar to P1 (male adults) and P3 (male adolescents; Table 2). A referral following a video consultation was needed in all prisons apart from P1 and P6. P3 had the highest number of referrals, but less than 10% of video consultations were conducted there. A hospital admission following a video consultation was required in all prisons except P3 and P4.

VC physicians D1 and D8 conducted more than 60% of all encounters (Table 6). However, referrals and admissions were rather scattered among physicians—except for D5, who initiated 6 (35%) of the referrals.

Table 5. Video consultations, referrals, and hospital admissions per prison.

Prison	Video consultations per prison, (n=435), n (%)	Referrals per prison, (n=17), n (%)	Hospital admissions per prison, (n=8), n (%)
P1	41 (9)	0 (0)	2 (25)
P2	194 (45)	5 (29)	1 (13)
P3	41 (9)	6 (35)	0 (0)
P4	75 (17)	4 (24)	0 (0)
P5	68 (16)	2 (12)	4 (50)
P6	16 (4)	0 (0)	1 (13)

Table 6. Video consultations, referrals, and hospital admissions per physician.

Physician	Video consultations per physician, (n=435), n (%)	Referrals per physician, (n=17), n (%)	Hospital admissions, (n=8), n (%)
D1	143 (33)	1 (6)	2 (25)
D2	9 (2)	0 (0)	2 (25)
D3	5 (2)	0 (0)	0 (0)
D4	18 (4)	0 (0)	1 (13)
D5	14 (3)	6 (35)	0 (0)
D6	8 (2)	0 (0)	0 (0)
D7	2 (1)	1 (6)	1 (13)
D8	124 (29)	3 (17)	1 (13)
D9	9 (2)	2 (12)	0 (0)
D10	97 (22)	4 (24)	0 (0)
D11	3 (1)	0 (0)	1 (13)
D12	3 (1)	0 (0)	0 (0)

XSL•FO
RenderX

Discussion

Principal Findings

Most scheduled and unscheduled video consultations did not require a subsequent patient transport to an extramural health care provider or facility. A referral was only needed in 4% (n=17) of the cases and hospital admission was only required in 2% (n=2) of 435 cases.

Comparison to Prior Work

To our knowledge, this is the first study that focused on referral and hospital admission rates in intramural health care when video consultations with family physicians and psychiatrists are offered. Another study evaluated telemedicine consultations with an emergency room (ER) physician and found that 36% of the patients from a correctional facility required transport to the ER after a video encounter [12]. This rate was higher than the referral and hospital admission rate of this study. If prison physicians were given the possibility to refer patients to a telemedicine satellite facility for subspecialty consults, outpatient visits would increase by 40% in the 2 years after implementation. In contrast, ER visits decreased. The authors interpreted the effect as better access to care [13]. The same authors examined telemedicine programs in the juvenile justice system by measuring, for example, outpatient costs, ER costs, and transportation costs [14], which are parameters that are associated with referrals and admissions. Other studies focused on the treatment of single diseases, such as hepatitis C or diabetes, when telemedicine consultations were offered in prison [15,16].

Other studies have shown that outside of the correctional system home telemonitoring could reduce hospital admission rates in people with chronic obstructive pulmonary disease [17,18] and congestive heart failure [19,20]. Telemedicine consultations with an emergency physician led to reduced transfers from skilled nursing facilities to ERs and subsequently to lower hospital admission rates [21], and Rosner et al [22] showed that using telemedicine reduced readmissions after hip and knee arthroplasties. In contrast, a meta-analysis that evaluated different remote management strategies for patients with inflammatory bowel disease demonstrated a reduction in physician visits but no significant effect on relapse or hospital admission rates [23].

Regarding a reduction of referral rates, previous studies outside of the correctional system generally focused on electronic solutions for primary care providers to contact other specialists and have demonstrated mixed results. Liddy et al [24] reported that implementing an electronic consultation service in Canada that allowed family physicians to communicate with secondary care providers regarding a patient's care reduced referral rates between 36% and 53%, which was in line with prior findings of that group [25]. However, their randomized controlled trial with 2 study arms (primary care physicians with and without access to the electronic consultation service), showed a significant referral reduction in both arms [26]. Furthermore, according to an online survey among primary care physicians, a phone consultation with an HIV specialist reduced the perceived need to refer the patient to a secondary care provider, although the authors acknowledged that actual referral rates had not been studied [27]. In another study, web-based consultation between primary care physicians and nephrologists did not affect referral rates [28].

In comparison to referral rates of (extramural) family medicine practices without the use of video consultations, a referral rate of 3.8% found in this study is at the lower end of the expected spectrum. Generally, referral rates to secondary care providers vary between practices. A recent study showed an average monthly rate of 20.3% with a range of 0.4%-67.1%. Outside of the correctional system, mental health services were the 10th most common specialty for referrals [29]. Older studies showed mean rates of 1.4% to 37% [30-33]. Variance can also be found regarding hospital admission rates of (extramural) family medicine practices. Mean rates of approximately 50-53 admissions per 1000 patients per year have been reported [34,35], which can only indirectly be compared with the data of this analysis (8 admissions per 435 encounters within 9 months).

International studies showed that telemedicine was able to deliver high-quality and timely primary care for adult and adolescent prison inmates [13,36], reduce costs [37], and facilitate mental health services [38,39]. Using telemedicine to improve access, cost, and quality of secondary care, for example, in the fields of ophthalmology [40,41], cardiology [42], and dentistry [10,43] has also been described before and is still under evaluation [44]. Especially, countries with remote regions outside of metropolitan areas, such as the United States or Australia, have reported the use of telemedicine in their correctional systems [45], but also more densely populated countries deemed the use as beneficial [10,46].

Strengths and Limitations

The pilot project was the first broad implementation of telemedicine in the German correctional system. But there are some limitations: first, a possible selection bias has to be considered: prison staff may have chosen to directly call an ambulance or organize a transport to an extramural facility instead of using video consultations for more severe medical cases. If only patients with less severe diseases were seen by VC physicians, the likelihood of required hospital admission was consequently decreased. Similarly, prison staff may have directly scheduled an appointment with a secondary care provider visiting the prison, which also might have reduced the likelihood of referrals. However, it was crucial that not only patients were given a choice to talk to a VC physician or receive usual care, but that prison staff was also free to choose whether or not to contact the VC—especially since 2 authors were VC founding members. Neither prison staff nor patients received incentives for contacting the VC. Second, there was no control group in this study, and therefore, referral and hospital admission rates cannot be compared with regular intramural care. Despite the low rates when using video consultations, a referral or an admission might not have been necessary if an in-person consultation had been done at that time. Third, only data from the VC portal was considered for this analysis. Fourth, it is unknown how many patients refused a video consultation. Fifth, at the beginning of the data collection, some VC physicians did

XSL•FO
RenderX

not complete the entire documentation template, therefore, some data were missing (the type of follow-up was not specified in 1 case, and in 2 cases, no diagnosis or reason for the referral was listed). The documentation improved throughout the project as physicians became more familiar with or received more training regarding the use of the VC portal. Finally, all consultations were carried out by a rather small group of physicians, and therefore, the influence of individual experience and working style cannot be excluded.

Implication for Practice and Research

The results show that a referral or admission was only required in a few cases after video consultations were offered in prison. Compared with extramural family medicine practices, the referral rates found in this study were at the lower end of the expected spectrum and support that the implementation of telemedicine in intramural systems on a larger scale should not be postponed or revised due to concerns of high referral and admission rates. Further research, including controlled studies, is needed to explore whether institutional factors contribute to the effectiveness and safety, such as training of staff, use of remote medical devices, and acceptance of telemedicine by inmates, prison (nursing) staff, and physicians working on-site—especially in the light of the fact that the data were generated prior to the COVID-19 pandemic and video consultations became much more common since then.

Acknowledgments

The authors thank all participating patients, prison staff, and VC physicians. No machine learning, artificial intelligence, language models, or similar technology has been used for the project or to draft the manuscript. The initial project was funded by the Ministry of Justice Baden-Württenberg. No funding was received for this analysis. The authors acknowledge the financial support from the Open Access Publication Fund of the University Medical Center Hamburg-Eppendorf.

Data Availability

The data sets generated or analyzed during this study are not publicly available due to privacy protection regulations but are available from the corresponding author on reasonable request.

Authors' Contributions

MS planned and acquired the project between the Ministry of Justice Baden-Württenberg and the A+ Videoclinic. MS and EEB monitored the telemedical processes and the quality management protocols. RK and MGC collected the data for the evaluation study of the pilot project. KS-B analyzed the data of this study. All authors helped to draft the manuscript and approved the final version.

Conflicts of Interest

MS is a founding member of the A+ Videoclinic and the head of the Medicine and Quality Committee. FV is the current chief executive officer. EEB is the acting medical head. JS represented the project to the Medical Association of Baden Württemberg as the medical leader. KS-B is a part-time employee of the A+ Videoclinic. RK and MGC declare no conflict of interest.

References

1. Number of prisoners and detainees in the prisons in Germany by type of detention in June 2022. Statista. URL: https://de.statista.com/statistik/daten/studie/158317/umfrage/gefangene-und-verwahrte-in-deutschland-nach-art-des-vollzugs/ [accessed 2024-06-03]
2. Number of prisoners in the prisons in Germany after the expected execution on 31. March 2022. Statista. URL: https://de.statista.com/statistik/daten/studie/75829/umfrage/strafgefangene-nach-vollzugsdauer-in-deutschland/#statisticContainer [accessed 2024-06-03]
3. Opitz-Welke A, Lehmann M, Seidel P, Konrad N. Medicine in the penal system. Dtsch Arztebl Int 2018;115(48):808-814 [FREE Full text] [doi: 10.3238/arztebl.2018.0808] [Medline: 30642429]
4. Davis DM, Bello JK, Rottnek F. Care of incarcerated patients. Am Fam Physician 2018;98(10):577-583 [FREE Full text] [Medline: 30365288]
5. Weinstein RS, Lopez AM, Joseph BA, Erps KA, Holcomb M, Barker GP, et al. Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers. Am J Med 2014;127(3):183-187. [doi: 10.1016/j.amjmed.2013.09.032] [Medline: 24384059]
6. Dissemination of telemedical methods and procedures. Bundesärztkammer - Verbreitung telemedizinischer Methoden und Verfahren. URL: https://www.bundesaerztekammer.de/themen/aerzte/digitalisierung/telemedizin-fernbehandlung/verbreitung-methoden-und-verfahren [accessed 2024-06-06]
7. Mateo M, Álvarez R, Cobo C, Pallas J, López AM, Gaite L. Telemedicine: contributions, difficulties and key factors for implementation in the prison setting. Rev Esp Sanid Penit 2019;21(2):95-105 [FREE Full text] [Medline: 31642860]
8. Anogianakis G, Ilonidis G, Milliaras S, Anogeianaki A, Vlachakis-Milliaras E. Developing prison telemedicine systems: the Greek experience. J Telemed Telecare 2003;9(Suppl 2):S4-S7. [doi: 10.1258/135763303322596101] [Medline: 14728746]

XSL•FO
RenderX

9. Edge C, Black G, King E, George J, Patel S, Hayward A. Improving care quality with prison telemedicine: the effects of context and multiplicity on successful implementation and use. J Telemed Telecare 2021;27(6):325-342. [doi: 10.1177/1357633X19869131] [Medline: 31640460]
10. Giraudeau N, Inquimbert C, Delafoy R, Tramini P, Valcarcel J, Meroueh F. Teledentistry, new oral care tool for prisoners. Int J Prison Health 2017;13(2):124-134. [doi: 10.1108/IJPH-04-2016-0011] [Medline: 28581375]
11. Colombo MG, Koch R, Joos S. Akzeptanz von videokonsultationen in justizvollzugsanstalten aus der perspektive von patient/-innen—ergebnisse der gemischt-methodischen evaluation eines modellprojekts aus baden-württemberg [Acceptance of video consultations in correctional facilities from the patients' perspective—results from the mixed-methods evaluation of a pilot project in baden-württemberg]. Psychiatr Prax 2022;49(2):80-88. [doi: 10.1055/a-1400-2388] [Medline: 33773499]
12. Ellis DG, Mayrose J, Jehle DV, Moscati RM, Pierluisi GJ. A telemedicine model for emergency care in a short-term correctional facility. Telemed J E Health 2001;7(2):87-92. [doi: 10.1089/153056201750279584] [Medline: 11421081]
13. Fox KC, Somes GW, Waters TM. Timeliness and access to healthcare services via telemedicine for adolescents in state correctional facilities. J Adolesc Health 2007;41(2):161-167. [doi: 10.1016/j.jadohealth.2007.05.001] [Medline: 17659220]
14. Fox KC, Somes G, Waters TM. The use of telemedicine technology and restructured referral patterns to reduce health care costs in a Juvenile Justice System. J Correct Health Care 2006;12(3):214-221. [doi: 10.1177/1078345806292667]
15. Kassar K, Roe C, Desimone M. Use of telemedicine for management of diabetes in correctional facilities. Telemed J E Health 2017;23(1):55-59. [doi: 10.1089/tmj.2016.0036] [Medline: 27223479]
16. Galán GJ, Alia CA, González MV, Berriguete R, González FF, Rodríguez CMF, et al. The contribution of telemedicine to hepatitis C elimination in a correctional facility. Rev Esp Enferm Dig 2019;111(7):550-555 [FREE Full text] [doi: 10.17235/reed.2019.6152/2018] [Medline: 31215210]
17. Holland A. Telehealth reduces hospital admission rates in patients with COPD. J Physiother 2013;59(2):129 [FREE Full text] [doi: 10.1016/S1836-9553(13)70168-1] [Medline: 23663801]
18. Cruz J, Brooks D, Marques A. Home telemonitoring effectiveness in COPD: a systematic review. Int J Clin Pract 2014;68(3):369-378. [doi: 10.1111/ijcp.12345] [Medline: 24472009]
19. Rosen D, McCall JD, Primack BA. Telehealth protocol to prevent readmission among high-risk patients with congestive heart failure. Am J Med 2017;130(11):1326-1330. [doi: 10.1016/j.amjmed.2017.07.007] [Medline: 28756266]
20. O'Connor M, Asdornwised U, Dempsey ML, Huffenberger A, Jost S, Flynn D, et al. Using telehealth to reduce all-cause 30-day hospital readmissions among heart failure patients receiving skilled home health services. Appl Clin Inform 2016;7(2):238-247 [FREE Full text] [doi: 10.4338/ACI-2015-11-SOA-0157] [Medline: 27437037]
21. Joseph JW, Kennedy M, Nathanson LA, Wardlow L, Crowley C, Stuck A. Reducing emergency department transfers from skilled nursing facilities through an emergency physician telemedicine service. West J Emerg Med 2020;21(6):205-209 [FREE Full text] [doi: 10.5811/westjem.2020.7.46295] [Medline: 33207167]
22. Rosner BI, Gottlieb M, Anderson WN. Effectiveness of an automated digital remote guidance and telemonitoring platform on costs, readmissions, and complications after hip and knee arthroplasties. J Arthroplasty 2018;33(4):988-996.e4 [FREE Full text] [doi: 10.1016/j.arth.2017.11.036] [Medline: 29229238]
23. Huang VW, Reich KM, Fedorak RN. Distance management of inflammatory bowel disease: systematic review and meta-analysis. World J Gastroenterol 2014;20(3):829-842 [FREE Full text] [doi: 10.3748/wjg.v20.i3.829] [Medline: 24574756]
24. Liddy C, Bello A, Cook J, Drimer N, Pilon MD, Farrell G, et al. Supporting the spread and scale-up of electronic consultation across Canada: cross-sectional analysis. BMJ Open 2019;9(5):e028888 [FREE Full text] [doi: 10.1136/bmjopen-2018-028888] [Medline: 31152043]
25. Liddy C, Moroz I, Keely E, Taljaard M, Fraser AM, Armstrong CD, et al. The use of electronic consultations is associated with lower specialist referral rates: a cross-sectional study using population-based health administrative data. Fam Pract 2018;35(6):698-705 [FREE Full text] [doi: 10.1093/fampra/cmy020] [Medline: 29635449]
26. Liddy C, Moroz I, Keely E, Taljaard M, Armstrong CD, Afkham A, et al. Understanding the impact of a multispecialty electronic consultation service on family physician referral rates to specialists: a randomized controlled trial using health administrative data. Trials 2019;20(1):348 [FREE Full text] [doi: 10.1186/s13063-019-3393-5] [Medline: 31182123]
27. Waldura JF, Neff S, Dehlendorf C, Goldschmidt RH. Teleconsultation improves primary care clinicians' confidence about caring for HIV. J Gen Intern Med 2013;28(6):793-800 [FREE Full text] [doi: 10.1007/s11606-013-2332-5] [Medline: 23371417]
28. van Gelder VA, Scherpbier-de Haan ND, van Berkel S, Akkermans RP, de Grauw IS, Adang EM, et al. Web-based consultation between general practitioners and nephrologists: a cluster randomized controlled trial. Fam Pract 2017;34(4):430-436 [FREE Full text] [doi: 10.1093/fampra/cmw131] [Medline: 28158576]
29. Ayadi HE, Desai A, Jones RE, Mercado E, Williams M, Rooks B, et al. Referral rates vary widely between family medicine practices. J Am Board Fam Med 2021;34(6):1183-1188 [FREE Full text] [doi: 10.3122/jabfm.2021.06.210213] [Medline: 34772773]
30. Crump WJ, Massengill P. Outpatient consultations from a family practice residency program: nine years' experience. J Am Board Fam Pract 1988;1(3):164-166. [Medline: 3177014]

XSL•FO
RenderX

31. Tabenkin H, Oren B, Steinmetz D, Tamir A, Kitai E. Referrals of patients by family physicians to consultants: a survey of the Israeli family practice research network. Fam Pract 1998;15(2):158-164 [FREE Full text] [doi: 10.1093/fampra/15.2.158] [Medline: 9613485]
32. Franks P, Zwanziger J, Mooney C, Sorbero M. Variations in primary care physician referral rates. Health Serv Res 1999;34(1 Pt 2):323-329 [FREE Full text] [Medline: 10199678]
33. Gwynne M, Page C, Reid A, Donahue K, Newton W. What's the right referral rate? specialty referral patterns and curricula across I3 collaborative primary care residencies. Fam Med 2017;49(2):91-96 [FREE Full text] [Medline: 28218933]
34. Duffy R, Neville R, Staines H. Variance in practice emergency medical admission rates: can it be explained? Br J Gen Pract 2002;52(474):14-17 [FREE Full text] [Medline: 11791810]
35. Blatchford O, Capewell S, Murray S, Blatchford M. Emergency medical admissions in Glasgow: general practices vary despite adjustment for age, sex, and deprivation. Br J Gen Pract 1999;49(444):551-554 [FREE Full text] [Medline: 10621990]
36. Swift C, Cain SM, Needham M. A primary care telehealth experience in a US army correctional facility in Germany. US Army Med Dep J 2016:76-80. [Medline: 26874102]
37. Brunicardi BO. Financial analysis of savings from telemedicine in Ohio's prison system. Telemed J 1998;4(1):49-54. [doi: 10.1089/tmj.1.1998.4.49] [Medline: 9599074]
38. Sullivan DH, Chapman M, Mullen PE. Videoconferencing and forensic mental health in Australia. Behav Sci Law 2008;26(3):323-331. [doi: 10.1002/bsl.815] [Medline: 18548513]
39. Deslich SA, Thistlethwaite T, Coustasse A. Telepsychiatry in correctional facilities: using technology to improve access and decrease costs of mental health care in underserved populations. Perm J 2013;17(3):80-86 [FREE Full text] [doi: 10.7812/TPP/12-123] [Medline: 24355894]
40. Aoki N, Dunn K, Fukui T, Beck JR, Schull WJ, Li HK. Cost-effectiveness analysis of telemedicine to evaluate diabetic retinopathy in a prison population. Diabetes Care 2004;27(5):1095-1101 [FREE Full text] [doi: 10.2337/diacare.27.5.1095] [Medline: 15111527]
41. Yogesan K, Henderson C, Barry CJ, Constable IJ. Online eye care in prisons in Western Australia. J Telemed Telecare 2001;7(Suppl 2):63-64. [doi: 10.1258/1357633011937173] [Medline: 11747663]
42. McCue MJ, Hampton CL, Malloy W, Fisk KJ, Dixon L, Neece A. Financial analysis of telecardiology used in a correctional setting. Telemed J E Health 2000;6(4):385-391. [doi: 10.1089/15305620050503852] [Medline: 11242546]
43. de Almeida Castro Morosini I, de Oliveira DC, de Morais Ferreira F, Fraiz FC, Torres-Pereira CC. Performance of distant diagnosis of dental caries by teledentistry in juvenile offenders. Telemed J E Health 2014;20(6):584-589. [doi: 10.1089/tmj.2013.0202] [Medline: 24693859]
44. Edge C, George J, Black G, Gallagher M, Ala A, Patel S, et al. Using telemedicine to improve access, cost and quality of secondary care for people in prison in England: a hybrid type 2 implementation effectiveness study. BMJ Open 2020;10(2):e035837 [FREE Full text] [doi: 10.1136/bmjopen-2019-035837] [Medline: 32075846]
45. Senanayake B, Wickramasinghe SI, Eriksson L, Smith AC, Edirippulige S. Telemedicine in the correctional setting: a scoping review. J Telemed Telecare 2018;24(10):669-675 [FREE Full text] [doi: 10.1177/1357633X18800858] [Medline: 30343655]
46. Brunetti ND, Dellegrottaglie G, Di Giuseppe G, De Gennaro L, Di Biase M. Prison break: remote tele-cardiology support for cardiology emergency in Italian penitentiaries. Int J Cardiol 2013;168(3):3138-3140. [doi: 10.1016/j.ijcard.2013.04.022] [Medline: 23623349]

Abbreviations

ER: emergency room
VC: video clinic
VCS: videoconferencing system

Edited by T Leung; submitted 08.12.22; peer-reviewed by H Leach, R Koppel, F Prazeres; comments to author 07.01.24; revised version received 29.02.24; accepted 01.03.24; published 28.06.24.

Please cite as:
Schmalstieg-Bahr K, Colombo MG, Koch R, Szecsenyi J, Völker F, Blozik EE, Scherer M
Intramural Health Care Through Video Consultations and the Need for Referrals and Hospital Admissions: Retrospective Quantitative Subanalysis of an Evaluation Study
Interact J Med Res 2024;13:e44906
URL: https://www.i-jmr.org/2024/1/e44906
doi:10.2196/44906
PMID:

XSL•FO
RenderX

©Katharina Schmalstieg-Bahr, Miriam Giovanna Colombo, Roland Koch, Joachim Szecsenyi, Friedrich Völker, Eva Elisabeth Blozik, Martin Scherer. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 28.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Behavioral Insights from Vaccine Adoption in Nigeria: Cross-Sectional Survey Findings

Sohail Agha[1,2], BA, PhD; Ifeanyi Nsofor[3], MSc, MD; Drew Bernard[2], BA; Sarah Francis[2], BA; Nandan Rao[4], MSc

[1]Behavior Design Lab, Stanford University, Stanford, CA, United States
[2]Behavioral Insights Lab, Seattle, WA, United States
[3]Centre for Family Health Initiative, Abuja, Nigeria
[4]Virtual Labs, Corvallis, OR, United States

Corresponding Author:
Sohail Agha, BA, PhD
Behavior Design Lab
Stanford University
581 Capistrano Way
Stanford, CA, 94305
United States
Phone: 1 1 206 351 9346
Email: sohailagha@gmail.com

Abstract

Background: To generate behavioral insights for the development of effective vaccination interventions, we need approaches that combine rapid and inexpensive survey data collection with instruments based on easy-to-use behavior models. This study demonstrates how an inexpensive digital survey helped identify the drivers of COVID-19 vaccination in Nigeria.

Objective: This study aims to illustrate how behavioral insights can be generated through inexpensive digital surveys.

Methods: We designed and conducted a cross-sectional survey with multistage sampling. Data were collected from Nigerians (aged ≥18 years) from 120 strata based on age, sex, state, and urban or rural location. Respondents were recruited via advertisements on Meta platforms (Facebook and Instagram) using the Virtual Lab open-source tool. We used a Meta Messenger chatbot for data collection; participants were compensated with 400 naira (US $0.87 cents). Data collection took 2 weeks. In total, 957 respondents completed the survey, at an advertising cost of US $1.55 per respondent. An 18-item instrument measuring core motivators, ability barriers, sociodemographic characteristics, and respondents' vaccination status was pretested before data collection. We ran separate logistic regression models to examine the relationships between vaccine uptake and core motivators, ability barriers, and sociodemographic variables. A final model that predicted vaccine uptake included all 3 sets of variables.

Results: About 56% (n=540) of respondents reported that they had received at least 1 COVID-19 vaccination. Three core motivators were positively associated with vaccine uptake: the belief that the COVID-19 vaccine promised a better life (adjusted odds ratio [aOR] 3.51, 95% CI 2.23-5.52), the belief that the vaccine would allow respondents to do more things they enjoyed (aOR 1.97, 95% CI 1.33-2.93), and respondents' perception that their friends and family members accepted their decision to get vaccinated (aOR 1.62, 95% CI 1.06-2.48). Two ability barriers were negatively associated with vaccine uptake: cost- or income-related concerns lowered the odds of being vaccinated (aOR 0.35, 95% CI 0.24-0.50) and the lack of availability of vaccines at places respondents routinely visited also lowered their odds of being vaccinated (aOR 0.29, 95% CI 0.21-0.40). After adjusting for other variables, the perceived fear of getting COVID-19 and the hardship associated with the disease were no longer associated with vaccine uptake.

Conclusions: These findings suggest that hope is more important for Nigerians than fear when it comes to vaccine adoption, enjoying life is more important than worrying about getting the disease, and approval from friends and family is more powerful than their disapproval. These findings suggest that emphasizing the benefits of leading a fuller life after being vaccinated is more likely to succeed than increasing Nigerians' fear of COVID-19. This study identifies a very different set of factors associated with COVID-19 vaccine adoption than previous Nigerian studies.

(Interact J Med Res 2024;13:e47817) doi:10.2196/47817

XSL•FO
RenderX

KEYWORDS

behavioral insights; COVID-19; Nigeria; surveys; vaccination

Introduction

Background

Globally, the COVID-19 pandemic has highlighted the importance of behavioral insights for increasing the use of preventive behaviors such as wearing a mask, social distancing, and getting vaccinated. However, recent studies on COVID-19 vaccination emerging from Nigeria and other low- and middle-income countries (LMICs) have primarily focused on identifying gaps in knowledge, attitudes, and beliefs associated with vaccine hesitancy. A number of these studies recommend educating health care workers (HCWs) and members of the general population on vaccine safety and efficacy and assume that equipping people with factually correct information will allay their concerns, increase their perceived risk of acquiring COVID-19, and lead to higher rates of vaccine adoption [1,2].

Data from other Nigerian studies, however, raise questions regarding the strength of the relationship between risk perception and vaccine acceptance. A recent hospital-based study in southern Nigeria found that, while most HCWs perceived themselves at risk of COVID-19, only about half were willing to receive the COVID-19 vaccine [3]. Substantial gaps between risk perception and willingness to accept a COVID-19 vaccine have been observed in the general population in northern Nigeria as well [4]. A study that interviewed over 5000 respondents across all states in Nigeria found that COVID-19 was not perceived as a threat by most respondents [5].

What seems to be an important determinant of vaccine acceptability in Nigeria is trust in the vaccine manufacturing process, health system, government [6], and institutions involved in risk communication on behalf of the government [7]. Yet, despite multiple studies showing a weak relationship between risk perception and willingness to adopt a COVID-19 vaccine in Nigeria, researchers continue to recommend the provision of factually correct information to fill “information gaps” [4].

A recent systematic review of the COVID-19 literature in Nigeria shows that vaccination rates among those at high risk of COVID-19, such as HCWs, were lower than among those at low risk of COVID-19 [8]. Given the weak relationship between risk perception and vaccine uptake in Nigeria, it is not surprising that a recent evaluation found that risk communication efforts in Nigeria were inadequate in sustaining changes in behavior observed at the beginning of the pandemic [9]. Lawal [9] showed that during the first 30 days of the discovery of COVID-19 in Nigeria, and until the national lockdown, public interest in learning about the disease surged. Visits to public places such as grocery stores declined during this period as stringent government policies resulted in reduced mobility of the population. The study by Lawal [9] found that, as the Nigerian population started becoming aware of the disease, there was a slight decline in the number of COVID-19 cases. However, this decline occurred for a relatively short period of time. The number of new cases started increasing again as the initial effects of risk communication interventions dissipated. Lawal [9] concluded that Nigerians listened to messages telling them to take preventive measures such as social distancing or masking for some time but eventually got tired of the messages and stopped responding to them. In part, this was because the recommended public health precautions did not fit well in the context in which they lived their lives [9].

A clear picture of the drivers of COVID-19 vaccine hesitancy does not emerge from the recent public health literature on Nigeria, in part because much of this work is not based on a behavioral framework. The importance of theory-based work to understand the drivers of vaccine acceptance and design appropriate interventions has been emphasized [10]. In the absence of a clearly articulated framework for understanding vaccine-related behavior, it is difficult to interpret the findings of individual studies and arrive at a clear picture of what drives vaccine uptake in Nigeria.

As a result, there is very limited guidance available to support Nigerian practitioners in designing interventions that might accelerate COVID-19 vaccine uptake. For example, although many recent studies emphasize the importance of implementing health promotion interventions or increasing HCWs’ ability to communicate more effectively with members of the general public, most of these studies do not provide any guidance on what the content of this communication should be or what strategy should be used to persuade adults to get vaccinated. Thus, the available research is at a standstill in terms of providing insights that would help in designing more effective behavioral interventions to accelerate vaccine uptake in Nigeria.

An important reason that the literature does not provide a clear direction for the design of behavioral interventions is the lack of use of behavioral frameworks in explaining vaccine acceptance and uptake. Of the more than 20 peer-reviewed publications reviewed for this paper, we found only 1 that used a behavior model to interpret its findings [5]. This is not surprising as researchers have noted the limited use of behavioral frameworks in public health research and practice for over a decade [11-14].

The need for a behavior model that can be used to explain the vaccine adoption process in simple terms that resonate with practitioners is urgent. While a broader discussion of what a practitioner-friendly behavior model should comprise of is merited, the characteristics of such a model have been proposed [15]. A minimum criterion should be that the use of the model leads to deliberate programmatic decisions, a greater emphasis on strengthening activities supported by behavioral research findings, and the elimination of activities that are not evidence based.

The practitioner-friendly model used in this study, the Fogg Behavior Model (FBM), was introduced in the public health literature in 2019 to explain the effects of a social marketing behavior change campaign on the adoption of condoms by married men in Pakistan [13]. More recently, it has been used to identify behavioral drivers associated with the (1) adoption

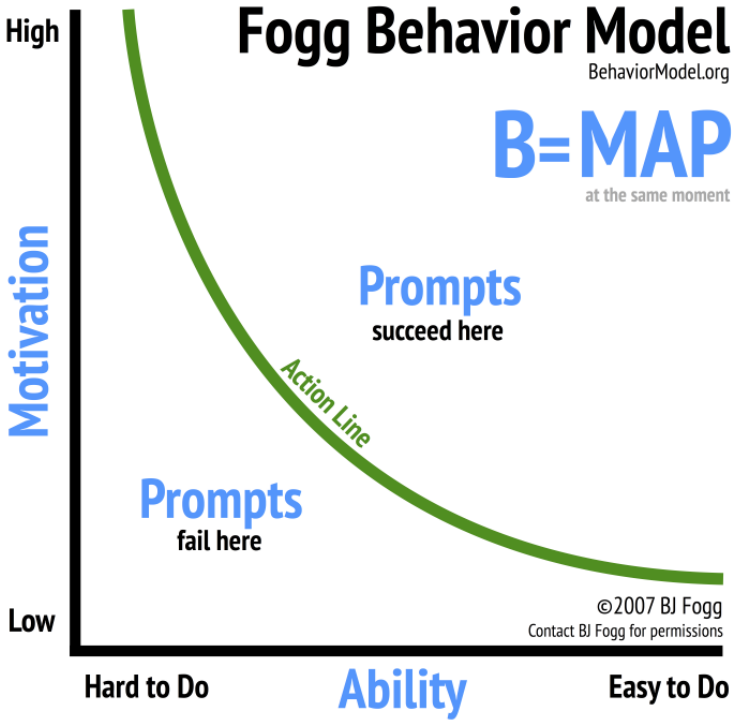
XSL•FO
RenderX

of COVID-19 vaccination by a low-income population in Cote d'Ivoire [16], (2) adoption of iron folate by pregnant women in India [15], (3) uptake of COVID-19 vaccination by HCWs in Nigeria [17,18], and (4) use of contraception by adolescent girls and young women in Nigeria [19]. A recent study also demonstrates the use of the FBM in making timely programmatic adjustments to a contraceptive social marketing intervention implemented in Nigeria [20]. To the best of our knowledge, this is the first time that the FBM is being applied to understand the dynamics of COVID-19 vaccine adoption in the general population of Nigeria.

FBM: Motivation and Ability as the Drivers of Behavior

The FBM is a model developed for use by practitioners to understand the drivers of human behavior and assist them in the design of behavior change interventions. Fogg states that behavior happens when motivation, ability, and a prompt happen at the same moment. As shown in Figure 1, the model can be visualized in 2 dimensions. Figure 1 shows motivation along the y-axis and ability along the x-axis. Motivation ranges from high to low for any behavior. Ability also ranges from high to low for any behavior. For simplicity, we describe a behavior as being easy to do or hard to do. For a prompt to work, a person needs to have sufficient motivation and ability. The motivation-ability threshold is reflected by an action line in the FBM. Behavior occurs when a person whose motivation and ability are above the action line is prompted. The prompt does not work if the person does not have sufficient motivation to undertake the behavior and finds the behavior hard to do, that is, they are below the action line [21].

Figure 1. Fogg Behavior Model (reproduced from Fogg [22], with permission from BJ Fogg).



Methods

Questionnaire Design: Core Motivators

Fogg defines motivation as having 3 components: anticipation, sensation, and belonging. Anticipation reflects the hopes and fears a person associates with a behavior. Sensation reflects the pleasure or pain a person associates with a behavior. Belonging is reflected by the acceptance or rejection of the behavior by people whose opinions a person considers important.

We conducted a review of the literature to identify relevant constructs and appropriate measures of motivation and ability. The survey questionnaire (Multimedia Appendix 1) was designed using instruments developed to test motivation and ability constructs in the FBM. These instruments had been previously tested in the general population in Cote d'Ivoire and with HCWs in Nigeria and had shown acceptable levels of reliability [16,18]. Respondents were allowed to answer how strongly they agreed or disagreed with 6-point Likert scale items measuring core motivators, such as hope, fear, pleasure, pain, acceptance, and rejection, associated with the adoption of the COVID-19 vaccine.

Ability Barriers

Fogg defines ability in terms of 5 barriers: time, money, the physical effort required to adopt a behavior, the mental effort required to adopt a behavior, and whether the behavior fits into the person's routine [21]. Fogg initially considered social

XSL•FO
RenderX

deviance (or social norms) as part of ability but did not include social norms as an ability barrier in later iterations of the model [22], possibly because of the complexity of the relationship between norms and behavior. To capture ability barriers, respondents could answer how strongly they agreed with items on a 5-point Likert scale. These items captured the 5 ability factors related to COVID-19 vaccine uptake: time, money, mental effort, physical effort, and routine. The instrument also contained questions on sociodemographic variables. In total, the instrument comprised 18 questionnaire items.

The instrument was pretested twice in Nigeria with samples of approximately 100 respondents. The first pretest showed that the relationship among variables measuring motivation, ability, and COVID-19 vaccine uptake was in the expected direction, with 1 exception. The variable measuring agreement or disagreement with the question "Many of my friends and family would think poorly of me if they knew I had taken the COVID-19 vaccine" did not demonstrate the expected relationship with vaccine uptake. This was replaced with the statement "Most people I know have obtained the COVID-19 vaccination." The second pretest showed that the latter response was associated with the outcome in the expected direction. All other items included in the instrument demonstrated the expected relationships with vaccine uptake in both pretests.

Survey Design and Sampling

We implemented a cross-sectional survey with multistage sampling. Nigeria is a large and diverse country with 36 states and the Federal Capital Territory of Abuja. These states and the Federal Capital Territory are grouped into 6 geopolitical zones: northwest, northeast, north central, southwest, southeast, and south-south. Although this survey was not designed to be a representative survey of Nigeria, we aimed to capture the diversity of the country's population by ensuring that respondents from 1 state within each geopolitical zone were sampled. Thus, the states were selected at the first stage. Sokoto state was selected from the northwest, Bauchi state from the northeast, Niger state from north central, Lagos state from the southwest, Anambra state from the southeast, and Rivers state from the south-south. The state selected within each zone reflects its socioeconomic, religious, and ethnic diversity. In addition, these 6 states had ongoing COVID-19 vaccination campaigns to help ensure that survey findings would be useful for COVID-19 program managers in Nigeria.

The sample was stratified by 6 states, 5 age groups, male and female sex, and urban or rural location. This resulted in 120 strata from which respondents were sampled. We ran a total of 120 different ad sets targeting respondents based on the characteristics mentioned above.

Respondents were recruited via advertisements on the Meta digital ad platforms (Facebook and Instagram) using the Virtual Lab open-source tool [23]. The Virtual Lab tool ran ads targeting respondents in all 120 strata. We used a Meta Messenger chatbot for the survey data collection, compensating respondents who completed the survey with 400 naira (US $0.87 cents) in mobile phone credit. Respondents who clicked on the ads were directed to the messenger bot. Of the 214,335 male and female respondents who reached through the ads, 3660 clicked on the link, 1367 started the survey, and 1011 answered most survey questions.

Respondents could complete the survey in one go or start the survey, stop, and return to complete it later. In total, 957 respondents answered all questions in the survey at an advertising cost of US $1.55 per person. Data from these 957 respondents are used for the analysis.

The outcome of interest for this study was having received at least 1 COVID-19 vaccination. The Government of Nigeria's data on the number of vaccinations provided in 2022 suggested that 50% of Nigerians had received at least 1 COVID-19 vaccination. A sample size calculation was made using an outcome value of 50% and a design effect of 1.5. We estimated that a sample size of 900 would provide a margin of error of 4 percentage points on the outcome of interest.

Statistical Analysis: Relationships Among Core Motivators, Ability Barriers, and Vaccine Uptake

Univariate analysis was conducted to provide the frequency distributions of core motivators, ability barriers, and sociodemographic characteristics of the sample. Bivariate analysis was conducted to explore the relationships between core motivators, ability factors, sociodemographic characteristics, and having received at least 1 COVID-19 vaccination. Multivariate logistic regression analysis was conducted to identify which core motivators, ability factors, and sample characteristics had a significant relationship with vaccine uptake [24].

We also ran a final multivariate model to determine whether there was any change in the relationship between individual variables and vaccine uptake after taking the 3 sets of variables (core motivators, ability factors, and sociodemographic characteristics) into account. Adjusted odds ratios (aORs) from the analyses are shown in the tables. P values were considered statistically significant at $P<.05$.

Ethical Considerations

The ethical approval for the study was obtained from Nigeria's National Health Research Ethics Committee (NHREC/01/01/2007). Informed consent was obtained from all respondents to the quantitative survey. Respondents were assured that all written and recorded data would be kept confidential by using codes to identify participants instead of names or any other personal identifiers. Respondents were informed about their right to refuse to participate in the study or withdraw at any time during the interview.

Results

Core Motivators and COVID-19 Vaccine Uptake

Table 1 shows the frequency distributions of core elements of motivation identified by the FBM, cross-tabulations between core motivators and COVID-19 vaccine uptake and the aOR of COVID-19 vaccination. The first column of Table 1 shows that about three-fourths (n=726, 75.9%) of survey respondents agreed or strongly agreed with the statement that getting vaccinated against COVID-19 allows a person to live a better life, a measure of hope. Nearly half (n=444, 46.4%) of respondents

XSL•FO
RenderX

reported that getting vaccinated protected them and their families from hardship, a measure of fear. Nearly two-thirds (n=610, 63.7%) of respondents agreed or strongly agreed with the statement that getting vaccinated allows a person to do the things they enjoy, a measure of pleasure. About 81.5% (n=780) reported that vaccination reduces the likelihood of getting or spreading COVID-19, a measure of pain. About 79.1% (n=757) of respondents reported that many of their family and friends approve of the COVID-19 vaccination, a variable measuring social acceptance. Consistent with the high social acceptance of the COVID-19 vaccination, rejection of the vaccine was much lower, that is, 37.1% (n=355) of respondents reported that most of their family and friends did not approve of the COVID-19 vaccination.

Table 1. Frequency distributions of core motivators, cross-tabulations, and the adjusted odds of COVID-19 vaccination in Nigeria.

Core motivators	Frequency distributions of core motivators (N=957, 100%), n (%)	Nigerian adults who obtained at least 1 vaccination (n=540, 56.4%), n (%)	*P* value	Adjusted odds of obtaining at least 1 COVID-19 vaccination, aOR (95% CI)[a]
Hope: Vaccination allows a person to live a better life			<.001	
Agree or strongly agree	726 (75.9)	483 (66.5)		3.40 (2.27-5.09)
Disagree or strongly disagree or do not know	231 (24.1)	57 (24.7)		1.00 (Reference)
Fear: Getting vaccinated protects people from hardship			<.001	
Agree or strongly agree	444 (46.4)	301 (67.8)		1.09 (0.78-1.51)
Disagree or strongly disagree or do not know	513 (53.6)	239 (46.6)		1.00 (Reference)
Pleasure: Getting vaccinated allows people to do more things they enjoy			<.001	
Agree or strongly agree	610 (63.7)	411 (67.4)		1.84 (1.30-2.62)
Disagree or strongly disagree or do not know	347 (36.3)	129 (37.2)		
Pain: Getting vaccinated reduces the likelihood of getting or spreading COVID-19			<.001	
Agree or strongly agree	780 (81.5)	477 (61.2)		1.15 (0.76-1.73)
Disagree or strongly disagree or do not know	177 (18.5)	63 (35.6)		1.00 (Reference)
Acceptance: Many friends and family approve vaccination			<.001	
Agree or strongly agree	757 (79.1)	472 (62.4)		1.76 (1.21-2.58)
Disagree or strongly disagree or do not know	200 (20.9)	68 (34)		1.00 (Reference)
Rejection: Most family and friends do not approve of vaccination			.56	
Agree or strongly agree	355 (37.1)	344 (57.1)		1.06 (0.79-1.43)
Disagree or strongly disagree or do not know	602 (62.9)	196 (55.2)		1.00 (Reference)

[a]Pseudo R^2=11.85%

The second column of Table 1 shows cross-tabulations between the core motivators of vaccination and vaccine uptake. There were large differences in vaccine uptake by core motivators at the bivariate level. Respondents who associated a COVID-19 vaccination with the hope of a better life had a 42-percentage point higher rate of vaccination (n=483, 66.5% vs n=57, 24.7%; *P*<.001). Those who feared the hardship that COVID-19 infection would bring had a 21-percentage point higher vaccination rate than other respondents (n=301, 67.8% vs n=239, 46.6%; *P*<.001). The pleasure that respondents associated with being able to do what they enjoyed doing because of being vaccinated was reflected by a 30-percentage point higher rate of vaccination (n=411, 67.4% vs n=129, 37.2%; *P*<.001). Respondents' concern that not being vaccinated would result in getting or spreading COVID-19 was associated with a 25-percentage point higher rate of vaccination (n=477, 61.2% vs n=63, 35.6%; *P*<.001). Acceptance of the vaccine by friends and family was associated with a 28-percentage point higher vaccination rate (n=472, 62.4% vs n=68, 34%; *P*<.001). It is interesting that social rejection, or the lack of approval of the vaccination by family members, was not associated with vaccine uptake.

The third column of Table 1 shows the adjusted odds of COVID-19 vaccination. With all core motivators in the model, Nigerians who believed that COVID-19 vaccination was associated with a better life, had 3 times higher odds of getting vaccinated (aOR 3.40, 95% CI 2.27-5.09). Those who believed that getting vaccinated would allow them to do more things that they enjoyed were more likely to be vaccinated (aOR 1.84, 95% CI 1.30-2.62). Friends and family members' acceptance of their

XSL•FO
RenderX

decision to get vaccinated was associated with a higher vaccination rate (aOR 1.76, 95% CI 1.21-2.58).

Ability Barriers and COVID-19 Vaccine Uptake

Table 2 shows the frequency distributions of ability factors identified by the FBM, cross-tabulations between ability factors and COVID-19 vaccine uptake, and the adjusted odds of COVID-19 vaccination. The first column of Table 2 shows that about 54% (n=516) of respondents felt that their family or work responsibilities made it difficult for them to get vaccinated. This variable measures the constraint of time. Over a third (n=347, 36.3%) of respondents felt that the cost or loss of income associated with getting vaccinated was a barrier. Nearly 40% (n=382) of respondents felt that the physical effort required made it difficult to get vaccinated. Nearly half (n=437, 46%) of the respondents felt that the decision to get vaccinated was difficult. This variable measures the mental effort required to get vaccinated. About 42% (n=401) of respondents reported that not having the vaccine available in places they routinely visited was a barrier to getting vaccinated. The latter measures the routine associated with adopting a behavior.

Table 2. Frequency distributions of ability factors, cross-tabulations, and the adjusted odds of COVID-19 vaccination in Nigeria.

Ability	Frequency distributions of ability factors (N=957, 100%), n (%)	Nigerian adults who obtained at least one vaccination (n=540, 56.4%), n (%)	*P* value	Adjusted odds of obtaining at least one COVID-19 vaccination[a]
Time: Family or work responsibilities make it difficult to find time				
Agree or strongly agree	516 (53.9)	277 (53.7)	.06	0.94 (0.69-1.27)
Disagree or strongly disagree or do not know	441 (46.1)	263 (59.6)		1.00 (Reference)
Money: Costs or loss of income make it difficult				
Agree or strongly agree	347 (36.3)	158 (45.5)	<.001	0.45 (0.33-0.62)
Disagree or strongly disagree or do not know	610 (63.7)	382 (62.6)		1.00 (Reference)
Physical effort: Physical effort makes it difficult				
Agree or strongly agree	382 (39.9)	218 (57.1)	.74	1.46 (1.06-2.01)
Disagree or strongly disagree or do not know	575 (60.1)	322 (56.0)		1.00 (Reference)
Mental effort: Decision to get vaccine is difficult				
Agree or strongly agree	437 (45.7)	237 (54.2)	.21	0.94 (0.70-1.27)
Disagree or strongly disagree or do not know	520 (54.3)	303 (58.3)		1.00 (Reference)
Routine: Vaccine not available where I routinely visit				
Agree or strongly agree	401 (41.9)	141 (35.2)	<.001	0.21 (0.16-0.28)
Disagree or strongly disagree or do not know	556 (58.1)	399 (71.8)		1.00 (Reference)

[a]Pseudo R^2=12%.

The second column of Table 2 shows cross-tabulations between ability factors and vaccine uptake. Respondents who agreed or strongly agreed with the statement that costs or loss of income were a barrier reported a 17-percentage point lower vaccination rate compared to others (n=158, 45.5% vs n=382, 62.6%; *P*<.001). The lack of availability of the COVID-19 vaccine in places that they routinely visited was associated with a 37-percentage point lower rate of vaccination (n=141, 35.2% vs n=399, 71.8%; *P*<.001).

The third column of Table 2 shows the adjusted odds of a COVID-19 vaccination. With all ability factors in the model, Nigerians who believed that the cost or the loss of income made it difficult to obtain a COVID-19 vaccination were less likely to get vaccinated (aOR 0.45, 95% CI 0.33-0.62). The lack of availability of vaccines at places respondents routinely visited was associated with a lower likelihood of vaccination (aOR 0.21, 95% CI 0.16-0.28). Contrary to our expectations, respondents who felt that physical effort makes it difficult to get vaccinated were more likely to be vaccinated (aOR 1.46, 95% CI 1.06-2.01).

Sociodemographic Factors and COVID-19 Vaccine Uptake

Table 3 shows the frequency distributions of sociodemographic characteristics of respondents in the sample, cross-tabulations between these characteristics and vaccine uptake, and the adjusted odds of COVID-19 vaccination. The first column of Table 3 shows that, as expected from a digital survey, the sample had relatively young participants: 56% (n=531) of respondents were aged 18-29 years and 16% (n=151) were aged 40 years and older. Male participants represented a higher proportion of the sample (n=592, 61.9%). About 39% (n=371) of respondents had a primary or secondary school certificate, one quarter

XSL•FO
RenderX

(n=231, 24.1%) had an ordinary national diploma (OND) or a higher national diploma (HND), and one-third of respondents (n=317, 33.4%) had a bachelor's or higher degree. A majority of respondents were from urban areas: 57% (n=544) were from cities, 34% (n=323) from towns, and 9% (n=90) from rural areas.

Table 3. Frequency distributions of sociodemographic variables, cross-tabulations, and the adjusted odds of COVID-19 vaccination in Nigeria.

Demographic	Sample characteristics (N=957, 100%), n (%)	Nigerian adults who obtained at least 1 vaccination (n=540, 56.4%), n (%)	*P* value	Adjusted odds of obtaining at least 1 COVID-19 vaccination, aOR (95% CI)[a]
Age (years)			.71	
18-29	531 (55.5)	300 (56.5)		1.31 (0.89-1.93)
30-39	275 (28.7)	159 (57.8)		1.28 (0.85-1.93)
≥40	151 (15.8)	81 (53.8)		1.00 (Reference)
Sex			.06	
Male	592 (61.9)	348 (58.8)		1.00 (Reference)
Female	365 (38.1)	192 (52.6)		0.73 (0.56-0.95)
Education			<.001	
Primary or secondary school certificate	371 (38.8)	188 (50.7)		1.00 (Reference)
Ordinary national diploma (OND)	111 (11.6)	72 (64.9)		1.90 (1.22-2.98)
Higher national diploma (HND)	120 (12.5)	87 (72.5)		2.62 (1.64-4.19)
Bachelors or higher	317 (33.4)	169 (53.3)		1.12 (0.81-1.55)
Other	38 (4)	24 (63.2)		1.79 (0.89-3.60)
Location			.02	
City	544 (56.8)	320 (58.8)		1.00 (Reference)
Town	323 (33.8)	181 (56.0)		0.88 (0.68-1.17)
Rural	90 (9.4)	39 (43.3)		0.56 (0.35-0.89)

[a]Pseudo R^2=2.79%.

The second column of Table 3 shows cross-tabulations between sociodemographic characteristics and vaccine uptake. There was no statistically significant difference in the COVID-19 vaccination rate by age or sex. Education was associated with vaccine uptake: respondents with an OND (n=72, 64.9% vs n=188, 50.7%; *P*<.001) or an HND (n=87, 72.5% vs n=188 50.7%; *P*<.001) were more likely to be vaccinated than respondents with a primary or secondary school certificate. Urban residence was also associated with higher vaccine uptake: respondents from rural areas were significantly less likely to have obtained the COVID-19 vaccination (n=39, 43.3% vs n=320, 58.8%; *P*=.02).

The third column of Table 3 shows the adjusted odds of COVID-19 vaccination. With all sociodemographic characteristics in the model, female participants were less likely to get vaccinated (aOR 0.73, 95% CI 0.56-0.95). Having an OND (aOR 1.90, 95% CI 1.22-2.98) or a HND (aOR 2.62, 95% CI 1.64-4.19) increased a respondent's likelihood of being vaccinated. Nigerians living in rural areas were less likely to be vaccinated (aOR 0.56, 95% CI 0.35-0.89).

Full Model: Core Motivators, Ability Barriers, and Sociodemographic Characteristics

Table 4 shows the adjusted odds of COVID-19 vaccine uptake in Nigeria. The 3 core motivators identified earlier remained significant after adjusting for ability factors and sociodemographic characteristics. Nigerians who believed that the COVID-19 vaccination was associated with the promise of a better life were more likely to be vaccinated (aOR 3.51, 95% CI 2.23-5.52). Nigerians who felt that the vaccination would allow them to do more things they enjoyed were more likely to be vaccinated (aOR 1.97, 95% CI 1.33-2.93). Respondents' friends' and family members' acceptance of their decision to get vaccinated was associated with a higher likelihood of their being vaccinated (aOR 1.62, 95% CI 1.06-2.48).

XSL•FO
RenderX

Table 4. Adjusted odds (aOR) of COVID-19 vaccination in Nigeria.

	Adjusted odds of obtaining at least 1 COVID-19 vaccination, aOR (95% CI)[a]
Core motivators	
Getting vaccinated allows a person to live a better life (hope)	3.51 (2.23-5.52)
Getting vaccinated protects people from hardship (fear)	1.03 (0.71-1.49)
Allows people to do more things they enjoy (pleasure)	1.97 (1.33-2.93)
Reduces the likelihood of getting or spreading COVID-19 (pain)	1.17 (0.74-1.84)
Many friends and family approve vaccination (acceptance)	1.62 (1.06-2.48)
Most family and friends do not approve vaccination (rejection)	1.18 (0.83-1.67)
Ability factors	
Family or work responsibilities make it difficult (time)	0.85 (0.61-1.19)
Costs or loss of income make it difficult (money)	0.35 (0.24-0.50)
Physical effort makes it difficult (physical effort)	1.45 (1.02-2.07)
The decision to get the vaccine is difficult (mental effort)	1.08 (0.77-1.51)
Vaccine not available where I routinely visit (routine)	0.29 (0.21-0.40)
Sociodemographic factors	
Age (years)	
18-29	1.09 (0.68-1.74)
30-39	1.14 (0.70-1.86)
≥40	1.00 (Reference)
Sex	
Male	1.00 (Reference)
Female	0.80 (0.58-1.10)
Education	
Primary or secondary school certificate	1.00 (Reference)
Ordinary national diploma (OND)	2.27 (1.34-3.84)
Higher national diploma (HND)	3.57 (2.04-6.24)
Bachelors or higher	1.26 (0.86-1.85)
Other	1.88 (0.82-4.33)
Location	
City	1.00 (Reference)
Town	0.90 (0.64-1.26)
Rural	0.54 (0.31-0.92)

[a]Pseudo $R^{2=}$23.51%.

The relationships between ability factors and vaccine uptake remained important after adjusting for sociodemographic characteristics and core motivators. Nigerians with cost- or income-related concerns were less likely to obtain a COVID-19 vaccination (aOR 0.35, 95% CI 0.24-0.50). The lack of availability of vaccines at places they routinely visited made them less likely to get vaccinated (aOR 0.29, 95% CI 0.21-0.40).

After adjusting for motivation and ability, female participants were no longer less likely to obtain a COVID-19 vaccination. Respondents with an OND (aOR 2.27, 95% CI 1.34-3.84) or HND (aOR 3.57, 95% CI 2.04-6.24) were more likely to be vaccinated than those with primary or secondary school certificates. Rural residents were less likely to be vaccinated than residents living in cities (aOR 0.54, 95% CI 0.31-0.92).

Discussion

Principal Findings

The findings of this study show that 56.4% (n=540) of Nigerian adults who responded to the digital survey had obtained at least 1 COVID-19 vaccination by October 2022. Several core motivators were associated with vaccine uptake, after adjusting

XSL•FO
RenderX

for ability factors and sociodemographic variables. The beliefs that COVID-19 vaccination allows a person to live a better life and that it allows them to do more things that they enjoy increases the likelihood of being vaccinated. The belief that many friends and family members approve of COVID-19 vaccination is also associated with a greater likelihood of being vaccinated. Several ability barriers were also correlated with vaccine uptake, after adjusting for other variables. Respondents who felt that costs or loss of income associated with getting vaccinated made it difficult to get vaccinated were less likely to get vaccinated. The lack of availability of the COVID-19 vaccine at places respondents routinely visited was also negatively associated with vaccine uptake. The study also found a higher likelihood of Nigerians with OND or HND being vaccinated compared with those with primary or secondary school certificates and a lower likelihood of being vaccinated among rural respondents.

Strengths and Limitations

While this study provides useful insights for program design, its limitations should be acknowledged. The first limitation of this study is that no causal inferences can be made from it because of its cross-sectional design. This design limitation may explain an unexpected study finding: after adjusting for other variables, the physical effort required to get vaccinated was associated with a higher rate of vaccination. This finding may reflect reverse causality; those who are vaccinated may be more aware of the physical effort required to obtain a COVID-19 vaccination. Further investigation is needed to determine whether the positive relationship between the perceived physical effort required to get vaccinated and receiving a vaccination holds only for those who have been vaccinated or for the full sample.

The second limitation of this study is that variables that were not measured may be responsible for the observed relationships. For example, while the relationship between the belief that vaccination allows a person to enjoy life more and vaccine uptake is powerful, there is a possibility that unmeasured factors are driving this relationship. Thus, developing messages around how vaccination can help a person lead a fuller life and testing them through relatively inexpensive digital campaigns would be important prior to implementing an at-scale campaign that focuses on this message.

The third limitation of the study is that it is not representative of all Nigerians in the 6 states in which it was conducted. This is reflected in the higher educational status of the survey sample: about 33% (n=317) of respondents had a bachelor's or higher education. The participants were also relatively young; about 56% (n=531) of respondents were between 18 and 29 years. Moreover, male participants comprised a higher proportion of the sample than female participants. These findings are not uncommon for digital surveys conducted in LMICs.

A strength of this study is its cost efficiency and the timeliness with which the survey was conducted compared to face-to-face household surveys. A major barrier to the use of behavioral insights by practitioners in LMICs is the cost of data collection. Behavioral research is not well-funded in LMICs. Inexpensive digital surveys could substantially increase the ability of practitioners in LMICs to use behavioral insights to develop interventions that increase vaccine uptake.

Future Directions

We do not know the extent to which the findings from surveys conducted by recruiting respondents through web-based advertising are comparable to the findings from population-based household surveys. Although some studies show broadly similar patterns between digital and population-based surveys [25], more research is needed to identify what types of systematic differences may exist between these 2 survey modalities. It is important, for example, to learn whether inferences from digital surveys apply to the behavior of individuals who are not on digital platforms.

Given our inability to generalize these findings beyond Nigerians who are on Facebook and Instagram, how can the findings of this survey be used? First, the findings may be used to design interventions on digital platforms as well as to evaluate the effectiveness of those interventions. Digital behavior change interventions may be evaluated by experimental studies on digital platforms that compare vaccine uptake between intervention and control groups. A growing proportion of the Nigerian population is now on Facebook and Instagram: between 31 million and 36 million Nigerians 13 and older use Facebook and Instagram each month. High exposure to messages that associate the COVID-19 vaccination with a better, more fulfilling life is achievable through advertising on digital platforms and at a fraction of the cost of advertising on traditional mass media channels such as television.

Future interventions could build upon the findings of this study by conducting qualitative research to determine which motivation or ability factors are relevant in locations where interventions are planned. A survey conducted in the most densely populated, low-income commune in Yugpognon, Cote d'Ivoire, using the FBM found broadly comparable findings: motivation and ability were powerful drivers of vaccine adoption, although the specific elements of motivation and ability that were relevant in Yugpognon were, not surprisingly, different [16].

Our findings raise several questions that should be answered through additional research. Answers to these questions may help in the design of more effective COVID-19 service delivery interventions. It would be useful, for example, to learn whether some of the places routinely visited by Nigerians are amenable to serving as COVID-19 vaccine delivery sites. Are such potential vaccine delivery sites likely to vary by age, sex, by urban or rural residence and are they suitable for cost-efficient provision of COVID-19 vaccinations?

Our sample consisted primarily of Nigerians living in cities and towns, with a minority of respondents living in rural areas. A larger proportion of rural respondents may be obtained from digital surveys that oversample rural areas. This may be done by capping the number of respondents from urban areas and allowing more time for responses to come in from rural areas. This would, however, have implications for the cost of the rural component of the survey.

XSL•FO
RenderX

Comparison to Prior Work

Overall, the findings of our study provide a very different perspective on vaccine adoption in Nigeria than what is available in the peer-reviewed literature. Recent Nigerian studies on COVID-19 vaccine uptake have primarily focused on identifying gaps in knowledge, attitudes, and beliefs associated with vaccine hesitancy [1-3]. These studies place emphasis on the perceived risk of disease as a driver of vaccine uptake, despite mixed evidence on the role of risk perception on vaccine uptake in Nigeria [5,6,9]. Several of these recent studies propose that equipping people with factually correct information will allay their concerns, increase their perceived risk of acquiring COVID-19, and lead to higher rates of vaccine adoption. By contrast, a small but rapidly growing body of work is putting vaccine adoption in a behavioral context [13-19,26]. These studies take motivation and ability for behavior change into account in explaining the range of barriers that influence immunization decisions and suggest how programs should help individuals overcome them.

Conclusions

These findings help us consider a very different approach to intervention design—one that builds upon what people want for their future, what gives them pleasure, and how they are influenced by the approval of their friends and family members. Our findings suggest that hope is more important for Nigerians than fear when it comes to vaccine adoption, social approval is more powerful than social disapproval, and enjoying life is more important than worrying about getting the disease.

These findings suggest that an approach that is based on increasing the perception of hope and pleasure associated with vaccine adoption as well as increasing network members' social approval is likely to increase COVID-19 vaccine adoption in Nigeria. Our analysis also suggests that financial considerations play an important role in the uptake of COVID-19 vaccination in Nigeria. The costs associated with reaching a vaccination site or the loss of income associated with being away from work are important determinants of vaccine adoption. Nigerians with limited flexibility at work may find it challenging to visit a vaccination site during the hours that it is open. Consistent with this finding, making COVID-19 vaccines available at places that Nigerians visit routinely may have a large impact on vaccine uptake.

The use of a behavior model to understand drivers of COVID-19 vaccine uptake in Nigeria has helped provide a different perspective on vaccine-related decision-making in Nigeria than what is currently available in the published literature. The FBM, a model of human behavior rather than a model of health behavior per se, considers a broad range of factors influencing motivation, including an individual's hopes and fears, the sensation of pleasure or pain that they get from a particular behavior, and the social influences on them associated with their identity. The model also measures ability constraints including bandwidth-related constraints such as time or cognitive constraints, financial constraints, physical effort–related constraints, and habits or routine-related constraints. By comparison to behavior models that focus on perceived risk of and susceptibility to disease, the FBM situates behavior within the broader context of a person's life.

Acknowledgments

The authors gratefully acknowledge the study participants for the information they provided. Funding for this study was provided by The Bill & Melinda Gates Foundation. The funders had no input on the interpretation or publication of the study results.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

IN developed and submitted the institutional review board protocol to Nigeria's National Health Research Ethics Committee. SA and DB conceptualized the study. SA conducted the statistical analysis and wrote the first draft of the report. IN provided important context on Nigeria that enabled the interpretation of the study findings. DB developed the survey instrument and interpreted the study findings. DB and SF were responsible for acquiring funding for the study. SF and NR designed the digital data collection instruments and implemented the data collection and quality control procedures. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questionnaire.
[DOCX File , 22 KB - ijmr_v13i1e47817_app1.docx]

References

XSL•FO
RenderX

1. Robinson ED, Wilson P, ELeki BJ, Wonodi W. Knowledge, acceptance, and hesitancy of COVID-19 vaccine among health care workers in Nigeria. MGM J Med Sci 2021;8(2):102-110 [FREE Full text] [doi: 10.4103/mgmj.mgmj_4_21]
2. Osuagwu UL, Mashige KP, Ovenseri-Ogbomo G, Envuladu EA, Abu EK, Miner CA, et al. The impact of information sources on COVID-19 vaccine hesitancy and resistance in sub-Saharan Africa. BMC Public Health 2023;23(1):38 [FREE Full text] [doi: 10.1186/s12889-022-14972-2] [Medline: 36609264]
3. Adejumo OA, Ogundele OA, Madubuko CR, Oluwafemi RO, Okoye OC, Okonkwo KC, et al. Perceptions of the COVID-19 vaccine and willingness to receive vaccination among health workers in Nigeria. Osong Public Health Res Perspect 2021;12(4):236-243 [FREE Full text] [doi: 10.24171/j.phrp.2021.0023] [Medline: 34289295]
4. Iliyasu Z, Umar AA, Abdullahi HM, Kwaku AA, Amole TG, Tsiga-Ahmed FI, et al. "They have produced a vaccine, but we doubt if COVID-19 exists": correlates of COVID-19 vaccine acceptability among adults in Kano, Nigeria. Hum Vaccin Immunother 2021;17(11):4057-4064 [FREE Full text] [doi: 10.1080/21645515.2021.1974796] [Medline: 34613864]
5. Babafunke F, Olajumoke O, Oluwatofunmi L, Oluyemi A, Oluseyi A, Oluwagbemisola F, et al. Perceived risk and protective behaviors related to COVID-19 among the Nigerian general population: an application of the extended parallel process model. Sci Afr 2022;16:e01259 [FREE Full text] [doi: 10.1016/j.sciaf.2022.e01259] [Medline: 35765588]
6. Allagoa DO, Oriji PC, Tekenah ES, Obagah L, Njoku C, Afolabi AS, et al. Predictors of acceptance of COVID-19 vaccine among patients at a tertiary hospital in South-South Nigeria. Int J Community Med Public Health 2021;8(5):2165-2172 [FREE Full text] [doi: 10.18203/2394-6040.ijcmph20211733]
7. Ilesanmi OS, Akosile PO, Afolabi AA, Ukwenya VO. Handling distrust on risk communication in Nigeria: a strategy to strengthening the COVID-19 outbreak response. Disaster Med Public Health Prep 2021;17:e27 [FREE Full text] [doi: 10.1017/dmp.2021.245] [Medline: 34308825]
8. Babatope T, Ilyenkova V, Marais D. COVID-19 vaccine hesitancy: a systematic review of barriers to the uptake of COVID-19 vaccine among adults in Nigeria. Bull Natl Res Cent 2023;47(1):45 [FREE Full text] [doi: 10.1186/s42269-023-01017-w] [Medline: 36970323]
9. Lawal O. Understanding risk communication effectiveness from public interest, mobility, and COVID-19 cases: a case study of COVID-19 in Nigeria. Front Commun 2022;7:921648 [FREE Full text] [doi: 10.3389/fcomm.2022.921648]
10. Olu-Abiodun O, Abiodun O, Okafor N. COVID-19 vaccination in Nigeria: a rapid review of vaccine acceptance rate and the associated factors. PLoS One 2022;17(5):e0267691 [FREE Full text] [doi: 10.1371/journal.pone.0267691] [Medline: 35544545]
11. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
12. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev 2015;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
13. Agha S, Tollefson D, Paul S, Green D, Babigumira JB. Use of the fogg behavior model to assess the impact of a social marketing campaign on condom use in Pakistan. J Health Commun 2019;24(3):284-292. [doi: 10.1080/10810730.2019.1597952] [Medline: 30945612]
14. Evans WD, Gerard R, Symington L, Shaikh H, Agha S. Implementation practice models for development in low- and middle-income countries: systematic review of peer-reviewed literature. BMC Public Health 2022;22(1):1157 [FREE Full text] [doi: 10.1186/s12889-022-13530-0] [Medline: 35681165]
15. Agha S. Use of a practitioner-friendly behavior model to identify factors associated with COVID-19 vaccination and other behaviors. Vaccines (Basel) 2022;10(8):1261 [FREE Full text] [doi: 10.3390/vaccines10081261] [Medline: 36016149]
16. Pedersen B, Thanel K, Kouakou AY, Zo JR, Ouattara ML, Gbeke D, et al. Identifying drivers of COVID-19 vaccine uptake among residents of Yopougon Est, Abidjan, Côte d'Ivoire. Vaccines (Basel) 2022;10(12):2101 [FREE Full text] [doi: 10.3390/vaccines10122101] [Medline: 36560511]
17. Agha S, Chine A, Lalika M, Pandey S, Seth A, Wiyeh A, et al. Drivers of COVID-19 vaccine uptake amongst healthcare workers (HCWs) in Nigeria. Vaccines (Basel) 2021;9(10):1162 [FREE Full text] [doi: 10.3390/vaccines9101162] [Medline: 34696270]
18. Evans WD, Bingenheimer JB, Long M, Ndiaye K, Donati D, Rao NM, et al. Development and evaluation of an index to measure the ability to get vaccinated for COVID-19. Vaccines (Basel) 2023;11(2):342 [FREE Full text] [doi: 10.3390/vaccines11020342] [Medline: 36851220]
19. Agha S, Morgan B, Archer H, Paul S, Babigumira JB, Guthrie BL. Understanding how social norms affect modern contraceptive use. BMC Public Health 2021;21(1):1061 [FREE Full text] [doi: 10.1186/s12889-021-11110-2] [Medline: 34088295]
20. Meekers D, Onuoha C, Olutola O. Applying the Fogg Behavior Model to improve contraceptive social marketing during the COVID-19 lockdown in Nigeria: a case study. Gates Open Res 2020;4:141 [FREE Full text] [doi: 10.12688/gatesopenres.13186.1]
21. Fogg B. A behavior model for persuasive design. New York, NY, US: Association for Computing Machinery; 2009 Presented at: Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, California, USA p. 1-7. [doi: 10.1145/1541948.1541999]

XSL•FO
RenderX

22. Fogg BJ. Tiny Habits: The Small Changes that Change Everything. Boston, MA: Houghton Mifflin Harcourt; 2020.
23. Rao N, Donati D, Orozco V. Conducting Surveys and Interventions Entirely Online: A Virtual Lab Practitioner's Manual. The World Bank. 2020. URL: https://documents1.worldbank.org/curated/en/819191635501106862/pdf/Conducting-Surveys-and-Interventions-Entirely-Online-A-Virtual-Lab-Practitioner-s-Manual.pdf [accessed 2024-01-23]
24. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC; 2017.
25. Shaver LG, Khawer A, Yi Y, Aubrey-Bassler K, Etchegary H, Roebothan B, et al. Using Facebook advertising to recruit representative samples: feasibility assessment of a cross-sectional survey. J Med Internet Res 2019;21(8):e14021 [FREE Full text] [doi: 10.2196/14021] [Medline: 31429409]
26. Pedersen B, Thompson G, Kouakou AY, Mujinga M, Nicholes S, Martinez A, et al. Development and assessment of a six-item index to gauge motivation to receive COVID-19 vaccination. Vaccines 2023;12(1):6 [FREE Full text] [doi: 10.3390/vaccines12010006]

Abbreviations

aOR: adjusted odds ratio
FBM: Fogg Behavior Model
HCW: health care worker
HND: higher national diploma
LMICs: low- and middle-income countries
OND: ordinary national diploma

Edited by T Leung; submitted 02.04.23; peer-reviewed by J Smith, OO Oyewole; comments to author 27.04.23; revised version received 24.05.23; accepted 02.01.24; published 26.02.24.

Please cite as:
Agha S, Nsofor I, Bernard D, Francis S, Rao N
Behavioral Insights from Vaccine Adoption in Nigeria: Cross-Sectional Survey Findings
Interact J Med Res 2024;13:e47817
URL: https://www.i-jmr.org/2024/1/e47817
doi:10.2196/47817
PMID:38407956

©Sohail Agha, Ifeanyi Nsofor, Drew Bernard, Sarah Francis, Nandan Rao. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 26.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study

Zhuoga Deji[1,2], MSc; Yuantao Tong[1], PhD; Honglian Huang[1,3], PhD; Zeyu Zhang[1], PhD; Meng Fang[4], PhD; M James C Crabbe[5,6,7], PhD; Xiaoyan Zhang[1], PhD; Ying Wang[3], PhD

[1]Research Center for Translational Medicine, Shanghai East Hospital, School of Life Sciences and Technology, Tongji University, Shanghai, China

[2]Information School, The University of Sheffield, Sheffield, United Kingdom

[3]Department of Clinical Laboratory Medicine Center, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[4]Department of Laboratory Medicine, Shanghai Eastern Hepatobiliary Surgery Hospital, Shanghai, China

[5]Wolfson College, Oxford University, Oxford, United Kingdom

[6]Institute of Biomedical and Environmental Science & Technology, University of Bedfordshire, Bedfordshire, United Kingdom

[7]School of Life Sciences, Shanxi University, Shanxi, China

Corresponding Author:
Ying Wang, PhD
Department of Clinical Laboratory Medicine Center
Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine
110 Ganhe Rd
Shanghai, 200437
China
Phone: 86 021 65161782
Fax: 86 021 65161782
Email: nadger_wang@139.com

Abstract

Background: The novel coronavirus SARS-CoV-2 caused the global COVID-19 pandemic. Emerging reports support lower mortality and reduced case numbers in highland areas; however, comparative studies on the cumulative impact of environmental factors and viral genetic diversity on COVID-19 infection rates have not been performed to date.

Objective: The aims of this study were to determine the difference in COVID-19 infection rates between high and low altitudes, and to explore whether the difference in the pandemic trend in the high-altitude region of China compared to that of the lowlands is influenced by environmental factors, population density, and biological mechanisms.

Methods: We examined the correlation between population density and COVID-19 cases through linear regression. A zero-shot model was applied to identify possible factors correlated to COVID-19 infection. We further analyzed the correlation of meteorological and air quality factors with infection cases using the Spearman correlation coefficient. Mixed-effects multiple linear regression was applied to evaluate the associations between selected factors and COVID-19 cases adjusting for covariates. Lastly, the relationship between environmental factors and mutation frequency was evaluated using the same correlation techniques mentioned above.

Results: Among the 24,826 confirmed COVID-19 cases reported from 40 cities in China from January 23, 2020, to July 7, 2022, 98.4% (n=24,430) were found in the lowlands. Population density was positively correlated with COVID-19 cases in all regions (ρ=0.641, P=.003). In high-altitude areas, the number of COVID-19 cases was negatively associated with temperature, sunlight hours, and UV index (P=.003, P=.001, and P=.009, respectively) and was positively associated with wind speed (ρ=0.388, P<.001), whereas no correlation was found between meteorological factors and COVID-19 cases in the lowlands. After controlling for covariates, the mixed-effects model also showed positive associations of fine particulate matter (PM2.5) and carbon monoxide (CO) with COVID-19 cases (P=.002 and P<.001, respectively). Sequence variant analysis showed lower genetic diversity among nucleotides for each SARS-CoV-2 genome (P<.001) and three open reading frames (P<.001) in high altitudes compared to 300 sequences analyzed from low altitudes. Moreover, the frequencies of 44 nonsynonymous mutations and 32 synonymous mutations were significantly different between the high- and low-altitude groups (P<.001, mutation frequency>0.1). Key nonsynonymous

XSL•FO
RenderX

mutations showed positive correlations with altitude, wind speed, and air pressure and showed negative correlations with temperature, UV index, and sunlight hours.

Conclusions: By comparison with the lowlands, the number of confirmed COVID-19 cases was substantially lower in high-altitude regions of China, and the population density, temperature, sunlight hours, UV index, wind speed, PM2.5, and CO influenced the cumulative pandemic trend in the highlands. The identified influence of environmental factors on SARS-CoV-2 sequence variants adds knowledge of the impact of altitude on COVID-19 infection, offering novel suggestions for preventive intervention.

(Interact J Med Res 2024;13:e43585) doi:10.2196/43585

KEYWORDS

COVID-19; environmental factors; altitude; population density; virus mutation

Introduction

Background

In recent years, the outbreak of COVID-19 has had substantial impacts on human health, social life, and economic trends worldwide. To comprehensively explore the impact of potential factors on the strength and speed of viral infection, we first used the zero-shot model to screen the literature related to broad respiratory infectious diseases, including COVID-19, to identify the major influencing factors through text mining [1]. A significant ranking list of environmental factors was obtained (Multimedia Appendix 1), in which temperature and atmospheric pressure were highly correlated with respiratory viruses. Earlier findings showed that the spread of COVID-19 was linked to various factors, including the environment, sequence variants of the virus, and government countermeasures to protect public health in the face of outbreaks [2,3].

Previous studies also investigated the correlation between COVID-19 and altitude. COVID-19 transmission in high-altitude regions appears to differ from the global pattern, with a lower number of cases reported at high altitudes [4-9]. Moreover, population density was identified as a basic factor that significantly catalyzed the spread of COVID-19 in numerous countries, including India, the United States, China, and Malaysia [10-13]. However, few studies have performed an in-depth analysis of the effects of population density, altitude, and environmental factors on the variation in the severity of the COVID-19 pandemic among regions at different altitudes.

Moreover, virus mutation is another important factor in escaping the immune protection derived from a previous infection or vaccination [14-16]. Several studies have shown that people living at high altitudes may be less susceptible to developing severe adverse effects from COVID-19, along with reduced case fatality rates [4,8]. Environmental factors have also been shown to actively influence virus mutation and to play regulatory roles in viral evolution [17-20]. Nevertheless, the underlying physiological mechanism linking virus mutation and altitude that could affect the rate of COVID-19 transmission remains unclear.

Purpose

To fill this research gap, we aimed to explore the potential factors contributing to the outcome of COVID-19 at high altitudes in China in comparison to the lowlands. Toward this end, we first assessed the contributions of altitude and population density on the total number of COVID-19 cases at different altitudes, and further explored the correlation between COVID-19 cases and environmental factors.

In addition, a more detailed correlation analysis among regions at different altitudes was performed at the city level using mixed-effects multiple linear regression models controlling for potential covariates, including meteorological and air quality factors. Furthermore, we studied the genome diversity, mutation frequency, and correlations with environmental factors in high- and low-altitude regions. Overall, we aim to provide a better understanding of the key factors that could influence the cumulative infection and transmission rate of COVID-19 in high-altitude regions of China, which can in turn help to inform establishing improved policies for preventive interventions.

Methods

Study Area

We focused on the high-altitude southwestern regions of China and several low-altitude regions from mainland China. A more detailed description of each province included in the study area along with city-level information is provided in Multimedia Appendix 2.

Collection of Confirmed COVID-19 Cases

Data on COVID-19, including total confirmed cases at the province level from January 23, 2020, to July 7, 2022, were collected from the Dingxiangyuan (DXY-DX Doctor) website [21]; historical cases at the city level over the same study period were collected using the R package “nCovid2019.” The population density for each region was calculated by the following equation:

$$\text{Population density} = \text{population size/area of the land } (km^2) \quad \textbf{(1)}$$

The normalized daily number of confirmed COVID-19 cases was calculated using the following equation:

$$\text{Normalized COVID-19 cases} = \text{total number of confirmed cases/population density} \quad \textbf{(2)}$$

To reflect the COVID-19 infection situation in high-altitude regions of China during the study period, descriptive statistics were compiled for the daily average meteorological and air quality parameters, as shown in Multimedia Appendix 3.

XSL•FO
RenderX

Collection of Population, Meteorological, and Air Quality Factors

To analyze the correlation between altitude and COVID-19 cases in China, altitude information of highland regions (>1500 m) and lowland regions (<1500 m) was collected from the topographic map [22]. Meteorological factors were collected from the World Weather Online site [23]. Air quality variables were collected from the Statistical Yearbook for each municipality or province, along with associated data on population and land areas. Multimedia Appendix 4 provides a more detailed description of all these variables.

Collection of SARS-CoV-2 Genome Data

SARS-CoV-2 whole-genome sequencing data were collected from the GISAID (Global Initiative on Sharing All Influenza Data) website [24], including 300 sequences from high-altitude regions and 300 sequences randomly selected over the same time period from low-altitude regions. The accession numbers of the sequences included in this analysis are listed in Multimedia Appendix 5.

Statistical Analysis

The Kolmogorow-Smirnov test was performed for each variable (see Multimedia Appendix 6) with the null hypothesis of a normal distribution; since the *P* value of each variable was less than .05, the null hypothesis was rejected, indicating that the test distribution was not normally distributed. Therefore, we used the nonparametric Spearman correlation coefficient for the correlation analysis of environmental factors and mutation frequency with COVID-19 cases using SPSS software.

After taking into account the correlations for each of the independent variables, a mixed-effects multiple linear regression was used for an adjusted correlation analysis, with the model defined as follows:

$$Y = X\beta + Z\mu + \in \quad \textbf{(3)}$$

where $X\beta$ represents the fixed-effects set in this study, including the covariates meteorological and air quality, and $Z\mu$ represents an N×M design matrix containing each individual group (N) for each covariate (M) of the random effects. Four sets of mixed-effects models were analyzed using R software (version 4.2.1). Separate models were first run for each of the 12 random effects (environmental covariates), followed by two separate models that included all 6 meteorological indicators or air quality indicators as fixed effects simultaneously. The final model included all 12 indicators simultaneously.

Genome Analysis

The nucleotide sequences of the whole genome of SARS-CoV-2 were aligned to the reference sequence Wuhan-1 (NC_045512.2) using minimap2 2.17-r974. All mapped sequences were merged back with all others in a single alignment bam file. Variant calling was performed using bcftools mpileup v1.91. Gene sequences of SARS-CoV-2 were extracted and translated into amino acid sequences, which were aligned to the reference sequence by ClustalW. Variant calling was computed using an in-house–developed R script.

Sequence Diversity Calculation

Sequence diversity was calculated using the Shannon entropy (Sn) index in R software (version 4.2.1), which measures the diversity of nucleotides, amino acids, and their respective variant frequencies. The diversity of each nucleotide position (nucleotide 1 to 29,903) was calculated as the Sn according to the following formula [25]:

$$Sn = i\sum(pilnpi)/lnN \quad \textbf{(4)}$$

where pi represents the relative frequency of nucleotides or deletion at this position and N represents the total number of sequences. All high-altitude samples were compared with the low-altitude samples to identify differential sequence variations at both the nucleotide and amino acid levels.

Ethical Considerations

This study is based on an analysis of existing data and therefore received exemption from ethical approval (reference number 041797). Some data sets were slightly modified after the study received exemption from ethical approval; however, the updated data were collected from the same publicly available database. According to the University Research Ethics Policy, “the project does not involve human participants which will only use publicly available anonymized data; a project which will only use existing clinical or research data that has been robustly anonymized such that it no longer constitutes personal data” [26] (pp. 3.1.10, p. 20-21); therefore, ethical approval was not required for the above circumstance. In addition, Article 32 of the Ethical Review Measures for Human Life Science and Medical Research of the People’s Republic of China declares that “research involving human life science and medicine under the circumstances of using anonymized information data, legally obtained publicly available data, or data generated through observation without interference with public behavior, which do not cause harm to the human body and do not involve sensitive personal information or commercial interests, can be exempt from ethical review to reduce unnecessary burdens on researchers and promote the development of human life science and medical research” [27] (chapter 3). Therefore, this study falls within the exempt category from the institutional review board. Further ethical approval by a review committee was not required since the data utilized in this study have been sourced from a publicly accessible database, ensuring full anonymization, and the research process involved no direct interaction with human subjects, solely relying on the analysis of pre-existing and publicly available data, as mentioned in the above legislation.

Results

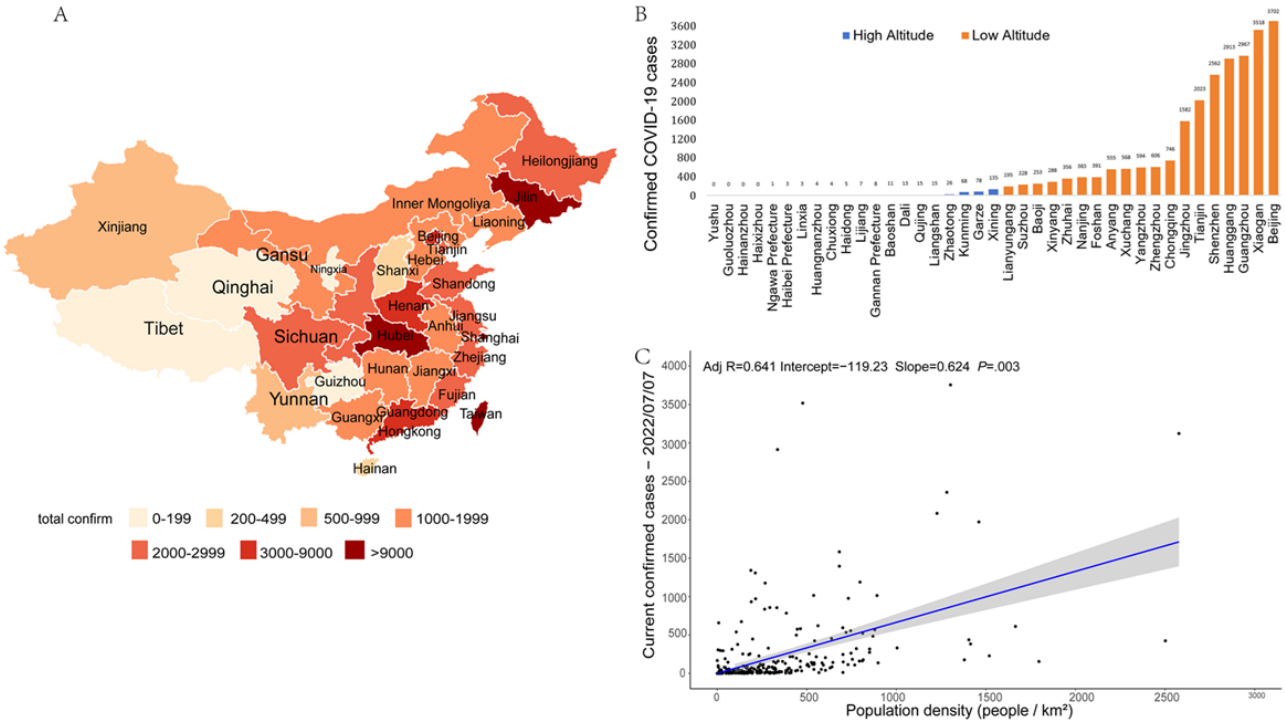
Correlation Between Altitude and COVID-19 Cases

During the study period from January 23, 2020, to June 7, 2022, a total of 732 COVID-19 cases were officially reported in the high-altitude region of China, which excluded one imported case in Tibet. The total numbers of confirmed COVID-19 cases were lower in the high-altitude regions, including Tibet, Qinghai, and Gansu, than in the low-altitude regions (Figure 1A); despite variation among regions, 98.4% (24,430/24,826) of the total COVID-19 cases were found in the lowland regions (Figure

XSL•FO
RenderX

1B). Taking into account the hidden factors of the dimension of population size (Multimedia Appendix 2) in different areas, a significant positive correlation between confirmed COVID-19 cases and population density (Figure 1C) was found during the study period.

Figure 1. Pattern of confirmed COVID-19 cases at high altitudes within the study period. (A) Geographic patterns of confirmed COVID-19 cases in China as of July 7, 2022. (B) Comparison of confirmed COVID-19 cases reported in different cities at different altitudes. (C) Linear regression analysis between population density and confirmed COVID-19 cases in China.

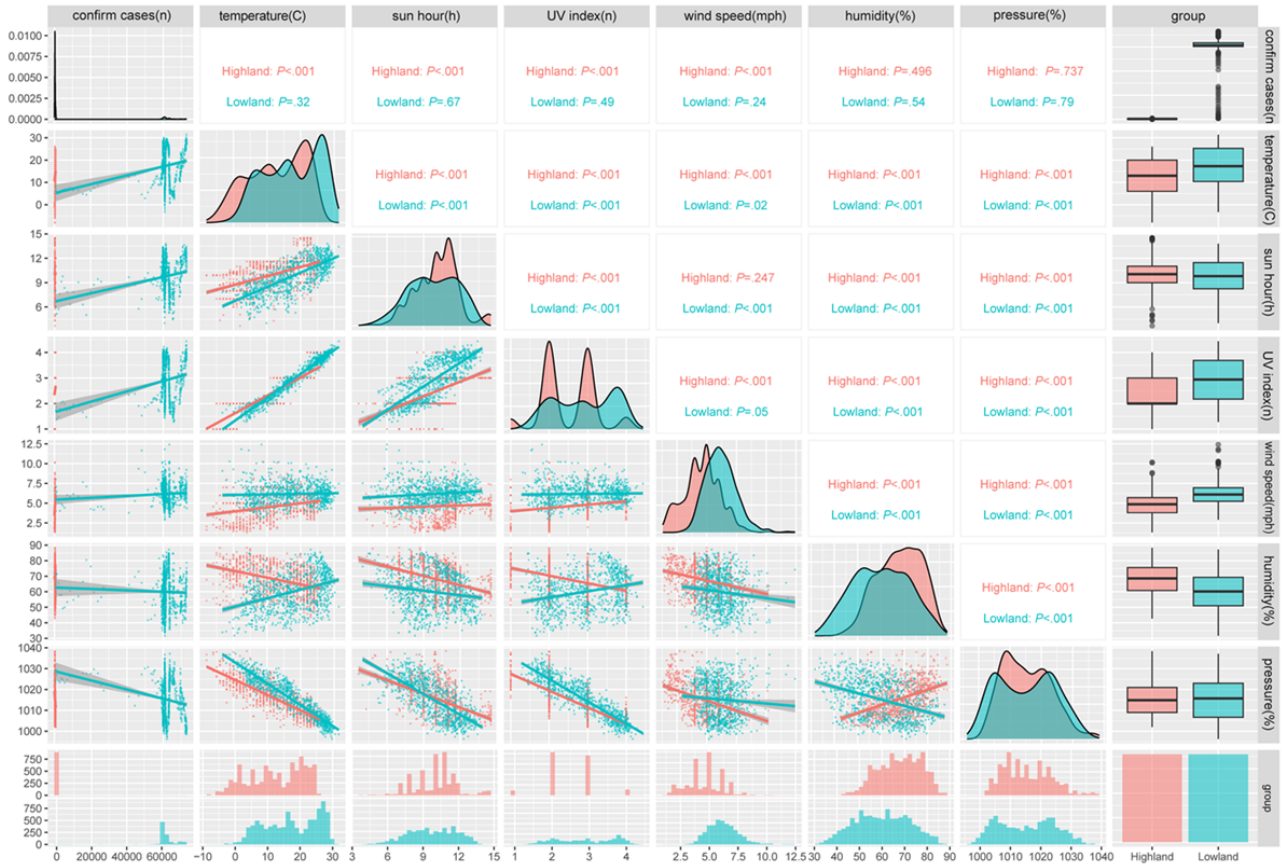


Correlation Between Meteorological Factors and Confirmed COVID-19 Cases in High- and Low-Altitude Regions

To gain a better understanding of the correlation of meteorological factors with the spread of COVID-19 in high-altitude regions when compared with that in the lowlands, we performed a comparative correlation analysis of each variable (Figure 2); the detailed numeric results are provided in Multimedia Appendix 7. In the highlands of China, average temperature (P=.003), sunlight hours (P<.001), and UV index (P=.009) were negatively correlated with the number of confirmed COVID-19 cases, whereas the wind speed (ρ=0.388, P<.001) was positively correlated with confirmed cases.

XSL•FO
RenderX

Figure 2. Spearman correlation analysis of confirmed COVID-19 cases and meteorological factors in high-altitude and low-altitude regions in China. The color represents each region group. Area charts represent frequencies and the scatterplots with lines represent the correlation. *P* values are based on 2-tailed tests; also see Multimedia Appendix 7.

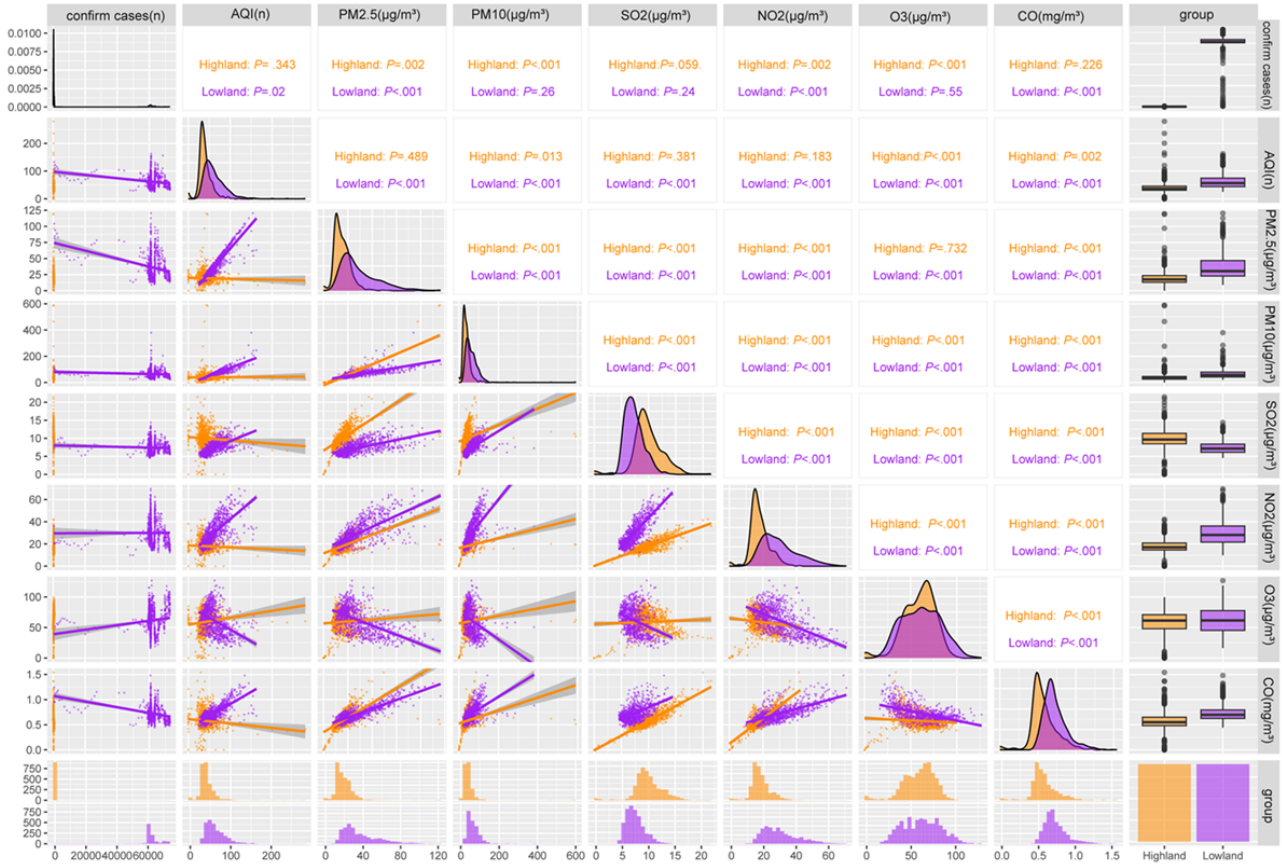


Correlation Between Air Quality Factors and Confirmed COVID-19 Cases in High- and Low-Altitude Regions

As shown in Figure 3, we found significant correlations between air quality factors and COVID-19 cases at high and low altitudes. In the highlands, COVID-19 cases were positively correlated with fine particulate matter (PM2.5; *P*=.002), coarse particulate matter (PM10; *P*<.001), nitrogen dioxide (NO2; *P*=.02), and ozone (O3; *P*<.001). In the lowlands, a negative correlation was found between confirmed COVID-19 cases and air quality index (AQI; *P*=.01), whereas positive correlations were found for PM2.5 (*P*<.001), NO2 (*P*<.001), and carbon monoxide (CO; *P*<.001); in contrast to the pattern in the highlands, the average daily concentrations of PM10 and O3 were not significantly correlated with COVID-19 cases in lowland regions of China. The detailed results are shown in Multimedia Appendix 7.

XSL•FO
RenderX

Figure 3. Spearman correlation analysis of confirmed COVID-19 cases and air quality factors in high- and low-altitude regions in China. The color represents each region group. Area charts represent frequencies and the scatterplot with lines represent correlations. AQI: air quality index; CO: carbon monoxide; NO2: nitrogen dioxide; O3: ozone; PM2.5: fine particulate matter; PM10: coarse particulate matter; SO2: sulfur dioxide. *P* values are based on 2-tailed tests; also see Multimedia Appendix 7.



Correlations of Meteorological and Air Quality Factors With Normalized Confirmed COVID-19 Cases After Covariate Adjustment

The results of the mixed-effects multiple linear regression among the 12 provinces (Multimedia Appendix 2) at different altitudes are summarized in Table 1. In model 1, including each of the fixed-effect indicators in 12 separate models, a lower number of confirmed COVID-19 cases was significantly associated with average temperature, sunlight hours, UV index, air pressure, and air quality factors, including PM2.5, O3, and CO. However, in model 2, including all of the meteorological covariates considered simultaneously, only average temperature, UV index, and air pressure were significantly associated with the normalized confirmed COVID-19 cases. In model 3, including all of the air quality covariates considered simultaneously, only the average concentrations of O3 and CO were significantly correlated with confirmed COVID-19 cases. Finally, in model 4, including all 12 environmental indicators considered simultaneously, only average temperature, wind speed, and CO were significantly associated with confirmed COVID-19 cases.

XSL•FO
RenderX

Table 1. Mixed-effects multiple linear regression of the associations of normalized COVID-19 cases with meteorological and air quality factors including adjustment for covariates.

Fixed effects	Model 1[a]		Model 2[b]		Model 3[c]		Model 4[d]	
	Estimate (SE)	*P* value	Estimate (SE)	*P* value	Estimate (SE)	*P* value	Estimate (SE)	*P* value
Average temperature (°C)	–15.41 (2.55)	.001	–21.44 (7.47)	.004	__[e]	—	–27.90 (7.72)	.001
Sunlight hours	–45.31 (9.50)	.001	–12.55 (15.04)	.40	—	—	–8.86 (15.31)	.56
UV index	–201.24 (27.53)	.001	–257.41 (71.06)	.001	—	—	249.03 (71.84)	.05
Wind speed	7.934 (10.81)	.46	16.72 (11.27)	.13	—	—	16.71 (12.00)	.02
Humidity	–0.071 (1.42)	.96	–1.99 (1.81)	.27	—	—	1.9 8 (1.94)	.31
Pressure	17.64 (2.51)	.001	20.35 (5.51)	.001	—	—	24.32 (5.58)	.05
AQI[f]	–1.45 (0.67)	.03	—	—	–0.05 (0.88)	.94	0.11 (0.89)	.90
PM2.5[g]	4.71 (0.99)	.001	—	—	4.28 (1.35)	.15	2.96 (1.36)	.05
PM10[h]	0.36 (0.47)	.43	—	—	0.49 (0.58)	.39	0.52 (0.58)	.36
SO2[i]	6.93 (6.01)	.24	—	—	18.41 (6.75)	.64	22.06 (6.94)	.14
NO2[j]	1.88 (1.612)	.24	—	—	8.69 (1.97)	.14	10.21 (2.07)	.07
O3[k]	2.29 (0.77)	.002	—	—	2.09 (0.8)	.008	0.56 (0.86)	.51
CO[l]	375.4 (75.5)	.001	—	—	471.52 (101.55)	.001	564.31 (107.99)	.001

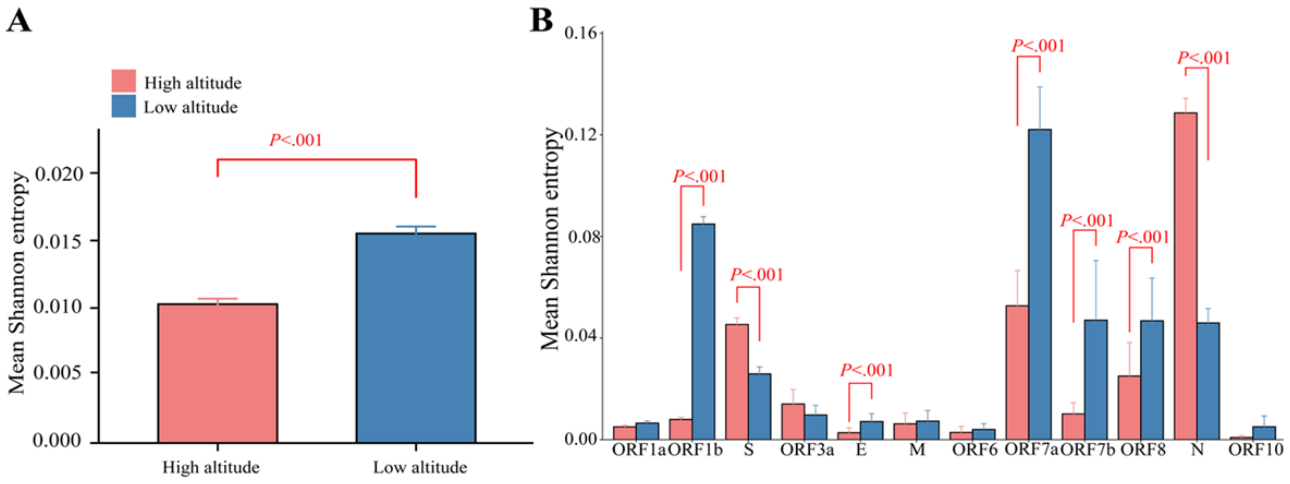
[a]Model 1 includes each of the fixed effects run in 12 separate models.
[b]Model 2 includes all meteorological covariates considered simultaneously.
[c]Model 3 includes all air quality covariates considered simultaneously.
[d]Model 4 includes all 12 environmental indicators considered simultaneously.
[e]Excluded from model.
[f]AQI: air quality index.
[g]PM2.5: fine particulate matter.
[h]PM10: coarse particulate matter.
[i]SO2: sulfur dioxide.
[j]NO2: nitrogen dioxide.
[k]O3: ozone.
[l]CO: carbon monoxide.

Diversity of SARS-CoV-2 Sequences in High- and Low-Altitude Groups

Comparable sequence diversity was found among the full-length SARS-CoV-2 genomes and 12 genes of SARS-CoV-2 in the high- and low-altitude groups (Figure 4). In the high-altitude group, the Sn values of nucleotides for each site were significantly lower than those in the low-altitude group (Figure 4A; $P<.001$). Likewise, the Sn values of amino acids for three open reading frames (ORFs; ORF1b, ORF7a, and ORF7b) in the high-altitude group were lower than those in the low-altitude group (Figure 4B; $P<.001$). Compared to those of the low-altitude group, the Sn values of the S and N genes were higher in the high-altitude group.

XSL•FO
RenderX

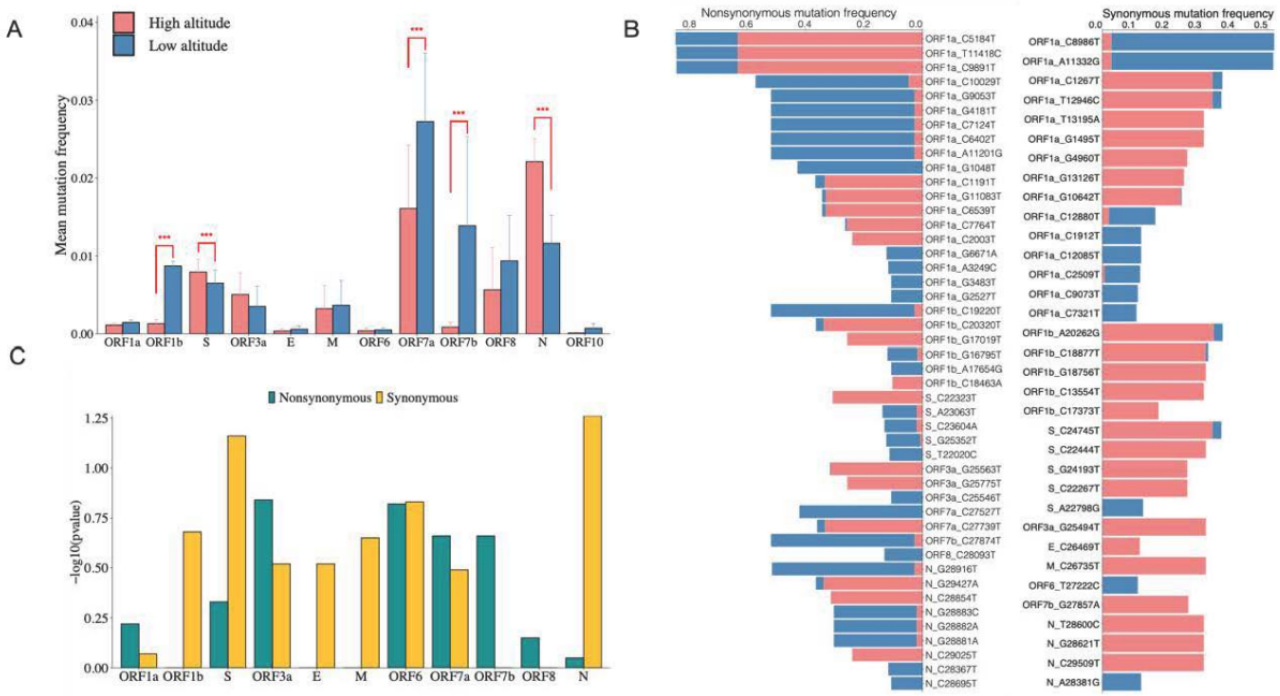
Figure 4. Mean sequence diversity across high-altitude (red) and low-altitude (blue) groups identified by Shannon entropy in nucleotides and different amino acids of the SARS-CoV-2 genome. (A) Mean sequence diversity identified by Shannon entropy in nucleotides of the whole SARS-CoV-2 genome. (B) Mean sequence diversity identified by Shannon entropy in different amino acids. Error bars indicate SEM. *P* values are based on the Wilcoxon signed rank test.



Comparison of SARS-CoV-2 Gene Variant Frequencies in the High- and Low-Altitude Groups

To determine the gene variant frequency difference between the high- and low-altitude groups, the Wilcoxon signed rank test was used to evaluate the significance of variant frequency in each of the ORFs (Figure 5A). The amino acid mutation frequencies for the three ORFs (ORF1b, ORF7a, and ORF7b) were significantly lower in the high-altitude group than in the low-altitude group (*P*<.001) but were higher for the S and N genes (*P*<.001). Among these, 44 nonsynonymous mutations and 32 synonymous mutations were found between the high- and low-altitude groups (Figure 5B; *P*<.001; mutation frequency>0.1). There was a greater proportion of nonsynonymous mutations in the low-altitude group than in the high-altitude group, whereas the opposite pattern was found for synonymous mutations. The –log10 *P* values for the differences in nonsynonymous and synonymous mutations for each ORF between the high- and low-altitude groups are shown in Figure 5C. Compared to the result in Figure 5A, the synonymous mutations of the S genes and N genes had relatively higher significance scores in the high-altitude group.

Figure 5. Distribution of mutations of SARS-CoV-2 in the high- and low-altitude groups. (A) Mean mutation frequency of open reading frames (ORFs) in the high- and low-altitude groups. (B) Comparison of frequencies of nonsynonymous and synonymous mutations between the high- and low-altitude groups. (C) Significance scores of nonsynonymous and synonymous mutations in each ORF. ***P<.001.

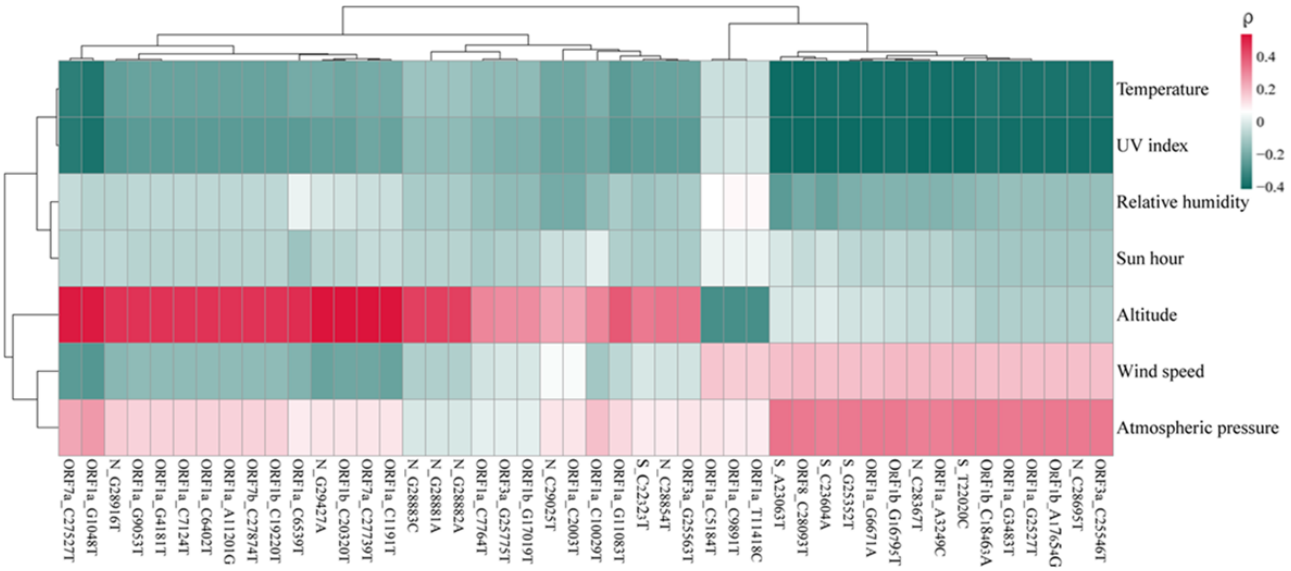


XSL•FO
RenderX

Correlation Analysis Between Mutation Frequency and Environmental Factors

The impact of 7 environmental factors on the nonsynonymous mutation frequency was analyzed for all sequences (Figure 6 and Multimedia Appendix 8). Most of the nonsynonymous mutations had a positive correlation with altitude, wind speed, and atmospheric pressure, but had negative correlations with UV index, relative humidity, and sunlight hours. The highest correlation between mutation frequency and environmental factors was found for altitude with N_G29427A, ORF1b_C20320T, and ORF1a_C1191T (ρ=0.53, *P*<.001); followed by UV index with ORF8_C28093T and S_A23063T (ρ=–0.42, *P*<.001); and temperature with ORF8_C28093T, S_C23604A, S_A23063T, and ORF1a_G6671A. Some significant nonsynonymous mutational events were also discovered in this study, including C22323T in the receptor-binding domain (RBD) of the S gene, which had a positive correlation with altitude (ρ=0.32, *P*<.001), indicating higher frequencies at high altitude. S gene mutations A23063T in the RBD region and C23604A in the fusion peptide (FP) region had higher frequencies at low altitude, and demonstrated a significant positive correlation with atmospheric pressure (ρ=0.34 and ρ=0.31, respectively; both *P*<.001) and a significant negative correlation with temperature and UV index (ρ=–0.42 and ρ=–0.41, respectively; both *P*<.001). Furthermore, co-occurring point mutations in the N gene, specifically G28881A, G28882A, and G28883C (R203K and G204R), had a higher mutation frequency at low altitude and were positively related to altitude (ρ=0.43, *P*<.001).

Figure 6. Spearman correlation coefficients of environmental factors and mutation frequencies among different sites in the high- and low-altitude groups. The color represents the strength of the Spearman correlation. Also see Multimedia Appendix 8.



Discussion

Principal Findings

This study found that the lower number of confirmed COVID-19 cases in high-altitude regions of China may be related to population density and environmental factors. By further exploring SARS-CoV-2 sequences, we found different mutation frequencies in the high- and low-altitude regions, which were also correlated with environmental factors.

Population density is one of the most effective predictors related to the regional pandemic; a larger population greatly increases the infection and transmission rates of COVID-19 [28]. Therefore, our results suggest that when analyzing the correlation between COVID-19 cases in regions with different population dimensions and other possible factors, it is important to take into account the *normalized* number of COVID-19 cases to study the pandemic trend.

Meteorological factors may also influence COVID-19 transmission and cumulative infection in high-altitude regions. Previous studies indicated that higher temperature and UV radiation could contribute to a decrease in new cases of COVID-19 infection at high altitude [29-31]. This is supported by our Spearman correlation analysis showing that the average daily temperature, sunlight hours, and UV index were negatively related to the normalized COVID-19 cases in high-altitude regions. A high-altitude environment is distinguished by lower temperatures compared with those of low-altitude regions, along with significant differences between daylight hours, air dryness, and levels of UV light radiation [32-34]. Importantly, our study further found that the wind speed in the high-altitude region of China was significantly associated with the reducing trend of COVID-19 cases, which is consistent with earlier studies conducted in Italy, New York, and Singapore [2,35]. Analyzing the correlation between covariates in regions at different altitudes with different population dimensions showed positive correlations of average temperature and wind speed with normalized confirmed COVID-19 cases, in line with previous research.

Air pollution is widely recognized as a major risk factor for respiratory infection in humans, which has also played a significant role in the spread of COVID-19. Previous research suggested that the concentrations of NO2, PM2.5, PM10, and O3 are positively correlated with the number of confirmed COVID-19 cases [33,34,36-38]. Our Spearman correlation results align with these previous findings. NO2 is often linked

XSL•FO
RenderX

with vehicle emissions and energy production [39], which is also an irritant of human respiratory diseases. Early research findings [40] support the results of this study, suggesting that travel restrictions should be among the specific actions implemented to reduce the spread of COVID-19. Previous studies indicated that PM2.5 and PM10 levels were positively correlated to the number of new daily confirmed COVID-19 cases in mainland China [41-43]. Consistently, our findings showed that lower average concentrations of PM2.5 and PM10 were related to a lower number of confirmed COVID-19 cases in the highlands.

It is clear that all of the above correlated factors may have influenced the cumulative infection trends of the pandemic in the highlands of China. The main sources of CO are motor vehicles and industrial source emissions; consequently, CO is closely related to population density and human activity. Therefore, taking into account the mixed correlations between all the covariates of focus in this study by including the related factors into a single mixed-effects model, among the air quality factors, only the average concentration of CO was positively associated with the normalized number of confirmed COVID-19 cases during the study period. This result also aligns with previous research conducted in China [43,44].

In terms of variations in the SARS-CoV-2 genome at high and low altitudes, the whole-genome sequences at high altitude showed lower diversity based on the mean Sn values of nucleotides and different amino acids compared to those in the low-altitude sites. Previous studies have suggested an impact of high altitude on the pathophysiology of COVID-19 [44,45], implying improved tolerance to SARS-CoV-2 infection for residents of high-altitude regions. Our study identified differences in the frequencies of SARS-CoV-2 sequence variants in various ORFs between high- and low-altitude regions, along with correlations between mutation frequency and environmental factors. Most of the nonsynonymous mutations identified have been reported previously and have important biological implications. We found that only the C22323T variant of the S gene had a positive correlation with altitude, with a greater frequency in the high-altitude group. Sivasubramanian et al [46] found that the S255F (C22323T) variant could reduce the affinity between the S protein and antibodies. We also identified the variant A23063T in the RBD region and the variant C23604A in the FP region of the S gene, with higher mutation frequencies detected at low altitude, showing a significantly positive correlation with atmospheric pressure and a strongly negative correlation with temperature and UV index. Previous studies found that the N501Y (A23063T) and Q493H variants enhanced the binding affinity to the human angiotensin converting-enzyme 2 receptor, thereby increasing infectivity [47], and that the Omicron peptide with variants N679K and P681H (C23604A) might increase viral infectivity and transmissibility [48]. These results suggest that atmospheric pressure, UV index, and temperature may affect the infectivity of COVID-19 driven by specific mutations of the S gene in low-altitude regions. In addition, three co-occurring variants (G28881A, G28882A, G28883C) in the N gene had higher mutation frequencies in the low-altitude group and were positively related to altitude; these mutations were previously reported to destabilize and decrease the overall structural flexibility of the SARS-CoV-2 genome [48].

Notably, the potential impact of the strategy implemented by the government in response to the pandemic in high-altitude regions on the observed differences should not be ignored. This study only included data collected from January 22, 2020, to March 19, 2021, during which time there was only one imported COVID-19 case in Tibet that was excluded from the analysis. However, after several consecutive months of no new cases, a newly confirmed case was reported in Tibet on August 7, 2022. Since then, eight new cases have been added with a total of 1437 confirmed cases reported to date [21]. Several reasons for this new outbreak are worth discussing. First, the virus underlying the epidemic in Tibet is Omicron BA.2.76, which has characteristics of strong, fast transmission ability and the potential to more readily escape from immune protection. Second, the first discovered infection on August 7 was a close contact of a family, with suspected transmission occurring at a family gathering, and August is the peak tourist season in Tibet. Thus, under the conditions of relatively relaxed travel restrictions, the close contact and population flow greatly increased the spread of the virus. Third, the lack of epidemic prevention and medical appliances in high-altitude regions limited access to health care systems, and the limited capacity for viral testing and contact tracing worsened the situation of the pandemic in Tibet. Therefore, effective and accurate government policies have played an important role in preventing Tibet from being affected by the epidemic, and the government policy in response to the pandemic must be considered; indeed, the response needs to be strengthened, although China indicated that they had ceased counting COVID-19 cases on December 23, 2022, and stated that the COVID-19 pandemic is effectively now over [49].

Limitations

Our study has some limitations. First, among the data available for confirmed COVID-19 cases, we only excluded the one known imported case in Tibet during the study period as of March 19, 2021, whereas the collected COVID-19 data from the other four areas may include both local and imported cases. This is mainly due to a lack of accurate publicly available COVID-19 data at the city level. Similarly, the incompleteness of some city-level information of genome sequences also led to difficulties in studying the differences in genetic variants among cities in the same province. Second, we aimed to evaluate whether virus mutations are associated with particular viral lineages. The sequences of SARS-CoV-2 in the high- and low-altitude groups represented 33 different virus lineages (Multimedia Appendix 9); B.1.617.2, AY.122, and B.1.36.16 were the major lineages, and the mutations of each lineage in which the mutation frequency was greater than 0.5 are listed in Multimedia Appendix 10. The number of different lineages between high- and low-altitude groups is also unbalanced. Third, we lacked the clinical information to study the potential relationships between mutations and clinical outcomes of infection at high and low altitudes. Therefore, future studies should focus on exploring the underlying mechanisms contributing to the links between patterns of SARS-CoV-2 mutation and case numbers at different altitudes.

XSL•FO
RenderX

Conclusions

Compared to that in the lowland area of China, the total number of confirmed COVID-19 cases in the highland was substantially lower. Population density and environmental factors, including average temperature, sunlight hours, UV index, wind speed, NO2, O3, PM2.5, and CO, were identified as indicators with a significant influence on the cumulative pandemic trend in the highlands. Among these factors, average temperature and CO were identified as the major meteorological and air quality factors associated with the spread of COVID-19 infection in China. Furthermore, we identified different mutations from SARS-CoV-2 isolates between high- and low-altitude regions, and there was a significant impact of environmental factors on virus mutation. Overall, this study adds important knowledge of the impact of altitude and related environmental factors on the cumulative infection rate of COVID-19, providing novel suggestions for preventive interventions.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (81972914, 81573023), Innovation Group Project of Shanghai Municipal Health Commission (2019CXJQ03), Fundamental Research Funds for the Central Universities (22120200014), and Shanghai “Rising Stars of Medical Talent” Youth Development Program (2019-72).

Data Availability

The data sets generated and analyzed during the study are provided in the Multimedia Appendix files.

Authors' Contributions

YW and XZ supervised the study and reviewed the draft of the paper. ZD and YT performed the data collection, main statistical analysis, and manuscript writing. HH, MF, and MJCC participated in manuscript verification, revisions, and editing. All authors had full access to all the data in the study and accept the responsibility to submit it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Key environmental factors associated with respiratory infectious diseases reported in the literature.
[DOCX File , 52 KB - ijmr_v13i1e43585_app1.docx]

Multimedia Appendix 2
Detailed characteristics of the regions included in the study area in the high- and low-altitude groups.
[DOCX File , 19 KB - ijmr_v13i1e43585_app2.docx]

Multimedia Appendix 3
Descriptive statistics of all variables collected during the study period.
[DOCX File , 14 KB - ijmr_v13i1e43585_app3.docx]

Multimedia Appendix 4
Description of all variables analyzed in this study.
[DOCX File , 14 KB - ijmr_v13i1e43585_app4.docx]

Multimedia Appendix 5
SARS-CoV-2 sequences downloaded from the GISAID database for analysis in this study.
[XLSX File (Microsoft Excel File), 86 KB - ijmr_v13i1e43585_app5.xlsx]

Multimedia Appendix 6
One-sample Kolmogorov-Smirnov test of normality for each variable.
[XLSX File (Microsoft Excel File), 11 KB - ijmr_v13i1e43585_app6.xlsx]

Multimedia Appendix 7
Spearman correlation analysis of meteorological factors and air quality factors with confirmed COVID-19 cases in high-altitude and low-altitude regions.
[XLSX File (Microsoft Excel File), 33 KB - ijmr_v13i1e43585_app7.xlsx]

XSL•FO
RenderX

Multimedia Appendix 8
Spearman correlation analysis of environmental factors and mutation frequency in the high- and low-altitude groups.
[XLSX File (Microsoft Excel File), 17 KB - ijmr_v13i1e43585_app8.xlsx]

Multimedia Appendix 9
Frequencies of nonsynonymous and synonymous mutations of the main lineages of SARS-CoV-2 in the high- and low-altitude groups.
[XLSX File (Microsoft Excel File), 10 KB - ijmr_v13i1e43585_app9.xlsx]

Multimedia Appendix 10
Number of samples of different SARS-CoV-2 lineages in the high-altitude (red) and low-altitude (blue) regions.
[DOCX File , 64 KB - ijmr_v13i1e43585_app10.docx]

References

1. Zhang Z, Fang M, Wu R, Zong H, Huang H, Tong Y, et al. Large-scale biomedical relation extraction across diverse relation types: model development and usability study on COVID-19. J Med Internet Res 2023 Sep 20;25:e48115 [FREE Full text] [doi: 10.2196/48115] [Medline: 37632414]
2. Bashir MF, Ma B, Bilal, Komal B, Bashir MA, Tan D, et al. Correlation between climate indicators and COVID-19 pandemic in New York, USA. Sci Total Environ 2020 Aug 01;728:138835 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138835] [Medline: 32334162]
3. Yao Y, Pan J, Liu Z, Meng X, Wang W, Kan H, et al. No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. Eur Respir J 2020 May;55(5):55 [FREE Full text] [doi: 10.1183/13993003.00517-2020] [Medline: 32269084]
4. Cano-Pérez E, Torres-Pacheco J, Fragozo-Ramos MC, García-Díaz G, Montalvo-Varela E, Pozo-Palacios JC. Negative correlation between altitude and COVID-19 pandemic in Colombia: a preliminary report. Am J Trop Med Hyg 2020 Dec;103(6):2347-2349 [FREE Full text] [doi: 10.4269/ajtmh.20-1027] [Medline: 33124543]
5. Chen H, Qin L, Wu S, Xu W, Gao R, Zhang X. Clinical characteristics and laboratory features of COVID-19 in high altitude areas: a retrospective cohort study. PLoS One 2021;16(5):e0249964 [FREE Full text] [doi: 10.1371/journal.pone.0249964] [Medline: 34003821]
6. Hakami AR, Dobie G. Studying the effect of particulate matter as SARS-CoV-2 transmitters. J Public Health Res 2021 Oct 21;11(1):2521 [FREE Full text] [doi: 10.4081/jphr.2021.2521] [Medline: 34674518]
7. Segovia-Juarez J, Castagnetto JM, Gonzales GF. High altitude reduces infection rate of COVID-19 but not case-fatality rate. Respir Physiol Neurobiol 2020 Oct;281:103494 [FREE Full text] [doi: 10.1016/j.resp.2020.103494] [Medline: 32679369]
8. Arias-Reyes C, Zubieta-DeUrioste N, Poma-Machicao L, Aliaga-Raduan F, Carvajal-Rodriguez F, Dutschmann M, et al. Does the pathogenesis of SARS-CoV-2 virus decrease at high-altitude? Respir Physiol Neurobiol 2020 Jun;277:103443 [FREE Full text] [doi: 10.1016/j.resp.2020.103443] [Medline: 32333993]
9. Ortiz-Prado E, Simbaña-Rivera K, Fernandez-Naranjo R, Vásconez JE, Henriquez-Trujillo AR, Vallejo-Janeta AP, UDLA-COVID-19 Team. SARS-CoV-2 viral load analysis at low and high altitude: a case study from Ecuador. Int J Environ Res Public Health 2022 Jun 28;19(13):7945 [FREE Full text] [doi: 10.3390/ijerph19137945] [Medline: 35805606]
10. Bhadra A, Mukherjee A, Sarkar K. Impact of population density on Covid-19 infected and mortality rate in India. Model Earth Syst Environ 2021;7(1):623-629 [FREE Full text] [doi: 10.1007/s40808-020-00984-7] [Medline: 33072850]
11. Ganasegeran K, Jamil MFA, Ch'ng ASH, Looi I, Pearlasamy KM. Influence of population density for COVID-19 spread in Malaysia: an ecological study. Int J Environ Res Public Health 2021 Sep 18;18(18):9866 [FREE Full text] [doi: 10.3390/ijerph18189866] [Medline: 34574790]
12. Sun Z, Zhang H, Yang Y, Wan H, Wang Y. Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. Sci Total Environ 2020 Dec 01;746:141347 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.141347] [Medline: 32755746]
13. Sy KTL, White LF, Nichols BE. Population density and basic reproductive number of COVID-19 across United States counties. PLoS One 2021 Apr 21;16(4):e0249271 [FREE Full text] [doi: 10.1371/journal.pone.0249271] [Medline: 33882054]
14. Mengist HM, Kombe Kombe AJ, Mekonnen D, Abebaw A, Getachew M, Jin T. Mutations of SARS-CoV-2 spike protein: implications on immune evasion and vaccine-induced immunity. Semin Immunol 2021 Jun;55:101533 [FREE Full text] [doi: 10.1016/j.smim.2021.101533] [Medline: 34836774]
15. Ramesh S, Govindarajulu M, Parise RS, Neel L, Shankar T, Patel S, et al. Emerging SARS-CoV-2 variants: a review of its mutations, its implications and vaccine efficacy. Vaccines 2021 Oct 18;9(10):1195 [FREE Full text] [doi: 10.3390/vaccines9101195] [Medline: 34696303]

XSL•FO
RenderX

16. Tong Y, Tan F, Huang H, Zhang Z, Zong H, Xie Y, et al. ViMRT: a text-mining tool and search engine for automated virus mutation recognition. Bioinformatics 2023 Jan 01;39(1):btac721 [FREE Full text] [doi: 10.1093/bioinformatics/btac721] [Medline: 36342236]
17. Hao Z, Li R, Hao C, Zhao H, Wan X, Guo D. Global evidence of temperature acclimation of COVID-19 D614G linage. Glob Chall 2021 Jun;5(6):2000132 [FREE Full text] [doi: 10.1002/gch2.202000132] [Medline: 33786200]
18. Sharif N, Dey SK. Impact of population density and weather on COVID-19 pandemic and SARS-CoV-2 mutation frequency in Bangladesh. Epidemiol Infect 2021 Jan 07;149:e16 [FREE Full text] [doi: 10.1017/S0950268821000029] [Medline: 33407987]
19. Sharif N, Sarkar MK, Ahmed SN, Ferdous RN, Nobel NU, Parvez AK, et al. Environmental correlation and epidemiologic analysis of COVID-19 pandemic in ten regions in five continents. Heliyon 2021 Mar;7(3):e06576 [FREE Full text] [doi: 10.1016/j.heliyon.2021.e06576] [Medline: 33817387]
20. Yang T, Yu P, Chang Y, Hsu SD. D614G mutation in the SARS-CoV-2 spike protein enhances viral fitness by desensitizing it to temperature-dependent denaturation. J Biol Chem 2021 Oct;297(4):101238 [FREE Full text] [doi: 10.1016/j.jbc.2021.101238] [Medline: 34563540]
21. COVID-19 Global Pandemic Real-time Report. DXY-DX Doctor. URL: https://ncov.dxy.cn/ncovh5/view/en_pneumonia?from=dxy&source=&link=&share= [accessed 2024-03-13]
22. Topographic map. URL: https://en-gb.topographic-map.com/map-d821h/China/ [accessed 2024-03-13]
23. World Weather API and Weather Forecast. World Weather Online. URL: https://www.worldweatheronline.com/ [accessed 2024-03-13]
24. Global Initiative on Sharing All Influenza Data (GISAID). URL: https://www.gisaid.org/ [accessed 2024-03-13]
25. Zagordi O, Däumer M, Beisel C, Beerenwinkel N. Read length versus depth of coverage for viral quasispecies reconstruction. PLoS One 2012 Oct 3;7(10):e47046 [FREE Full text] [doi: 10.1371/journal.pone.0047046] [Medline: 23056573]
26. Ethic policy governing research involving human participants personal data and human tissue: general principles and statements. The University of Sheffield. URL: https://www.sheffield.ac.uk/rpi/ethics-integrity/policy [accessed 2024-03-13]
27. Ethical review measures for human life science and medical research, Chapter 3, Article 32. People's Republic of China. URL: https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm [accessed 2024-03-13]
28. Rocklöv J, Sjödin H. High population densities catalyse the spread of COVID-19. J Travel Med 2020 May 18;27(3):taaa038 [FREE Full text] [doi: 10.1093/jtm/taaa038] [Medline: 32227186]
29. Li H, Xu X, Dai D, Huang Z, Ma Z, Guan Y. Air pollution and temperature are associated with increased COVID-19 incidence: a time series study. Int J Infect Dis 2020 Aug;97:278-282 [FREE Full text] [doi: 10.1016/j.ijid.2020.05.076] [Medline: 32502664]
30. Pirouz B, Shaffiee Haghshenas S, Pirouz B, Shaffiee Haghshenas S, Piro P. Development of an assessment method for investigating the impact of climate and urban parameters in confirmed cases of COVID-19: a new challenge in sustainable development. Int J Environ Res Public Health 2020 Apr 18;17(8):2801 [FREE Full text] [doi: 10.3390/ijerph17082801] [Medline: 32325763]
31. Ujiie M, Tsuzuki S, Ohmagari N. Effect of temperature on the infectivity of COVID-19. Int J Infect Dis 2020 Jun;95:301-303 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.068] [Medline: 32360939]
32. Song P, Han H, Feng H, Hui Y, Zhou T, Meng W, et al. High altitude relieves transmission risks of COVID-19 through meteorological and environmental factors: evidence from China. Environ Res 2022 Sep;212(Pt B):113214 [FREE Full text] [doi: 10.1016/j.envres.2022.113214] [Medline: 35405128]
33. Carleton T, Cornetet J, Huybers P, Meng KC, Proctor J. Global evidence for ultraviolet radiation decreasing COVID-19 growth rates. Proc Natl Acad Sci U S A 2021 Jan 07;118(1):e2012370118 [FREE Full text] [doi: 10.1073/pnas.2012370118] [Medline: 33323525]
34. Ratnesar-Shumate S, Williams G, Green B, Krause M, Holland B, Wood S, et al. Simulated sunlight rapidly inactivates SARS-CoV-2 on surfaces. J Infect Dis 2020 Jun 29;222(2):214-222 [FREE Full text] [doi: 10.1093/infdis/jiaa274] [Medline: 32432672]
35. Pani SK, Lin N, RavindraBabu S. Association of COVID-19 pandemic with meteorological parameters over Singapore. Sci Total Environ 2020 Oct 20;740:140112 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.140112] [Medline: 32544735]
36. Fiasca F, Minelli M, Maio D, Minelli M, Vergallo I, Necozione S, et al. Associations between COVID-19 incidence rates and the exposure to PM2.5 and NO: a nationwide observational study in Italy. Int J Environ Res Public Health 2020 Dec 13;17(24):9318 [FREE Full text] [doi: 10.3390/ijerph17249318] [Medline: 33322089]
37. Isaia G, Diémoz H, Maluta F, Fountoulakis I, Ceccon D, di Sarra A, et al. Does solar ultraviolet radiation play a role in COVID-19 infection and deaths? An environmental ecological study in Italy. Sci Total Environ 2021 Feb 25;757:143757 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.143757] [Medline: 33272604]
38. Pansini R, Fornacca D. COVID-19 higher mortality in Chinese regions with chronic exposure to lower air quality. Front Public Health 2020;8:597753 [FREE Full text] [doi: 10.3389/fpubh.2020.597753] [Medline: 33585383]
39. Ismail IN, Jalaludin J, Bakar SA, Hisamuddin NH, Suhaimi NF. Association of traffic-related air pollution (TRAP) with DNA damage and respiratory health symptoms among primary school children in Selangor. Asian J Atmosph Sci 2019 Jun 30;13(2):106-116. [doi: 10.5572/ajae.2019.13.2.106]

XSL•FO
RenderX

40. Pansini R, Fornacca D. Early spread of COVID-19 in the air-polluted regions of eight severely affected countries. Atmosphere 2021 Jun 21;12(6):795. [doi: 10.3390/atmos12060795]
41. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
42. Xie J, Zhu Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. Sci Total Environ 2020 Jul 01;724:138201 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138201] [Medline: 32408450]
43. Zhu Y, Xie J, Huang F, Cao L. Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. Sci Total Environ 2020 Jul 20;727:138704 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138704] [Medline: 32315904]
44. Srivastava S, Garg I, Bansal A, Kumar B. SARS-CoV-2 infection: physiological and environmental gift factors at high altitude. Virusdisease 2020 Dec;31(4):450-452 [FREE Full text] [doi: 10.1007/s13337-020-00626-7] [Medline: 32953947]
45. Stephens KE, Chernyavskiy P, Bruns DR. Impact of altitude on COVID-19 infection and death in the United States: a modeling and observational study. PLoS One 2021;16(1):e0245055 [FREE Full text] [doi: 10.1371/journal.pone.0245055] [Medline: 33444357]
46. Sivasubramanian S, Gopalan V, Ramesh K, Padmanabhan P, Mone K, Govindan K, et al. Phylodynamic pattern of genetic clusters, paradigm shift on spatio-temporal distribution of clades, and impact of spike glycoprotein mutations of SARS-CoV-2 isolates from India. J Glob Infect Dis 2021;13(4):164-171 [FREE Full text] [doi: 10.4103/jgid.jgid_97_21] [Medline: 35017872]
47. Sun S, Gu H, Cao L, Chen Q, Ye Q, Yang G, et al. Characterization and structural basis of a lethal mouse-adapted SARS-CoV-2. Nat Commun 2021 Sep 27;12(1):5654. [doi: 10.1038/s41467-021-25903-x] [Medline: 34580297]
48. Mustafa Z, Kalbacher H, Burster T. Occurrence of a novel cleavage site for cathepsin G adjacent to the polybasic sequence within the proteolytically sensitive activation loop of the SARS-CoV-2 Omicron variant: The amino acid substitution N679K and P681H of the spike protein. PLoS One 2022;17(4):e0264723 [FREE Full text] [doi: 10.1371/journal.pone.0264723] [Medline: 35436320]
49. Dyer O. Covid-19: China stops counting cases as models predict a million or more deaths. BMJ 2023 Jan 03;380:2. [doi: 10.1136/bmj.p2] [Medline: 36596574]

Abbreviations

AQI: air quality index
CO: carbon monoxide
FP: fusion peptide
GISAID: Global Initiative on Sharing All Influenza Data
NO2: nitrogen dioxide
O3: ozone
ORF: open reading frame
PM2.5: fine particulate matter
PM10: coarse particulate matter
RBD: receptor-binding domain
Sn: Shannon entropy

Edited by T de Azevedo Cardoso; submitted 18.10.22; peer-reviewed by M Kapsetaki, R Menhas, Y Zhu, D Pala; comments to author 17.05.23; revised version received 20.07.23; accepted 06.03.24; published 25.03.24.

Please cite as:
Deji Z, Tong Y, Huang H, Zhang Z, Fang M, Crabbe MJC, Zhang X, Wang Y
Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study
Interact J Med Res 2024;13:e43585
URL: https://www.i-jmr.org/2024/1/e43585
doi:10.2196/43585
PMID:38526532

©Zhuoga Deji, Yuantao Tong, Honglian Huang, Zeyu Zhang, Meng Fang, M James C Crabbe, Xiaoyan Zhang, Ying Wang. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 25.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/),

XSL•FO
RenderX

which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effect of Body Temperature Changes on the Course of Treatment in Patients With Pneumonia and Sepsis: Results of an Observational Study

Domen Guzelj[1*]; Anže Grubelnik[1*]; Nina Greif[1*]; Petra Povalej Bržan[2*], PhD; Jure Fluher[3*], MD; Žiga Kalamar[3*], MD; Andrej Markota[3*], MD, PhD

[1]Faculty of Medicine, University of Maribor, Maribor, Slovenia

[2]Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

[3]Medical Intensive Care Unit, University Medical Centre Maribor, Maribor, Slovenia

[*]all authors contributed equally

Corresponding Author:
Andrej Markota, MD, PhD
Medical Intensive Care Unit
University Medical Centre Maribor
Ljubljanska ulica 5
Maribor, 2000
Slovenia
Phone: 386 23212008
Email: andrej.markota@ukc-mb.si

Abstract

Background: Traditionally, patients who are critically ill with infection and fever have been treated with antipyretics or even physically cooled. Presumed benefits of the reduction of body temperature are mostly based on decreased metabolic demands. However, it has been shown that decreasing body temperature in patients who are critically ill is not associated with improvement in treatment outcomes. Additionally, there is some data to support the use of temperature modulation (therapeutic hyperthermia) as an adjuvant treatment strategy in patients with infection.

Objective: This study aims to determine the effect of body temperature on the course of intensive care unit (ICU) treatment of patients who are mechanically ventilated with pneumonia, sepsis, and positive tracheal aspirates on admission.

Methods: We performed a single-center retrospective study. Core body temperature was measured in all patients. We analyzed associations between average temperatures in the first 48 hours after admission to ICU and ICU treatment parameters. Additionally, patients were divided into three groups: patients with negative tracheal aspirates 1 week after ICU admission (P-N group), patients with a different pathogen in tracheal aspirates 1 week after ICU admission (P-HAP group), and patients with a persisting pathogen in tracheal aspirates 1 week after ICU admission (P-P group). Differences in body temperature and interventions aimed at temperature modulation were determined.

Results: We observed a significantly higher average temperature in the first 48 hours after admission to ICU in patients who survived to hospital discharge compared to nonsurvivors (mean 37.2 °C, SD 1 °C vs mean 36.9 °C, SD 1.6 °C; P=.04). We observed no associations between average temperatures in the first 48 hours after ICU admission and days of mechanical ventilation in the first 7 days of treatment (ρ=–0.090; P=.30), the average maximum daily requirement for noradrenaline in the first 7 days of treatment (ρ=–0.029; P=.80), average maximum FiO_2 in the first 7 days of ICU treatment (ρ=0.040; P=.70), and requirement for renal replacement therapy in the first 7 days of ICU treatment (mean 37.3 °C, SD 1.4 °C vs mean 37.0 °C, SD 1.3 °C; P=.23). In an additional analysis, we observed a significantly greater use of paracetamol in the P-N group (mean 1.0, SD 1.1 g vs mean 0.4, SD 0.7 g vs mean 0.4, SD 0.8 g; P=.009), a trend toward greater use of active cooling in the first 24 hours after ICU admission in the P-N group (n=11, 44% vs n=14, 33.3% vs n=16, 32%; P=.57), and no other significant differences in parameters of ICU treatment between patient groups.

Conclusions: We observed better survival in patients who developed higher body temperatures in the first 48 hours after admission to the ICU; however, we observed no changes in other treatment parameters. Similarly, we observed greater use of

XSL•FO
RenderX

paracetamol in patients with negative tracheal aspirates 1 week after ICU admission. Our results support the strategy of temperature tolerance in patients who are intubated with pneumonia and sepsis.

(Interact J Med Res 2024;13:e52590) doi:10.2196/52590

KEYWORDS

fever; targeted temperature management; pneumonia; sepsis; intensive care unit

Introduction

Increased body temperature has been recognized as a sign of illness for more than 2000 years and antipyretics have been used for at least 100 years with the aim of lowering body temperature in patients who are febrile [1]. However, a body of evidence exists ranging from studies in the preantibiotic era to a recently published pilot trial, where increased body temperature is investigated as a treatment option in patients with infection [2-4]. A fever-range increase of body temperature is a highly preserved response that is probably beneficial in patients who develop fever as a part of a normal immune response to infection [5]. Recently, we have observed that patients with lower body temperatures are at a higher risk of acquiring the presence of multidrug-resistant pathogens [6]. The aim of this study was to investigate the association between body temperature and the course of intensive care unit (ICU) treatment in patients with pneumonia and sepsis and to evaluate the effect of temperature on the persistence of positive control tracheal aspirates.

Methods

Study Design and Settings

We performed retrospective observational data collection from January 1, 2018, to December 31, 2021. The study was performed in a medical ICU in a tertiary center and supported by an institutional research grant (grant IRP-2022/01-01).

Ethical Considerations

The study was approved by University Medical Centre Maribor Ethical Committee, and informed consent was waived because of the observational and retrospective nature of the study (No. UKC-MB-KME-35/22). Patients' data have been anonymized and deidentified. No compensation was provided to the participants.

Study Population

We included adult (aged >18 years) patients with pneumonia, septic shock, positive tracheal aspirates within 24 hours after admission ("admission" tracheal aspirates), and tracheal aspirates withdrawn 8-14 days after admission to ICU ("week 2" tracheal aspirates). We excluded patients who were treated with targeted temperature management for accidental hypo- or hyperthermia, patients after cardiac arrest, and patients who are neurocritical (eg, patients with meningitis, encephalitis, ischemic or hemorrhagic stroke, or subarachnoid hemorrhage). Temperature management in our patient cohort was as per the treating physician.

Measurements

We collected basic demographic data and data related to ICU treatment, namely outcome of ICU treatment, ICU length of stay, core body temperature during ICU stay, use of renal replacement therapy, use of acetaminophen, maximum concentration of noradrenaline, maximum fraction of inspired oxygen, maximum level of positive end-expiratory pressure and maximum minute ventilation, and microbiological results of tracheal aspirates. Source data was paper based for temperature and therapeutic charts and electronic for other data. Core body temperature was used for study data. As per department policy, temperature was measured via thermal probe urinary catheters (Rüsch Sensor Urinary Catheters, Teleflex Medical, Athlone, Ireland), and temperature measurements were continuously displayed on ICU monitors (Philips IntelliVue MX800 Patient Monitoring System, Koninklijke Philips N.V., Amsterdam, Netherlands). If insertion of a urinary catheter is not possible, then an esophageal temperature probe is inserted, but temperature measurement in all of the included patients was performed via a urinary catheter. Temperature results from 2-hourly notations were used for statistical analysis. We compared the association between body temperature and course of treatment parameters between three groups of patients: patients with sterile week 2 tracheal aspirates (P-N group), patients with a different pathogen in week 2 tracheal aspirates (P-HAP group), and patients with a persistent pathogen presence in week 2 tracheal aspirates (P-P group).

Data Analysis

Statistical analyses were carried out using R (version 4.1.1; R Foundation for Statistical Computing). Nominal variables are presented with frequencies (percentages) and numerical variables with means (SDs) or medians (IQRs) when the normality assumption is violated. For the comparison of nominal dichotomous variables, the Fisher exact test was used. Continuous variables were first assessed for normality using the D'Agostino omnibus test. A comparison of continuous variables across groups was carried out using the Kruskal-Wallis test. Dunn post hoc test with Bonferroni correction was used to adjust for multiple comparisons. The association between average temperature in the first 48 hours and the duration of mechanical ventilation, the average maximum daily requirement for noradrenaline and the average maximum FiO_2 in the first 7 days was evaluated using Spearman correlation. Generalized linear models were used to additionally estimate the ICU and hospital survival in association with the abovementioned risk factors. A statistically significant observation was considered at $P<.05$.

XSL•FO
RenderX

Results

Baseline Characteristics

In all, 117 patients were included in the study analysis; 84 (71.8%) were male, and the mean age was 63.7 (SD 13.5) years. All patients were invasively mechanically ventilated in the ICU on day 1. Mean APACHE II and SOFA scores on admission were 21.7 (SD 6.6) and 10.1 (SD 2.6), respectively. A total of 77 (65.8%) were discharged alive from the ICU, and 46 (39.3%) patients were discharged alive from the hospital. P-N, P-HAP, and P-P groups consisted of 25, 42, and 50 patients, respectively. We observed no significant differences between the P-N, P-HAP, and P-P groups in all treatment parameters apart from the use of paracetamol on day 1, which was significantly greater in the P-N group (mean 1.0, SD 1.1 g vs mean 0.4, SD 0.7 g vs mean 0.4, SD 0.8 g; *P*=.009). General demographic data and parameters describing the course of the treatment in the ICU are described in Table 1. The study patient population and inclusion flowchart are presented in Figure 1.

Table 1. Demographic data and parameters describing the course of treatment in the intensive care unit (ICU).

Label and variable	Group			Total	*P* value
	P-N[a]	P-P[b]	P-HAP[c]		
Patients, n (%)	25 (21.4)	50 (42.7)	42 (35.9)	117 (100)	
Gender, n (%)					.26
Female	5 (15.2)	18 (54.5)	10 (30.3)	33 (28.2)	
Male	20 (23.8)	32 (38.1)	32 (38.1)	84 (71.8)	
Age (years)					.21
Median (IQR)	62.0 (52.0-74.0)	64.0 (57.2-71.8)	67.0 (62.0-76.0)	65.0 (57.0-74.0)	
Min-max	23.0-84.0	26.0-84.0	39.0-86.0	23.0-86.0	
BMI (kg/m^2)					.19
Mean (SD)	28.7 (7.0)	27.6 (4.3)	29.7 (6.4)	28.6 (5.8)	
Min-max	17.2-47.7	18.8-39.2	16.3-44.8	16.3-47.7	
Length of hospitalization (days)					.35
Median (IQR)	22.0 (17.0-29.0)	19.5 (15.0-26.8)	17.5 (13.0-26.8)	20.0 (14.0-26.9)	
Min-max	8.0-53.0	10.0-43.0	8.0-69.0	8.0-69.0	
ICU survival, n (%)					.26
Died	6 (15.0)	16 (40.0)	18 (45.0)	40 (34.2)	
Survived	19 (26.1)	34 (44.2)	24 (31.2)	77 (65.8)	
Hospital survival, n (%)					.59
Died	13 (18.3)	31 (43.7)	27 (38.0)	71 (60.7)	
Survived	12 (26.1)	19 (41.3)	15 (32.6)	46 (39.3)	
SOFA[d] at the time of admission					.76
Median (IQR)	9.5 (7.0-12.0)	10.0 (9.0-12.0)	10.0 (9.0-12.0)	10.0 (9.0-12.0)	
Min-max	5.0-16.0	5.0-15.0	4.0-15.0	4.0-16.0	
APACHE[e] II at the time of admission					.53
Mean (SD)	20.4 (6.3)	22.0 (6.4)	22.1 (7.2)	21.7 (6.6)	
Min-max	6.0-32.0	7.0-36.0	5.0-40.0	5.0-40.0	
Change of antibiotic in 7 days after admission (no), n (%)	7 (18.4)	17 (44.7)	14 (36.8)	38 (32.5)	.86

[a]P-N: patients with sterile week 2 tracheal aspirates.

[b]P-HAP: patients with a different pathogen in week 2 tracheal aspirates.

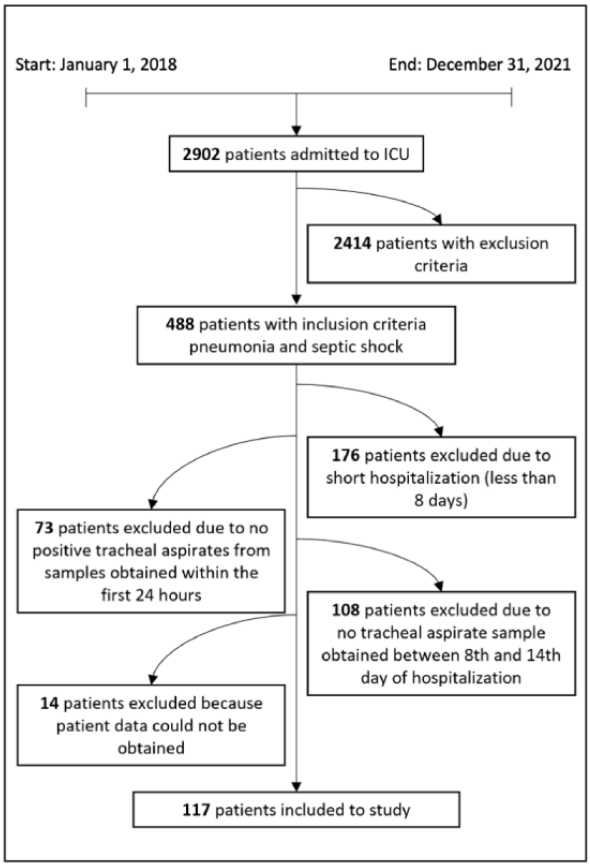
[c]P-P group: patients with a persistent pathogen presence in week 2 tracheal aspirates.

[d]SOFA: Sequential Organ Failure Assessment.

[e]APACHE: Acute Physiology and Chronic Health Evaluation.

XSL•FO
RenderX

Figure 1. Study patient population and inclusion flowchart. ICU: intensive care unit.

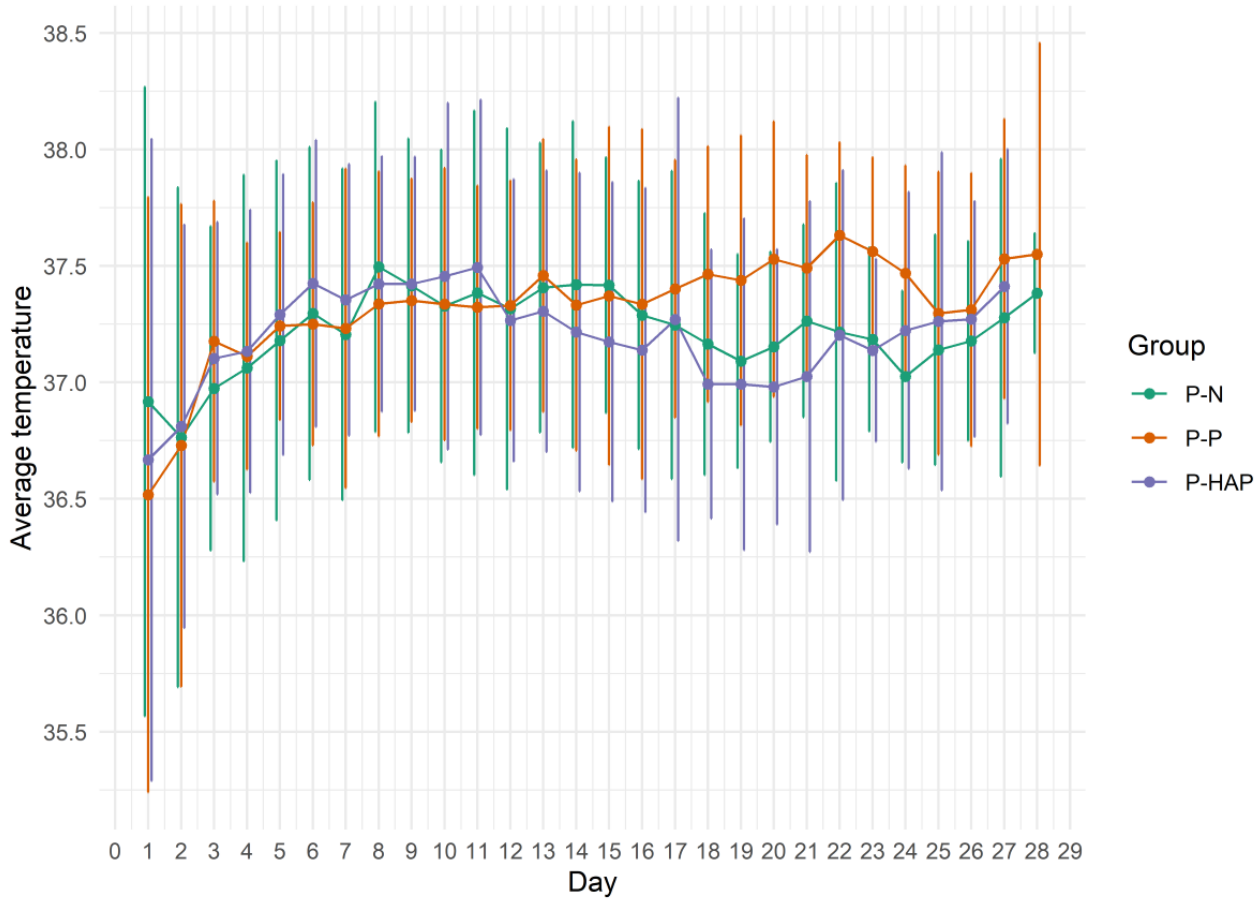


Main Results

Body temperature changes in the first 28 days of treatment are presented in Figure 2. We observed a significantly higher average temperature in the first 48 hours after admission to ICU in patients who survived to hospital discharge compared to nonsurvivors (mean 37.2 °C, SD 1 °C vs mean 36.9 °C, SD 1.6 °C; *P*=.04) and nonsignificant differences in average temperatures in the first 48 hours after admission to ICU between survivors and nonsurvivors to ICU discharge (mean 37.0 °C, SD 1.3 °C vs mean 37.0 °C, SD 1.2 °C; *P*=.60). We observed no associations between average temperatures in the first 48 hours after ICU admission and days of mechanical ventilation in the first 7 days of treatment (ρ=–0.090; *P*=.30), the average maximum daily requirement for noradrenaline in the first 7 days of treatment (ρ=–0.029; *P*=.80), average maximum FiO_2 in the first 7 days of ICU treatment (ρ=0.040; *P*=.70), and requirement for renal replacement therapy in the first 7 days of ICU treatment (mean 37.3 °C, SD 1.4 °C vs mean 37.0 °C, SD 1.3 °C; *P*=.23).

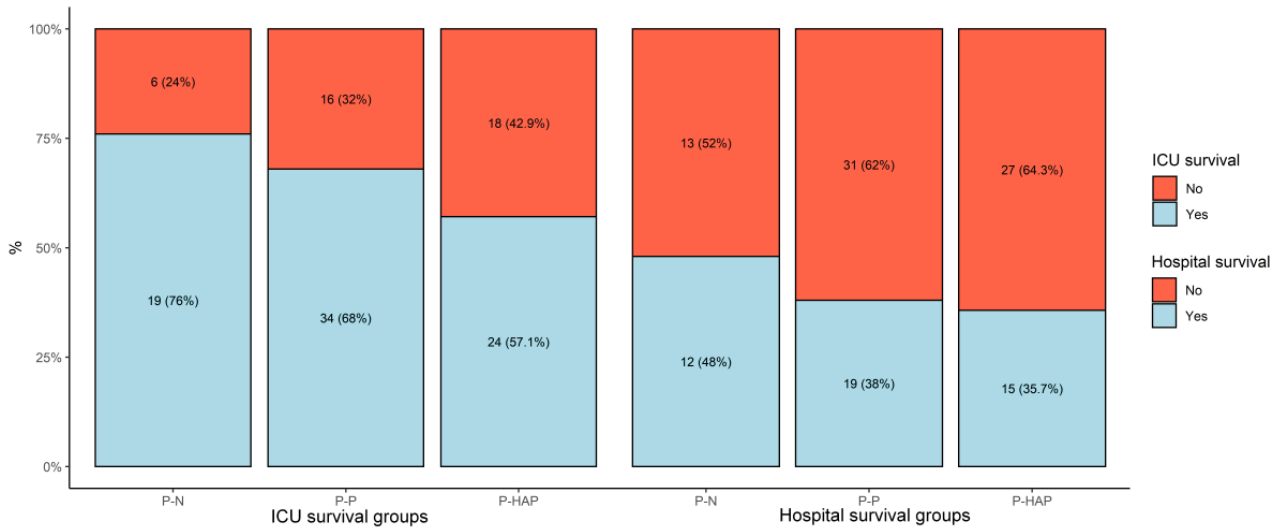
XSL•FO
RenderX

Figure 2. Average body temperature by groups in the first 28 days of treatment in the intensive care unit. P-HAP: patients with a different pathogen in week 2 tracheal aspirates; P-N: patients with sterile week 2 tracheal aspirates; P-P: patients with a persistent pathogen presence in week 2 tracheal aspirates.



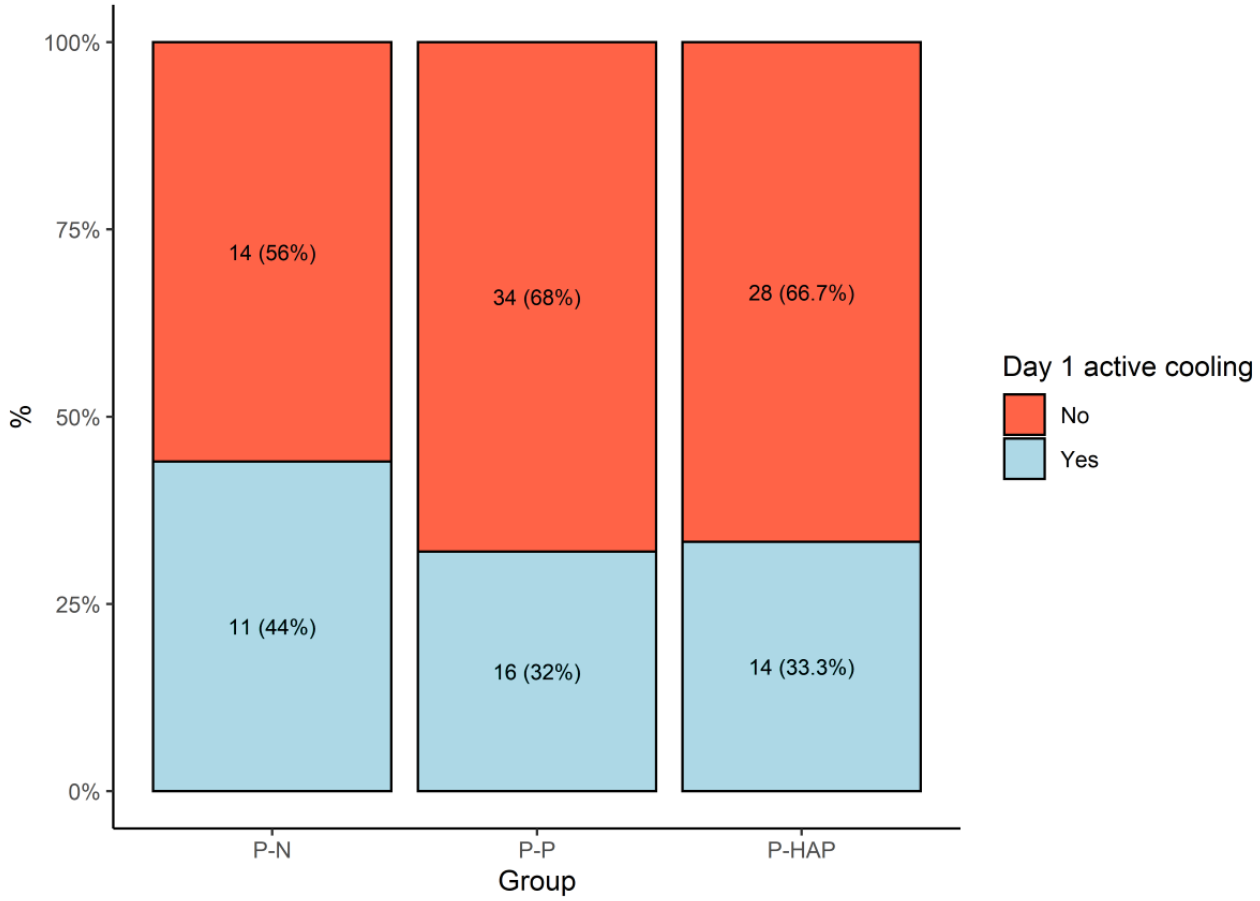
P-N, P-HAP, and P-P groups consisted of 25, 42, and 50 patients each, respectively. No significant differences in average body temperature in the first 48 hours were observed between the P-P, P-HAP, and P-N groups. Temperature variations increased after day 16 (Figure 2), however, and approximately two-thirds of patients were discharged from the ICU or died before day 16. Both hospital and ICU survival were higher in the P-N group compared to the P-HAP and P-P groups; however, the differences were not statistically significant (Figure 3). We observed statistically significant greater use of paracetamol in the first 24 hours after ICU admission in the P-N group compared to the P-HAP and P-P groups (mean 1.0, SD 1.1 g vs mean 0.4, SD 0.7 g vs mean 0.4, SD 0.8 g; *P*=.009). Active cooling was used more frequently on day 1 in the P-N group, but no significant differences were observed (Figure 4).

Figure 3. Intensive care unit (ICU) and hospital survival by groups. P-HAP: patients with a different pathogen in week 2 tracheal aspirates; P-N: patients with sterile week 2 tracheal aspirates; P-P: patients with a persistent pathogen presence in week 2 tracheal aspirates.



XSL•FO
RenderX

Figure 4. Active cooling on day one of hospitalization in the intensive care unit divided by groups. P-HAP: patients with a different pathogen in week 2 tracheal aspirates; P-N: patients with sterile week 2 tracheal aspirates; P-P: patients with a persistent pathogen presence in week 2 tracheal aspirates.



Microbiological Data

In all, 148 different causal pathogens were isolated. In admission tracheal aspirates, the majority (n=124, 66%) of pathogens were gram-negative bacteria, followed by gram-positive bacteria (n=56, 29.8%) and fungal pathogens (4.3%). In week 2 tracheal aspirates, the majority (n=128, 86.5%) of pathogens were gram-negative bacteria, followed by gram-positive bacteria (n=14, 9.5%) and fungal pathogens (n=6, 4.1%). We observed no differences between the P-N, P-HAP, and P-P groups in causative pathogens in admission tracheal aspirates (Table 2). There were no differences between the P-N, P-HAP, and P-P groups in the rate of change of antimicrobial therapy within the first week after ICU admission, and we observed no changes in the duration of initial (combined empiric and antibiogram-guided antimicrobial therapy) and cumulative ICU-stay antibiotic therapy (mean 7.6, SD 2.7 days vs mean 8.4, SD 3.1 days vs mean 8.2, SD 2.9 days; *P*=.55; mean 15.2, SD 5.6 days vs mean 14.8, SD 6.8 days vs mean 15.9, SD 6.2 days; *P*=.70; respectively).

XSL•FO
RenderX

Table 2. Isolated pathogens from the first and second samples divided by groups.

Genus	Group, n (%)			Total, n (%)
	P-P[a]	P-HAP[b]	P-N[c]	
Gram-positive bacteria				
Sample 1				
Staphylococcus	8 (4.3)	15 (8.0)	6 (3.2)	29 (15.4)
Streptococcus	5 (2.7)	11 (5.9)	5 (2.7)	21 (11.2)
Corynebacterium	3 (1.6)	2 (1.1)	1 (0.5)	6 (3.2)
Total	16 (8.5)	28 (14.9)	12 (6.4)	56 (29.8)
Sample 2				
Staphylococcus	5 (3.4)	2 (1.4)	N/A[d]	7 (4.7)
Corynebacterium	1 (0.7)	3 (2.0)	N/A	4 (2.7)
Streptococcus	1 (0.7)	1 (0.7)	N/A	2 (1.4)
Enterococcus	0 (0)	1 (0.7)	N/A	1 (0.7)
Total	7 (4.7)	7 (4.7)	N/A	14 (9.5)
Gram-negative bacteria				
Sample 1				
Klebsiella	17 (9.0)	6 (3.2)	8 (4.3)	31 (16.5)
Haemophilus	4 (2.1)	7 (3.7)	6 (3.2)	17 (9.0)
Escherichia	9 (4.8)	4 (2.1)	1 (0.5)	14 (7.4)
Enterobacter	9 (4.8)	3 (1.6)	2 (1.1)	14 (7.4)
Pseudomonas	10 (5.3)	0 (0)	1 (0.5)	11 (5.9)
Other	23 (12.2)	7 (3.7)	7 (3.7)	37 (19.7)
Total	72 (38.3)	27 (14.4)	25 (13.3)	124 (66)
Sample 2				
Klebsiella	16 (10.8)	12 (8.1)	N/A	28 (18.9)
Pseudomonas	16 (10.8)	8 (5.4)	N/A	24 (16.2)
Escherichia	11 (7.4)	4 (2.7)	N/A	15 (10.1)
Enterobacter	8 (5.4)	6 (4.1)	N/A	14 (9.5)
Acinetobacter	10 (6.8)	3 (2.0)	N/A	13 (8.8)
Other	18 (12.2)	16 (10.8)	N/A	34 (23.0)
Total	79 (53.4)	49 (33.1)	N/A	128 (86.5)
Fungi				
Sample 1				
Candida	1 (0.5)	6 (3.2)	1 (0.5)	8 (4.3)
Total	89 (47.3)	61 (32.4)	38 (20.2)	188 (100)
Sample 2				
Candida	5 (3.4)	1 (0.7)	N/A	6 (4.1)
Total	91 (61.5)	57 (38.5)	N/A	148 (100)

[a]P-P group: patients with a persistent pathogen presence in week 2 tracheal aspirates.

[b]P-HAP: patients with a different pathogen in week 2 tracheal aspirates.

[c]P-N: patients with sterile week 2 tracheal aspirates.

[d]N/A: not applicable.

XSL•FO
RenderX

Discussion

Principal Findings

We observed higher initial body temperatures in patients who survived to hospital discharge and a trend toward higher initial temperatures in patients who survived to ICU discharge. We did not observe any differences in body temperature between the P-N, P-HAP, and P-P groups of patients, but we did observe a significantly greater use of paracetamol in the P-N group of patients and a trend toward greater use of active cooling in the P-N group of patients. Additionally, we observed a trend toward better survival to ICU and hospital discharge in the P-N group of patients.

Traditionally, antipyretic therapy (mostly paracetamol), and in some cases active cooling, have been used to lower body temperature in patients who are febrile with infection. The presumed patient benefit comes from decreased metabolic demands associated with lower body temperature [7,8].

The survival of our patients is comparable to studies performed by Gursel and Demirtas [9], Depuydt et al [10], and Qiao et al [11] who reported survival to ICU discharge of patients who were critically ill with pneumonia in the range between 46% and 75%, compared to 65.8% survival to ICU discharge in our patients. They included patients with similar severity of illness, with SOFA scores in the range between 4 and 6 points and APACHE II scores around 20 points, compared to the SOFA score 10.1 (SD 2.6) and APACHE II score 21.7 (SD 6.6) in our study.

Our findings are in line with a number of other studies where better survival was observed in patients with infection and physiological-grade fever (ie, body temperature between around 37 °C and 39 °C). Both Kushimoto et al [12] in a prospective observational study and Shimazui et al [13] in a retrospective observational study observed higher mortality in patients with sepsis and lower body temperature on admission. Lee et al [14] also observed lower mortality in patients with physiological-grade fever in their prospective study. Rumbus et al [15] performed a meta-analysis of 42 studies reporting body temperature and mortality in patients with sepsis. They discovered a correlation between higher body temperature and lower mortality; higher body temperature predicted better survival, and hypothermia predicted lower survival. Similar to our findings the number of patients with body temperature over 39.5 °C was very low. Thomas-Rüddel et al [16] performed a secondary analysis of a large data set of patients with sepsis and discovered that initial body temperatures were distributed in two peaks: a smaller peak at around 35.5 °C and a larger peak approximately twice as large at around 38 °C. Again, the highest survival rate was observed in patients with hyperthermia and the lowest in patients with hypothermia. They also observed that ambient temperatures were significantly associated with body temperatures; lower ambient temperatures were associated with hypothermia, and higher outside temperatures were associated with hyperthermia. The potential benefits of active warming of patients with infection were further highlighted by Drewry et al [4], who performed a pilot randomized controlled study evaluating the use of therapeutic hyperthermia in patients with sepsis and observed a significantly better survival in the hyperthermia group; however, there were no differences in the primary results.

We can speculate that higher doses of paracetamol and greater use of active cooling probably decreased body temperatures in the P-N group of patients; however, this was not associated with any improvement during ICU treatment. We observed no differences in the use of noradrenalin, parameters of mechanical ventilation, requirement for renal replacement therapy, or organ dysfunction scores, and we observed a trend toward greater ICU and hospital survival in the P-N group. Our results suggest that there is probably no clinical benefit associated with the treatment of physiological-grade fever. Similarly, Zhang et al [17] and Ye et al [18] performed retrospective studies on patients with fever and sepsis, and they observed no beneficial effects of antipyretic therapy and possible harm associated with the use of external cooling. In a prospective randomized controlled trial, Young et al [19] observed approximately 0.5 °C lower body temperature in patients who received paracetamol (4 g daily) compared to patients who received a placebo. They also observed that patients who received more paracetamol experienced a longer ICU stay if they were nonsurvivors, and shorter ICU stay if they were survivors, which was explained by the probable effect of lower body temperature on clinicians' perception of the patients' prognosis. Lower body temperature in the paracetamol group was not associated with any improvement in the ICU course of treatment parameters [19,20].

To our knowledge, there is no other data to compare our results regarding the clearance or persistence of pathogens in tracheal aspirates. In 42.7% (n=50) of our patients (ie, the P-P group of patients), the same pathogen persisted in tracheal aspirates 1 week after initial samples were obtained. There were no differences in causative pathogens between the P-N, P-HAP, and P-P groups of patients, and we observed no significant changes in baseline data between the groups, including the rate of change of empiric antibiotic therapy. There was a trend toward higher temperature in the first 48 hours in the P-N group, compared to the P-HAP and P-P groups, possibly indicating that tolerance of hyperthermia could be beneficial for patients who are intubated with pneumonia and sepsis. There were no significant differences in the course of ICU treatment parameters despite greater use of paracetamol and active cooling, indicating that pharmacological or active cooling has no benefit in patients where fever is part of an appropriate response to infection.

There are a number of limitations to our study. We performed a single-center retrospective observational study with all inherent biases associated with this study design. Of the initial 488 patients with pneumonia, 249 were not included because there were no tracheal aspirates 1 week after ICU admission (because of extubation, death, or no clinical need for obtaining samples if the patients were still intubated). Additionally, target temperatures were defined by the treating physicians. However, there are no guidelines regarding target temperatures for this patient population, and we observed no differences in baseline data between the different groups.

XSL•FO
RenderX

Conclusions

To conclude, we observed better survival in patients who developed higher body temperatures in the first 48 hours after admission to the ICU; however, we observed no changes in other treatment parameters. Additionally, we observed greater use of interventions aimed at cooling the patients (use of paracetamol and a trend toward greater use of active cooling) in patients with negative tracheal aspirates 1 week after ICU admission. Additionally, in this group of patients, we observed a trend toward better survival. Our results speak against the use of interventions aimed at the reduction of body temperature and support the strategy of temperature tolerance in patients who are mechanically ventilated with pneumonia and sepsis.

Acknowledgments

The study was supported by an institutional research grant (grant IRP-2022/01-01).

Authors' Contributions

AM, JF, and ŽK conceptualized the study. AM and PPB developed the methodology. PPB used the software. AM and PPB performed the validation. PPB performed the formal analysis. DG, AG, and NG conducted the investigation and acquired the resources. AM curated the data curation. DG, AG, and NG prepared the original draft. AM, PPB, JF, ŽK, DG, AG, and NG reviewed and edited the paper. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Walter EJ, Hanna-Jumma S, Carraretto M, Forni L. The pathophysiological basis and consequences of fever. Crit Care 2016 Jul 14;20(1):200 [FREE Full text] [doi: 10.1186/s13054-016-1375-5] [Medline: 27411542]
2. Owens CA. The value of fever therapy for gonorrhea. JAMA 1936 Dec 12;107(24):1942. [doi: 10.1001/jama.1936.02770500008003]
3. Wagner-Jaureggj J. The history of the malaria treatment of general paralysis. Am J Psychiatry 1946 Mar;102:577-582. [doi: 10.1176/ajp.102.5.577] [Medline: 21026839]
4. Drewry A, Mohr N, Ablordeppey E, Dalton C, Doctor R, Fuller B, et al. Therapeutic hyperthermia is associated with improved survival in afebrile critically ill patients with sepsis: a pilot randomized trial. Crit Care Med 2022 Jun 01;50(6):924-934 [FREE Full text] [doi: 10.1097/CCM.0000000000005470] [Medline: 35120040]
5. Evans SS, Repasky EA, Fisher DT. Fever and the thermal regulation of immunity: the immune system feels the heat. Nat Rev Immunol 2015 Jun;15(6):335-349 [FREE Full text] [doi: 10.1038/nri3843] [Medline: 25976513]
6. Markota A, Skok K, Kalamar Ž, Fluher J, Gorenjak M. Better control of body temperature is not associated with improved hemodynamic and respiratory parameters in mechanically ventilated patients with sepsis. J Clin Med 2022 Feb 24;11(5):1211 [FREE Full text] [doi: 10.3390/jcm11051211] [Medline: 35268302]
7. Laupland K, Shahpori R, Kirkpatrick A, Ross T, Gregson D, Stelfox H. Occurrence and outcome of fever in critically ill adults. Crit Care Med 2008 May;36(5):1531-1535. [doi: 10.1097/CCM.0b013e318170efd3] [Medline: 18434882]
8. Doyle J, Schortgen F. Should we treat pyrexia? And how do we do it? Crit Care 2016 Oct 03;20(1):303 [FREE Full text] [doi: 10.1186/s13054-016-1467-2] [Medline: 27716372]
9. Gursel G, Demirtas S. Value of APACHE II, SOFA and CPIS scores in predicting prognosis in patients with ventilator-associated pneumonia. Respiration 2006;73(4):503-508. [doi: 10.1159/000088708] [Medline: 16205047]
10. Depuydt P, Vandijck D, Bekaert M, Decruyenaere J, Blot S, Vogelaers D, et al. Determinants and impact of multidrug antibiotic resistance in pathogens causing ventilator-associated-pneumonia. Crit Care 2008;12(6):R142 [FREE Full text] [doi: 10.1186/cc7119] [Medline: 19014695]
11. Qiao Q, Lu G, Li M, Shen Y, Xu D. Prediction of outcome in critically ill elderly patients using APACHE II and SOFA scores. J Int Med Res 2012;40(3):1114-1121 [FREE Full text] [doi: 10.1177/147323001204000331] [Medline: 22906285]
12. Kushimoto S, Gando S, Saitoh D, Mayumi T, Ogura H, Fujishima S, JAAM Sepsis Registry Study Group. The impact of body temperature abnormalities on the disease severity and outcome in patients with severe sepsis: an analysis from a multicenter, prospective survey of severe sepsis. Crit Care 2013 Nov 13;17(6):R271 [FREE Full text] [doi: 10.1186/cc13106] [Medline: 24220071]
13. Shimazui T, Nakada T, Walley K, Oshima T, Abe T, Ogura H, et al. Significance of body temperature in elderly patients with sepsis. Crit Care 2020 Jun 30;24(1):387 [FREE Full text] [doi: 10.1186/s13054-020-02976-6] [Medline: 32605659]
14. Lee B, Inui D, Suh G, Kim J, Kwon J, Park J, et al. Association of body temperature and antipyretic treatments with mortality of critically ill patients with and without sepsis: multi-centered prospective observational study. Crit Care 2012 Feb 28;16(1):R33 [FREE Full text] [doi: 10.1186/cc11211] [Medline: 22373120]

XSL•FO
RenderX

15. Rumbus Z, Matics R, Hegyi P, Zsiboras C, Szabo I, Illes A, et al. Fever is associated with reduced, hypothermia with increased mortality in septic patients: a meta-analysis of clinical trials. PLoS One 2017;12(1):e0170152 [FREE Full text] [doi: 10.1371/journal.pone.0170152] [Medline: 28081244]
16. Thomas-Rüddel DO, Hoffmann P, Schwarzkopf D, Scheer C, Bach F, Komann M, et al. Fever and hypothermia represent two populations of sepsis patients and are associated with outside temperature. Crit Care 2021 Oct 21;25(1):368 [FREE Full text] [doi: 10.1186/s13054-021-03776-2] [Medline: 34674733]
17. Zhang Z, Chen L, Ni H. Antipyretic therapy in critically ill patients with sepsis: an interaction with body temperature. PLoS One 2015;10(3):e0121919 [FREE Full text] [doi: 10.1371/journal.pone.0121919] [Medline: 25822614]
18. Ye S, Xu D, Zhang C, Li M, Zhang Y. Effect of antipyretic therapy on mortality in critically ill patients with sepsis receiving mechanical ventilation treatment. Can Respir J 2017;2017:3087505. [doi: 10.1155/2017/3087505] [Medline: 28386165]
19. Young P, Saxena M, Bellomo R, Freebairn R, Hammond N, van Haren F, et al. Acetaminophen for fever in critically ill patients with suspected infection. N Engl J Med 2015 Dec 03;373(23):2215-2224. [doi: 10.1056/NEJMoa1508375] [Medline: 26436473]
20. Young P, Bellomo R, Bernard G, Niven D, Schortgen F, Saxena M, et al. Fever control in critically ill adults. An individual patient data meta-analysis of randomised controlled trials. Intensive Care Med 2019 Apr;45(4):468-476. [doi: 10.1007/s00134-019-05553-w] [Medline: 30741326]

Abbreviations

APACHE: acute physiology and chronic health evaluation
ICU: intensive care unit
P-HAP: patients with a different pathogen in week 2 tracheal aspirates
P-N: patients with sterile week 2 tracheal aspirates
P-P: patients with a persistent pathogen presence in week 2 tracheal aspirates
SOFA: sequential organ failure assessment

Edited by T de Azevedo Cardoso; submitted 13.09.23; peer-reviewed by S Liu, E Kawamoto; comments to author 03.11.23; revised version received 24.11.23; accepted 06.12.23; published 01.03.24.

Please cite as:
Guzelj D, Grubelnik A, Greif N, Povalej Bržan P, Fluher J, Kalamar Ž, Markota A
The Effect of Body Temperature Changes on the Course of Treatment in Patients With Pneumonia and Sepsis: Results of an Observational Study
Interact J Med Res 2024;13:e52590
URL: https://www.i-jmr.org/2024/1/e52590
doi:10.2196/52590
PMID:38427413

©Domen Guzelj, Anže Grubelnik, Nina Greif, Petra Povalej Bržan, Jure Fluher, Žiga Kalamar, Andrej Markota. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 01.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Gamification and Oral Health in Children and Adolescents: Scoping Review

Rui Moreira[1,2*], MD; Augusta Silveira[3,4,5*], MSc, DMD, PhD; Teresa Sequeira[3,4,5*], PhD; Nuno Durão[2*], MD; Jessica Lourenço[2*], MD; Inês Cascais[6*], MD; Rita Maria Cabral[2*], MD; Tiago Taveira Gomes[3,7*], MSc, ME, MD, PhD

[1]School of Medicine and Biomedical Sciences, Instituto de Ciencias Biomédicas Abel Salazar, University of Porto, Porto, Portugal
[2]Department of Stomatology and Oral Maxillofacial Surgery, University Hospital Center of Santo António, Porto, Portugal
[3]Faculty of Health Sciences, Fernando Pessoa University, University of Porto, Porto, Portugal
[4]Institute for Research, Innovation and Development, Fernando Pessoa Foundation, University of Porto, Porto, Portugal
[5]Center for Innovative Biomedicine and Biotechnology, University of Coimbra, Coimbra, Portugal
[6]Paediatrics Department, Maternal & Child Center of the North, University Hospital Center of Santo António, Porto, Portugal
[7]Department of Community Medicine, Information and Decision in Health, Medical School of Porto, University of Porto, Porto, Portugal
[*]all authors contributed equally

Corresponding Author:
Rui Moreira, MD
School of Medicine and Biomedical Sciences
Instituto de Ciencias Biomédicas Abel Salazar
University of Porto
Rua de Jorge Viterbo Ferreira, 228
Porto, 4050-313
Portugal
Phone: 351 914180479
Email: ruidcmoreira@gmail.com

Abstract

Background: Oral health is a determinant of overall well-being and quality of life. Individual behaviors, such as oral hygiene and dietary habits, play a central role in oral health. Motivation is a crucial factor in promoting behavior change, and gamification offers a means to boost health-related knowledge and encourage positive health behaviors.

Objective: This study aims to evaluate the impact of gamification and its mechanisms on oral health care of children and adolescents.

Methods: A systematic search covered multiple databases: PubMed/MEDLINE, PsycINFO, the Cochrane Library, ScienceDirect, and LILACS. Gray literature, conference proceedings, and WHOQOL internet resources were considered. Studies from January 2013 to December 2022 were included, except for PubMed/MEDLINE, which was searched until January 2023. A total of 15 studies were selected following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The eligibility criteria were peer-reviewed, full-text, and empirical research related to gamification in oral health care, reports of impact, and oral health care outcomes. The exclusion criteria encompassed duplicate articles; unavailable full texts; nonoriginal articles; and non–digital game–related, non–oral health–related, and protocol studies. Selected studies were scrutinized for gamification mechanisms and outcomes. Two main questions were raised: “Does gamification in oral health care impact oral health?” and “Does oral health care gamification enhance health promotion and literacy?” The PICO (Patient, Intervention, Comparison, Outcome) framework guided the scoping review.

Results: Initially, 617 records were obtained from 5 databases and gray literature sources. After applying exclusion criteria, 15 records were selected. Sample size in the selected studies ranged from 34 to 190 children and adolescents. A substantial portion (11/15, 73%) of the studies discussed oral self-care apps supported by evidence-based oral health. The most clearly defined data in the apps were “brushing time” (11/11, 100%) and “daily amount brushing” (10/11, 91%). Most studies (11/15, 73%) mentioned oral health care behavior change techniques and included “prompt intention formation” (11/26, 42%), “providing instructions” (11/26, 42%), “providing information on the behavior-health link” (10/26, 38%), “providing information on consequences” (9/26, 35%), “modeling or demonstrating behavior” (9/26, 35%), “providing feedback on performance” (8/26, 31%), and “providing contingent rewards” (8/26, 31%). Furthermore, 80% (12/15) of the studies identified game design elements incorporating

XSL•FO
RenderX

gamification features in oral hygiene applications. The most prevalent gamification features were "ideological incentives" (10/12, 83%) and "goals" (9/16, 56%), which were found in user-specific and challenge categories, respectively.

Conclusions: Gamification in oral health care shows potential as an innovative approach to promote positive health behaviors. Most studies reported evidence-based oral health and incorporated oral health care behavior change techniques.

(Interact J Med Res 2024;13:e35132) doi:10.2196/35132

KEYWORDS

gamification; mechanisms of gamification; gamification components; intrinsic and extrinsic motivators; oral health care; health behavior; oral health care applications

Introduction

People's health behaviors, such as physical activity, diet, tobacco and alcohol use, recreational drug consumption, and adherence to chronic medications, directly influence their health risks and consequent diseases. To decrease the burden of preventable chronic diseases and enhance well-being in society, it is essential to bring about a change in behavior [1-4]. According to the World Health Organization, oral health is a strong indicator of general health, well-being, and quality of life. The treatment of oral pathologies is expensive and usually not covered under universal health care, accounting for 5% of health care expenditures and 20% of out-of-pocket expenses in wealthy countries [5-11]. Health care systems have limited resources, and making informed decisions based on data collection, focusing on individuals, can provide better health outcomes without incurring additional costs; this approach can deliver better value and at the same time reduce costs [12,13].

Motivation is a core target of a wide range of established behavior change techniques [1,14-17]. Computer games can be used to increase health-related knowledge and promote desirable health behaviors in children [18]. Games are designed to provide enjoyment, engagement, and satisfaction [1,19-22]. Mobile phones and mobile health technologies can address these issues at low costs [1,23,24]. Mobile devices are useful for delivering health interventions due to their widespread adoption, powerful technical capabilities, and portability [25]. The positive emotional attachment with the user may increase the benefit of health promotion via mobile devices, allowing health interventions to be delivered immediately, anytime, and anywhere [26]. The use of health care apps provides easy access to information and has the potential to improve patient engagement and treatment compliance [27]. Indeed, the number of health care apps available has been growing year after year, with over 200 billion app downloads worldwide from the Apple App Store and Google Play in 2020 [28].

Bohn et al [29] found that educational applications are valuable tools for enhancing patient-provider communication in dental settings. Studies have pointed out that the traditional educational approach, which relies mainly on reading and listening to standardized content, should be replaced with customizable and interactive involvement, using communication tools that are familiar to newer generations [30]. Gamification is a possible response to overcoming the challenges of communication and motivation in health care [31], as it can track individual behaviors and involve users in goal-chasing activities while displaying progress and feedback through personalized information apps [1,32].

Deterding et al [33] define gamification as the integration of game design elements into nongame contexts [34]. This process enriches products, services, and information systems with game design features to positively influence the motivation, productivity, and behavior of users [33,35,36]. Gamified systems commonly use motivational features such as immediate success feedback (reward mechanism), continuous progress feedback, and goal setting. These systems work through interface elements such as point-scoring systems, badges, levels, challenges, competitions, relatedness support, social feedback (engagement loops), recognition, comparison through leaderboards, teams, communication functions, autonomized support through customizable avatars and environments, and narratives providing emotional and value-based rationales for certain activities (customization) [1,34,37,38]. Users receive badges that represent success and can be shared in social networks and displayed in a digital trophy cabinet when new milestones are achieved [35]. Recently, the self-determination theory has become a key framework for health behavior interventions and research [1,39-43]. According to this theory, gamification induces 4 main intrinsic motivators: user satisfaction, conveyance of optimism, provision of meaning [35,44], and facilitation of social interactions [45,46].

Apart from financial aspects, extrinsic motivators systematically activate intrinsic motivators, such as social recognition, support of learning processes, and behavioral change. The evidence suggests that positive outcomes are stronger when gamification is used to target behavioral outcomes [1,23,24,47-50]. However, critics have pointed out the lack of high-quality effect studies on gamification [22,38,51,52]. Nevertheless, clinicians have an opportunity to promote engagement in health promotion through a motivating, fulfilling, and fun activity [53]. Therefore, this study aims to analyze the impact of gamification and its mechanisms on oral health care.

The objective of this scoping review was to analyze the impact of gamification and its mechanisms on oral health care in children and adolescents. We assessed the effectiveness of gamification in promoting changes in oral health behavior and enhancing oral health care outcomes for its users. Furthermore, we investigated the role of gamification in oral health care, including the integration of evidence-based oral health care concepts and gamification design elements in application design.

XSL•FO
RenderX

Methods

Objective

This scoping review aims to analyze the impact and efficiency of gamification mechanisms on oral health care, with a focus on promoting user engagement to expand oral health literacy and support oral health care policies, in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [54,55].

Scoping Review

This review used a systematic methodology to identify gamification on oral health care apps, evaluate their features, identify their gamification mechanisms, and follow their outcomes.

We did not apply any restrictions related to population samples or oral health care systems, but we excluded contexts outside oral health.

The following main questions guided our analysis:

- Does gamification in oral health care impact oral health?
- Does oral health care gamification enhance health promotion and literacy?

On the basis of these questions, we searched several domains, including oral health care outcomes, gamification design elements, evidence-based oral health care concepts used in app development, and oral health behavior changes related to gamification. We use the PICO (Patient, Intervention, Comparison, Outcome) framework to elucidate the systematic review questions:

- P (Patient, Population, or Problem): the population under investigation included children and adolescents: mother-preschooler (3-6 years old) dyads and adolescents (younger than 16 years).
- I (Intervention): the intervention under investigation was the gamification strategy and its usefulness for oral health promotion and literacy.
- C (Comparison): gamification was compared with its alternative, traditional learning methods.
- (Outcome): the outcome assessed was divided into 2 main sections: behavior change techniques and gamification mechanisms.

Search Strategy

The search was conducted across several electronic bibliographic databases, including PubMed/MEDLINE, PsycINFO, the Cochrane Library, ScienceDirect, and LILACS. In addition, gray literature, conference proceedings, and WHOQOL internet resources were assessed. The search strategy included terms related to gamification and oral health care such as gamification, oral health care, policies, games, digital, apps, and outcomes. The authors used a controlled and hierarchically organized vocabulary produced by the National Library of Medicine called the Medical Subject Headings to ensure that the search results accurately reflected the subject content of journal articles as they are published. The search strategy enabled us to identify both published and unpublished studies. All sources were last searched until July 2023, except for PubMed/MEDLINE, which was searched until January 2023. The references to gamification date back to nearly 2010, and articles exploring gamification in oral health were only mentioned or studied in the last decade. There were no language restrictions, and studies published between January 2013 and December 2022 (10 years) were included.

The search string used on PubMed was “Gamification” [mh] OR “oral healthcare” [tiab] OR “gamification” [tiab] OR ad [tiab] OR “applications” [tiab] OR “gamification” [tiab] OR “digital” [tiab] OR “games” [tiab] OR “outcomes” [mh] OR policies [tiab] OR gamification [tiab] OR “oral health” [mh:noexp] (see Multimedia Appendix 1).

Initially, all types of articles were considered eligible, including systematic reviews, research articles, and prospective and retrospective studies, as long as they met the following criteria: (1) peer-reviewed, (2) full-text papers, (3) empirical research (qualitative and quantitative), (4) explained research methods, (5) gamification as a research subject, (6) effect reported in terms of impact (affect, behavior, and cognition) and user experience, and (7) oral health care outcomes. Criteria 1 to 4 were implemented to ensure focus on high-quality work reporting original research. Criteria 3, 4, and 7 were also included to enable assessment of the quality of evidence. Criterion 5 admitted papers that studied gamification in a broader concept, even if it did not elicit game elements. Criteria 6 and 7 were chosen to assess reported health and well-being outcomes and potential mediators, with user experience being included given its prevalence as an outcome.

However, studies were excluded if they (1) were duplicated; (2) had full text not available; (3) were not the original article; (4) did not refer to a game; (5) were nondigital, such as conventional games like cards or board games; (6) did not concern oral health; and (7) were study protocols without outcomes to report.

Study Quality Assessment, Data Extraction, and Analysis Plan

All searched articles were filtered using broad selection criteria framed as questions:

- Does gamification in oral health care impact oral health?
- Does oral health care gamification enhance health promotion and literacy?

The study selection and data extraction were performed blindly. After the search, all references were imported into a reference management system (Mendeley), and duplicates were removed. The remaining articles’ titles and abstracts were assessed to identify eligible studies. To determine eligibility, three additional questions were asked and answered: (1) Is the topic relevant to the defined scope? (2) Does it meet the inclusion and exclusion criteria? and (3) Is the methodology appropriate? To ensure a comprehensive, transparent, and objective extraction process, a standardized prepiloted form was used to extract data from the included studies. Two reviewers independently extracted the data, and any discrepancies were resolved through discussion with a third author. Additionally, the third reviewer further scrutinized the data to verify the consistency of the extraction

XSL•FO
RenderX

process and resolved any remaining discrepancies. In case of missing or additional data, other researchers were contacted [56,57].

The eligibility assessment of each full-text paper was conducted by 2 independent raters. In cases where discrepancies arose, they were resolved through discussion and comparison of our evaluations. Articles were excluded from our review when both investigators unanimously concurred on their ineligibility due to inappropriate methodology or results that did not address the key research questions at hand. This rigorous and collaborative approach to eligibility assessment ensured the quality and relevance of the articles included in our study.

Each article was classified as having low, moderate, or high relevance. Articles were deemed highly relevant if they effectively demonstrated an impact on the considered items, whereas moderate relevance was attributed to those that projected such items. Low relevance was assigned to manuscripts that did not present any conclusions or perspectives in these domains.

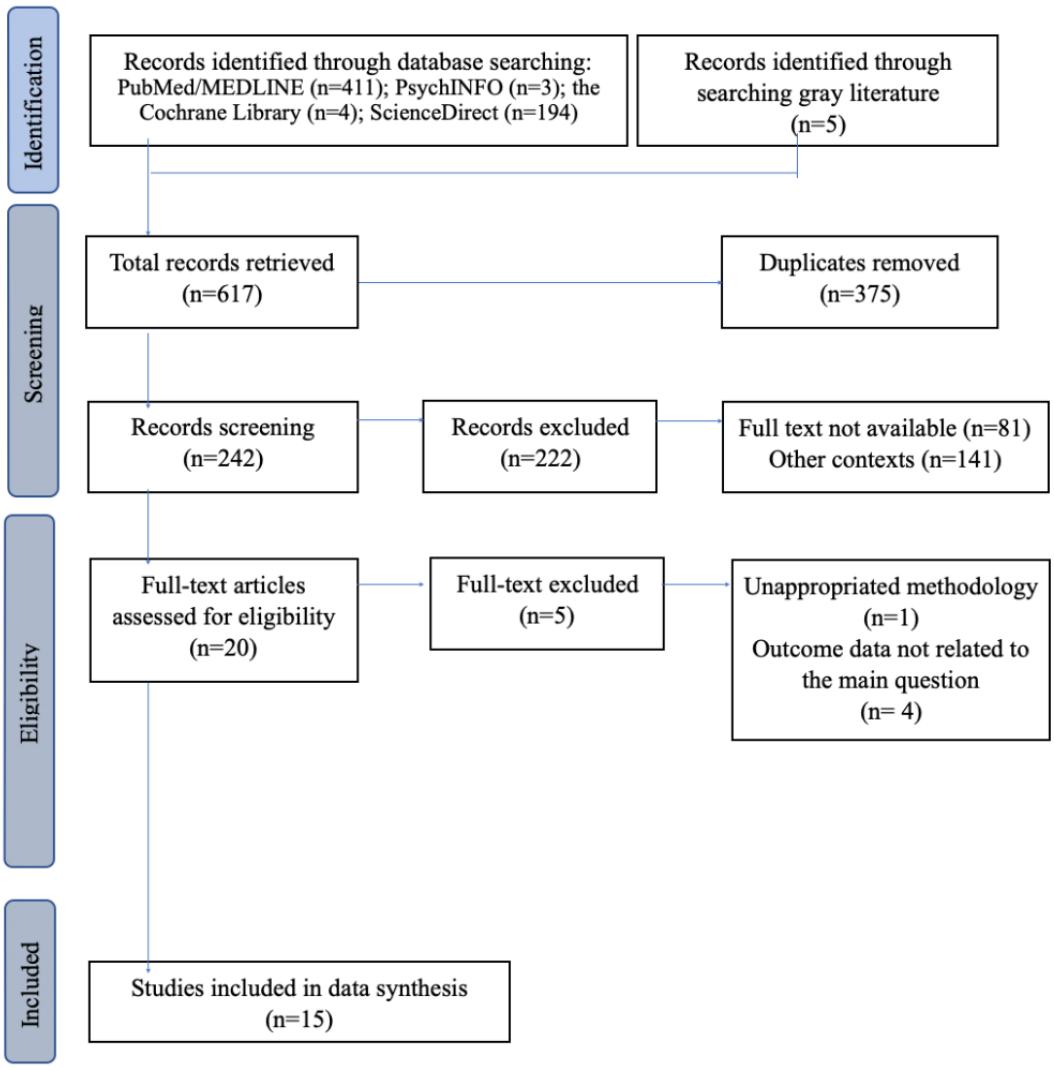
Data from the scoping literature were extracted into an Excel (Microsoft Corp) data sheet using a support checklist. The data sheet was divided into sections dedicated to a theory, area, concept, theme, or element from the framework of gamification and oral health care, including game mode, population sample, gamification components, behavior change techniques, and outcomes. After synthesizing the data and assessing the quality of the evidence, the writing of the scoping review article began.

Results

Process Selection

Figure 1 shows the PRISMA flowchart outlining the process of record identification, selection, eligibility, and inclusion. Initially, 617 records were retrieved from 5 databases and gray literature sources. After removing duplicates and records that did not meet the inclusion criteria, 15 records were selected for analysis.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



XSL•FO
RenderX

Sample Characteristics

Of the 15 studies available (see Multimedia Appendix 2 [9,18,26,27,30,58-67]), the majority (n=8, 53%) collected data at multiple time points (2 or more) from various points or conditions [18,30,58-62,68]. Of these 8 studies, 7 were blind, randomized control trials [18,30,58,59,61-63]. A content analysis of the applications for behavior change was adopted in 7 studies [9,27,60-62,64,65]. One-time cross-sectional questionnaires were applied in 3 studies [26,66,67].

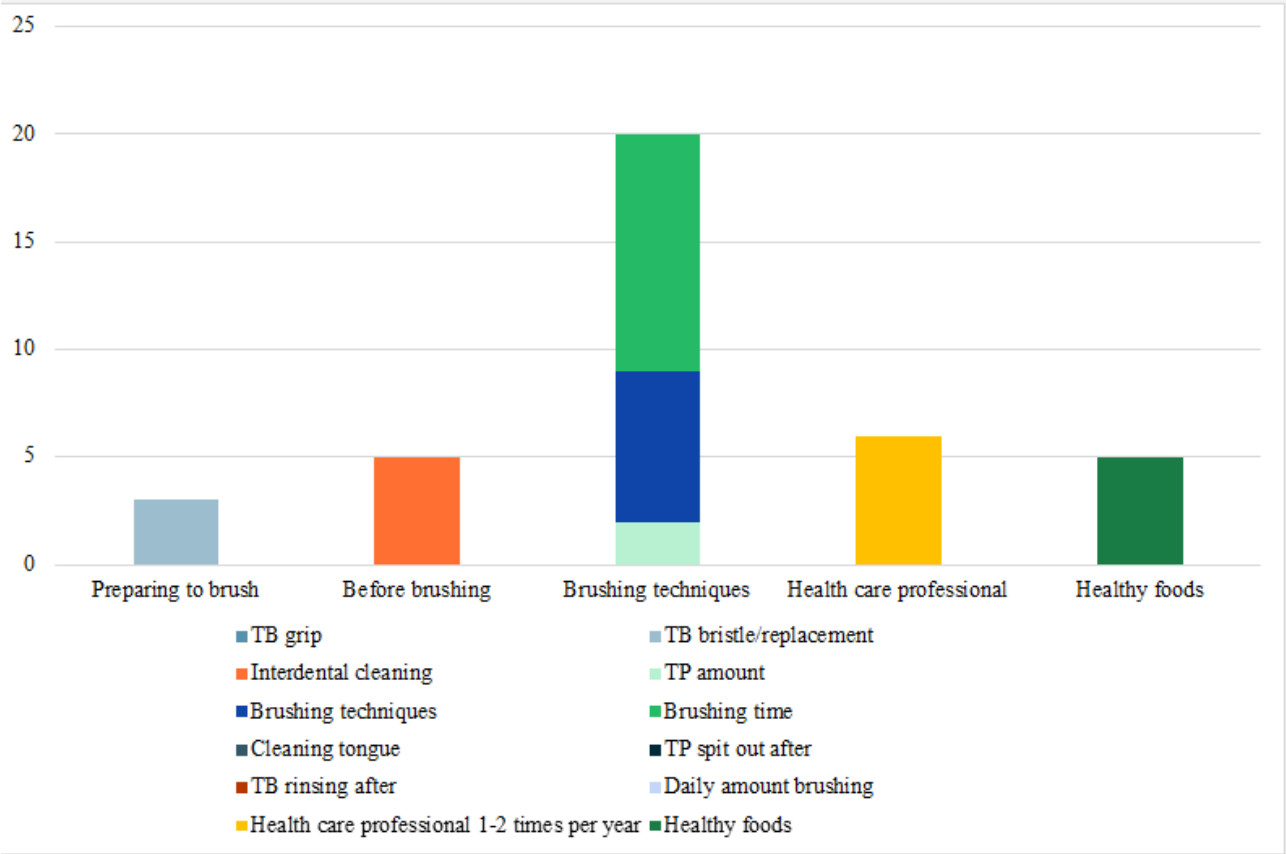
The sample size in the selected studies ranged from 34 to 190 individuals. All participants were children younger than 13 years, except for 4 studies that also included adults [26,61,62,65]. Mobile apps were the predominant modality used to change oral health care behavior in the studies (n=12), with 3 exceptions based on computer games. One of the games was available on a tablet and DVD for PC [63], whereas the other 2 apps were associated with a toothbrush sensor [58,59].

Oral Hygiene Evidence-Based Categories

The oral self-care applications supported by evidence-based oral health were mentioned in 73% (11/15) of the selected studies. The group categories and archetypes of evidence-based content related to oral hygiene were identified by analyzing the selected studies, as illustrated in Figure 2.

We found that the most clearly defined types of information in the applications were “brushing time” for at least 2 minutes (11/11, 100%) and “daily amount brushing” at least twice a day (10/11, 91%). However, “toothpaste spit out after” brushing was only found in 1 application (1/12, 8%), and “cleaning the tongue” and “toothbrush grip” were never mentioned.

Figure 2. Evidence-based content related to group categories and archetypes of oral hygiene, as recognized in the chosen studies. The y-axis displays the listed group categories and archetypes, whereas the x-axis signifies the frequency of their occurrence within the studies. The accompanying legend elucidates the color-coding system denoting distinct group categories and archetypes. TB: toothbrush; TP: toothpaste.



Behavior Change Techniques

Behavior change techniques for oral health care resulted from mechanisms of gamification, either implicit or explicit. Oral health care behavior change techniques were mentioned in 73% (11/15) of the studies. The data collected from these studies were used to create a behavior change score with 26 items [9]. The studies were analyzed to determine the frequency of behavior change techniques used in oral health applications, as shown in Table 1.

Among the behavior change techniques scrutinized, a set of 7 distinct components emerged as the most prevalently used across the array of surveyed studies. These components, along with their respective frequencies of use, are outlined as follows: “prompt intention formation” (11/26, 42%), “provide instructions” (11/26, 42%), “provide information on behavior-health link” (10/26, 38%), “provide information on consequences” (9/26, 35%), “model or demonstrate behavior” (9/26, 35%), “provide feedback on performance” (8/26, 31%), and “provide contingent rewards” (8/26, 31%).

XSL•FO
RenderX

Table 1. Frequency distribution of behavior change techniques observed in oral health applications identified in the selected studies (N=137).

Behavior change techniques	Value, n (%)
1. Provide information behavior health link	10 (7.3)
2. Provide information on consequences	8 (6.6)
3. Provide information seeking others' approval	1 (0.7)
4. Prompt intention formation	11 (8)
5. Prompt barrier identification	3 (2.2)
6. Provide general encouragement	7 (5.1)
7. Set graded tasks	2 (1.5)
8. Provide instruction	11 (8)
9. Model or demonstrate behavior	9 (6.6)
10. Prompt specific goal setting	4 (2.9)
11. Prompt review of the behavioral goals	5 (3.6)
12. Prompt self-monitoring behavior	2 (1.5)
13. Provide feedback on performance	8 (5.8)
14. Provide contingent rewards	8 (5.8)
15. Teach to use prompts or cues	4 (2.9)
16. Agree on behavioral contract	7 (5.1)
17. Prompt practice	7 (5.1)
18. Use follow-up prompts	4 (2.9)
19. Provide social comparison	2 (1.5)
20. Plan social support or change	2 (1.5)
21. Prompt identification as a role model	2 (1.5)
22. Prompt self-talk	2 (1.5)
23. Relapse prevention	4 (2.9)
24. Stress management	1 (0.7)
25. Motivational interviewing	5 (3.6)
26. Time management	7 (5.1)

Gamification Mechanisms

The game design elements based on the gamification features rating criteria for oral hygiene applications were recognized in 12 (80%) of the 15 studies. Table 2 illustrates the 26 gamification features [9] that were considered, along with the number of times each element was identified.

On average, the applications included in the study possessed an average of 10.6 of 31 potential gamification features. Notably, specific game design elements were prevalently used within various categories. Among the system design features, "meaning" (10/36, 28%) stood out prominently. Within the challenges category, "goals" (9/16, 56%) emerged as the most frequently incorporated element. For rewards, "ownership" (9/22, 41%) was notably prevalent. Among social influences, both "collaboration" and "reputation" (both 3/19, 16%) were prominent. Additionally, within the user-specific category, "ideological incentives" (10/12, 83%) exhibited a substantial presence. Interestingly, none of the applications used "badges," "conforming behavior," "virtual goods," or "self-expression" (Table 2).

XSL•FO
RenderX

Table 2. Frequency distribution of game design elements detected in the surveyed studies.

Game design elements	Value, n (%)
System design (n=36)	
Visual feedback	9 (25)
Audible feedback	4 (11)
Reminder	5 (14)
Meaning	10 (28)
Integration concepts	3 (8)
Visually resembling games	1 (3)
Fantasy	4 (11)
Challenges (n=16)	
Goals	9 (56)
Time pressure	4 (25)
Progressive disclosure	3 (19)
Rewards (n=22)	
Ownership	9 (41)
Achievement	5 (23)
Point system	4 (18)
Badges	0 (0)
Bonus	4 (18)
Social influences (n=19)	
Loss aversion	2 (11)
Status	2 (11)
Collaboration	3 (16)
Reputation	3 (16)
Competition	1 (5)
Envy	1 (5)
Shadowing	2 (11)
Social facilitation	2 (11)
Conforming behavior	0 (0)
Leaderboards	2 (11)
Altruism	1 (5)
Virtual goods	0 (0)
User specifics (n=12)	
User levels	1 (8)
Ideological incentives	10 (83)
Virtual characters	1 (8)
Self-expression	0 (0)

Discussion

Principal Findings

The oral health outcomes related to gamification interventions highlighted the role of gamification in promoting oral health care and literacy. This scoping review also highlights the limitations of currently available oral health care apps and points out the main areas to invest in for the future. A total of 11 (73%) of the 15 articles found positive impacts of using oral health apps, especially in children and adolescents. They facilitate the responsiveness of oral preventive care [66]; improve knowledge in high-risk populations; encourage dietary changes [63]; and promote a reduction of clinical plaque, gingival, and caries

XSL•FO
RenderX

indexes [26]. Additionally, they show a statistically significant improvement in health care indices [58], tooth brushing quality (duration and distribution) [60], and motivation to brush teeth for longer [26] and seem effective in adolescents with fixed orthodontic appliances by self-reported behavior and psychosocial factors [59]. Gamification structures augment oral health literacy, facilitate user alertness for oral health care themes and professional feedback, and engage commitment. A greater improvement in gingival status is commonly reported [27,61,62,69].

The feedback provided by participants showed a higher level of satisfaction in learning about oral health care through games rather than traditional noninteractive methods. Most studies reported a positive impact of gamification, particularly in children and adolescents, who are considered the main target audience of these apps [18,26,30,58-66].

The studied apps contained educational content with evidence-based dentistry and high-quality teaching for oral self-care. Some of these also feature gamification elements and behavior change techniques. The results of the studies demonstrate that these apps have excellent functionality, effectiveness, efficiency, and user satisfaction [9,64,66]. Several studies evaluating multiple oral hygiene apps have found evidence-based content, such as brushing time and daily amount of brushing. Fijačko et al [9], Parker et al [27], and Hotwani et al [64] all reported on these elements.

The health behavior change techniques found in the analyzed apps included prompt intention formation, shaping and demonstrating behavior, providing information about the link between behavior and health consequences, instructions, and contingent rewards [26,30,58,63,66,67]. A set of 7 distinct components emerged as the most prevalently employed across the array of surveyed studies. These components, along with their respective frequencies of use, are outlined as follows: “prompt intention formation,” “provide instructions,” “provide information on behavior-health link,” “provide information on consequences,” “model or demonstrate behavior,” “provide feedback on performance,” and “provide contingent rewards.” These components collectively represent the core elements of behavior change techniques that were consistently integrated into the analyzed oral health applications, aiming to enhance engagement and promote positive behavior change.

Regarding game design elements, these applications emphasized feedback, goal attainment, sense of ownership, and ideological incentives [27,61,65-67]. Parker et al [27] identified some recurring game design elements among 20 apps analyzed, such as knowledge provision, self-monitoring of frequency, and duration of toothbrushing. Hotwani et al [64] found that information provision, goal setting, feedback, progressive disclosure, and time pressure were frequently used in the 6 apps evaluated. Fijačko et al [9] analyzed 17 apps and identified time pressure, digital characters, and fantasy as key game design elements.

Comparison With Prior Work

Delivering trustworthy information to users is essential for promoting healthy habits. Health care apps should undergo validation by health care institutions and professionals before becoming public to ensure their accuracy and reliability. However, there is a risk of users becoming overly dependent, potentially compromising the need for regular appointments with oral health care professionals in real life [27,70].

Considering evidence-based oral health care, most apps emphasize brushing for at least 2 minutes and twice a day. Although this is an important core of oral hygiene recommendations, there is still room for improvement. To achieve holistic oral care, it would be advisable to incorporate other aspects such as oral hygiene techniques, the use of devices, dietary advice, sugar intake control, guidance on early childhood caries, baby oral hygiene, the effects of fluoride, the use of fluoride toothpaste, toothbrushing training videos, and regular dental visits. The development of apps should be based on theoretical models when designing educational content, and the accuracy of the content should be a priority to bring about real behavior change [61,64,70]. Sharif and Alkadhimi [65] went beyond the basics and included interdental cleaning, spitting out after brushing, avoiding mouth rinsing after brushing, characteristics of the toothbrush, and the quantity of fluoride content in toothpaste and mouthwash. Other advisable strategies include reporting about others’ approval, social interactions with oral health professionals and other application users, identifying barriers to oral hygiene and potential overcoming strategies, providing encouragement, setting graded tasks and goals, displaying tracked data and objectives, feedback on performance, setting a behavioral contract with oral health professionals, social comparison, and social support [1,9,27,35,63,65,71].

Patient adherence to a smartphone app is more effective because of the ability to customize reminders and prompts, constant accessibility, adjustability to the user, ability to provide tailored feedback, widespread use, and interactive features [63,72,73].

Strengths and Limitations

Despite the proven efficiency of gamification in health care promotion and prevention, it remains an unexplored territory in oral health, mainly applied to specific educational purposes and oral health promotion [74]. This scoping review highlights the limitations of currently available oral health care applications and points out the main areas to invest in for the future. Two major limitations were found in this study. The first is the limited availability of articles related to the main topics. Gamification, within the context of health care, is not a recent concept. However, its quantitative assessment, particularly in the field of oral health, remains relatively uncommon. To address such limitation and ensure a comprehensive review, we diligently accessed and explored 5 databases, along with incorporating gray literature sources. This approach was essential to include as many relevant studies as possible, aligning with our predefined inclusion criteria. Considering the diverse nature of our search strategy, we believe that we have made every effort to provide a reliable representation of the existing literature in the field of gamification in oral health care.

The second limitation is the heterogeneous studies with varying focus and dispersing attention. The diversity of approaches makes a rigorous comparison more challenging.

XSL•FO
RenderX

Taking into account the multiple aspects involved in gamification strategies and by transparently outlining these parameters, we believe that our work can serve as a valuable reference for future researchers seeking to design studies that address and overcome these challenges. Our hope is that this will pave the way for a more effective understanding of the underlying mechanisms behind the implementation of gamification in oral health context.

Future Directions

Future studies could focus on other age groups as well, such as the study by Zolfaghari et al [61], which developed applications for mothers of children, improving their oral health literacy and practice and promoting plaque control in children within just 1 month of use [61,74,75].

There is potential for future optimization of key gamification features, such as badges, encouragement of correct behavior, digital goods, self-expression, reminders, fantasy themes, time pressure, disclosure of progress, achievements, points systems, bonuses, loss aversion, status, collaboration, reputation, competition, shadowing, social facilitation, leaderboards, altruism, user levels, and digital characters [26,27,30,66,71,76].

Conclusions

Gamification in oral health care does have an impact; it enhances oral health promotion and literacy. It represents a potential new approach for oral health care providers to change people's oral health behavior. The most frequent game design mechanisms adopted were meaning, ideological incentives, feedback, goals, and ownership. Some authors have highlighted several factors for gamification success, including engagement strategy, applications aesthetics, evidence-based information content, behavioral change taxonomies, attention to psychological needs, evaluation, validation, quality assessment, and professional regulation standards for oral health care applications. More studies are needed to better understand the clinical, psychological, and social processes involved in selecting the most efficient gamification mechanisms. The process of mobile health in oral health care is in the initial stage, but gamification is crucial for improving individual health-related practices.

Acknowledgments

The authors would like to express their gratitude to Professor Tiago Taveira-Gomes for his mentorship, review of the manuscript, and unwavering support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search string.
[PNG File , 248 KB - ijmr_v13i1e35132_app1.png]

Multimedia Appendix 2
Full paper details.
[DOCX File , 29 KB - ijmr_v13i1e35132_app2.docx]

Multimedia Appendix 3
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 54 KB - ijmr_v13i1e35132_app3.pdf]

References

1. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
2. Stevens G, Mascarenhas M, Mathers C. Global health risks: progress and challenges. Bull World Health Organ 2009 Sep;87(9):646 [FREE Full text] [doi: 10.2471/blt.09.070565] [Medline: 19784438]
3. Glanz K, Rimer BK, Viswanath K. Introduction to health behavior theories that focus on individuals. In: Health Behavior and Health Education: Theory, Research, and Practice, 4th Edition. San Francisco, CA: Jossey-Bass; Jul 2008:67-74.
4. Schroeder SA. Shattuck Lecture. We can do better—improving the health of the American people. N Engl J Med 2007 Sep 20;357(12):1221-1228. [doi: 10.1056/NEJMsa073350] [Medline: 17881753]
5. Oral health: fact sheet. World Health Organization. 2023 Mar 20. URL: https://www.who.int/news-room/fact-sheets/detail/oral-health [accessed 2020-06-24]
6. Petersen PE. The World Oral Health Report 2003: continuous improvement of oral health in the 21st century—the approach of the WHO Global Oral Health Programme. Community Dent Oral Epidemiol 2003 Dec;31(Suppl 1):3-23. [doi: 10.1046/j..2003.com122.x] [Medline: 15015736]

XSL•FO
RenderX

7. Nunn JH. The burden of oral ill health for children. Arch Dis Child 2006 Mar;91(3):251-253 [FREE Full text] [doi: 10.1136/adc.2005.077016] [Medline: 16492889]
8. Rabiei S, Mohebbi SZ, Patja K, Virtanen JI. Physicians' knowledge of and adherence to improving oral health. BMC Public Health 2012 Oct 09;12:855 [FREE Full text] [doi: 10.1186/1471-2458-12-855] [Medline: 23046660]
9. Fijačko N, Gosak L, Cilar L, Novšak A, Creber RM, Skok P, et al. The effects of gamification and oral self-care on oral hygiene in children: systematic search in app stores and evaluation of apps. JMIR Mhealth Uhealth 2020 Jul 08;8(7):e16365 [FREE Full text] [doi: 10.2196/16365] [Medline: 32673235]
10. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
11. Political declaration of the high-level meeting of the general assembly on the prevention and control of non-communicable diseases: draft resolution / submitted by the President of the General Assembly. UN General Assembly. New York: United Nations URL: https://digitallibrary.un.org/record/710899 [accessed 2023-01-24]
12. Porter ME, Teisberg EO. How physicians can change the future of health care. JAMA 2007 Mar 14;297(10):1103-1111. [doi: 10.1001/jama.297.10.1103] [Medline: 17356031]
13. Porter ME. What is value in health care? N Engl J Med 2010 Dec 23;363(26):2477-2481. [doi: 10.1056/NEJMp1011024] [Medline: 21142528]
14. Reeve J. Understanding Motivation and Emotion, Sixth Edition. New York: Wiley; Nov 3, 2014.
15. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
16. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
17. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
18. Panic K, Cauberghe V, de Pelsmacker P. Promoting dental hygiene to children: comparing traditional and interactive media following threat appeals. J Health Commun 2014;19(5):561-576. [doi: 10.1080/10810730.2013.821551] [Medline: 24393019]
19. Wattanasoontorn V, Boada I, García R, Sbert M. Serious games for health. Entertainment Computing 2013 Dec;4(4):231-247. [doi: 10.1016/j.entcom.2013.09.002]
20. Baranowski T, Buday R, Thompson DI, Baranowski J. Playing for real: video games and stories for health-related behavior change. Am J Prev Med 2008;34(1):74-82 [FREE Full text] [doi: 10.1016/j.amepre.2007.09.027] [Medline: 18083454]
21. Kharrazi H, Lu AS, Gharghabi F, Coleman W. A scoping review of health game research: past, present, and future. Games Health J 2012;1(2):153-164 [FREE Full text] [doi: 10.1089/g4h.2012.0011] [Medline: 24416638]
22. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Human–Computer Interaction 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
23. Sawyer B. Games for health. MIT Media Lab Talks. Cambridge, MA: MIT; 2014. URL: https://www.media.mit.edu/video/view/wellbeing2014-11-05 [accessed 2023-01-25]
24. Munson SA, Poole E, Perry DB, Peyton T. Gamification and health. In: Walz SP, Deterding S, editors. The Gameful World: Approaches, Issues, Applications. Cambridge, London: MIT Press; 2015:597-623.
25. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
26. Underwood B, Birdsall J, Kay E. The use of a mobile app to motivate evidence-based oral hygiene behaviour. Br Dent J 2015 Aug 28;219(4):E2. [doi: 10.1038/sj.bdj.2015.660] [Medline: 26315196]
27. Parker K, Bharmal RV, Sharif MO. The availability and characteristics of patient-focused oral hygiene apps. Br Dent J 2019 Apr;226(8):600-604. [doi: 10.1038/s41415-019-0197-7] [Medline: 31028329]
28. Annual number of mobile app downloads worldwide 2020. Statista. URL: https://www.statista.com/statistics/271644/worldwide-free-and-paid-mobile-app-store-downloads/ [accessed 2021-03-27]
29. Bohn CE, McQuistan MR, McKernan SC, Askelson NM. Preferences related to the use of mobile apps as dental patient educational aids: a pilot study. J Prosthodont 2018;27(4):329-334. [doi: 10.1111/jopr.12667] [Medline: 28872732]
30. Zotti F, Dalessandri D, Salgarello S, Piancino M, Bonetti S, Visconti L, et al. Usefulness of an app in improving oral hygiene compliance in adolescent orthodontic patients. Angle Orthod 2016;86(1):101-107 [FREE Full text] [doi: 10.2319/010915-19.1] [Medline: 25799001]
31. Dubois D, Tamburrelli G. Understanding gamification mechanisms for software development. New York, NY, USA: Association for Computing Machinery; 2013 Presented at: Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering; August 18-26, 2013; Saint Petersburg, Russia p. 659-662. [doi: 10.1145/2491411.2494589]
32. Morschheuser BS, Rivera-Pelayo V, Mazarakis A, Zacharias V. Interaction and reflection with quantified self and gamification: an experimental study. J Lit Technol 2014;15(2):136-156 [FREE Full text]

XSL•FO
RenderX

33. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". New York, NY, US: Association for Computing Machinery; 2011 Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28-30, 2011; Tampere, Finland p. 9-15 URL: https://dl.acm.org/doi/abs/10.1145/2181037.2181040 [doi: 10.1145/2181037.2181040]
34. Marston HR, Hall AK. Gamification: applications for health promotion and health information technology engagement. In: Novák D, Tulu B, Brendryen H, editors. Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice. Hershey, PA: IGI Global; 2016:78-104.
35. Blohm I, Leimeister JM. Gamification. Wirtschaftsinf 2013;55(4):275-278. [doi: 10.1007/s11576-013-0368-0]
36. Huotari K, Hamari J. Defining gamification: a service marketing perspective. New York, NY, US: Association for Computing Machinery; 2012 Presented at: Proceeding of the 16th International Academic MindTrek Conference; October 3-5, 2012; Tampere, Finland p. 17-22. [doi: 10.1145/2393132.2393137]
37. Rigby S, Ryan RM. Glued to Games: How Video Games Draw Us in and Hold Us Spellbound. Santa Barbara, CA: Bloomsbury Academic; 2011.
38. Seaborn K, Fels DI. Gamification in theory and action: a survey. Int J Hum Comput 2015;74:14-31. [doi: 10.1016/j.ijhcs.2014.09.006]
39. Deci EL, Ryan RM. Motivation, personality, and development within embedded social contexts: an overview of self-determination theory. In: Ryan RM, editor. The Oxford Handbook of Human Motivation. New York: Oxford University Press; 2012:85-108.
40. Fortier MS, Duda JL, Guerin E, Teixeira PJ. Promoting physical activity: development and testing of self-determination theory-based interventions. Int J Behav Nutr Phys Act 2012;9:20 [FREE Full text] [doi: 10.1186/1479-5868-9-20] [Medline: 22385751]
41. Ng JYY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
42. Patrick H, Williams GC. Self-determination theory: its application to health behavior and complementarity with motivational interviewing. Int J Behav Nutr Phys Act 2012;9:18 [FREE Full text] [doi: 10.1186/1479-5868-9-18] [Medline: 22385676]
43. Teixeira PJ, Palmeira AL, Vansteenkiste M. The role of self-determination theory and motivational interviewing in behavioral nutrition, physical activity, and health: an introduction to the IJBNPA special series. Int J Behav Nutr Phys Act 2012;9:17 [FREE Full text] [doi: 10.1186/1479-5868-9-17] [Medline: 22385644]
44. McGonigal J. Reality Is Broken: Why Games Make Us Better and How They Can Change the World. New York: Penguin Press; 2011.
45. Nieborg DB. Am I mod or not?—an analysis of first person shooter modification culture. 2005 Presented at: Creative Gamers Seminar—Exploring Participatory Culture in Gaming; January 14-15, 2005; University of Tampere, Finland.
46. Sotamaa O. "Have fun working with our product!": critical perspectives on computer game mod competitions. 2005 Presented at: DiGRA Conference; 2005; Vancouver, British Columbia, Canada.
47. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
48. King D, Greaves F, Exeter C, Darzi A. 'Gamification': influencing health behaviours with games. J R Soc Med 2013;106(3):76-78 [FREE Full text] [doi: 10.1177/0141076813480996] [Medline: 23481424]
49. Pereira P, Duarte E, Rebelo F, Noriega P. A review of gamification for health-related contexts. In: Marcus A, editor. Design, User Experience, and Usability: User Experience Design for Diverse Interaction Platforms and Environments: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part II. Cham: Springer; 2014:742-753.
50. Sola D, Couturier J, Voyer B. Unlocking patient activation in chronic disease care. Br J Health Care Manag 2015;21(5):220-225. [doi: 10.12968/bjhc.2015.21.5.220]
51. Hamari J, Koivisto J, Sarsa H. Does gamification work?—a literature review of empirical studies on gamification. : IEEE; 2014 Presented at: 47th Hawaii International Conference on System Sciences; January 6-9, 2014; Waikoloa, HI, USA p. 3025-3034. [doi: 10.1109/hicss.2014.377]
52. Juul J. Gamification backlash roundup. The Ludologist. 2011. URL: https://www.jesperjuul.net/ludologist/2011/04/02/gamification-backlash-roundup/ [accessed 2024-01-25]
53. Kawachi I. It's all in the game—the uses of gamification to motivate behavior change. JAMA Intern Med 2017;177(11):1593-1594. [doi: 10.1001/jamainternmed.2017.4798] [Medline: 28973152]
54. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
55. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]
56. Center for Health Studies and Research of the University of Coimbra. URL: https://www.uc.pt/en/ceisuc/ [accessed 2022-12-15]

XSL•FO
RenderX

57. Repositório de Instrumentos de Medição e Avaliação em Saúde. RIMAS. 2022. URL: http://rimas.uc.pt/ [accessed 2024-01-25]
58. Alkilzy M, Midani R, Höfer M, Splieth C. Improving toothbrushing with a smartphone app: results of a randomized controlled trial. Caries Res 2019;53(6):628-635. [doi: 10.1159/000499868] [Medline: 31132765]
59. Scheerman JFM, van Meijel B, van Empelen P, Verrips GHW, van Loveren C, Twisk JWR, et al. The effect of using a mobile application ("WhiteTeeth") on improving oral hygiene: a randomized controlled trial. Int J Dent Hyg 2020;18(1):73-83 [FREE Full text] [doi: 10.1111/idh.12415] [Medline: 31291683]
60. Jacobson D, Jacobson J, Leong T, Lourenco S, Mancl L, Chi DL. Evaluating child toothbrushing behavior changes associated with a mobile game app: a single arm pre/post pilot study. Pediatr Dent 2019;41(4):299-303 [FREE Full text] [Medline: 31439090]
61. Zolfaghari M, Shirmohammadi M, Shahhosseini H, Mokhtaran M, Mohebbi SZ. Development and evaluation of a gamified smart phone mobile health application for oral health promotion in early childhood: a randomized controlled trial. BMC Oral Health 2021;21(1):18 [FREE Full text] [doi: 10.1186/s12903-020-01374-2] [Medline: 33413304]
62. Shirmohammadi M, Razeghi S, Shamshiri AR, Mohebbi SZ. Impact of smartphone application usage by mothers in improving oral health and its determinants in early childhood: a randomised controlled trial in a paediatric dental setting. Eur Arch Paediatr Dent 2022;23(4):629-639 [FREE Full text] [doi: 10.1007/s40368-022-00731-9] [Medline: 35841512]
63. Aljafari A, Gallagher JE, Hosey MT. Can oral health education be delivered to high-caries-risk children and their parents using a computer game?—a randomised controlled trial. Int J Paediatr Dent 2017;27(6):476-485 [FREE Full text] [doi: 10.1111/ipd.12286] [Medline: 28052471]
64. Hotwani K, Sharma K, Nagpal D, Lamba G, Chaudhari P. Smartphones and tooth brushing: content analysis of the current available mobile health apps for motivation and training. Eur Arch Paediatr Dent 2020;21(1):103-108. [doi: 10.1007/s40368-019-00457-1] [Medline: 31154620]
65. Sharif MO, Alkadhimi A. Patient focused oral hygiene apps: an assessment of quality (using MARS) and knowledge content. Br Dent J 2019;227(5):383-386 [FREE Full text] [doi: 10.1038/s41415-019-0665-0] [Medline: 31520041]
66. Ito C, Marinho Filho AV, Ito M, Azevedo MM, de Almeida MA. Preliminary evaluation of a serious game for the dissemination and public awareness on preschool children's oral health. Stud Health Technol Inform 2013;192:1034. [Medline: 23920808]
67. Campos LFXA, Cavalcante JP, Machado DP, Marçal E, de B Silva PG, Rolim JPML. Development and evaluation of a mobile oral health application for preschoolers. Telemed J E Health 2019;25(6):492-498. [doi: 10.1089/tmj.2018.0034] [Medline: 30036174]
68. Aljafari A, Rice C, Gallagher JE, Hosey MT. An oral health education video game for high caries risk children: study protocol for a randomized controlled trial. Trials 2015;16:237 [FREE Full text] [doi: 10.1186/s13063-015-0754-6] [Medline: 26016586]
69. de Oliveira LBJ, Jorge MSB. Uso da gamificação com vistas à prevenção e promoção da saúde bucal: protocolo de revisão de escopo. Use of gamification for the prevention and promotion of oral health: scope review protocol. Uso de la gamificación para la prevención y promoción de la saludbucal: protocolo de revisión del alcance [Article in Portuguese]. Res Soc Dev 2022;11(9):e31811931939. [doi: 10.33448/rsd-v11i9.31939]
70. Wang K, Yu KF, Liu P, Lee GHM, Wong MCM. Can mHealth promotion for parents help to improve their children's oral health? a systematic review. J Dent 2022;123:104185 [FREE Full text] [doi: 10.1016/j.jdent.2022.104185] [Medline: 35691452]
71. Moreira R. Gamificação na promoção da saúde oral. Gamification in oral health promotion [Article in Portuguese]. In: Conclusion of MBA Advanced Course on Management of Healthcare Organizations and Services. Porto, Portugal: Repository of the University Fernando Pessoa; 2021.
72. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
73. Wang VCX. Handbook of Research on Advancing Health Education through Technology. Hershey, PA: IGI Global; 2015.
74. Zaror C, Mariño R, Atala-Acevedo C. Current state of serious games in dentistry: a scoping review. Games Health J 2021;10(2):95-108 [FREE Full text] [doi: 10.1089/g4h.2020.0042] [Medline: 33818135]
75. Figueiredo M, Garcia M, Barone DAC, Roberta OR, Lurmmerz G. Gamification in oral health: experience with school of rural area. Rev ABENO 2015;15(3):98-108 [FREE Full text]
76. Scheerman JFM, van Meijel B, van Empelen P, Kramer GJC, Verrips GHW, Pakpour AH, et al. Study protocol of a randomized controlled trial to test the effect of a smartphone application on oral-health behavior and oral hygiene in adolescents with fixed orthodontic appliances. BMC Oral Health 2018;18(1):19 [FREE Full text] [doi: 10.1186/s12903-018-0475-9] [Medline: 29415697]

Abbreviations

PICO: Patient, Intervention, Comparison, Outcome

XSL•FO
RenderX

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T de Azevedo Cardoso; submitted 22.11.21; peer-reviewed by G Tobias, A Teles; comments to author 03.02.22; revised version received 30.04.22; accepted 08.11.23; published 04.04.24.

Please cite as:
Moreira R, Silveira A, Sequeira T, Durão N, Lourenço J, Cascais I, Cabral RM, Taveira Gomes T
Gamification and Oral Health in Children and Adolescents: Scoping Review
Interact J Med Res 2024;13:e35132
URL: https://www.i-jmr.org/2024/1/e35132
doi:10.2196/35132
PMID:38573750

©Rui Moreira, Augusta Silveira, Teresa Sequeira, Nuno Durão, Jessica Lourenço, Inês Cascais, Rita Maria Cabral, Tiago Taveira Gomes. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 04.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of the Accuracy, Credibility, and Readability of Statin-Related Websites: Cross-Sectional Study

Eunice Ling[1], BAppSc, MD; Domenico de Pieri[1*], MD; Evenne Loh[1*], MD; Karen M Scott[1], BEd, MA, PhD; Stephen C H Li[2], MBBS, MSci, MPH, MBA, MAACB; Heather J Medbury[1,3], BSc, PhD

[1]Sydney Medical School, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[2]Institute of Clinical Pathology and Medical Research, Westmead Hospital, Westmead, Australia

[3]Vascular Biology Research Centre, Surgery, Westmead Hospital, Westmead, Australia

[*]these authors contributed equally

Corresponding Author:
Heather J Medbury, BSc, PhD
Vascular Biology Research Centre
Surgery
Westmead Hospital
REN Building
Westmead, 2145
Australia
Phone: 61 8890 3668
Email: heather.medbury@sydney.edu.au

Abstract

Background: Cardiovascular disease (CVD) represents the greatest burden of mortality worldwide, and statins are the most commonly prescribed drug in its management. A wealth of information pertaining to statins and their side effects is on the internet; however, to date, no assessment of the accuracy, credibility, and readability of this information has been undertaken.

Objective: This study aimed to evaluate the quality (accuracy, credibility, and readability) of websites likely to be visited by the general public undertaking a Google search of the side effects and use of statin medications.

Methods: Following a Google web search, we reviewed the top 20 consumer-focused websites with statin information. Website accuracy, credibility, and readability were assessed based on website category (commercial, not-for-profit, and media), website rank, and the presence or absence of the Health on the Net Code of Conduct (HONcode) seal. Accuracy and credibility were assessed following the development of checklists (with 20 and 13 items, respectively). Readability was assessed using the Simple Measure of Gobbledegook scores.

Results: Overall, the accuracy score was low (mean 14.35 out of 20). While side effects were comprehensively covered by 18 websites, there was little information about statin use in primary and secondary prevention. None of the websites met all criteria on the credibility checklist (mean 7.8 out of 13). The median Simple Measure of Gobbledegook score was 9.65 (IQR 8.825-10.85), with none of the websites meeting the recommended reading grade of 6, even the media websites. A website bearing the HONcode seal did not mean that the website was more comprehensive or readable.

Conclusions: The quality of statin-related websites tended to be poor. Although the information contained was accurate, it was not comprehensive and was presented at a reading level that was too difficult for an average reader to fully comprehend. As such, consumers risk being uninformed about this pharmacotherapy.

(Interact J Med Res 2024;13:e42849) doi:10.2196/42849

KEYWORDS

statins; consumer health information; readability; credibility; accuracy; digital health, health information seeking; cardiovascular; mortality; management; pharmacotherapy; risk; medication

XSL•FO
RenderX

Introduction

Background

Cardiovascular diseases (CVDs) are the primary cause of death globally, with an estimated 17.9 million people dying of CVDs in 2021. This represents 31% of all global deaths. Of these deaths, 85% are due to heart attack and stroke, whose most common etiology is atherosclerosis [1]—the development of fatty plaque within artery walls. A key pharmacological treatment for atherosclerosis is statin therapy. It has a role in the primary and secondary prevention of vascular events, with a lowering of low-density lipoprotein cholesterol leading by 2 mmol/L, typically giving a 10% absolute benefit (the reduction in the probability of an event's occurrence within a population receiving treatment) for those diagnosed with vascular disease and a 5% absolute benefit for those with risk factors yet without having experienced a vascular event [2]. This creates issues when we consider that patients may be biochemically abnormal (with hypercholesterolemia) but asymptomatic. Such patients may doubt the use of the prescribed statin therapy as they determine the cost-benefit balance between tangible adverse effects and theoretical benefits. This may prompt the consultation of alternative sources of knowledge to aid decision-making. In this era of shared decision-making, where patients participate in the medical decisions that affect their health [3], it is essential that the information they access is high quality and easily understood.

Use of the Internet for Health Information Seeking

In this milieu, the internet has risen as a key source of health-related information, with 79% of adults seeking web-based health information in America and 79% to 86% in China, the Philippines, Hong Kong, Indonesia, and Vietnam [4,5]. Comparatively, seeking web-based health information is as popular as playing games or downloading music from the web [6]. Notably, the COVID-19 pandemic has presented unprecedented challenges, catapulting society further into a future dependence on telehealth and internet-assisted health care [7,8]. As such, traditional in-clinic and leaflet modes of health information delivery are being supplemented, and in some cases supplanted, by internet searches. With this dramatic change in the terrain upon which patients and their families are attaining information, it is crucial to determine the quality of web-based health information put forth to them.

Health literacy is defined as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" [9]. It requires a complex group of skills such as reading, listening, analyzing, and decision-making, as well as the ability to apply the aforementioned skills to health situations [10]. Those with poor health literacy are vulnerable to undertaking unnecessary tests and treatments or, conversely, refusing beneficial tests and treatments. In part, they may be misled into assessing the quality of web-based health information based on its search result ranking, image quality, celebrity endorsement, and website authorship rather than relying on the criteria of established quality guidelines [11,12]. Health literacy–related knowledge and skills are particularly deficient among vulnerable populations, who are also more likely to experience CVD [12,13]. Unfortunately, those with poor health literacy are susceptible to the influence of mass media and emotionally persuasive texts. This may explain the response to the television program "Catalyst" in Australia [14], whose criticism of statins resulted in 11% of patients (in a survey by the Australian National Heart Foundation of 1094 patients) who watched the program ceasing to take their cholesterol-lowering medication and significant and sustained changes in statin usage, with 2.6% fewer statins (equivalent to 14,005 dispensing) each week [15].

Internet Standards

There is considerable heterogeneity in the quality of web-based health information [16]. The quality of information can be examined in 3 domains: accuracy, credibility, and readability [17]. Each is defined as follows: accuracy is the intent to be evidence based and safe by adequately offering a complete, unbiased picture and its relevance [17]; credibility is the attribution of source and authorship and the disclosure of conflicts of interest for the presented information [18]; and readability is the ease of understanding due to the style of writing, describing the reading comprehension level a layperson requires to understand a text [19,20]. Complex wording reduces engagement with, and application of, content [21]. This leaves patients vulnerable to becoming ill-informed and at risk of adverse health outcomes [22]. The quality of web-based health information is, therefore, a pressing issue, which the Health on the Net Foundation aims to address by providing a Health on the Net Code of Conduct (HONcode) seal, an internet-based certification of medical and health websites that adheres to a set of publishing principles regarding the source and purpose of medical information. However, few consumers are aware of Health on the Net or the HONcode.

While there have been assessments of the quality of websites on many key areas of health care, such as diabetes [23], obesity [24], and hypertension [25], and on surgical interventions [26,27], there is a dearth of assessment about the quality of web-based information pertaining to medications. Such assessments are necessary to inform clinicians of the quality of content likely to be accessed by their patients, who are particularly interested in the likelihood and nature of adverse side effects. We aimed to assess the quality of consumer health information on websites about statins through consideration of accuracy, credibility, and readability.

Methods

Study Design

Overview

Through this cross-sectional study, we analyzed the accuracy, credibility, and readability of websites that were most commonly presented to patients searching for the keywords "statin" and "statin side effects." We determined each website's search engine ranking, category (commercial, not-for-profit, or media), and the presence or absence of the HONcode seal. Furthermore,

XSL•FO
RenderX

we determined the relationship between the accuracy, credibility, and readability of the websites found.

In selecting websites on statins to analyze, we aimed to emulate a typical consumer's search for web-based health information. A web search was conducted using Google, in keeping with evidence that it accounts for more than half of all web traffic [28-30] and is an increasingly preferred search engine by the general public [31]: 91% of American adults using the internet use a search engine, and of those, 83% use Google more often than other search engines [4,32]. To conduct the search, location filters, user information, search history, cached data, and cookies were disabled, and sponsored results were excluded to avoid inadvertent search bias. The search terms used were "statin" and "statin side effects," following the advice of our lipidologist coauthor (SL) that the generic term "statin," rather than specific medication names, was commonly used in discussion with patients in clinical practice and that "statin side effects" were a key concern of patients.

The first 20 ranked websites in the Google search results page were analyzed (after removing any duplication from search results of the 2 search terms). We did not identify further websites in the search results, given that, in general, websites returned on the first Google search results page generate 92% of all traffic from an average search [16,33]. This drops by 95% for the second page and by 78% and 58% for subsequent pages [31]. Thus, we did not aim to identify all websites on statins but rather to emulate an authentic consumer search.

Website Search Rank

The effect of the association between search result ranking and accuracy, credibility, and readability was considered. Given that engagement is highest with the first 5 websites in search results, garnering 67% of all clicks from a search results page [31,34], the websites were divided into 4 sets of 5 websites, each according to their ranking as per Google search result. Thus, websites ranked 1-5 were called quartile 1, websites 6-10 were called quartile 2, and so forth.

Website Categorization

In the interest of determining whether the nature of the authorship of the websites had a bearing on their accuracy, credibility, and readability, each of the 20 chosen websites was categorized into 3 types: commercial (defined as a website that generates revenue or cash and is not affiliated with the government), not-for-profit (a website that garners support for a cause rather than revenue, including government and charities), and media (a website that reports new findings or stories, with the primary purpose of the website being news reporting).

Presence of HONcode Seal

We assessed whether the quality of the websites was associated with the presence or absence of the HONcode seal.

Measures

Overview

In assessing the accuracy, credibility, and readability of the websites, we considered existing tools and developed study-specific tools where necessary.

Accuracy

For accuracy, we developed a statin-specific tool that took into account medical guidelines. The checklist (Table 1) was designed after referring to other studies purporting to assess the quality of web-based health information [17]. A key difference here was that we were looking at a specific treatment. Three features were considered: (1) the intent to be evidence based; (2) safety, in that a website should adequately offer users a complete, unbiased picture of statin treatment; and (3) relevance, in that it is reasonable to expect a website to address the criteria in the checklist [17]. The checklist was intrinsically linked to a website's comprehensiveness, consistent with other studies that have evaluated completeness as an integral part of accuracy [16].

To meet these features, guidelines from the American College of Cardiology and the Australian Heart Foundation [37] were synthesized into short statements, which formed the accuracy criteria. These statements formed a checklist that each website was required to address to be considered "accurate." The development of the statements was further informed and determined by a review of the treatment of cholesterol in light of its evidence base [38], as well as criteria from the treatment section of the validated DISCERN tool [39]. This section (items 9-16 of the DISCERN tool) addresses issues of risk, benefit, and how the website guides decision-making surrounding treatment options [39]. Combining these sources ensured that a higher score would be awarded to websites providing the most evidence-based information. In total, 20 equally weighted criteria were devised, and a score of 20 was arbitrarily defined as a minimum acceptable standard.

Each accuracy item was scored as: "present and complete" (2), "present but incomplete" (1), "absent" (0), or "inaccurate" (*). A maximum score of 40 could be awarded to each website assessed. "Incompleteness" was defined as a nonexact or indirect mention of a topic outlined in a criterion rather than an explicit statement. Two reviewers (DdP and E loh) completed the assessment.

Any hyperlinks that navigated to information within a website were followed and the data were included in the final assessment; links leading to external websites were not followed. Embedded videos were analyzed. Once each reviewer concluded their analysis, the results were compared. Discrepancies were resolved through discussion until reaching a consensus.

XSL•FO
RenderX

Table 1. Website accuracy checklist.

Accuracy criteria	Descriptor
1	Mentions that cholesterol is a modifiable risk factor for cardiovascular disease [35]
2	Mentions that consultation with a doctor is essential before and while taking statins and when ceasing them [36]
3	Lists conditions for which statins are used [36]
4	Defines the target population for statin therapy [37]
5	Mentions the importance of adherence to statin therapy [36]
6	Addresses the subtleties of primary prevention [38]
7	Mentions that statins are about reducing complications of high cholesterol rather than achieving a specific (low-density lipoprotein) cholesterol [38]
8	Describes or at least lists the benefits of statin therapy [39]
9	Describes or at least lists the side effects or risks of statin therapy [39]
10	Describes how treatment affects the overall quality of life [39]
11	Mentions low to moderate dose statin therapy is recommended in primary prevention [37]
12	Specifically addresses rhabdomyolysis [36]
13	Describes the approximate financial burden to the patient [36]
14	Describes the duration of treatment before an effect is measurable [39]
15	Describes how statins work or at least what they do [39]
16	Describes what may happen without treatment [39]
17	Explores the possibility of using alternative therapies to statins [39]
18	Mentions that statins must not be used during pregnancy [37]
19	Describes drug interactions or at least lists them [39]
20	Mentions that statins do not replace a healthy lifestyle [36]

Credibility

In developing the criteria to be included in the assessment of credibility (Table 2), DISCERN was chosen as a reference, as well as other studies that used DISCERN or another available tool for website assessment. However, as the 5-point Likert scale used in DISCERN can be subject to response style bias [40], a present (1) or absent (0) scale was adopted as it has been shown to improve the objectivity of data collection [41-43].

XSL•FO
RenderX

Table 2. Website credibility checklist.

Item	Criteria	Score criteria
1	Referencing or citations obtained from peer-reviewed journals	1 point if the articles for which the references are obtained are published in peer-reviewed journals [16,44]
2	Website updated within last 24 months	The latest update should be within the past 24 months [45]
3	Avoids anecdotal evidence for making claims	Does not use anecdotal evidence as a basis for claims; quoting a case study without using claims is acceptable [46]
4	Mailing address present	Physical contact address of the website clearly stated [45]
5	Contact information available	Contact information including name, position, telephone number, address, and email [47]
6	Sponsorship stated	Any sponsorship should be clearly stated
7	Organizational privacy policy stated	Organization privacy policy should be clearly stated [47]
8	Declaration of the author's qualification	Author's qualification should be health care related [16,44]
9	Paid access tab present	If paid access is available, the difference in the information obtained from paid vs unpaid access should be clearly stated [48]
10	Disclosure of funding or conflicts of interest	Conflicts of interest and funding disclosure should be clearly stated [44]
11	The presence of an HONcode seal or third-party certification	Presence of a HONcode[a] seal or any other third-party certification [16]
12	Advertisement neutral	Advertisements should steer clear from the website information (eg, no pop-ups related to the website content) [16]
13	Disclaimer regarding web-based health information	A disclaimer should be clearly stated that web-based health information does not replace a practitioner's advice [44]

[a]HONcode: Health on the Net Code of Conduct.

Each website was appraised according to this list. A score was allocated for each website's front page, with internal links explored only if relevant. Data (credibility scores) were undertaken as independent assessments by 2 assessors (E Loh and DdP). The results were compared, and if discrepancies arose, discussions were held to clarify the score, with external input from advisors (HM, SL, and KS) obtained where appropriate.

Readability

For readability, various tools are available, including the Flesch Kincaid Reading Ease, Flesch Kincaid Grade Level, Simple Measure of Gobbledygook (SMOG), and Average Grade Level. We used SMOG as it is considered the gold standard for assessing the readability of health care material and has a high correlation with the other scoring systems [24,49]. Importantly, the outcome measure is easy to understand as, for example, an SMOG readability grade of 6 represents a text comprehensible to all individuals with sixth-grade reading skills and above [50-53]. This grade level was set as the basis of readability, given that the available literature sets this as the standard for "superior" readability. To use this tool, texts from the 20 selected websites were copied and saved as separate Microsoft Word (Microsoft Corp) and plain text documents for analysis, deleting text unrelated to the health information topic (eg, author information or disclaimers) to prevent this from confounding the scoring. A single web-based readability calculator [54] was used to generate the scores.

Data Analysis

The website category and ranking findings were compared by ANOVA, and differences between websites with and without the HONcode seal were analyzed with 2-tailed *t* tests. In addition, the relationship between credibility and readability with accuracy was assessed by Pearson correlation.

Ethical Considerations

As the research was not conducted on human subjects, no ethics review was required.

Results

Selected Websites

The top 20 websites returned by the search are listed in Multimedia Appendix 1 [55-74]. Of the 20 websites chosen from the search, 45% (n=9) were categorized as commercial, 45% (n=9) not-for-profit, and 10% (n=2) media (Table 3). There was an even distribution of commercial and not-for-profit websites across the 4 quartiles, with both media websites found in the fourth quartile. Eight of the websites bore the HONcode seal.

XSL•FO
RenderX

Table 3. Top 20 statin websites' category, HONcode[a] presence or absence, accuracy, credibility, and readability.

Website rank	Category	HONcode Seal	Accuracy	Credibility	Readability (SMOG[b] score)
1	Commercial	Yes	20	6	10.7
2	Not-for-profit	Yes	9	11	11.3
3	Not-for-profit	No	23	9	12.5
4	Commercial	Yes	16	6	13
5	Commercial	No	10	4	7.7
6	Not-for-profit	No	15	7	13
7	Commercial	No	14	10	9.6
8	Not-for-profit	No	13	5	10.8
9	Not-for-profit	No	16	11	10.2
10	Commercial	Yes	11	10	8.6
11	Commercial	Yes	16	9	9.4
12	Commercial	No	13	6	11
13	Not-for-profit	No	17	9	8.9
14	Not-for-profit	No	8	7	9
15	Not-for-profit	No	8	7	10.3
16	Media	No	8	4	7.8
17	Not-for-profit	Yes	16	11	6.7
18	Commercial	Yes	19	8	9.7
19	Commercial	Yes	21	7	9.2
20	Media	No	14	2	8.1

[a]HONcode: Health on the Net Code of Conduct.

[b]SMOG: Simple Measure of Gobbledegook.

Accuracy

The mean website accuracy score was 14.35 (SD 4.43). In terms of accuracy, the 3 highest-scoring websites were Wikipedia (score of 23), Drugs.com (score of 21), and Medicine.net (score of 20). These were the only websites to achieve a score of 20 or above. No website contradicted any checklist criterion. The top 3 performing checklist criteria were related to side effects and statin mechanism of action (criteria 9, 12, and 15), with a score of "present and complete" for each of these criteria achieved by 18, 14, and 12 websites, respectively. Although side effects were covered to some degree in all websites, criteria about drug safety (criteria 18 and 19) were complete in only 8 and 7 websites, respectively. Other poorly performing criteria reflected the lack of detail about primary prevention (criteria 11 and 6), with a score of "absent" assigned to 19 and 17 of the websites, respectively.

Credibility

None of the sampled websites met all credibility criteria for a perfect score of 13. The mean score overall was 7.45, with a range of 2-11. Importantly, 12 websites referenced peer-reviewed journal articles as a source of information, and 15 avoided anecdotal evidence for making claims. Media and some commercial websites reported personal opinions. While only 6 websites provided an organization's contact details, the others provided an email address or feedback form for contact purposes. Sponsorship was explicitly stated in 12 websites, with reference to either government or private organizations. All websites declared their organization's privacy policy, including websites with lower overall credibility scores. Only 8 websites declared author qualifications, which were primarily health related. None of the websites required paid access. Funding sources were fully disclosed in 11 websites, with the remaining 9 not reporting their source of funding or conflicts of interest. Twelve websites either had no advertisements or non–health care advertisements; the 8 websites that did not meet this criterion were commercial or media websites. Only 5 websites did not include a disclaimer that web-based health information does not replace a practitioner's advice: all of these websites were commercial or media websites.

Readability

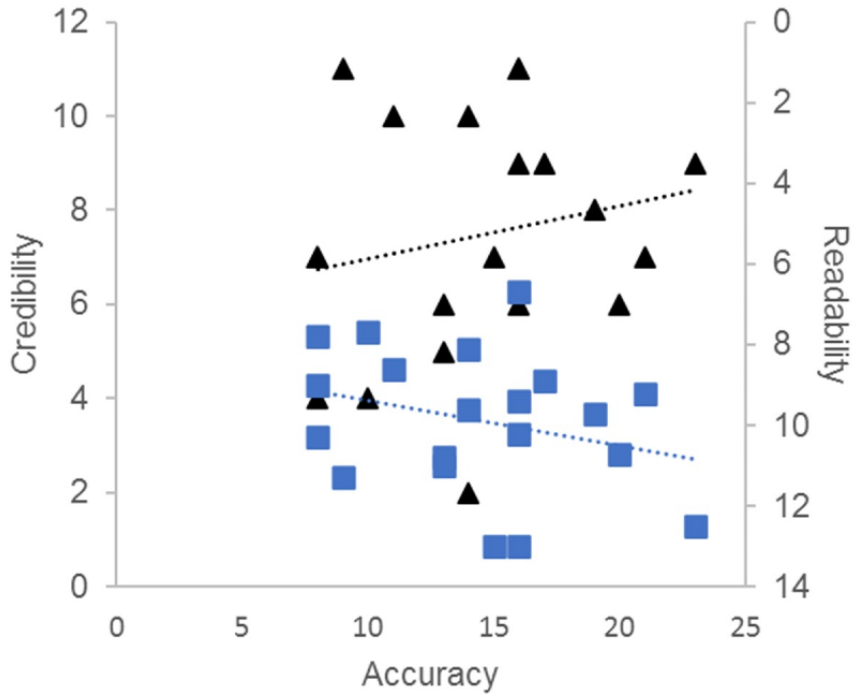
Overall, for SMOG readability, the median was 9.65 (IQR 8.825-10.85) and the average was 9.875 (SD 1.75), that is, above the ninth-grade level. None of the websites met the recommended grade level of 6; even the media websites required an eighth-grade level of comprehension.

XSL•FO
RenderX

Correlation Between Accuracy, Credibility, and Readability

No significant correlation was evident between the correlation between credibilty and accuracy (*P*=.23) (Figure 1).

Figure 1. Relationship between website credibility (black triangles) and correlation readability (blue squares) with accuracy (*P*=.23).



Website Search Rank

Websites that featured prominently in search results were not necessarily the most accurate, with no significant difference between the quartiles (*P*=.64). Indeed, of the 20 websites reviewed, the second highest scoring website for accuracy was ranked in the 19th position on search results. Similarly, there was no difference in credibility (*P*=.63) or readability (*P*=.06) between the quartiles.

Website Categorization

Comparing commercial with not-for-profit websites, 2-tailed *t* tests revealed there was no significant difference in terms of accuracy (*P*=.275), credibility (*P*=.83), or readability (*P*=.452). As there were only 2 media websites, a comparison with them was not made. Notably, they had the lowest scores for credibility, but both scored among the most readable.

Presence of HONcode Seal

Of the 20 websites, 8 were HONcode certified, with 6 of these categorized as commercial websites and 2 as not-for-profit. The mean accuracy scores for websites with and without the HONcode seal were 16 (SD 4.2) and 13.25 (SD 4.4), respectively, but this was not significantly different (*P*=.18). The presence or absence of the HONcode seal did not preclude a website from scoring at either end of the accuracy scale. Although the 8 websites with the HONcode seal scored higher in credibility (mean 8.5, SD 2) than websites that were not HONcode certified (mean 6.75, SD 2.7), this was not significantly different (*P*=.139). There was no significant difference in readability scores (*P*=.92) when comparing websites with HONcode seal status or lack thereof (mean 9.83, SD 1.9 and mean 9.91, SD 1.7, respectively).

Discussion

This study found that overall, the quality of websites with statin-related information tended to be poor. The website content was not sufficiently comprehensive, and the reading level was too difficult for the average reader to fully comprehend. The credibility of the websites varied, although overall websites bearing the HONcode seal had higher credibility than those without.

Here, we formally assessed the quality of websites addressing statins and their side effects. The finding that the quality of information is of variable caliber is consistent with studies investigating web-based health information on other topics [4,17,75]. Although the criteria used by Google's ranking algorithm is confidential, Google's guidelines state that it uses a series of algorithms that account for the words of the query, relevance and usability of web pages, the expertise of sources, ease of use on mobile device interfaces, as well as location and settings to determine the results displayed [76,77]. However, this study demonstrates that the most prominent websites in the Google search ranking are not necessarily of high quality.

The lack of correlation between accuracy and credibility or readability is a concern if patients are using the information to understand their condition and take action related to it. Patients with poor health literacy may use inaccurate and untrustworthy information in deciding whether to see a health professional following the onset of symptoms or whether to undertake tests and treatments that may be unnecessary or recommended by health professionals [12]. Vulnerable populations are at higher risk of having poor health literacy and experiencing CVD

XSL•FO
RenderX

[12,13], making them especially vulnerable to inaccurate, untrustworthy, and unreadable websites on statins.

While most websites analyzed as part of this study scored low in accuracy, this tended to be attributed to a lack of completeness of information rather than a lack of factual information. While the checklist developed here may be stringent, it would be reasonable to expect that websites dedicated to statins would be comprehensive. The lack of comprehensiveness in the information provided on the websites could result in consumers overlooking important details unless they browse through multiple websites. Furthermore, visits to multiple websites may not generate clarity but confusion. This is due to the increased likelihood of encountering inconsistent information, particularly as websites have different agendas based on the website type. That said, the commercial websites scored, on average, just as well as the not-for-profit websites, indicating that they can be a valuable source of information for consumers. It also indicates that government and other not-for-profit websites will be required to at least match the accuracy of commercial websites if they are to remain relevant in Google's search algorithm, as having information-rich content is a factor that contributes to higher search rankings [78].

When browsing the internet, one would expect government and other not-for-profit websites to provide credible information. However, some of these websites returned relatively low credibility scores and overall were not significantly more credible than commercial websites. Over half of the 20 websites analyzed provided evidence-based information and avoided anecdotal evidence, increasing their credibility rating [16,44]; however, it was concerning that 5 websites provided information based on anecdotal evidence. As expected, media websites received low credibility scores as news articles about statins contained personal views and anecdotes. Other indicators of credibility were lacking by a large proportion of websites, in particular author qualifications and details about sponsorship, funding, and conflicts of interest [16,44]. Furthermore, many of the commercial and media websites included advertisements, including health-related advertisements [16], and 5 of them did not include the disclaimer that web-based health information does not replace a practitioner's advice [44]. Thus, even patients with a degree of health literacy would find it difficult to accurately appraise the credibility of many of these websites on statins.

Given that the general public is unlikely to be fully equipped to gauge the credibility of web-based health information presented [12], clinicians could advise that patients identify the presence of the HONcode seal as this merits some confidence in the information presented [16]. However, the code does not necessarily imply that websites are comprehensive. Additionally, website developer application for the HONcode seal is voluntary, so high-quality websites may not bear the HONcode seal. The finding that the readability of the websites with the HONcode seal was not at a suitable level means that such websites may not represent digestible patient health information. Furthermore, a practical issue is that the HONcode seal is at the bottom of the web page and is thus not necessarily evident at first glance.

Many patient demographic groups have been found to read at a level more than 3 years below their completed educational years [79]. Thus, the study results may not be indicative of the severity of the problem posed by websites with high readability scores in terms of the general public's understanding of web-based information [79]. Those with limited literacy skills tend to have poorer health status due to a lack of knowledge and understanding of health care issues and a diminished ability to participate in shared decision-making in the clinical context [80]. They also tend to have poorer compliance with treatment recommendations and subsequent disease progression, as well as a higher risk for seeking emergency care and more frequent and longer inpatient admissions [27].

Additionally, other factors besides readability play into the way a text is received, including logical and sequential presentation of information. Additionally, alternative media such as images and graphs provide a well-documented "picture superiority" effect that boosts understanding of and engagement with a text [81], although some of these may also require interpretation by consumers.

A limitation of this study of these websites is that the internet is dynamic, with websites updated at any time. The search used in undertaking the study is constrained temporally in its noniterative nature, as well as its method, which used only the major search engine Google. While a metasearch capturing results from multiple search engines would provide a more comprehensive view of the information about statins on the internet, it is unlikely to represent the behavior of the public [82]. Additionally, only 8 of the websites in the study were updated in some way after completion of this study, and the information on some websites is dated as more than 10 years old.

Overall, this study has demonstrated that within the surfeit of information available on the internet regarding statin therapy, the quality of websites is of mixed caliber. The content of information is generally accurate but incomplete, while credibility is variable. Readability is generally of a level too difficult for the general public to comprehend. This suggests a need for guidance to website developers of health care websites in order to capitalize on the vast potential of the internet to equip patients with the empowerment of improved health information and health literacy. It also highlights that clinicians will need to be educated themselves about what is on the internet and what constitutes accuracy, credibility, and readability in order to impart this knowledge to their patients. During the COVID-19 pandemic, the methods through which patients seek information about their health have shifted toward increasingly internet-based means, making the quality of information on the internet of particular significance in the current climate and for the foreseeable future.

XSL•FO
RenderX

Acknowledgments

The findings were originally developed for the award of a Doctor of Medicine. Publication charges were funded by New South Wales (NSW) Health Pathology.

Authors' Contributions

E Ling attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. The project was planned by HM with input from all authors. Checklists were developed by E Loh and DdP with the assistance of SL and KS. Data were collected by E Ling, E Loh, and DdP. Data analysis was done by E Ling, E Loh, and DdP with assistance from HM. Data interpretation and manuscript review were done by all authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top 20 websites analyzed in the study.
[DOCX File , 14 KB - ijmr_v13i1e42849_app1.docx]

References

1. Cardiovascular diseases (CVDs). World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2023-12-29]
2. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet 2016;388(10059):2532-2561 [FREE Full text] [doi: 10.1016/S0140-6736(16)31357-5] [Medline: 27616593]
3. Grad R, Légaré F, Bell NR, Dickinson JA, Singh H, Moore AE, et al. Shared decision making in preventive health care: what it is; what it is not. Can Fam Physician 2017;63(9):682-684 [FREE Full text] [Medline: 28904031]
4. Mozafarpour S, Norris B, Borin J, Eisner BH. Assessment of readability, quality and popularity of online information on ureteral stents. World J Urol 2018;36(6):985-992 [FREE Full text] [doi: 10.1007/s00345-018-2179-9] [Medline: 29435639]
5. Krot K. Online information seeking and trust in doctors: an empirical study. 2017 Presented at: 24th International Scientific Conference on Economic and Social Development – "Managerial Issues in Modern Business"; October 13-14, 2017; Warsaw.
6. Kim H, Park SY, Bozeman I. Online health information search and evaluation: observations and semi-structured interviews with college students and maternal health experts. Health Info Libr J 2011;28(3):188-199 [FREE Full text] [doi: 10.1111/j.1471-1842.2011.00948.x] [Medline: 21831218]
7. Wosik J, Clowse MEB, Overton R, Adagarla B, Economou-Zavlanos N, Cavalier J, et al. Impact of the COVID-19 pandemic on patterns of outpatient cardiovascular care. Am Heart J 2021;231:1-5 [FREE Full text] [doi: 10.1016/j.ahj.2020.10.074] [Medline: 33137309]
8. Anderson J, Rainie L, Vogels EA. Experts say the 'new normal' in 2025 will be far more tech-driven, presenting more big challenges. Pew Research Center.: Elon University; 2021. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/02/PI_2021.02.18_New-Normal-2025_FINAL.pdf [accessed 2023-12-29]
9. Institute of Medicine, Board on Neuroscience and Behavioral Health, Committee on Health Literacy. In: Nielsen-Bohlman L, Panzer AM, Kindig DA, editors. Health Literacy: A Prescription to End Confusion. Washington (DC): National Academies Press; 2004.
10. Peerson A, Saunders M. Health literacy revisited: what do we mean and why does it matter? Health Promot Int 2009;24(3):285-296 [FREE Full text] [doi: 10.1093/heapro/dap014] [Medline: 19372101]
11. Mackert M, Kahlor L, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: addressing the obesity epidemic. Telemed J E Health 2009;15(7):672-677 [FREE Full text] [doi: 10.1089/tmj.2009.0012] [Medline: 19694596]
12. Keselman A, Smith CA, Murcko AC, Kaufman DR. Evaluating the quality of health information in a changing digital ecosystem. J Med Internet Res 2019;21(2):e11129 [FREE Full text] [doi: 10.2196/11129] [Medline: 30735144]
13. Havranek EP, Mujahid MS, Barr DA, Blair IV, Cohen MS, Cruz-Flores S, et al. Social determinants of risk and outcomes for cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2015;132(9):873-898 [FREE Full text] [doi: 10.1161/CIR.0000000000000228] [Medline: 26240271]
14. Demasi M, Arnott I, editors. Heart of the Matter. Australia: Australian Broadcasting Corporation; 2013.
15. Schaffer AL, Buckley NA, Dobbins TA, Banks E, Pearson SA. The crux of the matter: did the ABC's catalyst program change statin use in Australia? Med J Aust 2015;202(11):591-595 [FREE Full text] [doi: 10.5694/mja15.00103] [Medline: 26068693]

XSL•FO
RenderX

16. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700 [FREE Full text] [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
17. Ahmed OH, Sullivan SJ, Schneiders AG, McCrory PR. Concussion information online: evaluation of information quality, content and readability of concussion-related websites. Br J Sports Med 2012;46(9):675-683. [doi: 10.1136/bjsm.2010.081620] [Medline: 21504964]
18. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. J Med Internet Res 2017;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]
19. Klare GR. The measurement of readability: useful information for communicators. ACM J Comput Doc 2000;24(3):107-121 [FREE Full text] [doi: 10.1145/344599.344630]
20. Alamoudi U, Hong P. Readability and quality assessment of websites related to microtia and aural atresia. Int J Pediatr Otorhinolaryngol 2015;79(2):151-156 [FREE Full text] [doi: 10.1016/j.ijporl.2014.11.027] [Medline: 25554575]
21. Australia's Health 2018: In Brief. Canberra: Australian Institute of Health and Welfare; 2018.
22. McInnes N, Haglund BJA. Readability of online health information: implications for health literacy. Inform Health Soc Care 2011;36(4):173-189 [FREE Full text] [doi: 10.3109/17538157.2010.542529] [Medline: 21332302]
23. Lipari M, Berlie H, Saleh Y, Hang P, Moser L. Understandability, actionability, and readability of online patient education materials about diabetes mellitus. Am J Health Syst Pharm 2019;76(3):182-186 [FREE Full text] [doi: 10.1093/ajhp/zxy021] [Medline: 31408087]
24. El-Haddad N, Spooner C, Faruqi N, Denney-Wilson E, Harris M. Readability and content analysis of lifestyle education resources for weight management in Australian general practice. BMC Obes 2016;3:16 [FREE Full text] [doi: 10.1186/s40608-016-0097-1] [Medline: 26966546]
25. Gulati R, Nawaz M, Lam L, Pyrsopoulos NT. Comparative readability analysis of online patient education resources on inflammatory bowel diseases. Can J Gastroenterol Hepatol 2017;2017:3681989 [FREE Full text] [doi: 10.1155/2017/3681989] [Medline: 28740843]
26. Grewal P, Williams B, Alagaratnam S, Neffendorf J, Soobrah R. Quality of vascular surgery web sites on the internet. J Vasc Surg 2012;56(5):1461-1467 [FREE Full text] [doi: 10.1016/j.jvs.2012.04.058] [Medline: 22801107]
27. Badarudeen S, Sabharwal S. Assessing readability of patient education materials: current role in orthopaedics. Clin Orthop Relat Res 2010;468(10):2572-2580 [FREE Full text] [doi: 10.1007/s11999-010-1380-y] [Medline: 20496023]
28. Browser market share Australia. Statcounter. 2021. URL: https://gs.statcounter.com/browser-market-share/all/australia [accessed 2021-09-04]
29. Martin J. Best web browsers. Tech Advisor. 2021. URL: https://www.techadvisor.com/article/728377/best-web-browsers-2021.html [accessed 2021-09-04]
30. Baker J. Brafton 2017 content marketing benchmark report. Brafton. 2017. URL: https://www.brafton.com/blog/strategy/brafton-2017-content-marketing-benchmark-report/ [accessed 2021-09-04]
31. Chitika insights: the value of Google result positioning. Chitika. 2013. URL: https://research.chitika.com/wp-content/uploads/2022/02/chitikainsights-valueofgoogleresultspositioning.pdf [accessed 2023-12-29]
32. Purcell K, Brenner J, Rainie L. Search engine use 2012. Pew Research Centre. 2012. URL: https://www.ris.org/uploadi/editor/1341041853PIP_Search_Engine_Use_2012.pdf [accessed 2023-12-29]
33. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004;7(5):497-510. [doi: 10.1089/cpb.2004.7.497] [Medline: 15667044]
34. Google organic CTR history. Advanced Web Ranking. 2022. URL: https://www.advancedwebranking.com/ctrstudy/ [accessed 2023-12-29]
35. Wilson PWF. Overview of established risk factors for cardiovascular disease. UpToDate. 2021. URL: https://www.uptodate.com/contents/overview-of-established-risk-factors-for-cardiovascular-disease#topicContent [accessed 2021-09-04]
36. Davidson P, Chew D. Reducing Risk in Heart Disease: An Expert Guide to Clinical Practice for Secondary Prevention of Coronary Heart Disease. Australia: National Heart Foundation of Australia; 2012:18.
37. Stone NJ, Robinson JG, Lichtenstein AH, Merz CNB, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. J Am Coll Cardiol 2014;63(25 Pt B):2889-2934 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.002] [Medline: 24239923]
38. Krumholz HM. Treatment of cholesterol in 2017. JAMA 2017;318(5):417-418 [FREE Full text] [doi: 10.1001/jama.2017.6753] [Medline: 28738130]
39. Charnock D, Shepperd S. Discern Online. United Kingdom: Radcliffe Online; 2021. URL: http://www.discern.org.uk/background_to_discern.php [accessed 2021-09-03]
40. Beaunoyer E, Arsenault M, Lomanowska AM, Guitton MJ. Understanding online health information: evaluation, tools, and strategies. Patient Educ Couns 2017;100(2):183-189 [FREE Full text] [doi: 10.1016/j.pec.2016.08.028] [Medline: 27595436]
41. Peabody D. Two components in bipolar scales: direction and extremeness. Psychol Rev 1962;69:65-73. [doi: 10.1037/h0039737] [Medline: 14484721]

XSL•FO
RenderX

42. Komorita SS, Graham WK. Number of scale points and the reliability of scales. Educ Psychol Meas 2016;25(4):987-995 [FREE Full text] [doi: 10.1177/001316446502500404]
43. Dolnicar S, Grün B, Leisch F, Rossiter J. Three good reasons NOT to use five and seven point Likert items. 2011 Presented at: 21st CAUTHE National Conference; February 8-11, 2011; Adelaide, Australia URL: https://ro.uow.edu.au/commpapers/775/
44. Weymann N, Härter M, Dirmaier J. Quality of online information on type 2 diabetes: a cross-sectional study. Health Promot Int 2015;30(4):821-831 [FREE Full text] [doi: 10.1093/heapro/dau019] [Medline: 24688114]
45. Freeman KS, Spyridakis JH. An examination of factors that affect the credibility of online health information. Tech Commun 2004;51(2):239-263 [FREE Full text]
46. Sillence E, Briggs P, Harris P, Fishwick L. A framework for understanding trust factors in web-based health advice. Int J Hum Comput Stud 2006;64(8):697-713 [FREE Full text] [doi: 10.1016/j.ijhcs.2006.02.007]
47. Neumark Y, Flum L, Lopez-Quintero C, Shtarkshall R. Quality of online health information about oral contraceptives from Hebrew-language websites. Isr J Health Policy Res 2012;1(1):38 [FREE Full text] [doi: 10.1186/2045-4015-1-38] [Medline: 23006798]
48. Metzger MJ. Making sense of credibility on the web: models for evaluating online information and recommendations for future research. J Am Soc Inf Sci Technol 2007;58(13):2078-2091. [doi: 10.1002/asi.20672]
49. Fitzsimmons PR, Michael BD, Hulley JL, Scott GO. A readability assessment of online Parkinson's disease information. J R Coll Physicians Edinb 2010;40(4):292-296. [doi: 10.4997/JRCPE.2010.401] [Medline: 21132132]
50. McLaughlin GH. SMOG grading-a new readability formula. J Read 1969;12(8):639-646 [FREE Full text]
51. Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. Res Social Adm Pharm 2013;9(5):503-516 [FREE Full text] [doi: 10.1016/j.sapharm.2012.05.009] [Medline: 22835706]
52. Oliffe M, Thompson E, Johnston J, Freeman D, Bagga H, Wong PKK. Assessing the readability and patient comprehension of rheumatology medicine information sheets: a cross-sectional health literacy study. BMJ Open 2019;9(2):e024582 [FREE Full text] [doi: 10.1136/bmjopen-2018-024582] [Medline: 30813117]
53. Cheng C, Dunn M. Health literacy and the internet: a study on the readability of Australian online health information. Aust N Z J Public Health 2015;39(4):309-314 [FREE Full text] [doi: 10.1111/1753-6405.12341] [Medline: 25716142]
54. My Byline media. Readability Formulas. URL: https://readabilityformulas.com/ [accessed 2021-09-04]
55. MedicineNet. URL: https://www.medicinenet.com/statins/drug-class.htm#what_are_statins_and_how_do_they_work [accessed 2024-02-07]
56. Statins. Healthdirect. URL: https://www.healthdirect.gov.au/statins [accessed 2024-02-07]
57. Statin. Wikipedia. URL: https://en.wikipedia.org/wiki/Statin [accessed 2024-02-07]
58. Bruce DF. Statin side effects. WebMD. URL: https://www.webmd.com/cholesterol-management/side-effects-of-statin-drugs#1 [accessed 2024-02-07]
59. Mann D, Pallarito K. Are statins bad for you? 4 things to know. Health. URL: https://www.health.com/health/gallery/0,,20552165,00.html [accessed 2024-02-07]
60. Statins. MedlinePlus. URL: https://medlineplus.gov/statins.html [accessed 2024-02-07]
61. Statins. HealthEngine. URL: https://healthengine.com.au/info/statins [accessed 2024-02-07]
62. Statins. NHS. URL: https://www.nhs.uk/conditions/statins/ [accessed 2024-02-07]
63. Statin side effects: weigh the benefits and risks. Mayo Clinic. URL: https://www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/in-depth/statin-side-effects/art-20046013 [accessed 2024-02-07]
64. Side effects of atorvastatin. Healthline. URL: https://www.healthline.com/health/high-cholesterol/statins-list-of-common-types [accessed 2024-02-07]
65. The uses and risks of statins. Medical News Today. URL: https://www.medicalnewstoday.com/articles/8274.php [accessed 2024-02-07]
66. Mercola J. You've been genetically hijacked to produce Pfizer's Pfrankenstein proteins. Mercola. 2024. URL: https://articles.mercola.com/sites/articles/archive/2010/07/20/the-truth-about-statin-drugs-revealed.aspx [accessed 2024-02-07]
67. Atorvastatin. NHS. URL: https://www.nhs.uk/medicines/atorvastatin/ [accessed 2024-02-07]
68. Statin adverse effects. Statin Effects Study. URL: https://statineffects.com/info/adverse_effects.htm [accessed 2024-02-07]
69. Getting treatment: the most widely used medicine to lower cholesterol is a statin, but there are other medicines available too and some may only be prescribed in a specialist lipid clinic. Heart UK. URL: https://www.heartuk.org.uk/getting-treatment/treatment-intro [accessed 2024-02-07]
70. Gajraj H. Why I've ditched statins for good. The Telegraph. 2014. URL: https://www.telegraph.co.uk/news/health/10717431/Why-Ive-ditched-statins-for-good.html [accessed 2024-02-07]
71. Statin medicines: frequently asked questions. NPS MedicineWise. URL: https://www.nps.org.au/consumers/statin-medicines-frequently-asked-questions [accessed 2024-02-07]
72. Zocor. RxList. URL: https://www.rxlist.com/zocor-side-effects-drug-center.htm [accessed 2024-02-07]
73. Simvastatin. Drugs.com. 2023. URL: https://www.drugs.com/simvastatin.html [accessed 2024-02-07]

XSL•FO
RenderX

74. Nedelman M. Statins: what happens to patients who stop taking them after side effects? CNN Health. 2017. URL: https://edition.cnn.com/2017/07/24/health/statins-side-effects-outcomes-study/index.html [accessed 2024-02-07]
75. Bernard S, Cooke T, Cole T, Hachani L, Bernard J. Quality and readability of online information about type 2 diabetes and nutrition. JAAPA 2018;31(11):41-44 [FREE Full text] [doi: 10.1097/01.JAA.0000546481.02560.4e] [Medline: 30358679]
76. How search algorithms work mountain view. Google. URL: https://www.google.com/search/howsearchworks/algorithms/ [accessed 2023-12-29]
77. Makino T, Jung C, Phan D. Finding more mobile-friendly search results. Google Search Central Blog. 2015. URL: https://developers.google.com/search/blog/2015/02/finding-more-mobile-friendly-search] [accessed 2023-12-29]
78. Webmaster guidelines mountain view. Google. 2021. URL: https://developers.google.com/search/docs/essentials [accessed 2023-12-29]
79. Weiss BD. Progress in health literacy-encouraging, and not so encouraging. Health Lit Res Pract 2017;1(4):e231-e232 [FREE Full text] [doi: 10.3928/24748307-20171016-01] [Medline: 31294267]
80. Weiss BD, Schwartzberg JG. Health Literacy and Patient Safety: Help Patients Understand: Manual for Clinicians, Volume 1. Tucson: American Medical Association; 2007.
81. Kasabwala K, Agarwal N, Hansberry DR, Baredes S, Eloy JA. Readability assessment of patient education materials from the American Academy of Otolaryngology—Head and Neck Surgery Foundation. Otolaryngol Head Neck Surg 2012;147(3):466-471 [FREE Full text] [doi: 10.1177/0194599812442783] [Medline: 22473833]
82. Breakey VR, Harris L, Davis O, Agarwal A, Ouellette C, Akinnawo E, et al. The quality of information about sickle cell disease on the internet for youth. Pediatr Blood Cancer 2017;64(4):e26309 [FREE Full text] [doi: 10.1002/pbc.26309] [Medline: 27786409]

Abbreviations

CVD: cardiovascular disease
HONcode: Health on the Net Code of Conduct
SMOG: Simple Measure of Gobbledygook

Edited by G Eysenbach; submitted 21.09.22; peer-reviewed by L Daraz, WF Khaw; comments to author 24.12.22; revised version received 13.04.23; accepted 17.11.23; published 14.03.24.

Please cite as:
Ling E, de Pieri D, Loh E, Scott KM, Li SCH, Medbury HJ
Evaluation of the Accuracy, Credibility, and Readability of Statin-Related Websites: Cross-Sectional Study
Interact J Med Res 2024;13:e42849
URL: https://www.i-jmr.org/2024/1/e42849
doi:10.2196/42849
PMID:38483461

©Eunice Ling, Domenico de Pieri, Evenne Loh, Karen M Scott, Stephen C H Li, Heather J Medbury. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 14.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Visibility and Scientific Relevance of Strabismus Research: Bibliometric Analysis

Aleksander Stupnicki[1]; Basil Suresh[1], iBSc (Hons), MBBS; Saurabh Jain[1,2], MB BS

[1]University College London Medical School, London, United Kingdom

[2]Royal Free London NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Aleksander Stupnicki
University College London Medical School
4 Huntley Street
London, WC1E 6DE
United Kingdom
Phone: 44 02031088235
Email: zchatup@ucl.ac.uk

Abstract

Background: Quality and accuracy of online scientific data are crucial, given that the internet and social media serve nowadays as primary sources of medical knowledge.

Objective: This study aims to analyze the relationship between scientific relevance and online visibility of strabismus research to answer the following questions: (1) Are the most popular strabismus papers scientifically relevant? (2) Are the most high-impact strabismus studies shared enough online?

Methods: The Altmetric Attention Score (AAS) was used as a proxy for online visibility, whereas citations and the journal's impact factor (IF) served as a metric for scientific relevance. Using "strabismus" as a keyword, 100 papers with the highest AAS and 100 papers with the highest number of citations were identified. Statistical analyses, including the Spearman rank test, linear regression, and factor analysis, were performed to assess the relationship between AAS, citations, a journal's IF, and mentions across 18 individual Web 2.0 platforms.

Results: A weak, positive, statistically significant correlation was observed between normalized AAS and normalized citations ($P<.001$; $r=0.27$) for papers with high visibility. Only Twitter mentions and Mendeley readers correlated significantly with normalized citations ($P=.02$ and $P<.001$, respectively) and IF ($P=.04$ and $P=.009$, respectively), with Twitter being the strongest significant predictor of citation numbers ($r=0.53$). For high-impact papers, no correlation was found between normalized citations and normalized AAS ($P=.12$) or the IF of the journal ($P=.55$).

Conclusions: While clinical relevance influences online attention, most high-impact research related to strabismus is not sufficiently shared on the web. Therefore, researchers should make a greater effort to share high-impact papers related to strabismus on online media platforms to improve accessibility and quality of evidence-based knowledge for patients.

(Interact J Med Res 2024;13:e50698) doi:10.2196/50698

KEYWORDS

strabismus research; squint; social media; scientific relevance; altmetrics; accuracy; medical knowledge; metric; bibliometric analysis; research; strabismus; online visibility; platform; evidence-based information; accessibility

Introduction

Patients, health care professionals, and researchers increasingly use social media and online platforms as a source of knowledge, health care news, and scientific research [1]. Despite the worldwide prevalence of strabismus remaining stable at around 2% [2], the public's online interest in the topic has been rising, a trend reflected by the increasing popularity of queries related to the disease over the past 2 decades according to Google Trends. Due to this increasing reliance on online platforms, it is essential to ensure the quality and relevance of scientific data that are commonly accessed on the web, especially for lay members of the public who may lack the skills or time to assess that themselves.

To quantify the relevance of research within the field of medical science, the number of citations and impact factor (IF) of the

XSL•FO
RenderX

journal are used most frequently [3]. The dissemination of the same academic information through platforms used by the general public, on the other hand, can be most reliably quantified by the Altmetric Attention Score (AAS), a real-time weighted measure of mentions across all Web 2.0 social media platforms [4,5].

Bibliometric analyses using altmetrics and other scientometrics have been conducted previously in the field of ophthalmic research to evaluate publication trends [6], disruptiveness of papers [7], or research productivity [8]. To date, however, there have been no such analyses within the subspecialty of strabismus, despite the pervasiveness of the disease.

Therefore, we decided to analyze the relationship between the scientific relevance of strabismus research and its contributions to the online sphere, in order to answer the following questions: (1) Are the most popular strabismus papers scientifically relevant? (2) Are the most high-impact strabismus studies shared online enough?

Methods

In line with the best practice of literature searching [9], a thesaurus synonym search was performed to identify appropriate keywords for database search. As of January 2023, the thesaurus does not identify synonyms for "strabismus," and the *Cambridge Dictionary* confirms it is the only medical term to describe the condition [10]. No morphological variation of the term has been identified, eliminating the need for the use of truncation in keyword searches.

Therefore, a list of research papers including the keyword "strabismus" was generated on January 27, 2023, with Altmetric Explorer with no other restrictions (search period: January 2011 to January 2023). The keyword search engine in Altmetric Explorer yields comprehensive results including outputs that match the keyword across publication title, author name, or journal title [11]. Hence, the pooled list was then filtered by a consultant ophthalmologist according to relevance to include 100 papers with the highest AAS (a total of 255 titles and abstracts were analyzed to compile 100 relevant publications). Additional preliminary searches using lay synonyms of strabismus, including "squint" and "cross-eye," were performed but yielded no relevant or sufficiently high AAS results for inclusion, proving the keyword "strabismus" captures the bona fide core of publications in the field.

On the same day (January 27, 2023), for each of the papers, Web of Science (WoS) was used to add information on the number of citations, time since publication, and IF of the journal at the time of publication; other metrics traditionally used to assess the quality and relevance of scientific research [12]. Additional data on the source of AAS, including mentions across (1) news, (2) blogs, (3) Twitter, (4) peer review, (5) Facebook, (6) Wikipedia, (7) LinkedIn, (8) Weibo, (9) Google+, (10) Reddit, (11) Pinterest, (12) F100, (13) Q&A, (14) policy, (15) patent, (16) video, (17) syllabi, and (18) Mendeley, were pooled from Altmetric website and evaluated to characterize the field.

For systematic comparison, the same approach to searching was implemented to yield a list of papers with the highest number of citations: on January 27, 2023, WoS was used to generate a list of 100 papers including the keyword "strabismus" with the highest number of citations, excluding papers published before 2011, and the year Altmetric Explorer was founded and started tracking the AAS (search period: January 2011 to January 2023). No other filters were applied to the search. On the same day, the AAS for each of the papers was manually pooled from Altmetric website. Data on time since publication and the journal's IF at the time of publication were extracted from WoS. To account for temporal differences [13], the values for AAS and citations for both groups have been then normalized per year since publication.

Kolmogorov-Smirnov test was used to verify that the distribution of the data does not follow a normal distribution, and Spearman rank correlation coefficient was used to test for correlation between all variables. Following correlational calculations, linear regression analysis and factor analysis were performed to explain patterns among correlated variables, both of which are statistical techniques commonly used in altmetric research [14,15]. SPSS (IBM Corp) was used for all statistical calculations. Statistical significance was defined as $P<.05$.

Ethical Considerations

No ethics board approval was required, as the study did not involve any human participants.

Results

Correlation Analysis

The normalized AASs of the 100 papers with the highest online visibility (median AAS 11, IQR 6-16) correlated significantly with normalized citations ($P<.001$) but demonstrated a weak strength of the relationship ($r=0.27$) for papers with AAS <150. To achieve this result, we excluded 3 outlier papers with significantly higher AASs (922, 413, and 169, respectively, compared to median 11, IQR 6-16; z score>3), which would otherwise skew the statistical analysis. Spearman rank test demonstrated no correlation between the normalized AAS and the IF of the journal ($P=.15$) or time ($P=.37$).

For the 100 papers with the highest number of citations (median 30, IQR 15-45), no statistically significant correlation was found between normalized citations and the normalized AAS ($P=.12$) or IF of the journal ($P=.55$), but as expected, they correlated significantly with time ($P=.01$).

Upon analysis of AAS sources, we found a weak, positive, statistically significant correlation between normalized citations and Twitter mentions ($P=.02$; $r=0.27$), normalized citations and Mendeley readers ($P<.001$; $r=0.40$), and normalized citations and policy mentions ($P=.02$; $r=0.24$) for the 100 papers with highest AAS. The same variables showed a weak, positive, statistically significant correlation with the IF of the journal at the time of publication: Twitter and IF ($P=.04$; $r=0.25$), Mendeley readers and IF ($P=.009$; $r=0.32$), and policy mentions and IF ($P=.04$; $r=0.26$).

Correlations between the number of mentions in the news, on blogs, in peer-reviews, on Facebook, Wikipedia, LinkedIn, Reddit, Google+, Weibo, Pinterest, syllabi, or video and

XSL•FO
RenderX

normalized citations or IF were not significant at $P<.05$ (Multimedia Appendix 1).

Multivariate Analysis

To better understand variance among the correlated variables, a linear regression model was run with normalized citations as the dependent variable and Twitter mentions, Mendeley readers, and policy mentions as covariates. ANOVA test showed significant variance within the sample, confirming the suitability of the test ($P<.001$). We obtained an R^2 value of 0.31, indicating that 31% of the variance within citations can be explained cumulatively by the 3 AAS sources. Only Twitter and policy mentions, however, were significant predictors ($P<.001$ and $P=.004$ respectively), with Twitter mentions being the most important predictor as indicated by the highest standardized coefficient ($r=0.53$).

Factor Analysis

Factor analysis was performed on metrics with adequate data for the papers with the top 100 AAS normalized citations, normalized AAS, IF, time since publication (months), Twitter mentions, Mendeley readers, and policy documents. Bartlett test of sphericity indicated an approximate chi-square value of $\chi^2_{93}=124.4$ ($P<.001$) and a Kaiser-Meyer-Olkin adequacy value of 0.607, together indicating the suitability of the data set for factor analysis. Three factors were identified across these variables: factor 1 between Twitter, Mendeley, AAS, and citations; factor 2 between Mendeley, policy, and citations; and factor 3 between IF and time.

Discussion

Principal Findings

The significant, yet, weak correlation of AAS with citations for papers with the highest online visibility shows that the clinical relevance of strabismus-related publications (as measured by citations) can contribute to increased online popularity but is not the sole determining factor. Furthermore, the lack of correlation between AAS and IF demonstrates that the relative importance of a journal in the field (and consequently the paper) does not determine its online popularity, which raises questions about the quality of strabismus research receiving the most online attention.

Through correlational and multivariate analysis of mentions across individual Web 2.0 platforms, we have demonstrated Twitter mentions to be a significant and strong predictor of citations for the most popular strabismus papers. Although prior studies demonstrated a more significant, causative impact of tweets on citation numbers [16], the relationship can differ between fields [17] and seems to be statistically significant for strabismus research, albeit moderate compared with other research domains. Overall, our findings imply that dissemination of strabismus research through Twitter can have an impact on scholarly visibility and subsequently citation rates.

Furthermore, we demonstrated a lack of statistically significant correlations between traditional scientometrics (citations and IF) and mentions across other social or media platforms, which has been also observed in other fields of research [17]. This reveals existing gaps that require more references to research papers on strabismus, including social media platforms like Facebook or LinkedIn, as well as critical, knowledge-oriented pages such as Wikipedia.

Furthermore, in terms of factor analysis, factor 1 linking AAS, citations, Mendeley, and Twitter likely suggests that for strabismus research, there is a degree of overlap regarding the user bases or networks between Mendeley and Twitter, despite the former being considered a platform largely used by academic professionals as opposed to the more widely public microblogging service [18]. It may also indicate that highly cited papers are receiving engagement and being discussed in both academic and general public networks indicating that such papers may have a wider social impact. Factor 2 linking Mendeley, policy documents, and citations may suggest that for highly cited papers, there is increased interest and readership on Mendeley—a proportion of which may be faculty and departmental figures. This in turn may lead to policy mentions for impactful research papers. Therefore, this suggests that papers with high citation counts and academic impact may be influencing policies and organizational standards [19]. Factor 3 linking IF and time could be due to the overall increase and growth of the cited strabismus literature over time; however, this is less relevant to our research question.

For the 100 papers with the highest number of citations, the lack of correlation between citations or IF and the AAS suggests that clinical relevance or perceived prestige related to the publishing journal does not affect the online visibility of strabismus papers. Researchers publishing in the strabismus realm should, therefore, make a greater effort to share their high-impact papers on social media. In turn, this could increase the visibility and accessibility of their research, especially for the lay public who rarely browses journals for medical knowledge, enhance collaboration, and further enhance the overall impact of their research.

Strengths and Limitations

AAS itself is a useful tool for authors to get quick, up-to-date insight into the performance of their papers on the web. It is crucial, however, to bear in mind the inherent limitations of AAS due to the fact that it is ultimately only a metric of “mentions” or “posts” and is not an indicator of research quality or legitimacy [20]. In isolation, it may be deemed unreliable, as “viral” papers that do not exhibit robust research methodology or present sensible conclusions may still acquire a high AAS. Furthermore, the AAS does not account for following or website traffic, therefore providing no information on the actual number of viewers. As a result, a frequently mentioned paper can effectively have low visibility and reception, despite a high AAS. Although Twitter mentions and Mendeley readership seem to have some impact on citation numbers of the most popular strabismus papers, they only account for a small proportion of the variance within citation data (31%). Thus, using alternative metrics, like tweets, as predictors of scientific contribution and success does not constitute a comprehensive and precise appraisal method, as demonstrated before across several other fields [16,17].

XSL•FO
RenderX

Furthermore, citations take a longer time to accumulate, whereas AAS is updated almost in real time, so even when a new paper is published that can have high publicity in social media and among the scientific community, the citation numbers will lag months or sometimes years behind. Due to this phenomenon, there may be weaker or no association between citations and AAS for the latest research papers, which could skew the results of our 12-year view. Further statistical testing would be necessary to confirm that.

Additionally, in the case of strabismus research, both the AAS and citation numbers are characteristically low, which raises questions about the reliability of the data set, as it is analogical to having a small sample size. Various scientifically irrelevant factors can cause a high AAS, such as the topic of the paper, sensationalism, how easily it is understood by the general public, or the number of intersections of the topic with other branches of medicine. A good example of the effect of those confounding factors is a publication included in our data set entitled "Evidence That Leonardo da Vinci Had Strabismus," which had the highest AAS of 922 (over 83 times the median score) but only 6 citations. This demonstrates that especially for papers with high AAS scores, the virality of the topic can have a higher impact on the AAS than its scientific significance.

Conclusions and Future Directions

We have demonstrated that the clinical relevance of strabismus research contributes to the amount of online attention it receives. However, the most high-impact strabismus research is not sufficiently shared across online platforms. Therefore, we recommend that researchers make a greater effort to share high-impact studies on social media platforms to improve the quality of evidence-based information about strabismus and improve the accessibility of this knowledge. To maximize the societal impact of research, it is important to interact with both academic and general audiences, as shown by the overlap between Mendeley and Twitter engagement of strabismus publications.

Furthermore, we revealed Twitter mentions to be the strongest predictor of citation numbers for strabismus papers, highlighting the potential impact of social media on scholarly visibility. Our findings also highlight the need for engagement of strabismus researchers across a broader range of platforms, including Facebook, LinkedIn, or Wikipedia. However, due to its inherent biases and limitations, the AAS itself or mentions across specific platforms should only complement traditional metrics, such as IF and citations, to provide a broader picture of the publicity of the paper but should not act as a stand-alone metrics for assessing the quality and relevance of strabismus papers.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary table of Spearman rank test for normalized citations, impact factor, and all 18 Web 2.0 platforms analyzed.
[XLSX File (Microsoft Excel File), 11 KB - ijmr_v13i1e50698_app1.xlsx]

References

1. Surani Z, Hirani R, Elias A, Quisenberry L, Varon J, Surani S, et al. Social media usage among health care providers. BMC Res Notes 2017;10(1):654 [FREE Full text] [doi: 10.1186/s13104-017-2993-y] [Medline: 29187244]
2. Hashemi H, Pakzad R, Heydarian S, Yekta A, Aghamirsalim M, Shokrollahzadeh F, et al. Global and regional prevalence of strabismus: a comprehensive systematic review and meta-analysis. Strabismus 2019;27(2):54-65. [doi: 10.1080/09273972.2019.1604773] [Medline: 31012389]
3. Thompson DF, Walker CK. A descriptive and historical review of bibliometrics with applications to medical sciences. Pharmacotherapy 2015;35(6):551-559. [doi: 10.1002/phar.1586] [Medline: 25940769]
4. García-Villar C. A critical review on altmetrics: can we measure the social impact factor? Insights Imaging 2021;12(1):92 [FREE Full text] [doi: 10.1186/s13244-021-01033-2] [Medline: 34215935]
5. Priem J, Hemminger BH. Scientometrics 2.0: new metrics of scholarly impact on the social Web. First Monday 2010;15(7) [FREE Full text] [doi: 10.5210/fm.v15i7.2874]
6. Boudry C, Baudouin C, Mouriaux F. International publication trends in dry eye disease research: a bibliometric analysis. Ocul Surf 2018;16(1):173-179. [doi: 10.1016/j.jtos.2017.10.002] [Medline: 29031646]
7. Patel PA, Patel PN, Becerra AZ, Mehta MC. Bibliometric analysis of the 100 most-disruptive articles in ophthalmology. Clin Exp Ophthalmol 2022;50(6):690-695. [doi: 10.1111/ceo.14109] [Medline: 35593696]
8. Huang A, Kim S, Zhu H, Pathare N, Ooi XY, Kirby RP, et al. A bibliometric analysis of ophthalmology resident research productivity in the United States. medRxiv Preprint posted online on February 08, 2022. [doi: 10.1101/2022.02.06.22270335]
9. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc 2018;106(4):531-541 [FREE Full text] [doi: 10.5195/jmla.2018.283] [Medline: 30271302]
10. Strabismus. Cambridge Dictionary. 2024. URL: https://dictionary.cambridge.org/dictionary/english/strabismus [accessed 2024-03-27]

XSL•FO
RenderX

11. Advanced search. Altmetric. URL: https://help.altmetric.com/support/solutions/articles/6000240857-advanced-search [accessed 2024-05-09]
12. Myers BA, Kahn KL. Practical publication metrics for academics. Clin Transl Sci 2021;14(5):1705-1712 [FREE Full text] [doi: 10.1111/cts.13067] [Medline: 33982433]
13. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: how great is the impact? Scientometrics 2015;105(3):1809-1831 [FREE Full text] [doi: 10.1007/s11192-015-1645-z]
14. Thelwall M, Wilson P. Regression for citation data: an evaluation of different methods. J Informetr 2014;8(4):963-971 [FREE Full text] [doi: 10.1016/j.joi.2014.09.011]
15. Taherdoost H, Sahibuddin S, Jalaliyoon N. Exploratory factor analysis; concepts and theory. SSRN J 2022 [FREE Full text]
16. Eysenbach G. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. J Med Internet Res 2011;13(4):e123 [FREE Full text] [doi: 10.2196/jmir.2012] [Medline: 22173204]
17. Thelwall M, Nevill T. Could scientists use Altmetric.com scores to predict longer term citation counts? J Informetr 2018;12(1):237-248 [FREE Full text] [doi: 10.1016/j.joi.2018.01.008]
18. Haustein S, Larivière V, Thelwall M, Amyot D, Peters I. Tweets vs. Mendeley readers: how do these two social media metrics differ? Inf Technol 2014;56(5):207-215 [FREE Full text] [doi: 10.1515/itit-2014-1048]
19. Yu H, Murat B, Li J, Li L. How can policy document mentions to scholarly papers be interpreted? An analysis of the underlying mentioning process. Scientometrics 2023;128:6247-6266 [FREE Full text] [doi: 10.1007/s11192-023-04826-y]
20. Elmore SA. The Altmetric attention score: what does it mean and why should i care? Toxicol Pathol 2018;46(3):252-255 [FREE Full text] [doi: 10.1177/0192623318758294] [Medline: 29448902]

Abbreviations

AAS: Altmetric Attention Score
IF: impact factor
WoS: Web of Science

Edited by T de Azevedo Cardoso; submitted 10.07.23; peer-reviewed by A Santisteban, Y Jiang; comments to author 01.02.24; revised version received 27.03.24; accepted 29.04.24; published 12.06.24.

Please cite as:
Stupnicki A, Suresh B, Jain S
Online Visibility and Scientific Relevance of Strabismus Research: Bibliometric Analysis
Interact J Med Res 2024;13:e50698
URL: https://www.i-jmr.org/2024/1/e50698
doi:10.2196/50698
PMID:38865170

©Aleksander Stupnicki, Basil Suresh, Saurabh Jain. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 12.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effectiveness of a Smartphone App to Promote Physical Activity Among Persons With Type 2 Diabetes: Randomized Controlled Trial

Stephanie E Bonn[1], PhD; Madeleine Hummel[1], MD; Giulia Peveri[2,3], PhD; Helén Eke[1], MSc; Christina Alexandrou[4], PhD; Rino Bellocco[2,5], PhD; Marie Löf[4], PhD; Ylva Trolle Lagerros[1,6], MD, PhD

[1]Clinical Epidemiology Division, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[3]Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy
[4]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden
[5]Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[6]Center for Obesity, Academic Specialist Center, Stockholm Health Services, Stockholm, Sweden

Corresponding Author:
Stephanie E Bonn, PhD
Clinical Epidemiology Division, Department of Medicine Solna
Karolinska Institutet
Maria Aspmans gata 30A
Stockholm, 171 64
Sweden
Phone: 46 8 517 791
Email: stephanie.bonn@ki.se

Abstract

Background: Physical activity is well known to have beneficial effects on glycemic control and to reduce risk factors for cardiovascular disease in persons with type 2 diabetes. Yet, successful implementation of lifestyle interventions targeting physical activity in primary care has shown to be difficult. Smartphone apps may provide useful tools to support physical activity. The DiaCert app was specifically designed for integration into primary care and is an automated mobile health (mHealth) solution promoting daily walking.

Objective: This study aimed to investigate the effect of a 3-month-long intervention promoting physical activity through the use of the DiaCert app among persons with type 2 diabetes in Sweden. Our primary objective was to assess the effect on moderate to vigorous physical activity (MVPA) at 3 months of follow-up. Our secondary objective was to assess the effect on MVPA at 6 months of follow-up and on BMI, waist circumference, hemoglobin A_{1c}, blood lipids, and blood pressure at 3 and 6 months of follow-up.

Methods: We recruited men and women with type 2 diabetes from 5 primary health care centers and 1 specialized center. Participants were randomized 1:1 to the intervention or control group. The intervention group was administered standard care and access to the DiaCert app at baseline and 3 months onward. The control group received standard care only. Outcomes of objectively measured physical activity using accelerometers, BMI, waist circumference, biomarkers, and blood pressure were assessed at baseline and follow-ups. Linear mixed models were used to assess differences in outcomes between the groups.

Results: A total of 181 study participants, 65.7% (119/181) men and 34.3% (62/181) women, were recruited into the study and randomized to the intervention (n=93) or control group (n=88). The participants' mean age and BMI were 60.0 (SD 11.4) years and 30.4 (SD 5.3) kg/m^2, respectively. We found no significant effect of the intervention (group by time interaction) on MVPA at either the 3-month (β=1.51, 95% CI –5.53 to 8.55) or the 6-month (β=–3.53, 95% CI –10.97 to 3.92) follow-up. We found no effect on any of the secondary outcomes at follow-ups, except for a significant effect on BMI at 6 months (β=0.52, 95% CI 0.20 to 0.84). However, mean BMI did not differ between the groups at the 6-month follow-up.

Conclusions: We found no evidence that persons with type 2 diabetes being randomized to use an app promoting daily walking increased their levels of MVPA at 3 or 6 months' follow-up compared with controls receiving standard care. The effect of the

XSL•FO
RenderX

app on BMI was unclear, and we found nothing to support an effect on secondary outcomes. Further research is needed to determine what type of mHealth intervention could be effective to increase physical activity among persons with type 2 diabetes.

Trial Registration: ClinicalTrials.gov NCT03053336; https://clinicaltrials.gov/study/NCT03053336

(Interact J Med Res 2024;13:e53054) doi:10.2196/53054

KEYWORDS

behavior change; exercise; intervention; mHealth; smartphone app; self-monitoring

Introduction

Despite increased knowledge and public health initiatives, more than 460 million people, corresponding to over 6% of the world's population, are estimated to be diagnosed with type 2 diabetes today [1]. Persons with type 2 diabetes followed in primary care can be prescribed lifestyle interventions in combination with medications. Lifestyle interventions may include weight management, smoking cessation, stress reduction, and improved dietary habits or physical activity [2]. Physical activity is well known to have beneficial effects on glycemic control and to reduce risk factors for cardiovascular disease [3]. Yet, it has proven difficult to implement lifestyle interventions targeting physical activity in primary care [4]. Nevertheless, walking has been put forward as a useful therapeutic tool shown to improve glucose control, with clinically beneficial effects on blood glucose levels over time, and the potential to improve other clinical variables such as BMI and blood pressure in persons with type 2 diabetes [5].

During the past few decades, various telemonitoring, eHealth, and mobile health (mHealth) solutions targeting physical activity have been developed. Such solutions offer adaptable platforms for the delivery of self-management interventions that are easily accessible to both patients and health care practitioners, and users can engage with health information technology at their convenience. Smartphone apps may be useful in a health care setting to provide an additional tool to increase patients' engagement through the use of self-monitoring of, for example, physical activity between routine visits [6].

Today, there are many commercially developed and available smartphone apps targeting self-management of chronic conditions. Common features of available apps targeted toward persons with type 2 diabetes include self-tracking of blood glucose levels and components targeting physical activity or diet in different ways [7,8]. Nevertheless, there is a wide variety in type and number of features for diabetes management in available apps [7], making it difficult for patients to select the most appropriate one to use. There are many commercial apps targeting lifestyle among persons with type 2 diabetes; however, few solutions primarily target physical activity within this group, and even fewer have been developed specifically for implementation in primary care. Therefore, we developed a digital platform and a smartphone app specifically for targeting physical activity among persons with type 2 diabetes treated within primary care [9]. The app was built to be integrated into the existing digital infrastructure of primary care in Sweden, with the aim to provide care givers and patients with a scientifically evaluated self-care management tool.

Results from studies evaluating the effectiveness of mHealth solutions, including smartphone apps, targeting persons with type 2 diabetes, have been summarized previously [10-15]. Significant reductions in hemoglobin A_{1c} (HbA_{1c}) levels are generally shown after 3 months of follow-up. Most of the evaluated apps allowed the user to monitor their blood glucose levels and included physical activity or diet, either alone or in combination, as additional features. mHealth interventions targeting physical activity in adults in the general population have been shown to increase both minutes of physical activity and steps per day [16-18]. Nevertheless, apps primarily targeting physical activity, without including a component of glucose monitoring, in persons with type 2 diabetes are relatively uncommon. Poppe et al [19] evaluated a self-regulation–based eHealth and mHealth intervention primarily targeting physical activity in persons with type 2 diabetes and found positive results of the intervention on increased activity and decreased sedentary behavior, whereas Thorsen et al [20] found no effect of app-based interval walking on MVPA over 52 weeks compared with standard care. In summary, there is still a need to develop interventions targeting physical activity that are effective and can be implemented into primary care.

The aim of this study was to evaluate the effects of the DiaCert smartphone app promoting daily walking on moderate to vigorous physical activity (MVPA) and clinical variables in persons with type 2 diabetes. Our primary aim was to test the hypothesis that the app would lead to an increase in minutes of MVPA at 3 months compared with standard care only. Our secondary aim was to assess the effect of the app on MVPA after 6 months and on the clinical variables BMI, waist circumference, HbA_{1c}, cholesterol (total, low-density lipoprotein [LDL], and high-density lipoprotein [HDL]), triglycerides, and systolic and diastolic blood pressure at 3 and 6 months. We hypothesized that the app would lead to improvements in both MVPA and clinical variables.

Methods

Trial Design

We conducted a randomized controlled trial with 2 parallel groups between February 2017 and June 2019. The DiaCert study design [21] has been described in detail previously. Study participants were randomized 1:1 to the intervention or control group at baseline. The primary study outcome was MVPA (minutes/day) at 3 months of follow-up. Secondary outcomes included MVPA at 6 months of follow-up and the clinical variables BMI, waist circumference, HbA_{1c}, cholesterol (total, LDL, and HDL), triglycerides, and systolic and diastolic blood pressure at 3 and 6 months of follow-up. No changes to methods

XSL•FO
RenderX

were done after trial commencement. The study is reported according to the CONSORT (Consolidated Standards of Reporting Trials) statement [22] and the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth) checklist, which is developed for eHealth or mHealth interventions [23].

Ethical Considerations

The trial was approved by the ethics committee of the regional ethical review board in Stockholm, Sweden (Dnr: 2016/2041-31/2, 2016/99-32, 2017/1406-32) and registered at ClinicalTrials.gov (NCT03053336). All participants received both oral and written information about the study and provided their written informed consent to participate. Participants received no compensation for participation in the study. After data collection, data were anonymized.

Study Participants

The inclusion criteria were (1) having a diagnosis of type 2 diabetes, (2) being 18 years of age or older, (3) being able to communicate in Swedish, and (4) having access to and being able to use a smartphone. The exclusion criterion was not being able to walk.

Patients were recruited continuously from 5 primary health care centers and 1 specialized medical center in the Stockholm area, Sweden. Patients at the participating centers received initial information about the study from their physician or diabetes nurse, and those interested in participating were contacted by study personnel and given more detailed information. Thereafter, patients either agreed to participate and were scheduled for a baseline introductory meeting, or declined participation. We did not record the number of patients who did not agree to participate. All study participants met with study personnel at baseline and after 3 and 6 months. On each occasion, study outcomes including physical activity and other lifestyle factors, as well as clinical variables, were assessed.

Interventions

Study participants randomized to the intervention group continued to receive standard care as usual but also downloaded the DiaCert app during the baseline meeting. To access the individual user account in the app, a personal 6-digit code was entered. The code was given to participants by study personnel 7 days after the baseline meeting to avoid overlap with baseline accelerometer measurements. The intervention group was encouraged to use the app daily for 3 months. At the 3-month follow-up meeting, participants deleted the app from their phones together with study personnel. Participants then received no intervention during 3 months and were offered access to the app again at the 6-month follow-up.

The DiaCert app displayed daily steps that through connection to a digital platform were shared with study personnel. An individual step goal between 1000 and 10,000 steps [24] was set at baseline based on the participant's usual activity level. Participants in the intervention group were contacted by study personnel every second week by phone. During these follow-ups, the participant could revise his or her step goal with an even 500 steps increase or decrease. The maximum goal set at any time point was 10,000 steps. Users received automatic positive feedback messages including the user's name in the app on days when the goal was met. In addition to daily steps, information on HbA_{1c} taken during the study period was also displayed in the app.

The app design has previously been described in detail, including a presentation of the app screen by screen [9]. In brief, features of the app included a home screen displaying the daily steps in relation to an individual step goal during the past week. A circle was gradually filled as the user walked toward the step goal. It was completely filled and marked with a checkmark when the goal was reached. Through the home screen, the user was also able to access information on previous daily steps, questionnaires, and results of HbA_{1c}. The app was continuously updated to run with the current iOS and Android versions, but no changes were made to the content during the study. Users were asked to contact study personnel if they experienced malfunction of the app. The app was developed within the research project and is no longer available.

Study participants randomized to the control group received standard care, that is, continued their usual care as prescribed by their regular primary care physician and diabetes nurse, also after inclusion to the study. For ethical reasons, they were offered access to the DiaCert app at the 6-month follow-up.

Outcomes

Physical Activity

To assess the primary outcome of MVPA in this trial, physical activity was measured using the ActiGraph wGT3x-BT triaxial accelerometer (dynamic range: 8g) during 7 consecutive days. At each study meeting, the participants were asked to wear the accelerometer on their nondominant wrist day and night, starting at 4 PM the same day until 8 AM just over a week later. Participants wore the accelerometers on their wrist to increase feasibility and maximize compliance. Data were sampled at a frequency of 80 Hz.

We downloaded the collected data from each accelerometer using the manufacturer's program (ActiLife Software, version 6.13.3; ActiGraph), and thereafter, the raw data were extracted for data processing. As suggested by Migueles et al [25], processing of accelerometer data was performed using the open-source R package GGIR. GGIR version 2.0-0, R version 3.6.1, and RStudio version 1.2.5019 were used. Data collected before the first and after the last midnight were excluded in order to examine 7 complete days. As the first step of analysis, data were averaged over 5-second epochs and aggregated through application of Euclidean norm minus 1, with negative values rounded up to 0. Autocalibration was performed using local gravity to adjust for calibration errors and unreliable signals.

The default cut point for MVPA (100 milligravity) and default settings for the definition and management of nonwear time (ie, 4×15 minutes) in GGIR were applied [26]. Nonwear time was by default replaced with imputed averaged activity from the same time the other measured days. A valid day was defined as at least 16 hours of wear time and a valid week required 4 valid

XSL•FO
RenderX

days (including at least 1 weekend day) [25]. Variables were weighted 5:2 with data collected on weekdays and weekend days. For MVPA, bouts of consistent activity lasting for at least 1 minute were used, where 80% of the included epochs had to be above or equal to the cut point [27].

In addition to accelerometer measurements, we also assessed daily physical activity at baseline with 2 validated general questions used in routine health care [28]. Participants were asked to (1) report their usual time spent exercising during a week and (2) add up and report the total time estimated spent doing other types of leisure time physical activities of lower intensity in bouts of at least 10 minutes during a week. Walking, cycling, or gardening was presented to the respondent to exemplify activity level.

Clinical Variables

A detailed description of measurements has been published elsewhere [21]. In brief, HbA_{1c} (mmol/mol), total cholesterol (mmol/L), LDL cholesterol (mmol/L), HDL cholesterol (mmol/L), and triglycerides (mmol/L) were measured in fasting blood samples. HbA_{1c} was measured using the reference measurement procedure by the International Federation for Clinical Chemistry and Laboratory Medicine [29]. The enzymatic method [30] was used to measure total cholesterol and HDL cholesterol. LDL cholesterol was calculated using the Friedewald equation [31].

Height (to the nearest 0.5 cm), body weight (to the nearest 0.1 kg), and waist circumference (to the nearest 1.0 cm) were measured by study personnel who also performed 1 manual assessment of blood pressure (systolic and diastolic) after the participant had been sitting down for at least 5 minutes. BMI was calculated based on measured height and weight (kg/m^2).

Sample Size

Power calculations were performed a priori to determine the sample size need to detect a clinically significant difference of 8 minutes/day of MVPA [21]. A total of 250 participants (125 in each group) were estimated to provide 80% power at a 5% significance level. This included an expected dropout rate of 20%. Baseline data collection ended in June 2018 before reaching 250 participants because the DiaCert app was no longer compatible with the upgrades of iOS and Android.

Randomization and Blinding

A random allocation sequence list was generated by the first author (SEB) in Stata (version 14.0; StataCorp). Women and men were randomized separately in blocks of 10 within each participating primary health care center and the specialized medical center. Patients who agreed to participate in the study were continuously allocated to the next available spot on the list by study personnel (authors SEB and CA). Participants were informed about their group allocation at the end of the baseline meeting. Because of the nature of the intervention, neither participants nor study personnel were blinded to participants' group allocation.

Statistical Methods

Baseline characteristics of study participants are presented as mean (SD) or number (%) for continuous and categorical variables, respectively. Variables were checked for normality and outliers. The Student *t* test or the chi-square test was used to assess potential differences in baseline characteristics between the intervention and control group.

We used linear mixed models with fixed and random intercept and slope for the time variables to assess if there were longitudinal differences in MVPA at 3 months of follow-up (primary outcome) and secondary outcomes including MVPA at 6 months of follow-up and BMI, waist circumference, HbA_{1c}, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, and systolic and diastolic blood pressure at 3 and 6 months of follow-up between the intervention and control group. We have therefore included, in addition to time and group terms, a group×time interaction term to assess if any differences in the study outcomes were constant at 3 and 6 months. For outcomes that differed significantly between the study groups at baseline, that is, MVPA, additional sensitivity analysis adjusting the models for baseline levels of the outcome was performed according to the methods described by Twisk et al [32]. The analysis of intervention effect was made following the intention-to-treat approach [33]. Missing data were associated with a primary care center, with 17/25 (68%) participants with missing accelerometer data at baseline belonging to primary care center 1. Missing data were assumed to be missing at random and not depending on the study group as the degree of missing data was similar in both the intervention and control group during the intervention. Participants with complete data at baseline for each specific outcome variable were included in the analysis of intervention effect at 3 and 6 months.

Post hoc sensitivity analyses using self-reported data on physical activity from the baseline questionnaire were carried out to deal with the unblinded nature of our study and the potential bias that may have been present during baseline accelerometer measurements. Although participants were not connected to the DiaCert app until after the completion of baseline accelerometer measurements, they were aware of which group they were randomized to during measurements. This could potentially have affected baseline levels of MVPA, which might not have represented the "true" levels of MVPA before the start of the study. Multiple imputation based on MICE (multiple imputation by chained equations) was used to address this issue [34]. We first set all baseline values of MVPA for the intervention group to missing. To predict MVPA at baseline for participants in the intervention group in the hypothetical scenario in which participants were blind to the intervention, we implemented MICE using all relevant baseline variables. The variables included in the model were age, sex, height, weight, education, household income, marital status, smoking and snuff habits, the year of diabetes diagnosis, the center of recruitment, self-reported levels of physical activity, and the number of valid accelerometer days at baseline. We compared the imputed and the observed values of MVPA at baseline using a 2-tailed *t* test. Linear mixed models, as described above, were thereafter fitted

XSL•FO
RenderX

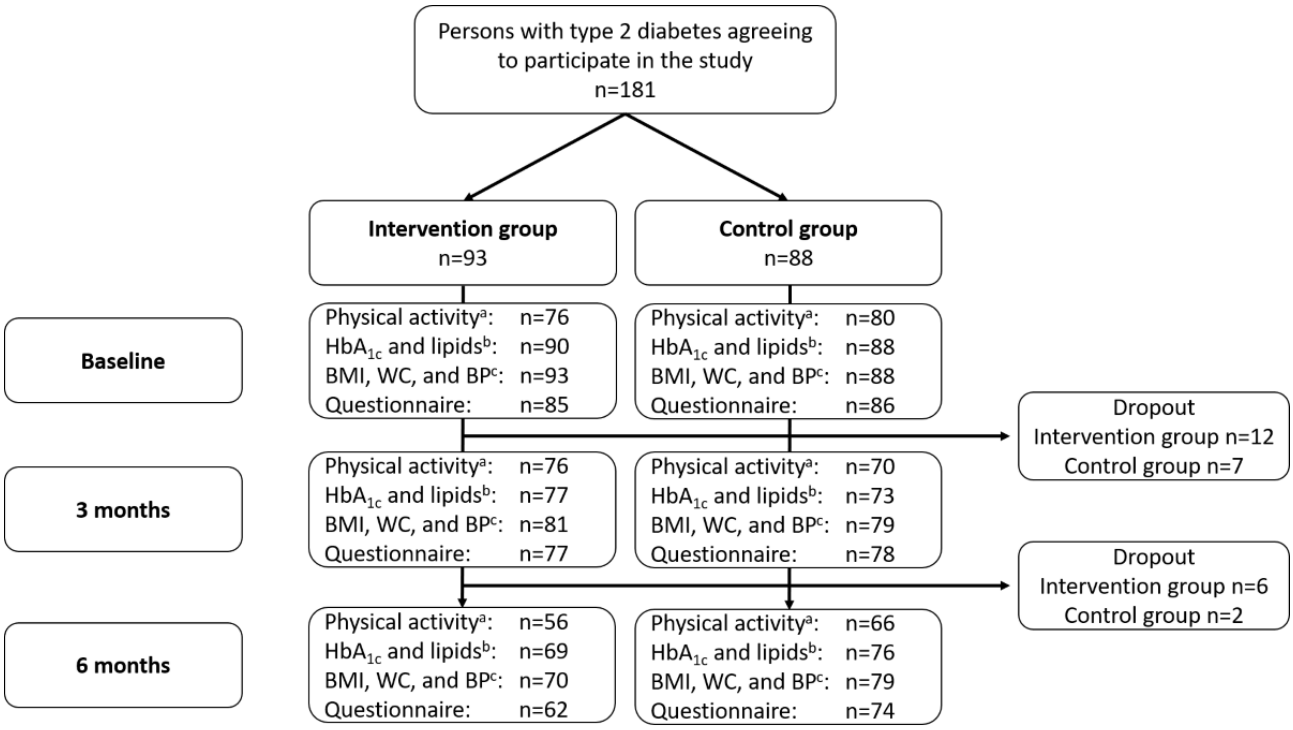
to contrast differences in MVPA between the intervention and control group using the imputed data for MVPA at baseline.

Statistical tests were 2-sided, and the significance level was set to $P<.05$. Statistical analyses were performed using Stata (version 17; StataCorp).

Results

A total of 181 persons with type 2 diabetes were included in the trial, of whom 93 were randomized to the intervention group and 88 were randomized to the control group. At 3 months, the dropout rate was 10.5% (19/181), and at 6 months it was 14.9% (27/181). Dropout was higher among participants in the intervention group (n=12 vs n=7 at 3 months, and n=6 vs n=2 at 6 months) than among those in the control group. In total, 156 participants had valid accelerometer data on physical activity at baseline. Of these, 137 (87.8%) also had valid data at the 3-month follow-up (primary outcome). At baseline, most participants (75.6%, 118/156) had valid accelerometer data from all 7 days, 14.7% (23/156) had valid data from 6 days, and 9.6% (15/156) from 5 or 4 days. A flowchart of participants with complete data from the different assessments is presented in Figure 1.

Figure 1. Flowchart of participation and data completeness at baseline and follow-up after 3 and 6 months in the DiaCert study. [a]Number of participants with valid accelerometer data, that is, at least 16 hours of wear time and a total of 4 valid days including at least 1 weekend day; [b]Triglycerides and cholesterol (total, low-density lipoprotein, and high-density lipoprotein). BP: blood pressure; HbA_{1c}: hemoglobin A_{1c}; WC: waist circumference.



Characteristics of all 181 study participants are presented in Table 1. The majority of participants were men (119/181, 65.8%), and the mean (SD) age at baseline among all participants was 60.0 (11.4) years. The mean (SD) BMI among all participants was 30.4 (5.3) kg/m^2. Overweight and obesity was common, 85.6% (155/181) had a BMI of ≥25 kg/m^2, and 48.1% (87/181) had a BMI of ≥30 kg/m^2. Mean (SD) HbA_{1c} was 53.6 (12.8) mmol/mol. There were no statistically significant differences between the study groups regarding age distribution, primary care center belonging, sex, smoking, time since diabetes diagnosis, or clinical variables including HbA_{1c} and lipid levels. In study participants with complete accelerometer data at baseline (n=156), there was a statistically significant difference in accelerometer-measured baseline levels of physical activity between the intervention and control group, with higher MVPA (38.3 vs 29.8 minutes/day, $P=.04$) measured in the intervention group. Using the imputed data for the intervention group, the baseline MVPA was estimated to be 32.3 minutes/day, which did not differ from the measured level in the control group ($P=.62$). Additionally, self-reported levels of physical activity at baseline, that is, time spent exercising and total leisure time activity, did not differ between the groups ($P=.20$ and $P=.20$, respectively). Baseline characteristics of participants with complete accelerometer data at baseline can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Characteristics of study participants by study group (N=181).

Characteristic	Intervention group (n=93)		Control group (n=88)		*P* value[a]
	n (%)	Mean (SD)	n (%)	Mean (SD)	
MVPA[b,c] (minutes/day)	76 (82)	38.3 (28.3)	80 (91)	29.7 (24.1)	*.04*
BMI (kg/m^2)	93 (100)	30.2 (5.5)	88 (100)	30.6 (5.2)	.61
Waist circumference					
Women	33 (35)	102 (12.6)	28 (32)	103 (15.8)	.75
Men	59 (63)	111 (15.6)	60 (68)	110 (13.3)	.75
Hemoglobin A_{1c} (mmol/mol)	89 (96)	53.6 (13.0)	88 (100)	53.5 (12.7)	.97
Total cholesterol (mmol/L)	74 (80)	4.56 (1.00)	70 (80)	4.50 (1.18)	.77
LDL[d] cholesterol (mmol/L)	73 (78)	2.70 (0.89)	70 (80)	2.64 (1.13)	.72
HDL[e] cholesterol (mmol/L)	74 (80)	1.25 (0.34)	70 (80)	1.23 (0.40)	.73
Triglycerides (mmol/L)	72 (77)	1.54 (0.76)	69 (78)	1.64 (0.94)	.45
Blood pressure (mm Hg)					
Systolic	92 (99)	138.7 (16.2)	88 (100)	137.0 (14.8)	.45
Diastolic	92 (99)	83.6 (9.9)	88 (100)	82.6 (9.5)	.50
Sex					.50
Male	34 (37)	—[f]	28 (32)	—	
Female	59 (63)	—	60 (68)	—	
Age (years)					.24
<50	20 (22)	—	16 (18)	—	
50-59	28 (30)	—	19 (22)	—	
60-69	31 (33)	—	30 (34)	—	
≥70	14 (15)	—	23 (26)	—	
Leisure time activity[g] (minutes/week)					.27
<60	6 (7)		11 (13)	—	
60-90	8 (10)	—	8 (10)	—	
90-150	17 (20)	—	15 (18)	—	
150-300	14 (17)	—	22 (26)	—	
>300	39 (46)	—	28 (33)	—	
Primary care centers					.88
1	35 (38)	—	31 (35)	—	
2	14 (15)	—	13 (15)	—	
3	6 (6)	—	8 (9)	—	
4	23 (25)	—	23 (26)	—	
5	10 (11)	—	11 (13)	—	
Specialized medical center	5 (5)	—	2 (2)	—	
Time spent exercising[h] (minutes/week)					.26
Never	35 (42)	—	38 (44)	—	
<30	12 (14)	—	12 (14)	—	
30-90	13 (15)	—	21 (24)	—	

XSL•FO
RenderX

Characteristic	Intervention group (n=93)		Control group (n=88)		*P* value[a]
	n (%)	Mean (SD)	n (%)	Mean (SD)	
>90	24 (29)	—	15 (17)	—	
Smoking[i]					.41
Yes	11 (13)	—	9 (11)	—	
No, ever smoker	31 (37)	—	40 (47)	—	
No, never smoker	42 (50)	—	36 (42)	—	
Time since diabetes diagnosis[j] (years)					.50
<1	8 (11)	—	13 (18)	—	
1-5	20 (29)	—	17 (24)	—	
>5	42 (60)	—	42 (58)	—	
Education[k] (years)					.69
≤12	45 (54)	—	44 (51)	—	
>12	38 (46)	—	42 (49)	—	

[a]2-tailed *t* test was used for continuous variables and the chi-square test was used for categorical variables. Italicized *P* values represent statistical significance.

[b]MVPA: moderate to vigorous physical activity.

[c]Missing data from n=17 (intervention) and n=8 (control).

[d]LDL: low-density lipoprotein.

[e]HDL: high-density lipoprotein.

[f]—not available.

[g]From questionnaire, missing data n=9 (intervention) and n=4 (control).

[h]From questionnaire, missing data n=9 (intervention) and n=2 (control).

[i]Missing data n=9 (intervention) and n=3 (control).

[j]Missing data n=23 (intervention) and n=16 (control).

[k]Missing data n=10 (intervention) and n=2 (control).

Effectiveness of the Intervention—MVPA

Results from between-group analysis and the intervention effect on minutes/day of MVPA are shown in Table 2. The mean change in minutes/day of MVPA from baseline to follow-ups is graphically shown in Figure 2. The statistically significant difference in minutes/day of MVPA seen between the groups at baseline, with participants in the intervention group being more active than participants in the control group, remained at the 3-month follow-up. The predicted mean difference between the groups after 3 months was 10.05 minutes (95% CI 1.66-18.44). At the 6-month follow-up, there was no statistically significant difference in MVPA between the groups (β=5.02, 95% CI –3.72 to 13.75).

XSL•FO
RenderX

Table 2. The intervention effect on the primary outcome of daily minutes of moderate to vigorous physical activity (MVPA) at 3 months of follow-up and on secondary outcomes including MVPA at 6 months of follow-up and clinical variables in the DiaCert study.

Characteristic	Group sample means				Model estimates[a]	
	Intervention (n=93)		Control (n=88)		Difference[b]	Group by time interaction
	n (%)	Mean (SD)	n (%)	Mean (SD)	Mean (95% CI)	β (95% CI)
MVPA (minutes/day)						
3 months	70 (75)	36.6 (25.5)	67 (76)	26.7 (21.1)	10.05 (1.66 to 18.44)	1.51 (–5.53 to 8.55)
6 months	55 (59)	34.2 (29.4)	63 (72)	31.1 (27.0)	5.02 (–3.72 to 13.75)	–3.53 (–10.97 to 3.92)
BMI (kg/m^2)						
Baseline	93 (100)	30.2 (5.5)	88 (100)	30.6 (5.2)	N/A[c]	N/A
3 months	81 (87)	30.1 (5.7)	79 (90)	30.3 (5.0)	–0.12 (–1.67 to 1.44)	0.29 (–0.02 to 0.61)
6 months	70 (75)	30.0 (6.0)	79 (90)	29.9 (5.0)	0.11 (–1.45 to 1.67)	0.52 (0.20 to 0.84)
Waist circumference (cm)						
Baseline	92 (99)	107 (15.1)	88 (100)	108 (14.4)	N/A	N/A
3 months	80 (86)	107 (15.6)	79 (90)	108 (13.6)	–0.61 (–4.90 to 3.69)	–0.46 (–1.75 to 0.83)
6 months	69 (74)	107 (17.0)	79 (90)	106 (13.6)	0.47 (–3.84 to 4.77)	0.61 (–0.71 to 1.94)
Hemoglobin A_{1c} (mmol/mol)						
Baseline	89 (96)	53.6 (13.0)	88 (100)	53.5 (12.7)	N/A	N/A
3 months	75 (81)	50.0 (9.9)	73 (83)	53.2 (13.4)	–2.45 (–6.08 to 1.17)	–2.54 (–5.36 to 0.29)
6 months	67 (72)	51.2 (10.7)	76 (86)	51.2 (10.6)	–0.21 (–3.87 to 3.44)	–0.30 (–3.16 to 2.57)
Total cholesterol (mmol/L)						
Baseline	74 (80)	4.56 (1.00)	70 (80)	4.50 (1.18)	N/A	N/A
3 months	62 (67)	4.39 (0.80)	57 (65)	4.48 (0.91)	–0.06 (–0.40 to 0.28)	–0.11 (–0.37 to 0.14)
6 months	55 (59)	4.38 (0.84)	59 (67)	4.27 (1.02)	0.10 (–0.24 to 0.44)	0.05 (–0.21 to 0.31)
LDL[d] cholesterol (mmol/L)						
Baseline	73 (78)	2.70 (0.89)	70 (80)	2.64 (1.13)	N/A	N/A
3 months	59 (63)	2.44 (0.65)	56 (64)	2.48 (0.89)	–0.01 (–0.33 to 0.30)	–0.07 (–0.30 to 0.15)
6 months	54 (58)	2.44 (0.68)	57 (65)	2.38 (0.85)	0.09 (–0.23 to 0.40)	0.02 (–0.20 to 0.25)
HDL[e] cholesterol (mmol/L)						
Baseline	74 (80)	1.25 (0.34)	70 (80)	1.23 (0.40)	N/A	N/A
3 months	62 (67)	1.22 (0.34)	57 (65)	1.24 (0.38)	–0.04 (–0.16 to 0.09)	–0.06 (–0.12 to 0.01)
6 months	55 (59)	1.26 (0.36)	57 (65)	1.24 (0.39)	–0.04 (–0.17 to 0.08)	–0.06 (–0.13 to 0.001)
Triglycerides (mmol/L)						
Baseline	72 (77)	1.54 (0.76)	69 (78)	1.65 (0.94)	N/A	N/A
3 months	60 (65)	1.71 (1.03)	58 (66)	1.76 (0.96)	–0.01 (–0.32 to 0.29)	0.10 (–0.15 to 0.34)
6 months	52 (56)	1.56 (0.83)	56 (64)	1.66 (0.83)	–0.07 (–0.38 to 0.24)	0.04 (–0.21 to 0.29)
Systolic BP[f] (mm Hg)						
Baseline	92 (99)	139 (16.2)	88 (100)	137 (14.8)	N/A	N/A
3 months	80 (86)	135 (13.3)	79 (90)	134 (11.8)	1.21 (–2.98 to 5.40)	–0.54 (–4.62 to 3.53)
6 months	70 (75)	136 (12.8)	78 (89)	136 (12.6)	0.70 (–3.60 to 4.99)	–1.05 (–5.23 to 3.13)
Diastolic BP (mm Hg)						
Baseline	92 (99)	83.6 (9.9)	88 (100)	82.6 (9. 5)	N/A	N/A
3 months	80 (86)	81.8 (9.1)	79 (90)	79.2 (10.5)	2.78 (–0.14 to 5.69)	1.79 (–1.06 to 4.64)

XSL•FO
RenderX

Characteristic	Group sample means				Model estimates[a]	
	Intervention (n=93)		Control (n=88)		Difference[b]	Group by time interaction
	n (%)	Mean (SD)	n (%)	Mean (SD)	Mean (95% CI)	β (95% CI)
6 months	70 (75)	81.2 (8.7)	78 (89)	79.9 (9.4)	1.87 (–1.12 to 4.85)	0.88 (–2.04 to 3.81)

[a]Results from linear mixed model analysis including participants with complete baseline data.

[b]Difference between groups at the specified time point based on predicted means from linear mixed models.

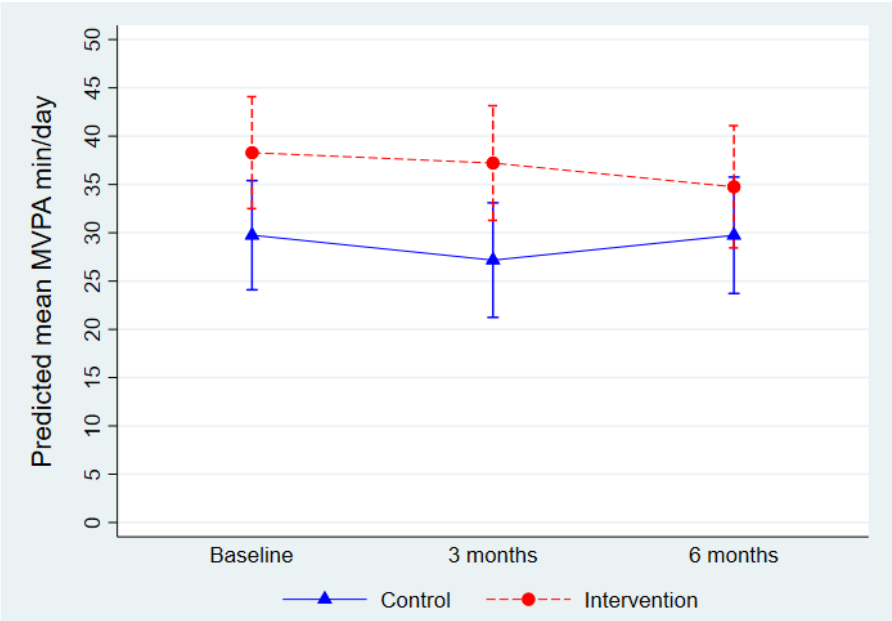
[c]—: not applicable.

[d]LDL: low-density lipoprotein.

[e]HDL: high-density lipoprotein.

[f]BP: blood pressure.

Figure 2. Changes over time in moderate to vigorous physical activity (MVPA, minutes/day) in the intervention and control group. Predicted group means from linear mixed model analysis. Results correspond to model estimates shown in Table 2.



We found no statistically significant effect of the intervention (group by time interaction) on MVPA at either the 3- or the 6-month follow-up (Table 2). When adjusting for baseline levels of MVPA, results remained nonsignificant at both 3 (β=4.38, 95% CI –2.11 to 10.88) and 6 (β=–0.65, 95% CI –7.58 to 6.29) months. Additionally, results from sensitivity analyses using imputed baseline data also remained nonsignificant at follow-up after both 3 (β=6.86, 95% CI –4.05 to 17.78) and 6 (β=1.44, 95% CI –9.87 to 11.76) months.

Effectiveness of Intervention—Clinical Variables

Detailed results from between-group analyses and the intervention effect on clinical variables included as secondary outcomes are shown in Table 2. The mean change in outcomes from baseline to follow-ups is graphically shown in Multimedia Appendix 2. We found no statistically significant differences in any of the secondary outcomes at the 3- or 6-month follow-ups, except for in the analysis of BMI where a statistically significant effect of the intervention was seen at 6 months (group by time interaction: 0.52, 95% CI 0.20-0.84). However, there was no difference in mean BMI between the groups at the 6-month follow-up (predicted difference in mean: 0.11, 95% CI –1.45 to 1.67). Participants in the control group had a slightly higher BMI at baseline than participants in the intervention group (30.6 vs 30.2 kg/m^2), although this difference was not statistically significant (*P*=.61).

Discussion

Principal Findings

In this 2-armed randomized controlled trial, we found no clear effect of the DiaCert app promoting daily walking in persons with type 2 diabetes. We found no increase in objectively measured MVPA after neither 3 (primary outcome) nor 6 months of follow-up compared with standard care when accounting for baseline levels of MVPA. This is in line with results from Thorsen et al [20]. They were not able to show an effect of app-based interval walking on MVPA over 52 weeks compared with standard care among persons with type 2 diabetes. On the contrary, Poppe et al [19] found that an eHealth and mHealth intervention that primarily targeted physical activity in persons with type 2 diabetes led to increased physical activity and decreased sedentary behavior. Participants in our study and the study by Thorsen et al [20] reported more time in MVPA at baseline compared with those in the study by Poppe et al [19]. This might have contributed to the difference in effect between

XSL•FO
RenderX

the studies, as a higher baseline level of MVPA may imply less room for improvement.

Other previous studies have also shown an effect of mHealth solutions to increase physical activity. Hochsmann et al [35] showed that an interactive smartphone game aimed at increasing daily steps in persons with type 2 diabetes had a significant effect on activity compared with a control group receiving standard lifestyle counseling. Glynn et al [36] evaluated the effect of an app that aimed to increase physical activity in primary care patients and found that the intervention group increased their number of daily steps compared with the control group. In the randomized Sophia step study, comprising persons with prediabetes or type 2 diabetes in Sweden [37], self-monitoring of daily physical activity using pedometers with registration of steps on the web in combination with counseling did not increase levels of physical activity but seemed to prevent the decrease in physical activity seen in the control group. Although our primary hypothesis was rejected and we could not show that the use of the DiaCert app led to increased MVPA among persons with type 2 diabetes, results from other studies still indicate that mHealth solutions can have positive effects on physical activity.

Although we found a statistically significant effect of the DiaCert intervention on BMI at the 6-month follow-up, there was no difference in mean BMI between the groups at any time point. Therefore, we cannot draw any conclusion regarding the effect on BMI. Not all studies evaluating interventions targeting physical activity include anthropometric outcomes, and the effect on BMI in previous studies are mixed, showing no effect or indicating a small reduction favoring the intervention group [10,12]. Thorsen et al [20] did not report BMI as an outcome, but found a nonsignificant reduction of waist circumference. Similar to the Sophia step study, our results provided no evidence for an effect on waist circumference [37].

Reviews and meta-analyses [11,13,15] provide strong evidence for a positive effect of mobile apps for lifestyle modification in persons with type 2 diabetes on HbA_{1c} levels. However, all of the evaluated apps included monitoring of blood glucose. The effects on other clinical markers including body weight, blood lipids, and blood pressure are less clear [38]. Our results are in line with the Sophia step study, where no effects on cardiometabolic variables, including HbA_{1c}, triglycerides, HDL cholesterol, and LDL cholesterol, were seen [37]. Neither our study nor the Sophia step study included a glucose-monitoring component, which may explain the lack of an effect on HbA_{1c}. Nevertheless, results from a meta-analysis by Lee et al [12] indicated that mHealth interventions in persons with type 2 diabetes >65 years of age may improve blood lipid profiles, which we found no evidence of. One explanation for the lack of an effect on cardiometabolic markers is that a changed lifestyle behavior for the better could have led to lowered medication, leaving HbA_{1c}, serum lipid levels, or blood pressure unchanged. We did not assess changes in medications during the intervention, which is a limitation of our study.

The use of different behavior change techniques may also explain differences in efficacy between apps. Fanning et al [39] investigated the effect of a basic self-monitoring app (tracking, feedback, information) and 2 theory-based tools (goal setting and point-based feedback) on MVPA among healthy but inactive adults. Four different groups received either a basic self-monitoring app only or the basic app together with (1) goal setting, or (2) feedback, or (3) goal setting and feedback. All groups increased their MVPA, but the feedback group showed the highest increase. While the DiaCert app comprised several behavioral change techniques that previously have been included in successful interventions, such as monitoring, goal setting, and positive feedback, it did not include other features associated with effective results, for example, frequent reminders or the option to share data with peers [40]. How the included components are designed may also affect results; for example, goal setting can be personalized or generic, and a goal can be more or less challenging.

Strengths and Limitations

The design of this randomized controlled study is one of the strengths of our study. The fact that study participants were recruited from 6 different care centers located in different areas with diverse populations and levels of socioeconomic status is another strength. This likely increased generalizability of our results. However, a limitation of our study is that we did not record the number of patients who turned down participation after being contacted by study personnel. Nevertheless, the mean age in our study was slightly lower than that of the average person with type 2 diabetes in Sweden, but levels of BMI and HbA_{1c} were similar [41]. Younger persons may be more inclined to participate in an app-based intervention, although internet access and smartphone usage are high also among older age groups in Sweden [42]. Another strength of our study is that both men and women were included. Earlier studies have shown that women participate in physical activity programs more than men [43]. The larger proportion of men in our study could partly be explained by the higher prevalence of type 2 diabetes among men; almost 60% of persons with type 2 diabetes within primary care in Sweden are men [41]. It could also be speculated that older men are more interested in using technology than older women and therefore more likely to participate in an mHealth intervention.

The fact that participants were not blinded to the intervention is a limitation. While baseline information regarding the clinical variables is unlikely to have been affected by participants being aware of their group allocation, this knowledge may have had an impact on accelerometer measurements. Participants can, intentionally or unintentionally, change their physical activity behaviors during the measuring period. The statistically significant differences between the intervention and control group in accelerometer-measured MVPA at baseline were not seen for physical activity assessed in the baseline questionnaire. This could be interpreted as an immediate effect of being randomized to the intervention group and thereby feeling encouraged to become more active. Future studies should be careful not to disclose allocation information to participants until all baseline measures have been completed. It is also a limitation that personnel working in the study were not blinded during measurements. Nevertheless, objective accelerometer measurement of physical activity and clinical biomarkers

XSL•FO
RenderX

including HbA_{1c} and lipid levels represent a strength in our study design as they are less prone to biased estimates resulting from unblinded study personnel. The continuous recruitment of study participants during the whole year also reduced the risk of results being biased due to season.

We did not reach our goal of including 250 participants in the study, which is a limitation. Because a digital solution must be continuously updated to run with the current iOS and Android versions, we had to end recruitment after 2.5 years for practical reasons, before reaching the goal. However, dropout rates were lower than the estimated 20% in our power calculation [21]. Lack of data on user engagement and adherence to the app is a limitation as higher user engagement has also been associated with more favorable outcomes [44]. Decreased app engagement over time during an intervention period has been suggested as a potential explanation for the lack of more long-term effects of physical activity promoting apps [45].

Conclusions

We found no evidence that persons with type 2 diabetes being randomized to use an app promoting daily walking increased their levels of MVPA at 3 or 6 months of follow-up compared with controls receiving standard care. Further, the effect on BMI was unclear, and we found nothing to support an effect of being randomized to use the DiaCert app on waist circumference, HbA_{1c}, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, and systolic or diastolic blood pressure compared with standard care. Further research is needed to determine what type of mHealth physical activity intervention could be effective to increase physical activity and improve cardiometabolic markers among persons with type 2 diabetes.

Acknowledgments

We would like to express our gratitude to the personnel at the participating health care centers for the help in recruiting participants to our study. This work was supported by Stockholm County Council and Karolinska Institutet (4D; YTL), as well as through grants from the Swedish Research Council for Health, Working life, and Welfare (Dnr: 2016-00985; YTL). The study was also financed by grants provided by Region Stockholm (Dnr: 20170784, the NSV project and a clinical research appointment; YTL) and funding from the strategic research area in care sciences (SFO-V; SEB). The funders had no role in the design, analysis, or writing of this paper. Generative AI tools have not been used in any portion of the presented work or manuscript writing.

Data Availability

The data sets generated during and analyzed during this study are not publicly available due ethical restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

YTL is responsible for the DiaCert study and has designed the study together with SEB and ML. SEB, MH and CA have contributed significantly to the data collection. HE processed accelerometer data and SEB performed the statistical analysis together with GP with critical input from RB. SEB prepared the initial draft of the manuscript together with MH, which has been critically reviewed by all authors. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of study participants with complete accelerometer data at baseline (n=156) by study group.
[DOCX File , 19 KB - ijmr_v13i1e53054_app1.docx]

Multimedia Appendix 2
Changes over time in secondary outcomes in the intervention and control group. Predicted group means from linear mixed model analysis; results correspond to model estimates shown in Table 2.
[PNG File , 255 KB - ijmr_v13i1e53054_app2.png]

Multimedia Appendix 3
CONSORT eHEALTH checklist.
[PDF File (Adobe PDF File), 50754 KB - ijmr_v13i1e53054_app3.pdf]

References

1. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes—global burden of disease and forecasted trends. J Epidemiol Glob Health 2020;10(1):107-111 [FREE Full text] [doi: 10.2991/jegh.k.191028.001] [Medline: 32175717]

XSL•FO
RenderX

2. American Diabetes Association. 5. Lifestyle management: standards of medical care in diabetes-2019. Diabetes Care 2019;42(Suppl 1):S46-S60 [FREE Full text] [doi: 10.2337/dc19-S005] [Medline: 30559231]
3. Kanaley JA, Colberg SR, Corcoran MH, Malin SK, Rodriguez NR, Crespo CJ, et al. Exercise/physical activity in individuals with type 2 diabetes: a consensus statement from the American College of Sports Medicine. Med Sci Sports Exerc 2022;54(2):353-368 [FREE Full text] [doi: 10.1249/MSS.0000000000002800] [Medline: 35029593]
4. Ried-Larsen M, MacDonald CS, Johansen MY, Hansen KB, Christensen R, Almdal TP, et al. Why prescribe exercise as therapy in type 2 diabetes? We have a pill for that!. Diabetes Metab Res Rev 2018;34(5):e2999. [doi: 10.1002/dmrr.2999] [Medline: 29488311]
5. Moghetti P, Balducci S, Guidetti L, Mazzuca P, Rossi E, Schena F, Italian Society of Diabetology (SID), Italian Association of Medical Diabetologists (AMD), Italian Society of Motor and Sports Sciences (SISMES). Walking for subjects with type 2 diabetes: a systematic review and joint AMD/SID/SISMES evidence-based practical guideline. Nutr Metab Cardiovasc Dis 2020;30(11):1882-1898 [FREE Full text] [doi: 10.1016/j.numecd.2020.08.021] [Medline: 32998820]
6. Karduck J, Chapman-Novakofski K. Results of the clinician apps survey, how clinicians working with patients with diabetes and obesity use mobile health apps. J Nutr Educ Behav 2018;50(1):62-69.e1. [doi: 10.1016/j.jneb.2017.06.004] [Medline: 29325664]
7. Jimenez G, Lum E, Car J. Examining diabetes management apps recommended from a Google search: content analysis. JMIR Mhealth Uhealth 2019;7(1):e11848 [FREE Full text] [doi: 10.2196/11848] [Medline: 30303485]
8. Sneha S, Thalla S, Rischie I, Shahriar H. Health internet technology for chronic conditions: review of diabetes management apps. JMIR Diabetes 2021;6(3):e17431 [FREE Full text] [doi: 10.2196/17431] [Medline: 34463627]
9. Bonn SE, Alexandrou C, Lagerros YT. A digital platform and smartphone app to increase physical activity in patients with type 2 diabetes: overview of a technical solution. JMIR Form Res 2023;7:e40285 [FREE Full text] [doi: 10.2196/40285] [Medline: 36917156]
10. Eberle C, Löhnert M, Stichling S. Effectiveness of disease-specific mHealth apps in patients with diabetes mellitus: scoping review. JMIR Mhealth Uhealth 2021;9(2):e23477 [FREE Full text] [doi: 10.2196/23477] [Medline: 33587045]
11. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
12. Lee JJN, Abdul Aziz A, Chan ST, Sahrizan RSFBRA, Ooi AYY, Teh YT, et al. Effects of mobile health interventions on health-related outcomes in older adults with type 2 diabetes: a systematic review and meta-analysis. J Diabetes 2023;15(1):47-57 [FREE Full text] [doi: 10.1111/1753-0407.13346] [Medline: 36649940]
13. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The effectiveness of smartphone apps for lifestyle improvement in noncommunicable diseases: systematic review and meta-analyses. J Med Internet Res 2018;20(5):e162 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]
14. Moschonis G, Siopis G, Jung J, Eweka E, Willems R, Kwasnicka D, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2023;5(3):e125-e143 [FREE Full text] [doi: 10.1016/S2589-7500(22)00233-3] [Medline: 36828606]
15. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
16. Feter N, Dos Santos TS, Caputo EL, da Silva MC. What is the role of smartphones on physical activity promotion? A systematic review and meta-analysis. Int J Public Health 2019;64(5):679-690. [doi: 10.1007/s00038-019-01210-7] [Medline: 30758514]
17. Mönninghoff A, Kramer JN, Hess AJ, Ismailova K, Teepe GW, Tudor Car L, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021;23(4):e26699 [FREE Full text] [doi: 10.2196/26699] [Medline: 33811021]
18. Pradal-Cano L, Lozano-Ruiz C, Pereyra-Rodríguez JJ, Saigí-Rubió F, Bach-Faig A, Esquius L, et al. Using mobile applications to increase physical activity: a systematic review. Int J Environ Res Public Health 2020;17(21):8238 [FREE Full text] [doi: 10.3390/ijerph17218238] [Medline: 33171871]
19. Poppe L, De Bourdeaudhuij I, Verloigne M, Shadid S, Van Cauwenberg J, Compernolle S, et al. Efficacy of a self-regulation-based electronic and mobile health intervention targeting an active lifestyle in adults having type 2 diabetes and in adults aged 50 years or older: two randomized controlled trials. J Med Internet Res 2019;21(8):e13363 [FREE Full text] [doi: 10.2196/13363] [Medline: 31376274]
20. Thorsen IK, Yang Y, Valentiner LS, Glümer C, Karstoft K, Brønd JC, et al. The effects of a lifestyle intervention supported by the InterWalk Smartphone App on increasing physical activity among persons with type 2 diabetes: parallel-group, randomized trial. JMIR Mhealth Uhealth 2022;10(9):e30602 [FREE Full text] [doi: 10.2196/30602] [Medline: 36170002]
21. Bonn SE, Alexandrou C, Steiner KH, Wiklander K, Östenson CG, Löf M, et al. App-technology to increase physical activity among patients with diabetes type 2—the DiaCert-study, a randomized controlled trial. BMC Public Health 2018;18(1):119 [FREE Full text] [doi: 10.1186/s12889-018-5026-4] [Medline: 29316905]

XSL•FO
RenderX

22. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;152(11):726-732 [FREE Full text] [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
23. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
24. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De Cocker K, Giles-Corti B, et al. How many steps/day are enough? For adults. Int J Behav Nutr Phys Act 2011;8:79 [FREE Full text] [doi: 10.1186/1479-5868-8-79] [Medline: 21798015]
25. Migueles JH, Cadenas-Sanchez C, Ekelund U, Nyström CD, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017;47(9):1821-1845 [FREE Full text] [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
26. Migueles JH, Huber F, Sabia S, Rowlands AV, van Hees VT. GGIR: a research community-driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. J Meas Phys Behav 2019;2(3):188-186 [FREE Full text] [doi: 10.1123/jmpb.2018-0063]
27. Menai M, van Hees VT, Elbaz A, Kivimaki M, Singh-Manoux A, Sabia S. Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. Sci Rep 2017;8:45772 [FREE Full text] [doi: 10.1038/srep45772] [Medline: 28367987]
28. Olsson SJG, Ekblom Ö, Andersson E, Börjesson M, Kallings LV. Categorical answer modes provide superior validity to open answers when asking for level of physical activity: a cross-sectional study. Scand J Public Health 2016;44(1):70-76. [doi: 10.1177/1403494815602830] [Medline: 26392418]
29. Jeppsson JO, Kobold U, Barr J, Finke A, Hoelzel W, Hoshino T, et al. Approved IFCC reference method for the measurement of HbA1c in human blood. Clin Chem Lab Med 2002;40(1):78-89. [doi: 10.1515/CCLM.2002.016] [Medline: 11916276]
30. Nilsson-Ehle P, Berggren So¨derlund M, Theodorsson E, Becker C, Laurell CB. Laurells Klinisk kemi i praktisk medicin. In: Laurells Klinisk kemi i praktisk medicin, 9. uppl [Laurell's Clinical Chemistry in Practical Medicine, 9th Edition]. Lund, Sweden: Studentlitteratur; 2012.
31. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18(6):499-502. [Medline: 4337382]
32. Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welten M, Heymans M. Different ways to estimate treatment effects in randomised controlled trials. Contemp Clin Trials Commun 2018;10:80-85 [FREE Full text] [doi: 10.1016/j.conctc.2018.03.008] [Medline: 29696162]
33. Li P, Stuart EA. Best (but oft-forgotten) practices: missing data methods in randomized controlled nutrition trials. Am J Clin Nutr 2019;109(3):504-508 [FREE Full text] [doi: 10.1093/ajcn/nqy271] [Medline: 30793174]
34. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011;20(1):40-49 [FREE Full text] [doi: 10.1002/mpr.329] [Medline: 21499542]
35. Höchsmann C, Müller O, Ambühl M, Klenk C, Königstein K, Infanger D, et al. Novel smartphone game improves physical activity behavior in type 2 diabetes. Am J Prev Med 2019;57(1):41-50 [FREE Full text] [doi: 10.1016/j.amepre.2019.02.017] [Medline: 31128953]
36. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014;64(624):e384-e391 [FREE Full text] [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
37. Rossen J, Larsson K, Hagströmer M, Yngve A, Brismar K, Ainsworth B, et al. Effects of a three-armed randomised controlled trial using self-monitoring of daily steps with and without counselling in prediabetes and type 2 diabetes-the Sophia Step Study. Int J Behav Nutr Phys Act 2021;18(1):121 [FREE Full text] [doi: 10.1186/s12966-021-01193-w] [Medline: 34496859]
38. Kim JE, Park TS, Kim KJ. The clinical effects of type 2 diabetes patient management using digital healthcare technology: a systematic review and meta-analysis. Healthcare (Basel) 2022;10(3):522 [FREE Full text] [doi: 10.3390/healthcare10030522] [Medline: 35327000]
39. Fanning J, Roberts S, Hillman CH, Mullen SP, Ritterband L, McAuley E. A smartphone "app"-delivered randomized factorial trial targeting physical activity in adults. J Behav Med 2017;40(5):712-729. [doi: 10.1007/s10865-017-9838-y] [Medline: 28255750]
40. Sporrel K, Nibbeling N, Wang S, Ettema D, Simons M. Unraveling mobile health exercise interventions for adults: scoping review on the implementations and designs of persuasive strategies. JMIR Mhealth Uhealth 2021;9(1):e16282 [FREE Full text] [doi: 10.2196/16282] [Medline: 33459598]
41. Nationella Diabetesregistret—ÅRsrapport 2020 [The Swedish National Diabetes Register, annual report 2020]. Nationella Diabetesregistret. 2020. URL: https://www.ndr.nu/pdfs/Arsrapport_NDR_2020.pdf [accessed 2024-02-13]
42. Annual report—The Swedes and the Internet 2021. Swedish Svenskarna och internet [The Swedish Internet Foundation]. 2021. URL: https://svenskarnaochinternet.se/english/earlier-years-reports/ [accessed 2024-02-13]
43. Stevens W, Hillsdon M, Thorogood M, McArdle D. Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women: a randomised controlled trial. Br J Sports Med 1998;32(3):236-241 [FREE Full text] [doi: 10.1136/bjsm.32.3.236] [Medline: 9773174]

XSL•FO
RenderX

44. Lavikainen P, Mattila E, Absetz P, Harjumaa M, Lindström J, Järvelä-Reijonen E, et al. Digitally supported lifestyle intervention to prevent type 2 diabetes through healthy habits: secondary analysis of long-term user engagement trajectories in a randomized controlled trial. J Med Internet Res 2022;24(2):e31530 [FREE Full text] [doi: 10.2196/31530] [Medline: 35200147]
45. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
HbA_{1c}: hemoglobin A_{1c}
HDL: high-density lipoprotein
LDL: low-density lipoprotein
mHealth: mobile health
MICE: multiple imputation by chained equations
MVPA: moderate to vigorous physical activity

Edited by T de Azevedo Cardoso; submitted 26.09.23; peer-reviewed by I Thorsen, A Mateo-Orcajada, G Wren; comments to author 15.11.23; revised version received 22.12.23; accepted 07.02.24; published 21.03.24.

Please cite as:
Bonn SE, Hummel M, Peveri G, Eke H, Alexandrou C, Bellocco R, Löf M, Trolle Lagerros Y
Effectiveness of a Smartphone App to Promote Physical Activity Among Persons With Type 2 Diabetes: Randomized Controlled Trial
Interact J Med Res 2024;13:e53054
URL: https://www.i-jmr.org/2024/1/e53054
doi:10.2196/53054
PMID:38512333

©Stephanie E Bonn, Madeleine Hummel, Giulia Peveri, Helén Eke, Christina Alexandrou, Rino Bellocco, Marie Löf, Ylva Trolle Lagerros. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 21.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Preliminary Checklist (METRICS) to Standardize the Design and Reporting of Studies on Generative Artificial Intelligence–Based Models in Health Care Education and Practice: Development Study Involving a Literature Review

Malik Sallam[1,2,3], MD, PhD; Muna Barakat[4], PhD; Mohammed Sallam[5], MSc, PharmD

[1]Department of Pathology, Microbiology and Forensic Medicine, School of Medicine, The University of Jordan, Amman, Jordan

[2]Department of Clinical Laboratories and Forensic Medicine, Jordan University Hospital, Amman, Jordan

[3]Department of Translational Medicine, Faculty of Medicine, Lund University, Malmo, Sweden

[4]Department of Clinical Pharmacy and Therapeutics, Faculty of Pharmacy, Applied Science Private University, Amman, Jordan

[5]Department of Pharmacy, Mediclinic Parkview Hospital, Mediclinic Middle East, Dubai, United Arab Emirates

Corresponding Author:
Malik Sallam, MD, PhD
Department of Pathology, Microbiology and Forensic Medicine
School of Medicine
The University of Jordan
Queen Rania Al-Abdullah Street-Aljubeiha
Amman, 11942
Jordan
Phone: 962 0791845186
Fax: 962 65353388
Email: malik.sallam@ju.edu.jo

Abstract

Background: Adherence to evidence-based practice is indispensable in health care. Recently, the utility of generative artificial intelligence (AI) models in health care has been evaluated extensively. However, the lack of consensus guidelines on the design and reporting of findings of these studies poses a challenge for the interpretation and synthesis of evidence.

Objective: This study aimed to develop a preliminary checklist to standardize the reporting of generative AI-based studies in health care education and practice.

Methods: A literature review was conducted in Scopus, PubMed, and Google Scholar. Published records with "ChatGPT," "Bing," or "Bard" in the title were retrieved. Careful examination of the methodologies employed in the included records was conducted to identify the common pertinent themes and the possible gaps in reporting. A panel discussion was held to establish a unified and thorough checklist for the reporting of AI studies in health care. The finalized checklist was used to evaluate the included records by 2 independent raters. Cohen κ was used as the method to evaluate the interrater reliability.

Results: The final data set that formed the basis for pertinent theme identification and analysis comprised a total of 34 records. The finalized checklist included 9 pertinent themes collectively referred to as METRICS (Model, Evaluation, Timing, Range/Randomization, Individual factors, Count, and Specificity of prompts and language). Their details are as follows: (1) Model used and its exact settings; (2) Evaluation approach for the generated content; (3) Timing of testing the model; (4) Transparency of the data source; (5) Range of tested topics; (6) Randomization of selecting the queries; (7) Individual factors in selecting the queries and interrater reliability; (8) Count of queries executed to test the model; and (9) Specificity of the prompts and language used. The overall mean METRICS score was 3.0 (SD 0.58). The tested METRICS score was acceptable, with the range of Cohen κ of 0.558 to 0.962 ($P<.001$ for the 9 tested items). With classification per item, the highest average METRICS score was recorded for the "Model" item, followed by the "Specificity" item, while the lowest scores were recorded for the "Randomization" item (classified as suboptimal) and "Individual factors" item (classified as satisfactory).

Conclusions: The METRICS checklist can facilitate the design of studies guiding researchers toward best practices in reporting results. The findings highlight the need for standardized reporting algorithms for generative AI-based studies in health care, considering the variability observed in methodologies and reporting. The proposed METRICS checklist could be a preliminary

XSL•FO
RenderX

helpful base to establish a universally accepted approach to standardize the design and reporting of generative AI-based studies in health care, which is a swiftly evolving research topic.

(Interact J Med Res 2024;13:e54704) doi:10.2196/54704

KEYWORDS

guidelines; evaluation; meaningful analytics; large language models; decision support

Introduction

The integration of generative artificial intelligence (AI) models into health care education and practice holds promising perspectives with numerous possibilities for continuous improvement [1-5]. Examples of generative AI-based conversational models characterized by ease of use and perceived usefulness include ChatGPT by OpenAI, Bing by Microsoft, and Bard by Google [6-8].

The vast potential of generative AI-based models in health care can be illustrated as follows. First, generative AI-based models can facilitate streamlining of the clinical workflow, with subsequent improvement in efficiency manifested in reduced time for care delivery and reduced costs [1,9-11]. Second, generative AI-based models can enhance personalized medicine with a huge potential to achieve refined prediction of disease risks and outcomes [1,12,13]. Third, generative AI-based models can be implemented to improve health literacy among lay individuals through the provision of easily accessible and understandable health information [1,14,15].

Despite the aforementioned advantages of generative AI-based models in health care, several valid concerns were raised, which should be considered carefully owing to their serious consequences [1,4,16]. For example, the lack of clarity on how generative AI-based models are trained raises ethical concerns [17,18]. Additionally, these models have an inherent bias in the generated content based on the modality of training used for their development and updates [17,18]. Importantly, the generation of inaccurate or misleading content, which might appear scientifically plausible to nonexperts (referred to as “hallucinations”), could have profound negative impacts in health care settings [1,19-21]. Furthermore, the integration of generative AI-based models in health care could raise complex medicolegal and accountability questions, compounded by the issues of data privacy and cybersecurity risks [1,4,22,23].

Similarly, the use of generative AI-based models can cause a paradigm shift in information acquisition, particularly in health care education [1,24-26]. However, careful consideration of the best policies and practices to incorporate AI-based models in health care education is needed [27]. This issue involves the urgent need to address the issues of inaccuracies, possible academic dishonesty, decline in critical thinking development, and deterioration of practical training skills [1].

Recently, a remarkable number of studies investigated the applicability and disadvantages of prominent generative AI-based conversational models, such as ChatGPT, Microsoft Bing, and Google Bard, in various health care and educational settings [1,2,4,28-34]. However, synthesizing evidence from such studies can be challenging owing to several reasons. Variations in methodologies implemented in various studies as well as in the reporting standards is a major limitation. This issue could hinder the efforts aiming to compare and contrast the results of generative AI-based studies in health care, contributing to the complexity in this domain. This variability arises from several factors, including different settings of the tested models, prompt variability, varying approaches used to evaluate the generated content of generative AI-based models, varying range of tested topics, and possible bias in selecting the tested subjects. Additionally, variability can be related to the different number and varying expertise of individual raters of content quality, as well as the variable number of queries executed, among other factors [35-37].

Therefore, it is important to initiate and develop an approach that can aid in establishing standardized reporting practices for studies aiming to evaluate the content of generative AI-based models, particularly in health care. This standardization can be crucial to facilitate the design of generative AI-based studies in health care, ensuring rigor and achieving precise comparison and credible synthesis of findings across different studies. Thus, we aimed to propose a preliminary framework (checklist) to establish proper guidelines for the design and reporting of findings of generative AI-based studies that address health care–related topics.

Methods

Study Design

The study was based on a literature review to highlight the key methodological aspects in studies that investigated 3 generative AI-based models (ChatGPT, Bing, and Bard) in health care education and practice. The literature review was conducted to identify relevant literature indexed in databases up to November 11, 2023 [38]. The databases used for this literature search were Scopus, PubMed/MEDLINE, and Google Scholar.

Ethical Considerations

This study did not involve human subjects, and thus, the requirement of ethical permission was waived.

Literature Search to Identify Relevant Records

The Scopus string query was as follows: (TITLE-ABS-KEY(“artificial intelligence” OR “AI”) AND TITLE-ABS-KEY (“healthcare” OR “health care”) AND TITLE-ABS-KEY (“education” OR “practice”)) AND PUBYEAR > 2022 AND DOCTYPE (ar OR re) AND (LIMIT-TO (PUBSTAGE , “final”)) AND (LIMIT-TO (SRCTYPE , “j”)) AND (LIMIT-TO (LANGUAGE , “English”)). The Scopus search yielded a total of 843 documents.

XSL•FO
RenderX

The PubMed advanced search tool was used as follows: ("artificial intelligence"[Title/Abstract] OR "AI"[Title/Abstract]) AND ("healthcare"[Title/Abstract] OR "health care"[Title/Abstract]) AND ("education"[Title/Abstract] OR "practice"[Title/Abstract]) AND ("2022/12/01"[Date - Publication] : "2023/11/11"[Date - Publication]). The PubMed search yielded a total of 564 records.

In Google Scholar and using the Publish or Perish software (version 8), in the title words and in the years 2022-2023, the search was as follows: "artificial intelligence" OR "AI" AND "healthcare" OR "health care" AND "education" OR "practice," with a maximum of 999 records retrieved [39].

Criteria for Record Inclusion

The records from the 3 databases were merged using EndNote 20.2.1 software. This was followed by removal of duplicate records and removal of preprints by using the following function: ANY FIELD preprint OR ANY FIELD rxiv OR ANY FIELD SSRN OR ANY FIELD Researchgate OR ANY FIELD researchsquare OR ANY FIELD. The retrieved records were eligible for the final screening step given the following inclusion criteria: (1) Original article; (2) English record; (3) Published (peer reviewed); and (4) Assessment in health care practice or health care education. Finally, the imported references were subjected to the search function in EndNote as follows: Title contains ChatGPT OR Title contains Bing OR Title contains Bard. The selection of the included records was performed by the first author (Malik Sallam).

Development of the Initial Checklist Items

Initial development of the proposed checklist began with the assessment of the methodology and results sections of the included records, a majority of which were regarded as cross-sectional descriptive studies. Then, we referred to 2 commonly used reporting and quality guidelines to proactively explore pertinent themes for the proposed checklist based on the nature of the included records: (1) STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) Statement: guidelines for reporting observational studies (checklist: cross-sectional studies) and (2) CASP (Critical Appraisal Skills Programme) Checklist (CASP qualitative studies checklist) [40,41]. This facilitated the allocation of ethical considerations, including transparency, methodological rigor, and issues related to bias, in the proposed checklist. Then, the 3 authors conducted an independent content review process to identify all the possible essential themes and best practices in generative AI-based health care studies among the included records. Finally, the authors had a collaborative discussion to refine the selected themes and classify these themes into specific components relevant to the study objectives. Special attention was given to aspects of method and result reporting that were perceived to impact the quality and reproducibility of the records, as identified by the 3 authors.

Establishing the Final Checklist Criteria

Careful examination of the included records resulted in the compilation of 3 independent lists of "pertinent themes," which are herein defined as being critical or recurring in the reporting of results of generative AI-based studies. A thorough discussion among the authors followed to reach a consensus on the pertinent themes. Recurring themes were defined as those found in the methods of at least three separate records. Critical aspects were defined as those that would impact the conclusions of the included records as agreed by the 3 authors.

The final pertinent themes were selected based on their author-perceived significance in the quality and reproducibility of the findings. A final list of 9 themes was agreed upon by the authors as follows: (1) the "Model" of the generative AI-based tool or tools used in the included record and the explicit mention of the exact settings employed for each tool; (2) the "Evaluation" approach to assess the quality of the content generated by the generative AI-based model in terms of objectivity to reach unbiased findings and subjectivity; (3) the exact "Timing" of generative AI-based model testing and its duration; (4) the "Transparency" of data sources used to generate queries for the generative AI-based model testing, including the permission to use copyrighted content; (5) the "Range" of topics tested (single topic, multiple related topics, or various unrelated topics, as well as the breadth of intertopic and intratopic queries tested); (6) the degree of "Randomization" of topics selected to be tested to consider the potential bias; (7) the "Individual" subjective role in evaluating the content and the possibility of interrater reliability concordance or discordance; (8) the number ("Count") of queries executed on each generative AI model entailing the sample size of queries tested; and (9) the "Specificity" of prompts used on each generative AI-based model, including the exact phrasing of each prompt and the presence of feedback and learning loops, and the "Specificity" of the language or languages used in testing, besides any other cultural issues. Thus, the final checklist was termed METRICS (Model, Evaluation, Timing, Range/Randomization, Individual factors, Count, and Specificity of prompts and language).

Scoring the METRICS Items and Classification of the METRICS Score

Testing of the included records was performed by 2 independent raters (Malik Sallam and Mohammed Sallam) independently, with each METRICS item scored using a 5-point Likert scale as follows: 5=excellent, 4=very good, 3=good, 2=satisfactory, and 1=suboptimal. For the items that were deemed "not applicable" (eg, individual factors for studies that employed objective methods for evaluation), no score was given. The scores for the 2 raters were then averaged. The average METRICS score was calculated as the sum of average scores for each applicable item divided by 10 minus the number of items deemed not applicable.

The subjective assessment of the 2 raters was performed based on predefined criteria as a general guide. For example, if the exact dates of model queries were mentioned, the "Timing" item was scored as excellent. The count of queries was agreed to be categorized as excellent if it was more than 500, while a single case or no mention of the count was considered suboptimal. For the prompt attributes, scores were assigned based on the availability of exact prompts, explicit mention of the language used, and details of prompting. Thus, prompts and language specificity were appraised positively if the study clearly and explicitly made the exact prompts available and if

XSL•FO
RenderX

there was an explicit mention of the language employed in the prompts. The evaluation method was agreed to be rated higher for objective assessments with full details and lower for subjective assessments. The explicit mention of the method of interrater reliability testing was agreed to be scored higher for the "Individual" item. Transparency was assessed based on the comprehensiveness of the data source, and the presence of full database disclosure and permission to use the data was agreed to be given an excellent score. Randomization was agreed to be scored the lowest for the absence of details and the highest for explicit detailed descriptions.

Finally, we decided to highlight the records that scored the highest for each METRICS item. The decision to take this approach was based on an attempt to refrain from providing examples for the other quality categories to avoid premature conclusions regarding the quality of the included studies owing to the preliminary pilot nature of the METRICS tool.

Statistical and Data Analysis

The average METRICS scores were classified into distinct categories of equal weights as follows: excellent (score 4.21-5.00), very good (3.41-4.20), good (2.61-3.40), satisfactory (1.81-2.60), and suboptimal (1.00-1.80).

The Cohen κ measure was used to assess the interrater reliability by 2 independent raters. The Cohen κ measure was categorized as follows: <0.20, poor agreement; 0.21-0.40, fair agreement; 0.41-0.60, good agreement; 0.61-0.80, very good agreement; and 0.81-1.00, excellent agreement.

For statistical analysis, we used IBM SPSS Statistics for Windows, Version 26.0 (IBM Corp). A *P* value <.05 was considered significant.

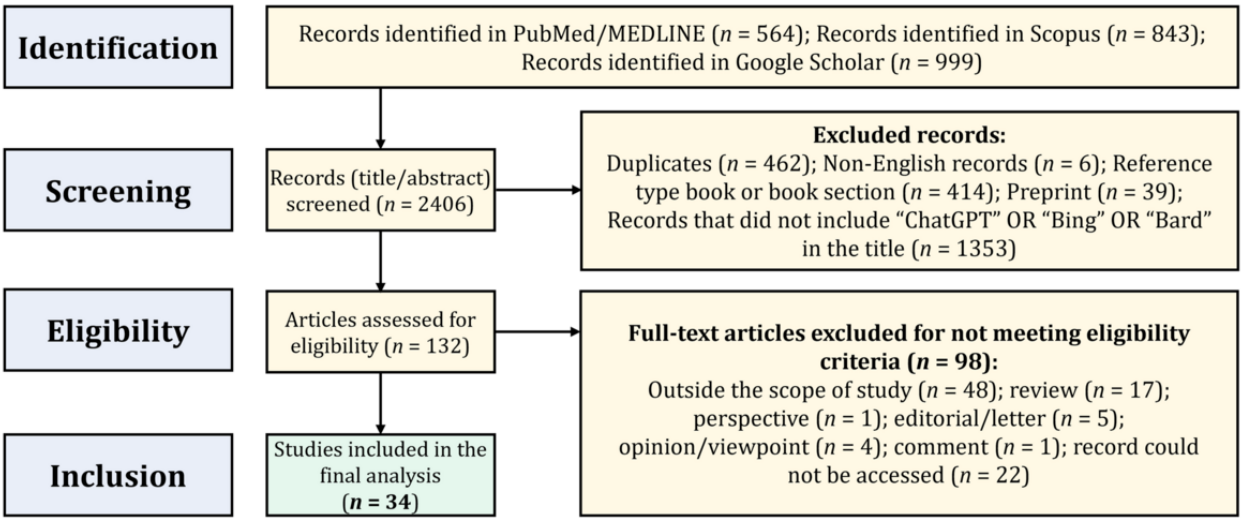
Results

Description of the Included Studies

A total of 34 studies were included in the final analysis that aimed to establish the METRICS criteria (Figure 1).

The most common source of records was Cureus, with 9 of the 34 records (27%), followed by BMJ Neurology Open, with 2 of the 34 records (6%). The remaining 23 records were published in 23 different journals.

Figure 1. The approach for selecting articles.



Evaluation of the Included Records Based on MED-METRICS Items

The METRICS checklist was divided into the following 3 parts: "Model" attributes, "Evaluation" approach, and features of "Data" (MED-METRICS).

The complete details of the model attributes of the included studies are presented in Table 1.

ChatGPT was tested in all the records included (34/34, 100%), followed by Google Bard (5/34, 15%) and Bing Chat (5/34, 15%). The exact dates of generative AI-based model queries were explicitly mentioned in 13 of the 34 records (38%). The count of cases or questions that were tested in the studies ranged from a single case to 2576 questions. The majority of studies (23/34, 68%) tested the AI models based on queries in the English language.

The complete details of the evaluation approach of the content generated by the AI-based models in the included studies are presented in Table 2.

Objective evaluation of the content generated by the generative AI-based model was noted in 15 of the 34 records (44%).

The complete details of the features of data used to generate queries for testing on the generative AI-based models, including the range of topics and randomization, are presented in Table 3.

Explicit mention of the randomization process was only noted in 4 of the 34 included studies (9%). Of the 34 records, 6 (18%) involved broad multidisciplinary medical exam questions (18%). Moreover, 2 studies (6%) explicitly mentioned the permission to use the data for the studies.

XSL•FO
RenderX

Table 1. Details of the model attributes of the included studies.

Authors	Model	Timing	Count	Specificity of the prompts and language
Al-Ashwal et al [42]	ChatGPT-3.5, ChatGPT-4, Bing, and Bard with unclear settings	One month, May 2023	255 drug-drug pairs	The exact prompts were provided in English.
Alfertshofer et al [43]	ChatGPT with unclear settings	Not provided	1800 questions	The exact prompt was used for each question. The questions were taken from United States, United Kingdom, Italian, French, Spanish, and Indian exams. A new session was used for each question.
Ali et al [44]	ChatGPT Feb 9 free version with unclear settings	Not provided	50 items	Information was provided fully in the supplementary file of the article in English.
Aljindan et al [45]	ChatGPT-4 with unclear settings	Not provided	220 questions	Initial prompting that involved role playing as a medical professional. The language was English.
Altamimi et al [46]	ChatGPT-3.5	Single day not otherwise specified	9 questions	The exact prompts were provided in English.
Baglivo et al [47]	Bing, ChatGPT, Chatsonic, Bard, and YouChat with full details of the mode and large language model, including plugins	Exact dates were provided for each model (April 12, 13, and 14, 2023, and July 13, 2023)	15 questions	Italian was used.
Biswas et al [48]	ChatGPT-3.5 with unclear settings	Exact date provided (March 16, 2023)	11 questions	The exact prompts were provided in English, with a new session for each question.
Chen et al [49]	ChatGPT-4 with unclear settings	Not provided	560 questions	The exact prompts were provided in English.
Deiana et al [50]	ChatGPT-3.5 and ChatGPT-4 with unclear settings	Not provided	11 questions	The exact prompts were not explicitly provided. English was used. A new session was used for each question. Up to three iterations were allowed for incorrect responses.
Fuchs et al [51]	ChatGPT-3 and ChatGPT-4 with unclear settings	Exact dates provided (February 19, 2023, and March 25, 2023)	60 questions	Dental medicine questions were translated from German to English, while the other questions were already present in English. Exact prompts were provided in English with prompting in 2 groups for the same questions: one group was primed with instructions, while the other was not primed. A total of 20 trials were conducted per group, and chat history was cleared after each trial.
Ghosh & Bir [52]	ChatGPT (version March 14, 2023) with unclear settings	March 14 and 16, 2023	200 questions	The exact prompts and language were not explicitly provided. The first response was taken as final, and the option of “regenerate response” was not used.
Giannos [53]	ChatGPT-3 and ChatGPT-4 with unclear settings	Not provided	69 questions	Not explicitly provided.
Gobira et al [54]	ChatGPT-4 with unclear settings	Not provided	125 questions	Portuguese was used.
Grewal et al [55]	ChatGPT-4 with unclear settings	The first week of May 2023	Not clear	The exact prompts were provided in English. One follow-up prompt was used for enhancement of some prompts.
Guerra et al [56]	ChatGPT-4 with unclear settings	Not provided	591 questions	The exact prompts were provided, while the language was not explicitly provided.
Hamed et al [57]	ChatGPT-4 with unclear settings	Not provided	Not clear	The exact prompts and language were not explicitly provided. Different prompts were tried to identify the most suitable.
Hoch et al [58]	ChatGPT (May 3rd version) with unclear settings	May 5 and 7, 2023	2576 questions	The exact prompts were provided, while the language was not explicitly provided.

XSL•FO
RenderX

Authors	Model	Timing	Count	Specificity of the prompts and language
Juhi et al [59]	ChatGPT with unclear settings	February 20, 2023, to March 5, 2023	40 drug-drug pairs	The exact prompts were provided in English.
Kuang et al [60]	ChatGPT with unclear settings	Not provided	Not clear	The exact prompts were not explicitly provided. English was used.
Kumari et al [61]	ChatGPT-3.5, Bard, and Bing with unclear settings	July 30, 2023	50 questions	The exact prompts were not explicitly provided. English was used.
Kung et al [62]	ChatGPT-3.5 and ChatGPT-4 with unclear settings	July 2023	215 questions	Not clear
Lai et al [63]	ChatGPT-4 (May 24; Version 3.5) with unclear settings	Not provided	200 questions	The exact prompts and language were not explicitly provided. Three attempts to answer the complete set of questions over 3 weeks (once per week), with a new session for each question.
Lyu et al [64]	ChatGPT with unclear settings	Mid-February 2023	Not clear	The exact prompts were provided in English.
Moise et al [65]	ChatGPT-3.5 with unclear settings	Not provided	23 questions	The exact prompts were provided in English, with a new session for each question.
Oca et al [66]	ChatGPT, Bing, and Bard with unclear settings	April 11, 2023	20 queries for each model	The exact prompts were provided in English.
Oztermeli & Oztermeli [67]	ChatGPT-3.5 with unclear settings	Not provided	1177 questions	The exact prompts were not explicitly provided. Turkish was used, with a new session for each question.
Pugliese et al [68]	ChatGPT with unclear settings	March 25, 2023	15 questions	The exact prompts were provided in English, with a new session for each question.
Sallam et al [69]	ChatGPT (default model) with unclear settings	February 25, 2023	Not provided	The exact prompts were provided in English.
Seth et al [70]	ChatGPT-3.5, Bard, and Bing AI	Not provided	6 questions	The exact prompts were provided in English.
Suthar et al [71]	ChatGPT-4 with unclear settings	Not provided	140 cases	The exact prompts were not explicitly provided. English was used.
Walker et al [72]	ChatGPT-4 with unclear settings	Not provided	5 cases	The exact prompts were not explicitly provided. English was used, with a new session for each question.
Wang et al [73]	ChatGPT-3.5 and ChatGPT-4 with unclear settings	February 14, 2023, for ChatGPT-3.5 and May 14-16, 2023, for ChatGPT-4	300 questions	The exact prompts were provided. Chinese and English were used. The prompts were enhanced though role play.
Wang et al [74]	ChatGPT-3.5 with unclear settings	March 5-10, 2023	Not clear	Chinese (Mandarin) and English were used. Examples of prompts were provided.
Zhou et al [75]	ChatGPT-3.5 with unclear settings	April 24-25, 2023	Single case and multiple poll questions	The exact prompts were provided in English.

XSL•FO
RenderX

Table 2. Classification based on the evaluation approach of the content generated by the artificial intelligence–based models.

Authors	Evaluation of performance	Individual role and interrater reliability
Al-Ashwal et al [42]	Objective via 2 different clinical reference tools	Not applicable
Alfertshofer et al [43]	Objective based on the key answers, with the questions screened independently by 4 investigators	Not applicable
Ali et al [44]	Objective for multiple-choice questions and true or false questions, and subjective for short-answer and assay questions	Assessment by 2 assessors independently with intraclass correlation coefficient for agreement
Aljindan et al [45]	Objective based on key answers and historical performance metrics	Not applicable
Altamimi et al [46]	Subjective	Not clear; Assessment for accuracy, informativeness, and accessibility by clinical toxicologists and emergency medicine physicians
Baglivo et al [47]	Objective based on key answers and comparison with 5th year medical students' performance	Not applicable
Biswas et al [48]	Subjective by a 5-member team of optometry teaching and expert staff with over 100 years of clinical and academic experience between them; Independent evaluation on a 5-point Likert scale ranging from very poor to very good	The median scores across raters for each response were studied; The score represented rater consensus, while the score variance represented disagreements between the raters
Chen et al [49]	Objective based on key answers	Not applicable
Deiana et al [50]	Subjective based on qualitative assessment of correctness, clarity, and exhaustiveness; Each response rated using a 4-point Likert scale scored from strongly disagree to strongly agree	Independent assessment by 2 raters with experience in vaccination and health communication topics
Fuchs et al [51]	Objective based on key answers	Not applicable
Ghosh & Bir [52]	Objective based on key answers; Subjectivity by raters' assessment	Scoring by 2 assessors on a scale of 0 to 5, with 0 being incorrect and 5 being fully correct, based on a preselected answer key
Giannos [53]	Objective based on key answers	Not applicable
Gobira et al [54]	Objective based on key answers, with an element of subjectivity through classifying the responses as adequate, inadequate, or indeterminate	Two raters independently scored the accuracy; After individual evaluations, the raters performed a third assessment to reach a consensus on the questions with differing results
Grewal et al [55]	Not clear	Not clear
Guerra et al [56]	Subjective through comparison with the results of a previous study on the average performance of users, and a cohort of medical students and neurosurgery residents	Not applicable
Hamed et al [57]	Subjective	Not clear
Hoch et al [58]	Objective based on key answers	Not applicable
Juhi et al [59]	Subjective and the use of Stockley's Drug Interactions Pocket Companion 2015 as a reference key	Two raters reached a consensus for categorizing the output
Kuang et al [60]	Subjective	Not clear
Kumari et al [61]	Subjective; Content validity checked by 2 experts of curriculum design	Three independent raters scored content based on their correctness, with an accuracy score ranging from 1 to 5
Kung et al [62]	Objective based on key answers	Not applicable
Lai et al [63]	Objective based on key answers	Not applicable
Lyu et al [64]	Subjective	Two experienced radiologists (with 21 and 8 years of experience) evaluated the quality of the ChatGPT responses
Moise et al [65]	Subjective through comparison with the latest American Academy of Otolaryngology–Head and Neck Surgery Foundation Clinical Practice Guideline: Tympanostomy Tubes in Children (Update)	Two independent raters evaluated the output; The interrater reliability was assessed using the Cohen κ test; To confirm consensus, responses were reviewed by the senior author
Oca et al [66]	Not clear	Not clear

XSL•FO
RenderX

Authors	Evaluation of performance	Individual role and interrater reliability
Oztermeli & Oztermeli [67]	Objective based on key answers	Not applicable
Pugliese et al [68]	Subjective using the Likert scale for accuracy, completeness, and comprehensiveness	Multirater: 10 key opinion leaders in nonalcoholic fatty liver disease and 1 nonphysician with expertise in patient advocacy in liver disease independently rating the AI[a] content
Sallam et al [69]	Subjective based on correctness, clarity, and conciseness	Fleiss multirater κ
Seth et al [70]	Subjective through comparison with the current health care guidelines for rhinoplasty, and evaluation by a panel of plastic surgeons through a Likert scale to assess the readability and complexity of the text and the education level required for understanding, and the modified DISCERN[b] score	Not clear
Suthar et al [71]	Subjective by 3 fellowship-trained neuroradiologists, using a 5-point Likert scale, with 1 indicating "highly improbable" and 5 indicating "highly probable"	Not applicable
Walker et al [72]	Modified EQIP[c] Tool with comparison with UK National Institute for Health and Care Excellence guidelines for gallstone disease, pancreatitis, liver cirrhosis, or portal hypertension, and the European Association for Study of the Liver guidelines	All answers were assessed by 2 authors independently, and in case of a contradictory result, resolution was achieved by consensus; The process was repeated 3 times per EQIP item; Wrong or out of context answers, known as "AI hallucinations," were recorded
Wang et al [73]	Subjective	Unclear
Wang et al [74]	Objective based on key answers	Not applicable
Zhou et al [75]	Subjective	Unclear

[a]AI: artificial intelligence.

[b]DISCERN is an instrument for judging the quality of written consumer health information on treatment choices.

[c]EQIP: Ensuring Quality Information for Patients.

XSL•FO
RenderX

Table 3. Classification of the included studies based on the features of data used to generate queries for testing on the generative artificial intelligence–based models.

Authors	Transparency	Range	Randomization
Al-Ashwal et al [42]	Full description using 2 tools for assessment: Micromedex, a subscription-based drug-drug interaction screening tool, and Drugs.com, a free database	Narrow; Drug-drug interaction prediction	Nonrandom; Purposeful selection of the drugs by a clinical pharmacist; 5 drugs paired with the top 51 prescribed drugs
Alfertshofer et al [43]	Full description using the question bank AMBOSS, with official permission for the use of the AMBOSS USMLE step 2CK practice question bank for research purposes granted by AMBOSS	Broad	Randomly extracted
Ali et al [44]	Developed by the researchers and reviewed by a panel of experienced academics for accuracy, clarity of language, relevance, and agreement on correct answers; Evaluation of face validity, accuracy, and suitability for undergraduate dental students	Narrow intersubject (dentistry); Broad intrasubject in restorative dentistry, periodontics, fixed prosthodontics, removable prosthodontics, endodontics, pedodontics, orthodontics, preventive dentistry, oral surgery, and oral medicine	Not clear
Aljindan et al [45]	Saudi Medical Licensing Exam questions extracted from the subscription CanadaQBank website	Broad in medicine, with 30% of the questions from medicine, 25% from obstetrics and gynecology, 25% from pediatrics, and the remaining 20% from surgery	Randomized through 4 researchers to ensure comprehensive coverage of questions and eliminate potential bias in question selection
Altamimi et al [46]	Snakebite management information guidelines derived from the World Health Organization, Centers for Disease Control and Prevention, and the clinical literature	Narrow	Not clear
Baglivo et al [47]	The Italian National Medical Residency test	Narrow; Vaccination-related questions from the Italian National Medical Residency Test	Not clear
Biswas et al [48]	Constructed based on the frequently asked questions on the myopia webpage of the Association of British Dispensing Opticians and the College of Optometrists	Narrow involving 9 categories: 1 each for disease summary, cause, symptom, onset, prevention, complication, natural history of untreated myopia, and prognosis, and 3 involving treatments	Not clear
Chen et al [49]	BoardVitals, which is an online question bank accredited by the Accreditation Council for Continuing Medical Education	Neurology-based; Broad intrasubject: basic neuroscience; behavioral, cognitive, psychiatry; cerebrovascular; child neurology; congenital; cranial nerves; critical care; demyelinating disorders; epilepsy and seizures; ethics; genetic; headache; imaging or diagnostic studies; movement disorders; neuro-ophthalmology; neuro-otology; neuroinfectious disease; neurologic complications of systemic disease; neuromuscular; neurotoxicology, nutrition, metabolic; oncology; pain; pharmacology; pregnancy; sleep; and trauma	Not clear
Deiana et al [50]	Questions concerning vaccine myths and misconceptions by the World Health Organization	Narrow on vaccine myths and misconceptions	Not clear
Fuchs et al [51]	Digital platform self-assessment questions tailored for dental and medical students at the University of Bern's Institute for Medical Education	Broad with multiple-choice questions designed for dental students preparing for the Swiss Federal Licensing Examination in Dental Medicine, and allergists and immunologists preparing for the European Examination in Allergy and Clinical Immunology	Not clear
Ghosh & Bir [52]	Department question bank, which is a compilation of first and second semester questions from various medical universities across India	Biochemistry	Random without details of randomization
Giannos [53]	Specialty Certificate Examination Neurology Web Questions bank	Neurology and neuroscience	Not clear

XSL•FO
RenderX

Authors	Transparency	Range	Randomization
Gobira et al [54]	National Brazilian Examination for Revalidation of Medical Diplomas issued by Foreign Higher Education Institutions (Revalida)	Preventive medicine, gynecology and obstetrics, surgery, internal medicine, and pediatrics	Not clear
Grewal et al [55]	Not clear	Radiology	Not clear
Guerra et al [56]	Questions released by the Congress of Neurological Surgeons in the self-assessment neurosurgery exam	Neurosurgery across 7 subspecialties: tumor, cerebrovascular, trauma, spine, functional, pediatrics, and pain or nerve	Not clear
Hamed et al [57]	Guidelines adapted from Diabetes Canada Clinical Practice Guidelines Expert Committee, the Royal Australian College of General Practitioners, Australian Diabetes Society position statement, and the Joint British Diabetes Societies	Management of diabetic ketoacidosis	Not clear
Hoch et al [58]	Question database of an online learning platform funded by the German Society of Oto-Rhino-Laryngology, Head and Neck Surgery	Otolaryngology with a range of 15 distinct otolaryngology subspecialties, including allergology, audiology, ear, nose and throat, tumors, face and neck, inner ear and skull base, larynx, middle ear, oral cavity and pharynx, nose and sinuses, phoniatrics, salivary glands, sleep medicine, vestibular system, and legal aspects	Not clear
Juhi et al [59]	A list of drug-drug interactions from previously published research	Narrow on drug-drug interaction	Not clear
Kuang et al [60]	Not clear	Neurosurgery	Not clear
Kumari et al [61]	Designed by experts in hematology-related cases	Hematology with the following intrasubject aspects: case solving, laboratory calculations, disease interpretations, and other relevant aspects of hematology	Not clear
Kung et al [62]	Orthopedic In-Training Examination	Orthopedics	Not clear
Lai et al [63]	The United Kingdom Medical Licensing Assessment, which is a newly derived national undergraduate medical exit examination	Broad in medicine with the following aspects: acute and emergency, cancer, cardiovascular, child health, clinical hematology, ear, nose and throat, endocrine and metabolic, gastrointestinal including liver, general practice and primary health care, genetics and genomics, infection, medical ethics and law, medicine of older adults, mental health, musculoskeletal, neuroscience obstetrics and gynecology, ophthalmology, palliative and end of life care, perioperative medicine and anesthesia, renal and urology, respiratory, sexual health, social and population health, and surgery	Not clear
Lyu et al [64]	Chest computed tomography and brain magnetic resonance imaging screening reports collected from the Atrium Health Wake Forest Baptist clinical database	Radiology	Not clear
Moise et al [65]	Statements published in the latest American Academy of Otolaryngology–Head and Neck Surgery Foundation Clinical Practice Guideline: Tympanostomy Tubes in Children (Update)	Narrow involving otolaryngology	Not clear
Oca et al [66]	Not clear	Narrow involving only queries on accurate recommendation of close ophthalmologists	Not clear

XSL•FO
RenderX

Authors	Transparency	Range	Randomization
Oztermeli & Oztermeli [67]	Turkish medical specialty exam, prepared by the Student Selection and Placement Center	Broad: basic sciences including anatomy, physiology-histology-embryology, biochemistry, microbiology, pathology, and pharmacology; clinical including internal medicine, pediatrics, general surgery, obstetrics and gynecology, neurology, neurosurgery, psychiatry, public health, dermatology, radiology, nuclear medicine, otolaryngology, ophthalmology, orthopedics, physical medicine and rehabilitation, urology, pediatric surgery, cardiovascular surgery, thoracic surgery, plastic surgery, anesthesiology and reanimation, and emergency medicine	Not clear
Pugliese et al [68]	Expert selection of 15 questions commonly asked by patients with nonalcoholic fatty liver disease	Narrow involving nonalcoholic fatty liver disease aspects	Not clear
Sallam et al [69]	Panel discussion of experts in health care education	Broad on health care education, medical, dental, pharmacy, and public health	Not clear
Seth et al [70]	Devised by 3 fellows of the Royal Australasian College of Surgeons, with experience in performing rhinoplasty and expertise in facial reconstructive surgery	Narrow involving technical aspects of rhinoplasty	Not clear
Suthar et al [71]	Quizzes from the Case of the Month feature of the American Journal of Neuroradiology	Narrow involving radiology	Not clear
Walker et al [72]	Devised based on the Global Burden of Disease tool	Narrow involving benign and malignant hepatopancreaticobiliary-related conditions	Not clear
Wang et al [73]	Medical Exam Help. A total of 10 inpatient and 10 outpatient medical records to form a collection of Chinese medical records after desensitization	Clinical medicine, basic medicine, medical humanities, and relevant laws	Not clear
Wang et al [74]	The Taiwanese Senior Professional and Technical Examinations for Pharmacists downloaded from the Ministry of Examination website	Broad involving pharmacology and pharmaceutical chemistry, pharmaceutical analysis and pharmacognosy (including Chinese medicine), pharmaceutics and biopharmaceutics, dispensing pharmacy and clinical pharmacy, therapeutics, pharmacy administration, and pharmacy law	Not clear
Zhou et al [75]	A single clinical case from OrthoBullets, a global clinical collaboration platform for orthopedic surgeons; Written permission to use their clinical case report	Very narrow involving a single orthopedic case	Not clear

Examples of Optimal Reporting of Each Criterion Within the METRICS Checklist

The records with the highest scores for each METRICS item, as determined by the average subjective interrater assessments, are shown in Table 4.

XSL•FO
RenderX

Table 4. Included records that had the highest METRICS score per item.

Item	Issues considered in each item	Excellent or very good reporting examples
#1 Model	What is the model of the generative AI[a] tool used for generating content, and what are the exact settings for each tool?	Baglivo et al [47]: Bing, ChatGPT, Chatsonic, Bard, and YouChat, with full details of the mode and large language model, including plugins
#2 Evaluation	What is the exact approach used to evaluate the content generated by the generative AI-based model and is it an objective or subjective evaluation?	Al-Ashwal et al [42]: Objective via 2 different clinical reference tools; Alfertshofer et al [43]: Objective based on key answers, with the questions screened independently by 4 investigators; Ali et al [44]: Objective for multiple-choice questions and true or false questions, and subjective for short-answer and assay questions; Aljindan et al [45]: Objective based on key answers and historical performance metrics; and Baglivo et al [47]: Objective based on key answers and comparison with 5th year medical students' performance
#3a Timing	When is the generative AI model tested exactly and what are the duration and timing of testing?	Baglivo et al [47]; Biswas et al [48]; Fuchs et al [51]; Ghosh & Bir [52]; Hoch et al [58]; Juhi et al [59]; Kumari et al [61]; Kung et al [62]; Oca et al [66]; Pugliese et al [68]; Sallam et al [69]; Wang et al [73]; and Zhou et al [75]
#3b Transparency	How transparent are the data sources used to generate queries for the generative AI-based model?	Alfertshofer et al [43]
#4a Range	What is the range of topics tested and are they intersubject or intrasubject with variability in different subjects?	Ali et al [44]; Chen et al [49]; Hoch et al [58]; and Wang et al [73]
#4b Randomization	Was the process of selecting the topics to be tested on the generative AI-based model randomized?	Alfertshofer et al [43] and Aljindan et al [45]
#5 Individual	Is there any individual subjective involvement in generative AI content evaluation? If so, did the authors describe the details in full?	Ali et al [44] and Moise et al [65]
#6 Count	What is the count of queries executed (sample size)?	Alfertshofer et al [43]; Chen et al [49]; Guerra et al [56]; Hoch et al [58]; and Oztermeli & Oztermeli [67]
#7 Specificity of the prompt or language	How specific are the exact prompts used? Were those exact prompts provided fully? Did the authors consider the feedback and learning loops? How specific are the language and cultural issues considered in the generative AI model?	Alfertshofer et al [43]; Biswas et al [48]; Fuchs et al [51]; Grewal et al [55]; Wang et al [73]; Moise et al [65]; and Pugliese et al [68]

[a]AI: artificial intelligence.

Interrater Assessment of the Included Records Based on METRICS Scores

The overall mean METRICS score was 3.0 (SD 0.58). For each item, the κ interrater reliability ranged from 0.558 to 0.962 (P<.001 for the 9 tested items), indicating good to excellent agreement (Table 5).

With classification per item, the highest average METRICS score was recorded for the "Model" item, followed by the "Specificity" item, while the lowest scores were recorded for the "Randomization" item (classified as suboptimal) and "Individual factors" item (classified as satisfactory) (Table 5).

XSL•FO
RenderX

Table 5. Interrater reliability per METRICS item.

METRICS[a] item	Score		Quality	Cohen κ	Asymptotic standard error	Approximate T	*P* value
	Mean[b] (SD)	Range					
Model	3.72 (0.58)	2.5-5.0	Very good	0.820	0.090	6.044	<.001
Timing	2.90 (1.93)	1.0-5.0	Good	0.853	0.076	6.565	<.001
Count	3.04 (1.32)	1.0-5.0	Good	0.962	0.037	10.675	<.001
Specificity of prompts and language	3.44 (1.25)	1.0-5.0	Very good	0.765	0.086	8.083	<.001
Evaluation	3.31 (1.16)	1.0-5.0	Good	0.885	0.063	9.668	<.001
Individual factors	2.50 (1.42)	1.0-5.0	Satisfactory	0.865	0.087	6.860	<.001
Transparency	3.24 (1.01)	1.0-5.0	Good	0.558	0.112	5.375	<.001
Range	3.24 (1.07)	2.0-5.0	Good	0.836	0.076	8.102	<.001
Randomization	1.31 (0.87)	1.0-4.0	Suboptimal	0.728	0.135	5.987	<.001
Overall	3.01 (0.58)	1.5-4.1	Good	0.381	0.086	10.093	<.001

[a]METRICS: Model, Evaluation, Timing, Range/Randomization, Individual factors, Count, and Specificity of prompts and language.

[b]The mean scores represent the results of evaluating the included studies averaged for the 2 rater scores.

Discussion

Principal Findings

The interpretation and synthesis of credible scientific evidence based on studies evaluating commonly used generative AI-based conversational models (eg, ChatGPT, Bing, and Bard) can be challenging. This is related to the discernible variability in the methods used for the evaluation of such models, as well as the varying styles of reporting. Such variability is fathomable considering the emerging nature of this research field with less than a year of reporting at the time of writing. Therefore, a standardized framework to guide the design of such studies and to delineate the best reporting practices can be beneficial, since rigorous methodology and clear reporting of findings are key attributes of science to reach reliable conclusions with real-world implications.

In this study, a preliminary checklist referred to as "METRICS" was formulated, which can help researchers aspiring to test the performance of generative AI-based models in various aspects of health care education and practice. It is crucial to explicitly state that the proposed METRICS checklist in this study cannot be claimed to be comprehensive or flawless. Nevertheless, this checklist could provide a solid base for future studies and much needed efforts aiming to standardize the reporting of AI-based studies in health care.

The principal finding of this study was the establishment of 9 key themes that are recommended to be considered in the design, testing, and reporting of generative AI-based models in research, particularly in the health care domain. These features included the model of the generative AI tool, evaluation approach, timing of testing and transparency, range of topics tested and randomization of queries, individual factors in the design and assessment, count of queries, and specific prompts and languages used. The relevance of these themes in the design and reporting of generative AI-based model content testing can be illustrated as follows.

First, the variability in generative AI model types used to conduct queries and the variability in settings pose significant challenges for cross-study comparisons. The significant impact of generative AI models on the resultant output is related to the distinct architectures and capabilities of various generative AI models, with expected variability in the performance and quality of the content generated by generative AI [76]. Additionally, various options to configure the models further affect the content generated by generative AI. Consequently, it is important to consider these variations when evaluating research using different generative AI models [77-81]. These issues can be illustrated clearly in the records included in this study that conducted contemporary analysis of at least two models. For example, Al-Ashwal et al [42] showed that Bing had the highest accuracy and specificity for predicting drug-drug interaction, outperforming Bard, ChatGPT-3.5, and ChatGPT-4. Moreover, Baglivo et al [47] showed not only intermodel variability but also intramodel variability in performance in the domain of public health. Additionally, in the context of providing information on rhinoplasty, Seth et al [70] showed that this intermodel variability in performance was the most comprehensible with the content of Bard, followed by ChatGPT and Bing.

Second, the continuous updating of generative AI models introduces significant temporal variability, which would influence the comparability of studies conducted at different times. The updates of generative AI models enhance their capabilities and performance [82]. Consequently, this temporal variability can lead to inconsistencies in synthesizing evidence, as the same model may demonstrate different outputs over time. Therefore, when analyzing or comparing studies involving AI models, it is crucial to consider the specific version and state of the model at the time of each study to accurately interpret and compare results. In this context, it is important to conduct

XSL•FO
RenderX

future longitudinal studies to discern the exact effect of changes in performance of commonly used generative AI models over time.

Third, the count of queries in the evaluation of generative AI models was identified among the pertinent themes of assessment. This appears understandable since studies employing a larger number of queries can provide a more comprehensive evaluation of the tested model. An extensive number of queries can reveal minor weaknesses, despite the difficulty to establish what constitutes an "extensive" number of queries of the minimum number of queries needed to reveal the real performance of a generative AI model in a particular topic. In this study, the number of queries varied from a single case to more than 2500 questions, showing the need for standardization and establishing a clear guide on the number of queries deemed suitable [58,75].

Fourth, a key theme identified in this study was the nature and language of the prompts used to conduct the studies. This involved the imperative of explicitly stating the used prompts and the language in which they were framed. The exact prompting approach and the presence of cultural and linguistic biases appear to be critical factors that can influence the quality of content generated by generative AI-based models [83]. Slight differences in wording or context in the prompt used to generate the generative AI content can lead to recognizable differences in the content generated [36,84,85]. Additionally, feedback mechanisms and learning loops that allow generative AI-based models to learn from interactions can change the model performance for the same query, which might not be consistently accounted for in all studies. These minor variations in prompts across different studies can also complicate the synthesis of evidence, highlighting the need for standardizing such an aspect. Additionally, as highlighted above, generative AI-based models may exhibit biases based on their training data, affecting performance across various cultural or linguistic contexts [86-88]. Thus, studies conducted in different regions or involving various languages might yield varying results. In this study, we found that a majority of the included records tested generative AI-based models using the English language, highlighting the need for more studies on other languages to reveal the possible variability in performance based on language. Comparative studies involving multiple languages can reveal such inconsistencies, for example, the study by Wang et al [74]. In the aforementioned study assessing ChatGPT performance in the Taiwanese pharmacist licensing exam, the performance in the English test was better than that in the Chinese test across all tested subjects [74]. Another comprehensive study by Alfertshofer et al [43] that assessed the performance of ChatGPT on 6 different national medical licensing exams highlighted the variability in performance per country and language. Based on the previous points, more studies that encompass diverse language and cultural perspectives are essential to assess and address possible cultural and language biases in generative AI-based models. Additionally, the design of generative AI-based models trained on a more diverse set of languages and cultural contexts is important to ensure that the training data sets are representative of different linguistic groups, which is of paramount importance in health care. Furthermore, cross-cultural validation studies appear valuable for understanding the performance of generative AI-based models in various language and cultural settings. These approaches could enhance the broad applicability of generative AI-based models in health care to ensure the fair distribution of the benefits of generative AI technologies.

Fifth, an important theme highlighted in this study was the approach of evaluating the content generated by generative AI-based models. Variable methods of assessment can introduce a discernible methodological variability. Specifically, the use of objective assessment ensures consistency in assessment. On the other hand, subjective assessment, even by experts, can vary despite the professional judgment and deep understanding provided by such an expert opinion [89].

Similarly, the number and expertise of evaluators or raters involved in constructing and evaluating the generative AI-based studies were identified as a pertinent theme in this study [90,91]. Variations in rater numbers across studies can lead to inconsistencies in synthesized evidence [68,69,92]. Additionally, the method used to establish agreement (eg, kappa statistics and consensus meetings) might differ in various studies, affecting the comparability of results.

Finally, data-pertinent issues were identified as key themes in this study. These involved the need for full transparency regarding the sources of data used to create the queries (eg, question banks, credible national and international guidelines, clinical reports, etc) [93,94]. Additionally, ethical considerations, such as consent to use copyrighted material and consent or anonymization of the clinical data, should be carefully stated in the evaluation studies of AI-based models. An important aspect that appeared suboptimal in the majority of included records was randomization to reduce or eliminate potential bias in query selection. Thus, this important issue should be addressed in future studies to allow unbiased evaluation of the content of generative AI-based models. Another important aspect is the need to carefully select the topics to be tested, which can belong to a broad domain (eg, medical examination) or a narrow domain (eg, a particular specialty) [95-97]. A comprehensive description of topics is essential to reveal subtle differences in generative AI performance across various domains. Biased query coverage per topic may result in unreliable conclusions regarding generative AI model performance.

The value of the METRICS checklist in guiding researchers toward more rigorous design and transparent reporting of health care studies assessing the performance of generative AI-based models can be highlighted as follows. Two studies exemplified the practical utility of the METRICS checklist (presented in its preprint form) across different research scenarios (health care education and health care practice) [98]. The first study conducted a detailed assessment of ChatGPT-3.5 performance in medical microbiology multiple choice questions compared with dental students [99]. By applying the METRICS checklist retrospectively, the study quality was delineated, including the identification of potential limitations, such as the absence of randomization, thus offering a more critical evaluation of the research design [99].

XSL•FO
RenderX

The second study investigating the performance of both ChatGPT-3.5 and ChatGPT-4 in the context of diagnostic microbiology case scenarios was conceived based on the METRICS checklist [100]. The prospective incorporation of the METRICS checklist was particularly instrumental in refining the study design and the reporting of results [100].

Thus, the aforementioned studies emphasize the effectiveness of the METRICS checklist as a tool to standardize research practice in a rapidly growing research field. The real-world application of the METRICS checklist has been valuable in identifying potential research limitations and in enhancing the overall structure and clarity of the reporting of results. These studies also demonstrate the value of the METRICS checklist for guiding researchers toward more rigorous design and transparent reporting of generative AI-based studies in health care [99,100].

Limitations

It is crucial to explicitly mention the need for careful interpretation of the findings based on the following limitations. First, the initial search process involved the broad term "artificial intelligence" and was conducted by a single author, which may have inadvertently resulted in missing relevant references. The record selection process was further limited by the screening of record titles for only 3 generative AI models (ChatGPT, Bing, and Bard) using the EndNote search function. Additionally, the reliance on including published English records, indexed in Scopus, PubMed, or Google Scholar, could raise concerns about potential selection bias and the exclusion of relevant studies. However, it is important to consider this limitation in light of the context of our study, which represents a preliminary report that needs to be validated by future comprehensive and exhaustive studies. Second, it is important to acknowledge that a few pertinent themes could have been overlooked despite our attempt to achieve a thorough analysis, given the limited number of authors. Therefore, future follow-up studies can benefit from inclusion of authors with diverse backgrounds, including different health care disciplines, computer scientists, researchers in the field of human-AI interaction, and AI developers. Additionally, the subjective nature of the pertinent theme selection can be considered as another important caveat in this study. This shortcoming extended to involve the raters' subjective assessments in assigning different METRICS scores. Moreover, the equal weight given to each item of the checklist in the METRICS score might not be a suitable approach, given the possibility of varying importance of each component. Thus, future comprehensive studies should focus on the relative importance of each METRICS component and its possible impact on the reporting of generative AI model performance. Third, the focus on a few specific generative AI-based conversational models (ie, ChatGPT, Bing, and Bard) can potentially overlook the nuanced aspects of other generative AI models. Nevertheless, our approach was justified by the popularity and widespread use of these particular generative AI-based models. However, it is important for future studies to expand the scope to include models, such as Llama or Claude, which could provide a more comprehensive evaluation of the utility of the METRICS checklist. Lastly, we fully and unequivocally acknowledge that the METRICS checklist is preliminary and needs further verification to ensure its valid applicability.

Future Perspectives

The METRICS checklist proposed in this study could be a helpful step toward establishing useful guidelines to design and report the findings of generative AI-based studies. The integration of generative AI models in health care education and practice necessitates a collaborative approach involving health care professionals, researchers, and AI developers. Synthesis of evidence with critical appraisal of the quality of each element in the METRICS checklist is recommended for continuous enhancement of AI output, which would result in successful implementation of AI models in health care while avoiding possible concerns. Regular multidisciplinary efforts and iterative revisions are recommended to ensure that the METRICS checklist properly reflects its original intended purpose of improving the quality of study design and result reporting in this swiftly evolving research field. Future studies would benefit from expanding the scope of literature review and data inclusion, with the incorporation of a wider range of databases, languages, and AI models. This is crucial for reaching the ultimate aim of standardization of the design and reporting of generative AI-based studies.

Conclusions

The newly devised "METRICS" checklist may represent a key initial step to motivate the standardization of reporting of generative AI-based studies in health care education and practice. Additionally, the establishment of this algorithm can motivate collaborative efforts to develop universally accepted reporting guidelines for the design and reporting of results of generative AI-based studies. In turn, this can enhance the comparability and reliability of evidence synthesis from these studies. The METRICS checklist, as presented by the findings of this study, can help to elucidate the strengths and limitations of the content generated by generative AI-based models, guiding their future development and application. The standardization offered by the METRICS checklist can be important to ensure the reporting of reliable and replicable results. Subsequently, this can result in the exploitation of the promising potential of generative AI-based models in health care while avoiding its possible concerns. The METRICS checklist could mark the significant progress in the evolving research field. Nevertheless, there is room for refinement through revisions and updates to verify its validity.

Data Availability

The data used in this study are available upon request from the corresponding author (Malik Sallam).

XSL•FO
RenderX

Authors' Contributions

Malik Sallam contributed to conceptualization; Malik Sallam, MB, and Mohammed Sallam contributed to methodology; Malik Sallam, MB, and Mohammed Sallam contributed to formal analysis; Malik Sallam, MB, and Mohammed Sallam performed the investigation; Malik Sallam, MB, and Mohammed Sallam contributed to data curation; Malik Sallam contributed to writing—original draft preparation; Malik Sallam, MB, and Mohammed Sallam contributed to writing—review and editing; Malik Sallam contributed to visualization; Malik Sallam contributed to supervision; and Malik Sallam contributed to project administration. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel) 2023 Mar 19;11(6):887 [FREE Full text] [doi: 10.3390/healthcare11060887] [Medline: 36981544]
2. Garg RK, Urs VL, Agarwal AA, Chaudhary SK, Paliwal V, Kar SK. Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. Health Promot Perspect 2023;13(3):183-191 [FREE Full text] [doi: 10.34172/hpp.2023.22] [Medline: 37808939]
3. Alam F, Lim MA, Zulkipli IN. Integrating AI in medical education: embracing ethical usage and critical understanding. Front Med (Lausanne) 2023;10:1279707 [FREE Full text] [doi: 10.3389/fmed.2023.1279707] [Medline: 37901398]
4. Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: A taxonomy and systematic review. Comput Methods Programs Biomed 2024 Jan 15;245:108013 [FREE Full text] [doi: 10.1016/j.cmpb.2024.108013] [Medline: 38262126]
5. Mesko B. The ChatGPT (Generative Artificial Intelligence) Revolution Has Made Artificial Intelligence Approachable for Medical Professionals. J Med Internet Res 2023 Jun 22;25:e48392 [FREE Full text] [doi: 10.2196/48392] [Medline: 37347508]
6. Rudolph J, Tan S, Tan S. War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. JALT 2023 Apr 25;6(1):771. [doi: 10.37074/jalt.2023.6.1.23]
7. Choudhury A, Shamszare H. Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. J Med Internet Res 2023 Jun 14;25:e47184 [FREE Full text] [doi: 10.2196/47184] [Medline: 37314848]
8. Shahsavar Y, Choudhury A. User Intentions to Use ChatGPT for Self-Diagnosis and Health-Related Purposes: Cross-sectional Survey Study. JMIR Hum Factors 2023 May 17;10:e47564 [FREE Full text] [doi: 10.2196/47564] [Medline: 37195756]
9. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194 [FREE Full text] [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
10. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad AK, et al. Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study. J Med Internet Res 2023 Aug 22;25:e48659 [FREE Full text] [doi: 10.2196/48659] [Medline: 37606976]
11. Giannos P, Delardas O. Performance of ChatGPT on UK Standardized Admission Tests: Insights From the BMAT, TMUA, LNAT, and TSA Examinations. JMIR Med Educ 2023 Apr 26;9:e47737 [FREE Full text] [doi: 10.2196/47737] [Medline: 37099373]
12. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689 [FREE Full text] [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]
13. Miao H, Li C, Wang J. A Future of Smarter Digital Health Empowered by Generative Pretrained Transformer. J Med Internet Res 2023 Sep 26;25:e49963 [FREE Full text] [doi: 10.2196/49963] [Medline: 37751243]
14. Liu T, Xiao X. A Framework of AI-Based Approaches to Improving eHealth Literacy and Combating Infodemic. Front Public Health 2021;9:755808 [FREE Full text] [doi: 10.3389/fpubh.2021.755808] [Medline: 34917575]
15. Choudhury A, Elkefi S, Tounsi A. Exploring factors influencing user perspective of ChatGPT as a technology that assists in healthcare decision making: A cross sectional survey study. medRxiv. 2023. URL: https://www.medrxiv.org/content/10.1101/2023.12.07.23299685v1.full [accessed 2024-02-02]
16. Hsu H, Hsu K, Hou S, Wu C, Hsieh Y, Cheng Y. Examining Real-World Medication Consultations and Drug-Herb Interactions: ChatGPT Performance Evaluation. JMIR Med Educ 2023 Aug 21;9:e48433 [FREE Full text] [doi: 10.2196/48433] [Medline: 37561097]
17. Chen Z. Ethics and discrimination in artificial intelligence-enabled recruitment practices. Humanit Soc Sci Commun 2023 Sep 13;10(1):567. [doi: 10.1057/s41599-023-02079-x]
18. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical Considerations of Using ChatGPT in Health Care. J Med Internet Res 2023 Aug 11;25:e48009 [FREE Full text] [doi: 10.2196/48009] [Medline: 37566454]
19. Emsley R. ChatGPT: these are not hallucinations - they're fabrications and falsifications. Schizophrenia (Heidelb) 2023 Aug 19;9(1):52 [FREE Full text] [doi: 10.1038/s41537-023-00379-4] [Medline: 37598184]

XSL•FO
RenderX

20. Gödde D, Nöhl S, Wolf C, Rupert Y, Rimkus L, Ehlers J, et al. A SWOT (Strengths, Weaknesses, Opportunities, and Threats) Analysis of ChatGPT in the Medical Literature: Concise Review. J Med Internet Res 2023 Nov 16;25:e49368 [FREE Full text] [doi: 10.2196/49368] [Medline: 37865883]
21. Kleesiek J, Wu Y, Stiglic G, Egger J, Bian J. An Opinion on ChatGPT in Health Care-Written by Humans Only. J Nucl Med 2023 May 13;64(5):701-703. [doi: 10.2967/jnumed.123.265687] [Medline: 37055219]
22. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Med Ethics 2021 Sep 15;22(1):122 [FREE Full text] [doi: 10.1186/s12910-021-00687-3] [Medline: 34525993]
23. Mijwil M, Aljanabi M, Ali AH. ChatGPT: Exploring the Role of Cybersecurity in the Protection of Medical Information. MJCS 2023 Feb 1:18-21. [doi: 10.58496/mjcs/2023/004]
24. Sun L, Yin C, Xu Q, Zhao W. Artificial intelligence for healthcare and medical education: a systematic review. Am J Transl Res 2023;15(7):4820-4828 [FREE Full text] [Medline: 37560249]
25. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. JMIR Med Educ 2023 Feb 08;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
26. Sallam M, Salim NA, Al-Tammemi AB, Barakat M, Fayyad D, Hallit S, et al. ChatGPT Output Regarding Compulsory Vaccination and COVID-19 Vaccine Conspiracy: A Descriptive Study at the Outset of a Paradigm Shift in Online Search for Information. Cureus 2023 Feb;15(2):e35029 [FREE Full text] [doi: 10.7759/cureus.35029] [Medline: 36819954]
27. Scherr R, Halaseh FF, Spina A, Andalib S, Rivera R. ChatGPT Interactive Medical Simulations for Early Clinical Education: Case Study. JMIR Med Educ 2023 Nov 10;9:e49877 [FREE Full text] [doi: 10.2196/49877] [Medline: 37948112]
28. Doshi R, Amin K, Khosla P, Bajaj S, Chheang S, Forman H. Utilizing Large Language Models to Simplify Radiology Reports: a comparative analysis of ChatGPT3.5, ChatGPT4.0, Google Bard, and Microsoft Bing. medRxiv. 2023. URL: https://www.medrxiv.org/content/10.1101/2023.06.04.23290786v2 [accessed 2024-02-02]
29. Ray PP. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems 2023;3:121-154. [doi: 10.1016/j.iotcps.2023.04.003]
30. Qarajeh A, Tangpanithandee S, Thongprayoon C, Suppadungsuk S, Krisanapan P, Aiumtrakul N, et al. AI-Powered Renal Diet Support: Performance of ChatGPT, Bard AI, and Bing Chat. Clin Pract 2023 Sep 26;13(5):1160-1172 [FREE Full text] [doi: 10.3390/clinpract13050104] [Medline: 37887080]
31. Zúñiga Salazar G, Zúñiga D, Vindel CL, Yoong AM, Hincapie S, Zúñiga A, et al. Efficacy of AI Chats to Determine an Emergency: A Comparison Between OpenAI's ChatGPT, Google Bard, and Microsoft Bing AI Chat. Cureus 2023 Sep;15(9):e45473 [FREE Full text] [doi: 10.7759/cureus.45473] [Medline: 37727841]
32. Fijačko N, Prosen G, Abella BS, Metličar, Štiglic G. Can novel multimodal chatbots such as Bing Chat Enterprise, ChatGPT-4 Pro, and Google Bard correctly interpret electrocardiogram images? Resuscitation 2023 Dec 24;193:110009. [doi: 10.1016/j.resuscitation.2023.110009] [Medline: 37884222]
33. Aiumtrakul N, Thongprayoon C, Suppadungsuk S, Krisanapan P, Miao J, Qureshi F, et al. Navigating the Landscape of Personalized Medicine: The Relevance of ChatGPT, BingChat, and Bard AI in Nephrology Literature Searches. J Pers Med 2023 Sep 30;13(10):1457 [FREE Full text] [doi: 10.3390/jpm13101457] [Medline: 37888068]
34. Stephens LD, Jacobs JW, Adkins BD, Booth GS. Battle of the (Chat)Bots: Comparing Large Language Models to Practice Guidelines for Transfusion-Associated Graft-Versus-Host Disease Prevention. Transfus Med Rev 2023 Jul;37(3):150753 [FREE Full text] [doi: 10.1016/j.tmrv.2023.150753] [Medline: 37704461]
35. Eysenbach G. The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers. JMIR Med Educ 2023 Mar 06;9:e46885 [FREE Full text] [doi: 10.2196/46885] [Medline: 36863937]
36. Meskó B. Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. J Med Internet Res 2023 Oct 04;25:e50638 [FREE Full text] [doi: 10.2196/50638] [Medline: 37792434]
37. Hristidis V, Ruggiano N, Brown EL, Ganta SRR, Stewart S. ChatGPT vs Google for Queries Related to Dementia and Other Cognitive Decline: Comparison of Results. J Med Internet Res 2023 Jul 25;25:e48966 [FREE Full text] [doi: 10.2196/48966] [Medline: 37490317]
38. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
39. Harzing AW. The Publish or Perish Book: Your guide to effective and responsible citation analysis. Melbourne, Australia: Tarma Software Research; 2010.
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche P, Vandenbroucke JP, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808 [FREE Full text] [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
41. CASP Qualitative Studies Checklist. Critical Appraisal Skills Programme. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2023-11-10]

XSL•FO
RenderX

42. Al-Ashwal FY, Zawiah M, Gharaibeh L, Abu-Farha R, Bitar AN. Evaluating the Sensitivity, Specificity, and Accuracy of ChatGPT-3.5, ChatGPT-4, Bing AI, and Bard Against Conventional Drug-Drug Interactions Clinical Tools. Drug Healthc Patient Saf 2023;15:137-147 [FREE Full text] [doi: 10.2147/DHPS.S425858] [Medline: 37750052]
43. Alfertshofer M, Hoch CC, Funk PF, Hollmann K, Wollenberg B, Knoedler S, et al. Sailing the Seven Seas: A Multinational Comparison of ChatGPT's Performance on Medical Licensing Examinations. Ann Biomed Eng 2023 Aug 08:03338-3. [doi: 10.1007/s10439-023-03338-3] [Medline: 37553555]
44. Ali K, Barhom N, Tamimi F, Duggal M. ChatGPT-A double-edged sword for healthcare education? Implications for assessments of dental students. Eur J Dent Educ 2024 Feb 07;28(1):206-211. [doi: 10.1111/eje.12937] [Medline: 37550893]
45. Aljindan FK, Al Qurashi AA, Albalawi IAS, Alanazi AMM, Aljuhani HAM, Falah Almutairi F, et al. ChatGPT Conquers the Saudi Medical Licensing Exam: Exploring the Accuracy of Artificial Intelligence in Medical Knowledge Assessment and Implications for Modern Medical Education. Cureus 2023 Sep;15(9):e45043 [FREE Full text] [doi: 10.7759/cureus.45043] [Medline: 37829968]
46. Altamimi I, Altamimi A, Alhumimidi AS, Altamimi A, Temsah M. Snakebite Advice and Counseling From Artificial Intelligence: An Acute Venomous Snakebite Consultation With ChatGPT. Cureus 2023 Jun;15(6):e40351 [FREE Full text] [doi: 10.7759/cureus.40351] [Medline: 37456381]
47. Baglivo F, De Angelis L, Casigliani V, Arzilli G, Privitera GP, Rizzo C. Exploring the Possible Use of AI Chatbots in Public Health Education: Feasibility Study. JMIR Med Educ 2023 Nov 01;9:e51421 [FREE Full text] [doi: 10.2196/51421] [Medline: 37910155]
48. Biswas S, Logan NS, Davies LN, Sheppard AL, Wolffsohn JS. Assessing the utility of ChatGPT as an artificial intelligence-based large language model for information to answer questions on myopia. Ophthalmic Physiol Opt 2023 Nov 21;43(6):1562-1570. [doi: 10.1111/opo.13207] [Medline: 37476960]
49. Chen TC, Multala E, Kearns P, Delashaw J, Dumont A, Maraganore D, et al. Assessment of ChatGPT's performance on neurology written board examination questions. BMJ Neurol Open 2023;5(2):e000530 [FREE Full text] [doi: 10.1136/bmjno-2023-000530] [Medline: 37936648]
50. Deiana G, Dettori M, Arghittu A, Azara A, Gabutti G, Castiglia P. Artificial Intelligence and Public Health: Evaluating ChatGPT Responses to Vaccination Myths and Misconceptions. Vaccines (Basel) 2023 Jul 07;11(7):1217 [FREE Full text] [doi: 10.3390/vaccines11071217] [Medline: 37515033]
51. Fuchs A, Trachsel T, Weiger R, Eggmann F. ChatGPT's performance in dentistry and allergy-immunology assessments: a comparative study. Swiss Dent J 2023 Oct 06;134(5):1-25 [FREE Full text] [Medline: 37799027]
52. Ghosh A, Bir A. Evaluating ChatGPT's Ability to Solve Higher-Order Questions on the Competency-Based Medical Education Curriculum in Medical Biochemistry. Cureus 2023 Apr;15(4):e37023 [FREE Full text] [doi: 10.7759/cureus.37023] [Medline: 37143631]
53. Giannos P. Evaluating the limits of AI in medical specialisation: ChatGPT's performance on the UK Neurology Specialty Certificate Examination. BMJ Neurol Open 2023;5(1):e000451 [FREE Full text] [doi: 10.1136/bmjno-2023-000451] [Medline: 37337531]
54. Gobira M, Nakayama LF, Moreira R, Andrade E, Regatieri CVS, Belfort R. Performance of ChatGPT-4 in answering questions from the Brazilian National Examination for Medical Degree Revalidation. Rev Assoc Med Bras (1992) 2023;69(10):e20230848 [FREE Full text] [doi: 10.1590/1806-9282.20230848] [Medline: 37792871]
55. Grewal H, Dhillon G, Monga V, Sharma P, Buddhavarapu VS, Sidhu G, et al. Radiology Gets Chatty: The ChatGPT Saga Unfolds. Cureus 2023 Jun;15(6):e40135 [FREE Full text] [doi: 10.7759/cureus.40135] [Medline: 37425598]
56. Guerra GA, Hofmann H, Sobhani S, Hofmann G, Gomez D, Soroudi D, et al. GPT-4 Artificial Intelligence Model Outperforms ChatGPT, Medical Students, and Neurosurgery Residents on Neurosurgery Written Board-Like Questions. World Neurosurg 2023 Nov;179:e160-e165. [doi: 10.1016/j.wneu.2023.08.042] [Medline: 37597659]
57. Hamed E, Eid A, Alberry M. Exploring ChatGPT's Potential in Facilitating Adaptation of Clinical Guidelines: A Case Study of Diabetic Ketoacidosis Guidelines. Cureus 2023 May;15(5):e38784 [FREE Full text] [doi: 10.7759/cureus.38784] [Medline: 37303347]
58. Hoch CC, Wollenberg B, Lüers J, Knoedler S, Knoedler L, Frank K, et al. ChatGPT's quiz skills in different otolaryngology subspecialties: an analysis of 2576 single-choice and multiple-choice board certification preparation questions. Eur Arch Otorhinolaryngol 2023 Sep;280(9):4271-4278 [FREE Full text] [doi: 10.1007/s00405-023-08051-4] [Medline: 37285018]
59. Juhi A, Pipil N, Santra S, Mondal S, Behera JK, Mondal H. The Capability of ChatGPT in Predicting and Explaining Common Drug-Drug Interactions. Cureus 2023 Mar;15(3):e36272 [FREE Full text] [doi: 10.7759/cureus.36272] [Medline: 37073184]
60. Kuang Y, Zou M, Niu H, Zheng B, Zhang T, Zheng B. ChatGPT encounters multiple opportunities and challenges in neurosurgery. Int J Surg 2023 Oct 01;109(10):2886-2891 [FREE Full text] [doi: 10.1097/JS9.0000000000000571] [Medline: 37352529]
61. Kumari A, Kumari A, Singh A, Singh SK, Juhi A, Dhanvijay AKD, et al. Large Language Models in Hematology Case Solving: A Comparative Study of ChatGPT-3.5, Google Bard, and Microsoft Bing. Cureus 2023 Aug;15(8):e43861 [FREE Full text] [doi: 10.7759/cureus.43861] [Medline: 37736448]

XSL•FO
RenderX

62. Kung JE, Marshall C, Gauthier C, Gonzalez TA, Jackson JB. Evaluating ChatGPT Performance on the Orthopaedic In-Training Examination. JB JS Open Access 2023;8(3):e23.00056 [FREE Full text] [doi: 10.2106/JBJS.OA.23.00056] [Medline: 37693092]
63. Lai UH, Wu KS, Hsu T, Kan JKC. Evaluating the performance of ChatGPT-4 on the United Kingdom Medical Licensing Assessment. Front Med (Lausanne) 2023 Sep 19;10:1240915 [FREE Full text] [doi: 10.3389/fmed.2023.1240915] [Medline: 37795422]
64. Lyu Q, Tan J, Zapadka ME, Ponnatapura J, Niu C, Myers KJ, et al. Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: results, limitations, and potential. Vis Comput Ind Biomed Art 2023 May 18;6(1):9 [FREE Full text] [doi: 10.1186/s42492-023-00136-5] [Medline: 37198498]
65. Moise A, Centomo-Bozzo A, Orishchak O, Alnoury MK, Daniel SJ. Can ChatGPT Guide Parents on Tympanostomy Tube Insertion? Children (Basel) 2023 Sep 30;10(10):1634 [FREE Full text] [doi: 10.3390/children10101634] [Medline: 37892297]
66. Oca MC, Meller L, Wilson K, Parikh AO, McCoy A, Chang J, et al. Bias and Inaccuracy in AI Chatbot Ophthalmologist Recommendations. Cureus 2023 Sep;15(9):e45911 [FREE Full text] [doi: 10.7759/cureus.45911] [Medline: 37885556]
67. Oztermeli AD, Oztermeli A. ChatGPT performance in the medical specialty exam: An observational study. Medicine (Baltimore) 2023 Aug 11;102(32):e34673 [FREE Full text] [doi: 10.1097/MD.0000000000034673] [Medline: 37565917]
68. Pugliese N, Wai-Sun Wong V, Schattenberg JM, Romero-Gomez M, Sebastiani G, NAFLD Expert Chatbot Working Group, et al. Accuracy, Reliability, and Comprehensibility of ChatGPT-Generated Medical Responses for Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2023 Sep 15 (forthcoming) [FREE Full text] [doi: 10.1016/j.cgh.2023.08.033] [Medline: 37716618]
69. Sallam M, Salim N, Barakat M, Al-Tammemi A. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. Narra J 2023 Mar 29;3(1):e103. [doi: 10.52225/narra.v3i1.103]
70. Seth I, Lim B, Xie Y, Cevik J, Rozen WM, Ross RJ, et al. Comparing the Efficacy of Large Language Models ChatGPT, BARD, and Bing AI in Providing Information on Rhinoplasty: An Observational Study. Aesthet Surg J Open Forum 2023;5:ojad084 [FREE Full text] [doi: 10.1093/asjof/ojad084] [Medline: 37795257]
71. Suthar PP, Kounsal A, Chhetri L, Saini D, Dua SG. Artificial Intelligence (AI) in Radiology: A Deep Dive Into ChatGPT 4.0's Accuracy with the American Journal of Neuroradiology's (AJNR) "Case of the Month". Cureus 2023 Aug;15(8):e43958 [FREE Full text] [doi: 10.7759/cureus.43958] [Medline: 37746411]
72. Walker HL, Ghani S, Kuemmerli C, Nebiker CA, Müller B, Raptis DA, et al. Reliability of Medical Information Provided by ChatGPT: Assessment Against Clinical Guidelines and Patient Information Quality Instrument. J Med Internet Res 2023 Jun 30;25:e47479 [FREE Full text] [doi: 10.2196/47479] [Medline: 37389908]
73. Wang H, Wu W, Dou Z, He L, Yang L. Performance and exploration of ChatGPT in medical examination, records and education in Chinese: Pave the way for medical AI. Int J Med Inform 2023 Sep;177:105173. [doi: 10.1016/j.ijmedinf.2023.105173] [Medline: 37549499]
74. Wang Y, Shen H, Chen T. Performance of ChatGPT on the pharmacist licensing examination in Taiwan. J Chin Med Assoc 2023 Jul 01;86(7):653-658. [doi: 10.1097/JCMA.0000000000000942] [Medline: 37227901]
75. Zhou Y, Moon C, Szatkowski J, Moore D, Stevens J. Evaluating ChatGPT responses in the context of a 53-year-old male with a femoral neck fracture: a qualitative analysis. Eur J Orthop Surg Traumatol 2023 Sep 30 (forthcoming). [doi: 10.1007/s00590-023-03742-4] [Medline: 37776392]
76. Malhotra R, Singh P. Recent advances in deep learning models: a systematic literature review. Multimed Tools Appl 2023 Apr 25;82(29):44977-45060. [doi: 10.1007/s11042-023-15295-z]
77. Hirosawa T, Kawamura R, Harada Y, Mizuta K, Tokumasu K, Kaji Y, et al. ChatGPT-Generated Differential Diagnosis Lists for Complex Case-Derived Clinical Vignettes: Diagnostic Accuracy Evaluation. JMIR Med Inform 2023 Oct 09;11:e48808 [FREE Full text] [doi: 10.2196/48808] [Medline: 37812468]
78. Levkovich I, Elyoseph Z. Suicide Risk Assessments Through the Eyes of ChatGPT-3.5 Versus ChatGPT-4: Vignette Study. JMIR Ment Health 2023 Sep 20;10:e51232 [FREE Full text] [doi: 10.2196/51232] [Medline: 37728984]
79. Flores-Cohaila JA, García-Vicente A, Vizcarra-Jiménez S, De la Cruz-Galán J, Gutiérrez-Arratia J, Quiroga Torres BG, et al. Performance of ChatGPT on the Peruvian National Licensing Medical Examination: Cross-Sectional Study. JMIR Med Educ 2023 Sep 28;9:e48039 [FREE Full text] [doi: 10.2196/48039] [Medline: 37768724]
80. Fraser H, Crossland D, Bacher I, Ranney M, Madsen T, Hilliard R. Comparison of Diagnostic and Triage Accuracy of Ada Health and WebMD Symptom Checkers, ChatGPT, and Physicians for Patients in an Emergency Department: Clinical Data Analysis Study. JMIR Mhealth Uhealth 2023 Oct 03;11:e49995 [FREE Full text] [doi: 10.2196/49995] [Medline: 37788063]
81. Huang RS, Lu KJQ, Meaney C, Kemppainen J, Punnett A, Leung F. Assessment of Resident and AI Chatbot Performance on the University of Toronto Family Medicine Residency Progress Test: Comparative Study. JMIR Med Educ 2023 Sep 19;9:e50514 [FREE Full text] [doi: 10.2196/50514] [Medline: 37725411]
82. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 2021 Mar 31;8(1):53 [FREE Full text] [doi: 10.1186/s40537-021-00444-8] [Medline: 33816053]

XSL•FO
RenderX

83. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion 2023 Nov;99:101805. [doi: 10.1016/j.inffus.2023.101805]
84. Giray L. Prompt Engineering with ChatGPT: A Guide for Academic Writers. Ann Biomed Eng 2023 Dec;51(12):2629-2633. [doi: 10.1007/s10439-023-03272-4] [Medline: 37284994]
85. Khlaif ZN, Mousa A, Hattab MK, Itmazi J, Hassan AA, Sanmugam M, et al. The Potential and Concerns of Using AI in Scientific Research: ChatGPT Performance Evaluation. JMIR Med Educ 2023 Sep 14;9:e47049 [FREE Full text] [doi: 10.2196/47049] [Medline: 37707884]
86. Varsha PS. How can we manage biases in artificial intelligence systems – A systematic literature review. International Journal of Information Management Data Insights 2023 Apr;3(1):100165. [doi: 10.1016/j.jjimei.2023.100165]
87. Yanagita Y, Yokokawa D, Uchida S, Tawara J, Ikusaka M. Accuracy of ChatGPT on Medical Questions in the National Medical Licensing Examination in Japan: Evaluation Study. JMIR Form Res 2023 Oct 13;7:e48023 [FREE Full text] [doi: 10.2196/48023] [Medline: 37831496]
88. Taira K, Itaya T, Hanada A. Performance of the Large Language Model ChatGPT on the National Nurse Examinations in Japan: Evaluation Study. JMIR Nurs 2023 Jun 27;6:e47305 [FREE Full text] [doi: 10.2196/47305] [Medline: 37368470]
89. Sallam M, Barakat M, Sallam M. Pilot Testing of a Tool to Standardize the Assessment of the Quality of Health Information Generated by Artificial Intelligence-Based Models. Cureus 2023 Nov;15(11):e49373 [FREE Full text] [doi: 10.7759/cureus.49373] [Medline: 38024074]
90. Sezgin E, Chekeni F, Lee J, Keim S. Clinical Accuracy of Large Language Models and Google Search Responses to Postpartum Depression Questions: Cross-Sectional Study. J Med Internet Res 2023 Sep 11;25:e49240 [FREE Full text] [doi: 10.2196/49240] [Medline: 37695668]
91. Wilhelm TI, Roos J, Kaczmarczyk R. Large Language Models for Therapy Recommendations Across 3 Clinical Specialties: Comparative Study. J Med Internet Res 2023 Oct 30;25:e49324 [FREE Full text] [doi: 10.2196/49324] [Medline: 37902826]
92. Ferreira AL, Chu B, Grant-Kels JM, Ogunleye T, Lipoff JB. Evaluation of ChatGPT Dermatology Responses to Common Patient Queries. JMIR Dermatol 2023 Nov 17;6:e49280 [FREE Full text] [doi: 10.2196/49280] [Medline: 37976093]
93. Thirunavukarasu AJ, Hassan R, Mahmood S, Sanghera R, Barzangi K, El Mukashfi M, et al. Trialling a Large Language Model (ChatGPT) in General Practice With the Applied Knowledge Test: Observational Study Demonstrating Opportunities and Limitations in Primary Care. JMIR Med Educ 2023 Apr 21;9:e46599 [FREE Full text] [doi: 10.2196/46599] [Medline: 37083633]
94. Lakdawala N, Channa L, Gronbeck C, Lakdawala N, Weston G, Sloan B, et al. Assessing the Accuracy and Comprehensiveness of ChatGPT in Offering Clinical Guidance for Atopic Dermatitis and Acne Vulgaris. JMIR Dermatol 2023 Nov 14;6:e50409 [FREE Full text] [doi: 10.2196/50409] [Medline: 37962920]
95. Borchert RJ, Hickman CR, Pepys J, Sadler TJ. Performance of ChatGPT on the Situational Judgement Test-A Professional Dilemmas-Based Examination for Doctors in the United Kingdom. JMIR Med Educ 2023 Aug 07;9:e48978 [FREE Full text] [doi: 10.2196/48978] [Medline: 37548997]
96. Sun H, Zhang K, Lan W, Gu Q, Jiang G, Yang X, et al. An AI Dietitian for Type 2 Diabetes Mellitus Management Based on Large Language and Image Recognition Models: Preclinical Concept Validation Study. J Med Internet Res 2023 Nov 09;25:e51300 [FREE Full text] [doi: 10.2196/51300] [Medline: 37943581]
97. Ettman CK, Galea S. The Potential Influence of AI on Population Mental Health. JMIR Ment Health 2023 Nov 16;10:e49936 [FREE Full text] [doi: 10.2196/49936] [Medline: 37971803]
98. Sallam M, Barakat M, Sallam M. METRICS: Establishing a Preliminary Checklist to Standardize the Design and Reporting of Generative Artificial Intelligence-Based Studies in Healthcare Education and Practice. Interact J Med Res 2024 Jan 26:54704 [FREE Full text] [doi: 10.2196/54704] [Medline: 38276872]
99. Sallam M, Al-Salahat K. Below average ChatGPT performance in medical microbiology exam compared to university students. Front. Educ 2023 Dec 21;8:1333415. [doi: 10.3389/feduc.2023.1333415]
100. Sallam M, Al-Salahat K, Al-Ajlouni E. ChatGPT Performance in Diagnostic Clinical Microbiology Laboratory-Oriented Case Scenarios. Cureus 2023 Dec;15(12):e50629 [FREE Full text] [doi: 10.7759/cureus.50629] [Medline: 38107211]

Abbreviations

AI: artificial intelligence
CASP: Critical Appraisal Skills Programme
METRICS: Model, Evaluation, Timing, Range/Randomization, Individual factors, Count, and Specificity of prompts and language

XSL•FO
RenderX

Edited by T de Azevedo Cardoso; submitted 19.11.23; peer-reviewed by I Albalawi, A Spina, F Halaseh; comments to author 13.12.23; revised version received 18.12.23; accepted 26.01.24; published 15.02.24.

Please cite as:
Sallam M, Barakat M, Sallam M
A Preliminary Checklist (METRICS) to Standardize the Design and Reporting of Studies on Generative Artificial Intelligence–Based Models in Health Care Education and Practice: Development Study Involving a Literature Review
Interact J Med Res 2024;13:e54704
URL: https://www.i-jmr.org/2024/1/e54704
doi:10.2196/54704
PMID:38276872

©Malik Sallam, Muna Barakat, Mohammed Sallam. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 15.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study

Zhuoga Deji[1,2*], MSc; Yuantao Tong[1*], PhD; Honglian Huang[1,3], PhD; Zeyu Zhang[1], PhD; Meng Fang[4], PhD; M James C Crabbe[5,6,7], PhD; Xiaoyan Zhang[1], PhD; Ying Wang[3], PhD

[1]Research Center for Translational Medicine, Shanghai East Hospital, School of Life Sciences and Technology, Tongji University, Shanghai, China

[2]Information School, The University of Sheffield, Sheffield, United Kingdom

[3]Department of Clinical Laboratory Medicine Center, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[4]Department of Laboratory Medicine, Shanghai Eastern Hepatobiliary Surgery Hospital, Shanghai, China

[5]Wolfson College, Oxford University, Oxford, United Kingdom

[6]Institute of Biomedical and Environmental Science & Technology, University of Bedfordshire, Bedfordshire, United Kingdom

[7]School of Life Sciences, Shanxi University, Shanxi, China

[*]these authors contributed equally

Corresponding Author:
Ying Wang, PhD
Department of Clinical Laboratory Medicine Center
Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine
110 Ganhe Rd
Shanghai, 200437
China
Phone: 86 021 65161782
Fax: 86 021 65161782
Email: nadger_wang@139.com

Related Article:

Correction of: https://www.i-jmr.org/2024/1/e43585

(Interact J Med Res 2024;13:e59070) doi:10.2196/59070

In “Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study” (Interact J Med Res 2024;13:e43585) the authors noted one error.

In the originally published manuscript, author Zhuoga Deji was noted as having contributed to the manuscript as the first author.

This has been corrected to note that both authors Zhuoga Deji and Yuantao Tong contributed equally to the manuscript, and they should be regarded as joint first authors.

The correction will appear in the online version of the paper on the JMIR Publications Interactive Journal of Medical Research publication website on April 10, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

XSL•FO
RenderX

Submitted 01.04.24; this is a non–peer-reviewed article;accepted 03.04.24; published 10.04.24.

Please cite as:
Deji Z, Tong Y, Huang H, Zhang Z, Fang M, Crabbe MJC, Zhang X, Wang Y
Correction: Influence of Environmental Factors and Genome Diversity on Cumulative COVID-19 Cases in the Highland Region of China: Comparative Correlational Study
Interact J Med Res 2024;13:e59070
URL: https://www.i-jmr.org/2024/1/e59070
doi:10.2196/59070
PMID:38598823

©Zhuoga Deji, Yuantao Tong, Honglian Huang, Zeyu Zhang, Meng Fang, M James C Crabbe, Xiaoyan Zhang, Ying Wang. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 10.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Associated With Worsened Mental Health of Health Care Workers in Canada During the COVID-19 Pandemic: Cross-Sectional Survey Study

Ali AbdulHussein[1], BEng, MEng, MBA; Zahid Ahmad Butt[2], PhD; Stanko Dimitrov[1], PhD; Brian Cozzarin[1], PhD

[1]Department of Management Science, Faculty of Engineering, University of Waterloo, Waterloo, ON, Canada

[2]School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Ali AbdulHussein, BEng, MEng, MBA
Department of Management Science
Faculty of Engineering
University of Waterloo
200 University Ave W
Waterloo, ON, N2L 3G1
Canada
Phone: 1 (519) 888 4567
Fax: 1 (519) 888 4567
Email: a9abdulh@uwaterloo.ca

Abstract

Background: Health care workers (HCWs) in Canada have endured difficult conditions during the COVID-19 pandemic. Many worked long hours while attending to patients in a contagious environment. This introduced an additional burden that may have contributed to worsened mental health conditions.

Objective: In this study, we examine the factors associated with worsened mental health conditions of HCWs as compared to before the start of the pandemic.

Methods: We use data from a survey of HCWs by Statistics Canada. A regression model is used to estimate the odds ratios (ORs) of worsened mental health after the start of the pandemic. The estimated odds ratio (OR) is associated with different independent variables that include demographics (age, sex, immigration status, and geographic area), occupational factors (work status, occupational group, and exposure category), and different access levels to personal protective equipment (PPE).

Results: Of 18,139 eligible participants surveyed, 13,990 (77.1%) provided valid responses. We found that HCWs younger than 35 years old were more likely (OR 1.14, 95% CI 1.03-1.27; P=.01) to exhibit worsened mental health as compared to the reference group (35-44 years old). As for sex, male HCWs were less likely (OR 0.76, 95% CI 0.67-0.86; P<.001) to exhibit worsened mental health as compared to female HCWs. Immigrant HCWs were also less likely (OR 0.57, 95% CI 0.51-0.64; P<.001) to exhibit worsened mental health as compared to nonimmigrant HCWs. Further, HCWs working in Alberta had the highest likelihood of exhibiting worsened mental health as compared to HCWs working elsewhere (Atlantic provinces, Quebec, Manitoba, Saskatchewan, Ontario, British Columbia, and Northern Territories). Frontline workers were more likely (OR 1.26, 95% CI 1.16-1.38; P<.001) to exhibit worsened mental health than nonfrontline HCWs. Part-time HCWs were less likely (OR 0.85, 95% CI 0.76-0.93; P<.001) to exhibit worsened mental health than full-time HCWs. HCWs who reported encountering COVID-19 cases were more likely (OR 1.55, 95% CI 1.41-1.70; P<.001) to exhibit worsened mental health as compared to HCWs who reported no contact with the disease. As for PPE, HCWs who never had access to respirators, eye protection, and face shields are more likely to exhibit worsened mental health by 1.31 (95% CI 1.07-1.62; P<.001), 1.51 (95% CI 1.17-1.96; P<.001), and 1.41 (95% CI 1.05-1.92; P=.02) than those who always had access to the same PPE, respectively.

Conclusions: Different HCW groups experienced the pandemic differently based on their demographic and occupational backgrounds as well as access to PPE. Such findings are important to stakeholders involved in the planning of personalized support programs and aid mental health mitigation in future crises. Certain groups require more attention.

(Interact J Med Res 2024;13:e50064) doi:10.2196/50064

XSL•FO
RenderX

KEYWORDS

health care workers; COVID-19; mental health; demographic factors; occupational factors; access to PPE; pandemic; health care system; psychological trauma; psychological; trauma; respirators; eye protection; face shields; support

Introduction

On March 11, 2020, the World Health Organization declared COVID-19 a global pandemic. The pandemic resulted in devastating health impacts on populations and a crisis within the health care system [1]. This health care system was tasked to handle the unprecedented inflow of patients. Functions within the system that were impacted include emergency departments, intensive care units, physician services, and long-term care units [2]. Health care workers (HCWs) of different occupational groups battled the pandemic. Overwhelmed hospitals in Canada canceled less urgent surgeries by up to 80% by June 2020 [3]. These patterns shifted the workload on HCWs and had an impact on the overall health care system. These conditions not only demanded more hospital capacity but also put an overwhelming strain on HCWs [4]. In addition to the operational pressure, HCWs also suffered from the lack of personal protection equipment (PPE), especially at the beginning of the pandemic [5]. HCWs were at the frontline in battling this pandemic. This battle has put pressure on their mental health conditions [6-8].

In this paper, we assess the various factors associated with worsened HCWs' mental health conditions as compared to before the start of the pandemic. Although experiencing mental health conditions may be a daily occurrence for some HCWs, the duration and severity during the pandemic were different. HCWs were also at a higher risk of infection, adding to the risk of further mental health conditions [9]. In the past, HCWs have experienced mental health problems during other outbreaks, including the Middle East respiratory syndrome and the severe acute respiratory syndrome [10]. Such conditions have been studied in the literature [11]. In this study, we use a recent data set by Statistics Canada from a national cross-sectional survey that was conducted in the fall of 2020 to assess the impact of COVID-19 on HCWs. Unlike work in the literature, we comprehensively assess the impact of demographic and occupational factors as well as the availability of PPE on HCWs' mental health conditions [9,12,13].

In the global literature, authors assessed the mental health conditions of HCWs during such outbreaks globally [14-16]. The impact of demographic, social, and occupational factors was reportedly linked to various mental health conditions. Researchers also assessed the prevalence of stress, anxiety, and other psychological well-being indicators of HCWs in Oman during the pandemic [12]. The focus was on young female HCWs who encountered confirmed or suspected COVID-19 cases during their work. Another study in Turkey examined the relationship between the perceived risk of infection and the mental health conditions of HCWs during the COVID-19 pandemic [13]. In addition, a study in the Chinese province of Hubei was conducted early in the pandemic to assess the psychological impact of the pandemic on the frontline medical staff [9]. The study measured the association of factors including professional group, age, and sex factors with work stress.

Our study is among the first to highlight the association of diverse demographic and occupational factors with the mental health condition of HCWs during the pandemic in Canada. We also considered the role of access to PPE on HCWs' mental health conditions. Unlike existing literature, we studied the individual impact of each of the following factors on the mental health of HCWs while holding the rest constant. Demographic factors include age group, sex, province of the workplace, and immigration status. Occupational factors include work status, frontline category, and exposure to confirmed or suspected COVID-19 cases. Finally, we also considered access to several PPE. The findings of the study will be of prime importance to key stakeholders, including mental health support program planners, health care policy makers, HCWs themselves, and researchers in the area. The goal of this paper is to understand which factors were associated with worsened mental health conditions in HCWs after the start of the pandemic.

Methods

Data Sources, Study Procedure, and Participants

We used a data set from a recent cross-sectional survey by Statistics Canada on the impacts of COVID-19 on HCWs. Unlike other traditional Statistics Canada surveys, a random selection of participants was not used. Instead, Statistics Canada sent an email invitation to HCWs across Canada. Then, a snowball sampling procedure was used. The invitation included a link to a web-based survey that was available through Statistics Canada's web page. Accordingly, 18,139 responses were collected between November 24 and December 13, 2020, across 7 provincial and territorial regions in Canada. No data were collected for this survey beyond these dates. Only responses by HCWs were included in the data set. The responses were completely anonymized by Statistics Canada.

The questionnaire asked HCWs for information related to the job environment, demographics, geography, and information on access to PPE as background information. Adaptive questioning was used. Our study was limited to these 3 factors categories only: demographics, occupational, and access to PPE based on the available data from Statistics Canada. In assessing mental health well-being, the survey asked HCWs: compared to before the COVID-19 pandemic, how would you say your mental health is now? HCWs self-reported their perceived mental health on a 5-point Likert scale: much better now, somewhat better now, about the same, somewhat worse now, much worse now.

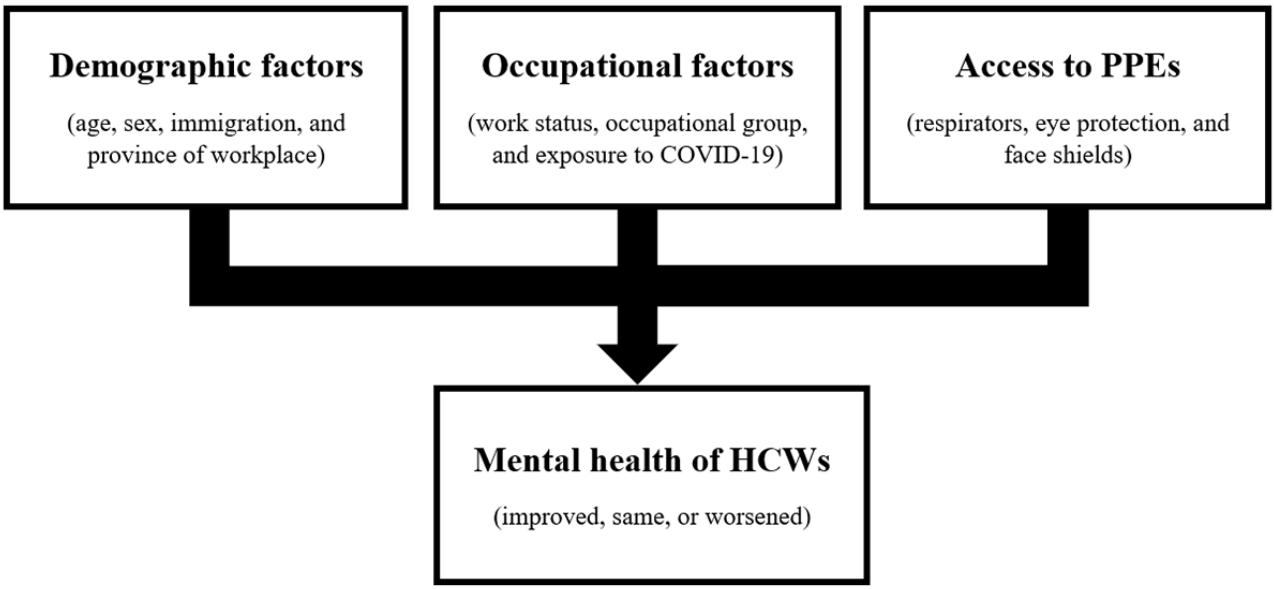
Statistical Analysis Strategy

Our study uniquely studied the association of various interesting factors, as depicted in Figure 1, with HCW mental health conditions as compared to before the start of the pandemic. To consider each factor separately, we used a multivariate ordinal logistic regression [17,18]. The model is defined in Multimedia Appendix 1. The dependent variable is the state of mental health

XSL•FO
RenderX

of the respondents. We reduced the mental health state categories from the 5 mentioned above to 3 (improved, same, or worsened) to yield statistically significant model estimates. The independent variables include demographics such as age, sex, immigration status, and the province of the workplace. Occupational variables include work status, occupational group, and exposure to COVID-19 cases. The model also estimated the association of HCWs' access to a variety of PPE, such as respirators, eye protection, and face shields. Access to PPE is also reported by the HCW on a 5-point Likert scale, which we reduced to 3 categories (always available, sometimes available, or never available) to yield statistically significant model estimates. All responses are self-reported. Analyses were conducted in RStudio (version 1.4.1717; Posit, PBC).

Figure 1. Factors associated with the mental health of HCWs. HCW: health care worker; PPE: personal protective equipment.



Ethical Considerations

Data collected by Statistics Canada were reviewed based on the following principles: privacy, security, transparency, accountability, trust, sustainability, data quality, and fairness, as well as well vetted to be harmless to participants or the public. All ethical approvals were considered by Statistics Canada. No further ethical approval to use these data was required by the authors and the survey is made available to the public. This study is exempt from further ethical approval. Furthermore, the authors also did not have access to information related to the development and testing of the survey. The authors are also not aware of any compensation provided to survey participants.

Results

Participant Profile

From the full data set of 18,139 responses, we considered nonduplicate responses that provided valid answers to all questions of interest to this study (N=13,990). The remaining responses contained missing or invalid responses. For the considered population, Table 1 outlines the number of observations and percentages for different demographic factors such as age group, sex, and immigration status. The number of observations and percentages were also calculated for occupational factors including the province of the primary workplace, work status (full-time vs part-time), frontline work status, occupational group, and exposure to confirmed or suspected COVID-19 cases. Additionally, the last 3 rows of Table 1 outline PPE access levels for 3 different PPE: face shields, eye protection, and respirators. Overall 4261 (31%) HCWs were younger than 35 years. The majority were female participants (n=12,682, 91%). In total, 12,510 (89%) HCWs were nonimmigrant. HCW workers were distributed across 7 regions, with the largest group of 6626 (47%) working in Ontario. The majority (n=10,152, 73%) of HCWs worked full-time and 5511 (39%) were nonfrontline allied health professionals. A total of 8188 (59%) HCWs did not report exposures to confirmed or suspected COVID-19 cases. As for access to PPE, 10,758 (77%) HCWs always had access to face shields, 10,118 (72%) always had access to eye protection, and 6898 (49%) always had access to respirators.

XSL•FO
RenderX

Table 1. Participant demographic, occupational, and PPE[a] access characteristics.

Characteristics	Value, n (%)
Age (years)	
<35	4261 (30.5)
35-44	3962 (28.3)
45-54	3415 (24.4)
≥55	2352 (16.8)
Sex	
Male	1308 (9.3)
Female	12,682 (90.7)
Immigration status	
Nonimmigrant	12,510 (89.4)
Immigrant	1480 (10.6)
Province of workplace	
Atlantic provinces	1972 (14.1)
Quebec	832 (5.9)
Ontario	6626 (47.4)
Manitoba and Saskatchewan	1308 (9.3)
Alberta	2018 (14.4)
British Columbia	1194 (8.5)
Northern territories	40 (0.3)
Work status	
Full-time	10,152 (72.6)
Part-time	3838 (27.4)
Frontline occupation	
Physician	397 (2.8)
Nurse	4689 (33.5)
Emergency medical personnel	223 (1.6)
Nonfrontline occupation	
Personal support worker	454 (3.2)
Allied health professional	5511 (39.4)
Laboratory worker	1267 (9.1)
Pharmacist	169 (1.2)
Dental professional	1280 (9.1)
Exposure to confirmed or suspected cases	
Yes	5802 (41.5)
No	8188 (58.5)
Access to PPE	
Face shields	
Always or usually available	10,758 (76.9)
Sometimes available	1070 (7.6)
Never available	357 (2.6)
Skipped answer	1805 (12.9)
Eye protection	

XSL•FO
RenderX

Characteristics	Value, n (%)
Always or usually available	10,118 (72.3)
Sometimes available	1132 (8.1)
Never available	556 (4.0)
Skipped answer	2184 (15.6)
Respirators	
Always or usually available	6898 (49.3)
Sometimes available	1369 (9.8)
Never available	676 (4.8)
Skipped answer	5047 (36.1)

[a]PPE: personal protective equipment.

Analysis Results

In the following, we present the results for estimating the associations between the various factors and the likelihood of worsened mental health conditions for HCWs.

Demographic Factors

Model estimates are expressed as odds ratios (ORs), as presented in Table 2. These ORs indicate the odds of worsened mental health conditions as compared to before the start of the pandemic. Based on the ORs in Table 2, HCWs who are younger than 35 years old were more likely (OR 1.14, 95% CI 1.03-1.27; *P*=.01) to exhibit worsened mental health conditions than the reference group (35 to 44 years old). Furthermore, those aged 45-54 years and 55 years and older were less likely to exhibit worsened mental health conditions than the reference group (OR 0.71, 95% CI 0.64-0.78; *P*<.001; and OR 0.55, 95% CI 0.49-0.61; *P*<.001, respectively). Hence, the older the HCW, the lower the likelihood of worsened mental health conditions. As for sex, male HCWs were less likely (OR 0.76, 95% CI 0.67-0.86; *P*<.001) to exhibit worsened mental health conditions than their female counterparts. Immigrant HCWs were also less likely (OR 0.57, 95% CI 0.51-0.64; *P*<.001) to exhibit worsened mental health conditions than nonimmigrants. Geographically, HCWs working in Alberta have the highest likelihood of worsened mental health conditions. HCWs living in Alberta were most likely to exhibit worsened mental health conditions.

XSL•FO
RenderX

Table 2. Estimates for the ordinal regression model for various factors associated with the mental health conditions of HCWs[a].

Independent variables	β	SE	OR[b] (95% CI)	*P* value
Age group (years; reference: 35 to 44 years)				
Less than 35	0.13	0.05	1.14 (1.03-1.27)	.01
45-54	–0.35	0.05	0.71 (0.64-0.78)	<.001
55 and older	–0.60	0.06	0.55 (0.49-0.61)	<.001
Sex (reference: female)				
Male	–0.28	0.06	0.76 (0.67-0.86)	<.001
Immigration status (reference: nonimmigrant)				
Immigrant	–0.56	0.06	0.57 (0.51-0.64)	<.001
Work location (reference: Alberta)				
Atlantic provinces	–0.41	0.07	0.66 (0.57-0.77)	<.001
Quebec	–0.30	0.09	0.74 (0.61-0.89)	<.001
Manitoba and Saskatchewan	–0.09	0.08	0.91 (0.77-1.08)	.27
Ontario	–0.24	0.06	0.78 (0.70-0.88)	<.001
British Columbia	–0.09	0.09	0.92 (0.77-1.09)	.31
North Territories	–0.56	0.34	0.57 (0.30-1.15)	.01
Work status (reference: full-time)				
Part-time	–0.16	0.04	0.85 (0.76-0.93)	<.001
Job setting (reference: nonfrontline)				
Frontline	0.23	0.05	1.26 (1.16-1.38)	<.001
Contact with patients with COVID-19 (reference: no contact)				
Exposure	0.44	0.05	1.55 (1.41-1.70)	<.001
Access to respirators (reference: did not need)				
Always	–0.03	0.05	0.97 (0.88-1.07)	.13
Sometimes	0.30	0.08	1.34 (1.14-1.59)	<.001
Never	0.27	0.10	1.31 (1.07-1.62)	<.001
Access to eye protection (reference: did not need)				
Always	–0.03	0.06	0.97 (0.85-1.10)	.17
Sometimes	0.29	0.11	1.33 (1.08-1.65)	<.001
Never	0.41	0.13	1.51 (1.17-1.96)	<.001
Access to face protection (reference: did not need)				
Always	0.26	0.07	1.30 (1.14-1.48)	<.001
Sometimes	0.40	0.11	1.50 (1.21-1.86)	<.001
Never	0.35	0.15	1.41 (1.05-1.92)	.02

[a]HCW: health care worker.

[b]OR: odds ratio.

Occupational Factors

Frontline HCWs, such as physicians, nurses, and emergency medical personnel, were more likely (OR 1.26, 95% CI 1.16-1.38; *P*<.001) to exhibit worsened mental health conditions than nonfrontline workers (personal support workers, allied health professionals, laboratory workers, pharmacists, and professionals). Part-time HCWs, however, were less likely (OR 0.85, 95% CI 0.76-0.93; *P*<.001) to exhibit worsened mental health conditions. Furthermore, HCWs who reported encountering suspected or confirmed COVID-19 cases were more likely (OR 1.55, 95% CI 1.41-1.70; *P*<.001) to exhibit worsened mental health conditions.

Access to PPE

HCWs who never had access to PPE such as respirators, eye protection, and face shields exhibited the highest likelihood of worsened mental health conditions than those who always had

XSL•FO
RenderX

access to such PPE. For instance, HCWs who never had access to respirators were more likely (OR 1.31, 95% CI 1.07-1.62; *P*<.001) to exhibit worsened mental health conditions than those who did not need this PPE. Similar trends were exhibited with access to eye protection (OR 1.51, 95% CI 1.17-1.96; *P*<.001) and face shields (OR 1.41, 95% CI 1.05-1.92; *P*=.02).

Discussion

Principal Findings

The statistical analysis in this study found that HCWs who are younger than 35 years old were found to be more likely to exhibit worsened mental health conditions than HCWs aged 35-44 years. Male HCWs were less likely to exhibit worsened mental health conditions than female HCWs. Immigrant HCWs were also less likely to exhibit worsened mental health conditions than nonimmigrant HCWs. In contrast, HCWs working in Alberta had a higher likelihood of worsened mental health conditions than HCWs working elsewhere (Atlantic provinces, Quebec, Manitoba, Saskatchewan, Ontario, British Columbia, and Northern Territories). Frontline workers were more likely to exhibit worsened mental health conditions than nonfrontline HCWs. Part-time HCWs were less likely to exhibit worsened mental health conditions than full-time HCWs. HCWs who reported encountering COVID-19 cases were more likely to exhibit worsened mental health conditions than HCWs who reported no contact with the disease. As for PPE, HCWs who never had access to respirators, eye protection, and face shields were more likely to exhibit worsened mental health conditions than those who always had access to the same PPE, respectively.

Comparison With Prior Work

Previous research has shown that shock events can result in psychological trauma to HCWs [19-21]. To the best of our knowledge, our study is the first to assess the association of a comprehensive variety of factors (demographics, occupational factors, and access to PPE) independently with the mental health conditions of HCWs.

With regard to demographic factors, a study found sex to be associated with worsened mental health conditions of HCWs. Female medical staff exhibited a higher incidence of severe anxiety than their male counterparts [22]. Sex was also found as a predictor of increased anxiety and distress in a study that found female HCWs more vulnerable to such conditions than male HCWs [23]. Our study findings corroborate these previous findings. Similarly, for age, our results demonstrated that older HCWs exhibited less likelihood of worsened mental health conditions. This agrees with findings from a study in the Middle East that found older workers enjoyed better mental health conditions than younger workers [24]. Older HCWs with longer work experience seemed to have handled the pandemic better than their younger counterparts. Another study found that HCWs 40 years or older were less likely to report higher anxiety during the pandemic than younger HCWs [25]. Our results confirm such findings as well.

As for occupational factors, our findings align with a cross-sectional study conducted in Oman, which found frontline workers to be more likely to have anxiety and sleep problems [26]. This was attributed to frontline workers' increased awareness of the mortality rate of COVID-19 and their fear of contracting the virus. Similar findings were reported by a study conducted in China as well [27]. Exposure to COVID-19 cases has been recognized as a risk factor associated with an increased likelihood of mental health issues in the literature. For instance, in France, female urologists working on the frontline were 1.41 times more likely to feel a degree of stress during their duties. Of this group, those who worked in a department where patients with COVID-19 are treated were 1.85 times more likely to report a degree of stress during work duties [28]. This was attributed to the workers' fear of infection and the spread of the virus. Direct involvement with COVID-19 care is also found to be highly associated with fear, depression, and anxiety as compared with those working under lower risk conditions [29]. The authors attributed this to the workers' fear of bringing the virus to their families at home.

As for the work setting, contrary to our findings, the literature reports higher levels of fear and anxiety among part-time HCWs than among full-time [30]. Another study in France found part-time HCWs to exhibit a greater association with distress [31]. Given the small number of part-time HCWs in our study, caution should be observed in interpreting the results. Regarding factors related to access to PPE, access to PPE was found to be associated with better health and less stress [24]. Lack of access to PPE was found to be a major source of HCW stress in a study that surveyed emergency physicians [32]. In Canada, a study found that inadequate PPE supply is associated with increased symptoms of anxiety and depression among HCWs [33]. Workers were concerned about the ability to access sufficient PPE during work hours. These findings are in line with our results. Compared to studies in the literature, our study is the first comprehensive study that assesses the association of a diverse pool of factors including demographics, occupational, and access to PPE, with HCWs' mental health conditions.

Implications: Mental Health Programming

Our comprehensive study sheds light on the association of various factors with HCWs' mental health after the start of the pandemic. It helps in understanding the vulnerability of various HCW groups to mental health during such events. Certain groups were at substantially higher risk of exhibiting worsened mental health conditions after the pandemic, hence the need for a specialized support program to target this group. As a direct implication, the findings can be used to inform guidelines for mental health support for HCWs during future public health emergencies. Such mental health support may be directed more specifically to more vulnerable groups. Literature has pointed to the importance of an evidence-based approach to designing mental health support programs for HCWs [34,35]. For instance, a study in Alberta during the pandemic pointed to the importance of understanding HCWs' occupational settings and mental health mitigation techniques [16]. Others have discussed the importance of support programs for HCWs during the pandemic based on need and background [36,37]. More importantly, many studies highlighted the importance of designing personalized mental health support programs based on various factors including demographic and occupational factors [38,39]. Some of these programs used mobile technology to offer mental health

XSL•FO
RenderX

support to HCWs [35,40]. As such, our study aims to inform stakeholders of such factors associated with HCW groups most vulnerable to future events such as the pandemic and health crises.

Limitations and Future Directions

This study has several limitations. First, the scope of the study was limited to assessing mental health in general without measuring different mental condition forms such as depression, anxiety, and fear. This is due to the limited data offered by the survey we used. Second, participants were asked, "Compared to before the COVID-19 pandemic, how would you say your mental health is now?" That is, HCWs had to objectively compare their mental health state from the time they received the survey to sometime in the past. There may be some variation in how HCWs perceive this comparison. Third, data collected were at a single point in time, which limits the ability to compare mental health assessments to later periods in the pandemic and during subsequent pandemic waves. Fourth, no information about HCW medical history is collected in this survey. In future research, we recommend collecting data at different points in the given event to allow for a time-series analysis and comparison at different crucial points of the event. We also suggest the collection of distress information and whether the conditions persisted beyond the initial shockwave. HCWs' years of work experience and ethnicity information can also be useful in future studies. We also believe medical history, particularly a preexisting history of mental health disorders, is an important factor associated with the mental health conditions of HCWs and should be considered in future studies.

Conclusions

This study investigated the factors associated with worsened mental health conditions of HCWs in Canada during the pandemic. Our study suggested the association of various factors with the likelihood of HCWs exhibiting worsened mental health conditions as compared to feeling neutral and better. In agreement with the literature, our findings concluded that younger (vs older), female (vs male), nonimmigrant (vs immigrant), full-time (vs part-time), and frontline (vs nonfrontline) HCWs living in Alberta (vs other provinces) exhibited a higher likelihood of worsened mental health conditions than those who felt neutral or better. Those who reported concerns about access to PPE also exhibited the same trend. Such findings can guide the future development of health care programming and inform mental health support planning for HCWs. COVID-19 is a shocking event that introduced uncertainty to the health care system.

Acknowledgments

We would like to thank Statistics Canada for providing the data used in this study.

Data Availability

The data that support the findings of this study is available through Statistics Canada.

Authors' Contributions

AA processed and cleaned the data and wrote the first draft of the manuscript with guidance from ZAB. BC and SD conducted iterative reviews and suggested revisions to the draft. All authors offered substantial contributions to the discussion and analysis presented in this manuscript. All authors agreed to publish the final draft of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Logistic model.
[DOCX File , 21 KB - ijmr_v13i1e50064_app1.docx]

References

1. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020. World Health Organization. 2020. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2023-03-01]
2. Overview: COVID-19's impact on health care systems. Canadian Institute for Health Information. 2021. URL: https://www.cihi.ca/en/covid-19-resources/impact-of-covid-19-on-canadas-health-care-systems/the-big-picture [accessed 2023-05-11]
3. Dudevich A, Frood J. Impact of the COVID-19 pandemic on health system use in Canada. Healthc Q 2021;24(2):12-14 [FREE Full text] [doi: 10.12927/hcq.2021.26552] [Medline: 34297658]
4. Carazo S, Pelletier M, Talbot D, Jauvin N, De Serres G, Vézina M. Psychological distress of healthcare workers in Québec (Canada) during the second and the third pandemic waves. J Occup Environ Med 2022;64(6):495-503 [FREE Full text] [doi: 10.1097/JOM.0000000000002487] [Medline: 35051960]
5. Brophy JT, Keith MM, Hurley M, McArthur JE. Sacrificed: Ontario healthcare workers in the time of COVID-19. New Solut 2021;30(4):267-281 [FREE Full text] [doi: 10.1177/1048291120974358] [Medline: 33174768]

XSL•FO
RenderX

6. Stuijfzand S, Deforges C, Sandoz V, Sajin CT, Jaques C, Elmers J, et al. Psychological impact of an epidemic/pandemic on the mental health of healthcare professionals: a rapid review. BMC Public Health 2020;20(1):1230 [FREE Full text] [doi: 10.1186/s12889-020-09322-z] [Medline: 32787815]
7. Vizheh M, Qorbani M, Arzaghi SM, Muhidin S, Javanmard Z, Esmaeili M. The mental health of healthcare workers in the COVID-19 pandemic: a systematic review. J Diabetes Metab Disord 2020;19(2):1967-1978 [FREE Full text] [doi: 10.1007/s40200-020-00643-9] [Medline: 33134211]
8. Liu JJW, Nazarov A, Plouffe RA, Forchuk CA, Deda E, Gargala D, et al. Exploring the well-being of health care workers during the COVID-19 pandemic: protocol for a prospective longitudinal study. JMIR Res Protoc 2021;10(9):e32663 [FREE Full text] [doi: 10.2196/32663] [Medline: 34477557]
9. Cai H, Tu B, Ma J, Chen L, Fu L, Jiang Y, et al. Psychological impact and coping strategies of frontline medical staff in hunan between january and march 2020 during the outbreak of coronavirus disease 2019 (COVID-19) in Hubei, China. Med Sci Monit 2020;26:e924171 [FREE Full text] [doi: 10.12659/MSM.924171] [Medline: 32291383]
10. Peeri NC, Shrestha N, Rahman MS, Zaki R, Tan Z, Bibi S, et al. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? Int J Epidemiol 2020;49(3):717-726 [FREE Full text] [doi: 10.1093/ije/dyaa033] [Medline: 32086938]
11. De Brier N, Stroobants S, Vandekerckhove P, De Buck E. Factors affecting mental health of health care workers during coronavirus disease outbreaks (SARS, MERS and COVID-19): a rapid systematic review. PLoS One 2020;15(12):e0244052 [FREE Full text] [doi: 10.1371/journal.pone.0244052] [Medline: 33320910]
12. Badahdah A, Khamis F, Al Mahyijari N, Al Balushi M, Al Hatmi H, Al Salmi I, et al. The mental health of health care workers in Oman during the COVID-19 pandemic. Int J Soc Psychiatry 2021;67(1):90-95 [FREE Full text] [doi: 10.1177/0020764020939596] [Medline: 32635837]
13. Yıldırım M, Arslan G, Özaslan A. Perceived risk and mental health problems among healthcare professionals during COVID-19 pandemic: exploring the mediating effects of resilience and coronavirus fear. Int J Ment Health Addict 2022;20(2):1035-1045 [FREE Full text] [doi: 10.1007/s11469-020-00424-8] [Medline: 33223977]
14. Serrano-Ripoll MJ, Meneses-Echavez JF, Ricci-Cabello I, Fraile-Navarro D, Fiol-deRoque MA, Pastor-Moreno G, et al. Impact of viral epidemic outbreaks on mental health of healthcare workers: a rapid systematic review and meta-analysis. J Affect Disord 2020;277:347-357 [FREE Full text] [doi: 10.1016/j.jad.2020.08.034] [Medline: 32861835]
15. Gupta B, Sharma V, Kumar N, Mahajan A. Anxiety and sleep disturbances among health care workers during the COVID-19 pandemic in India: cross-sectional online survey. JMIR Public Health Surveill 2020;6(4):e24206 [FREE Full text] [doi: 10.2196/24206] [Medline: 33284784]
16. Mrklas K, Shalaby R, Hrabok M, Gusnowski A, Vuong W, Surood S, et al. Prevalence of perceived stress, anxiety, depression, and obsessive-compulsive symptoms in health care workers and other workers in Alberta during the COVID-19 pandemic: cross-sectional survey. JMIR Ment Health 2020;7(9):e22408 [FREE Full text] [doi: 10.2196/22408] [Medline: 32915764]
17. Hadley J, Waidmann T. Health insurance and health at age 65: implications for medical care spending on new Medicare beneficiaries. Health Serv Res 2006;41(2):429-451 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00491.x] [Medline: 16584457]
18. Yunchao C, Yusof SA, Amin RM, Arshad MNM. Household debt and household spending behavior: evidence from Malaysia. J Ekon Malays 2020;54(1):111-120. [doi: 10.17576/jem-2020-5401-8]
19. Lancee WJ, Maunder RG, Goldbloom DS, Coauthors for the Impact of SARS Study. Prevalence of psychiatric disorders among Toronto hospital workers one to two years after the SARS outbreak. Psychiatr Serv 2008;59(1):91-95 [FREE Full text] [doi: 10.1176/ps.2008.59.1.91] [Medline: 18182545]
20. Maunder RG, Lancee WJ, Balderson KE, Bennett JP, Borgundvaag B, Evans S, et al. Long-term psychological and occupational effects of providing hospital healthcare during SARS outbreak. Emerg Infect Dis 2006;12(12):1924-1932 [FREE Full text] [doi: 10.3201/eid1212.060584] [Medline: 17326946]
21. Park J, Lee E, Park N, Choi YH. Mental health of nurses working at a government-designated hospital during a MERS-CoV outbreak: a cross-sectional study. Arch Psychiatr Nurs 2018;32(1):2-6 [FREE Full text] [doi: 10.1016/j.apnu.2017.09.006] [Medline: 29413067]
22. Huang JZ, Han MF, Luo TD, Ren AK, Zhou XP. [Mental health survey of medical staff in a tertiary infectious disease hospital for COVID-19]. Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi 2020;38(3):192-195 [FREE Full text] [doi: 10.3760/cma.j.cn121094-20200219-00063] [Medline: 32131151]
23. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
24. Zhang SX, Liu J, Jahanshahi AA, Nawaser K, Yousefi A, Li J, et al. At the height of the storm: healthcare staff's health conditions and job satisfaction and their associated predictors during the epidemic peak of COVID-19. Brain Behav Immun 2020;87:144-146 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.010] [Medline: 32387345]

XSL•FO
RenderX

25. Van Wert MJ, Gandhi S, Gupta I, Singh A, Eid SM, Haroon Burhanullah M, et al. Healthcare worker mental health after the initial peak of the COVID-19 pandemic: a US Medical Center cross-sectional survey. J Gen Intern Med 2022;37(5):1169-1176 [FREE Full text] [doi: 10.1007/s11606-021-07251-0] [Medline: 34993856]
26. Alshekaili M, Hassan W, Al Said N, Al Sulaimani F, Jayapal SK, Al-Mawali A, et al. Factors associated with mental health outcomes across healthcare settings in Oman during COVID-19: frontline versus non-frontline healthcare workers. BMJ Open 2020;10(10):e042030 [FREE Full text] [doi: 10.1136/bmjopen-2020-042030] [Medline: 33040019]
27. Cai Q, Feng H, Huang J, Wang M, Wang Q, Lu X, et al. The mental health of frontline and non-frontline medical workers during the coronavirus disease 2019 (COVID-19) outbreak in China: a case-control study. J Affect Disord 2020;275:210-215 [FREE Full text] [doi: 10.1016/j.jad.2020.06.031] [Medline: 32734910]
28. Abdessater M, Rouprêt M, Misrai V, Matillon X, Gondran-Tellier B, Freton L, et al. COVID19 pandemic impacts on anxiety of French urologist in training: outcomes from a national survey. Prog Urol 2020;30(8-9):448-455 [FREE Full text] [doi: 10.1016/j.purol.2020.04.015] [Medline: 32376208]
29. Lu W, Wang H, Lin Y, Li L. Psychological status of medical workforce during the COVID-19 pandemic: a cross-sectional study. Psychiatry Res 2020;288:112936 [FREE Full text] [doi: 10.1016/j.psychres.2020.112936] [Medline: 32276196]
30. Labrague LJ, de Los Santos JAA. Fear of COVID-19, psychological distress, work satisfaction and turnover intention among frontline nurses. J Nurs Manag 2021;29(3):395-403 [FREE Full text] [doi: 10.1111/jonm.13168] [Medline: 32985046]
31. Fournier A, Laurent A, Lheureux F, Ribeiro-Marthoud MA, Ecarnot F, Binquet C, et al. Impact of the COVID-19 pandemic on the mental health of professionals in 77 hospitals in France. PLoS One 2022;17(2):e0263666 [FREE Full text] [doi: 10.1371/journal.pone.0263666] [Medline: 35171915]
32. Marco CA, Larkin GL, Feeser VR, Monti JE, Vearrier L, ACEP Ethics Committee. Post-traumatic stress and stress disorders during the COVID-19 pandemic: survey of emergency physicians. J Am Coll Emerg Physicians Open 2020;1(6):1594-1601 [FREE Full text] [doi: 10.1002/emp2.12305] [Medline: 33392568]
33. Oudyk J, Smith P. Occupational Health Clinics for Ontario workers mayday, mayday workplace mental health series 2020. Occupational Health Clinics for Ontario. 2020. URL: https://www.ohcow.on.ca/mayday-mayday-2020.html [accessed 2023-06-05]
34. Greenberg N. Mental health of health-care workers in the COVID-19 era. Nat Rev Nephrol 2020;16(8):425-426 [FREE Full text] [doi: 10.1038/s41581-020-0314-5] [Medline: 32561871]
35. Lewis M, Palmer VJ, Kotevski A, Densley K, O'Donnell ML, Johnson C, et al. Rapid design and delivery of an experience-based co-designed mobile app to support the mental health needs of health care workers affected by the COVID-19 pandemic: impact evaluation protocol. JMIR Res Protoc 2021;10(3):e26168 [FREE Full text] [doi: 10.2196/26168] [Medline: 33635823]
36. Vizheh M, Qorbani M, Arzaghi SM, Muhidin S, Javanmard Z, Esmaeili M. The mental health of healthcare workers in the COVID-19 pandemic: a systematic review. J Diabetes Metab Disord 2020;19(2):1967-1978 [FREE Full text] [doi: 10.1007/s40200-020-00643-9] [Medline: 33134211]
37. David E, DePierro JM, Marin DB, Sharma V, Charney DS, Katz CL. COVID-19 pandemic support programs for healthcare workers and implications for occupational mental health: a narrative review. Psychiatr Q 2022;93(1):227-247 [FREE Full text] [doi: 10.1007/s11126-021-09952-5] [Medline: 34606067]
38. Siskind D, Harris M, Pirkis J, Whiteford H. Personalised support delivered by support workers for people with severe and persistent mental illness: a systematic review of patient outcomes. Epidemiol Psychiatr Sci 2012;21(1):97-110. [doi: 10.1017/s2045796011000734] [Medline: 22670417]
39. Deif R, Salama M. Depression from a precision mental health perspective: utilizing personalized conceptualizations to guide personalized treatments. Front Psychiatry 2021;12:650318 [FREE Full text] [doi: 10.3389/fpsyt.2021.650318] [Medline: 34045980]
40. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A mobile phone-based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021;9(5):e27039 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]

Abbreviations

HCW: health care worker
OR: odds ratio
PPE: personal protective equipment

XSL•FO
RenderX

Edited by T de Azevedo Cardoso; submitted 18.06.23; peer-reviewed by S Alwan; comments to author 08.11.23; revised version received 26.11.23; accepted 12.01.24; published 15.02.24.

Please cite as:
AbdulHussein A, Butt ZA, Dimitrov S, Cozzarin B
Factors Associated With Worsened Mental Health of Health Care Workers in Canada During the COVID-19 Pandemic: Cross-Sectional Survey Study
Interact J Med Res 2024;13:e50064
URL: https://www.i-jmr.org/2024/1/e50064
doi:10.2196/50064
PMID:38358785

©Ali AbdulHussein, Zahid Ahmad Butt, Stanko Dimitrov, Brian Cozzarin. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 15.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Tool to Assess Social Inclusion and Support Care Planning in Mental Health Supported Accommodation: Development and Preliminary Test Study

Sharon Eager[1], MSc; Helen Killaspy[1,2], MBBS, PhD; Joanna C[1], MSc; Gillian Mezey[3], MBBS; Peter McPherson[1], DPsych; Megan Downey[2], BSc; Georgina Thompson[1], MSc; Brynmor Lloyd-Evans[1], PhD

[1]Division of Psychiatry, University College London, London, United Kingdom

[2]Camden and Islington NHS Foundation Trust, London, United Kingdom

[3]Population Health Research Institute, St George's, University of London, London, United Kingdom

Corresponding Author:
Brynmor Lloyd-Evans, PhD
Division of Psychiatry
University College London
Maple House
149 Tottenham Court Road
London, W1T 7NF
United Kingdom
Phone: 44 (0)20 7679 9428
Email: b.lloyd-evans@ucl.ac.uk

Abstract

Background: Individuals with severe mental illness living in supported accommodation are often socially excluded. Social inclusion is an important aspect of recovery-based practice and quality of life. The Social Inclusion Questionnaire User Experience (SInQUE) is a measure of social inclusion that has been validated for use with people with mental health problems. Previous research has suggested that the SInQUE could also help support care planning focused on enabling social inclusion in routine mental health practice.

Objective: This study aims to develop a web-based version of the SInQUE for use in mental health supported accommodation services, examine its acceptability and perceived usefulness as a tool to support care planning with service users, determine the extent of uptake of the tool in supported accommodation settings, and develop a program theory and logic model for the online SInQUE.

Methods: This study involved a laboratory-testing stage to assess the acceptability of the SInQUE tool through “think-aloud” testing with 6 supported accommodation staff members and a field-testing stage to assess the acceptability, utility, and use of the SInQUE tool over a 5-month period. An implementation strategy was used in 1 London borough to encourage the use of the SInQUE. Qualitative interviews with 12 service users and 12 staff members who used the tool were conducted and analyzed using thematic analysis. The use of the SInQUE was compared with that in 2 other local authority areas, 1 urban and 1 rural, where the tool was made available for use but no implementation strategy was used.

Results: Overall, 17 staff members used the SInQUE with 28 different service users during the implementation period (approximately 10% of all service users living in supported accommodation in the study area). The staff and service users interviewed felt that the SInQUE was collaborative, comprehensive, user-friendly, and relevant. Although some staff were concerned that particular questions might be too personal, service users did not echo this view. Participants generally felt that the SInQUE could help identify individuals’ priorities regarding different aspects of social inclusion by prompting in-depth conversations and tailoring specific support to address service users’ inclusion goals. Some interviewees also suggested that the tool could highlight areas of unmet or unmeetable needs across the borough that could feed into service planning. The SInQUE was not used in the comparison areas that had no implementation strategy.

Conclusions: The online SInQUE is an acceptable and potentially useful tool that can be recommended to assess and support care planning to enable social inclusion of people living in mental health supported accommodation services. Despite this, uptake rates were modest during the study period. A concerted implementation strategy is key to embedding its use in usual care, including proactive endorsement by senior leaders and service managers.

XSL•FO
RenderX

(Interact J Med Res 2024;13:e45987) doi:10.2196/45987

KEYWORDS

social inclusion; supported accommodation; mental health; digital health; care planning

Introduction

Background

Social inclusion refers to an individual's ability to participate in important societal activities and their sense of community belonging [1,2]. Someone may feel socially excluded if they do not have opportunities for societal involvement and integration, often because of external factors that are beyond their control [3]. Social exclusion is a multifaceted continuum [2], typically signified by poverty, unemployment, inequality, and poor health [4].

People with serious mental illness are thought to be among the most socially excluded groups in society [5]. Individuals with this type of mental health problems often have smaller and less satisfying social networks [6], lower household income [7], and lower levels of employment [8,9] and experience more criminal and violent victimization [10,11] than those in the general population. Social exclusion can be conceptualized as both a cause and a consequence of mental illness [12]. Furthermore, greater social inclusion is associated with better quality of life and lower levels of loneliness among those with severe mental illness, suggesting that social exclusion is an important area for mental health practitioners to try to address [13,14].

Mental health supported accommodation services provide care and support to individuals with particularly severe and complex mental health problems as a way of supporting recovery in the community [15]. It is estimated that there are approximately 100,000 people living in mental health supported accommodation in England. Services are typically staffed by support workers, with additional specialist clinical input provided by National Health Service (NHS) community mental health teams [16]. In England, three main types of supported accommodation are provided: (1) residential care homes for those with the highest needs that comprise 24-hour–staffed communal facilities where placements are not time limited, with meals, supervision of medication, cleaning, and activities provided to service users; (2) supported housing services that provide shared or individual, self-contained, and time-limited tenancies with staff based on-site up to 24 hours a day to assist service users in gaining skills to move on to less supported accommodation; and (3) floating outreach services that provide visiting support for a few hours per week to people living in permanent, self-contained, and individual tenancies with the aim of reducing support over time to zero [16].

Service users living in supported accommodation are often socially isolated, with low levels of employment and little involvement in civil and political processes [17]. Many report feeling lonely and isolated and experiencing a high level of stigma that causes them to become more socially isolated [18]. There is evidence that users of mental health supported accommodation services report a variety of unmet needs, such as accessing employment opportunities and forming intimate relationships [19,20]. However, relatively little research has been conducted to determine the precise needs of service users living in supported accommodation [21], and a greater focus on this group is needed to identify and implement interventions that are likely to be the most useful for them [22].

Supporting service users to work toward desired goals and community engagement is highly congruent with recovery-based practice in mental health. Recovery-based practice recognizes and builds on service users' strengths and promotes empowerment through collaboration between them and staff to identify and work toward specific goals [23]. Many of the identified goals are markers of social inclusion, such as employment, social network development, and participation in community activities [24]. There is qualitative evidence from a large national research program suggesting that staff working in mental health supported accommodation services operate with a considerable degree of recovery orientation [24,25], and the more recovery-orientated these services are, the more likely people are to move on successfully to more independent settings [26].

People living in mental health supported accommodation have expressed a strong preference for individually tailored services that offer choice and promote autonomy, consistent with a recovery-based approach [18]. Patient-reported outcome measures have been recommended to inform this individualized approach by directly capturing service users' perspectives on constructs such as goal attainment, quality of life, and social inclusion [27]. Such measures enable service users to make informed decisions about their own support and care planning, in line with World Health Organization recommendations for recovery-based practice in community care provision [28]. Resources delivered across web-based platforms, particularly those that offer guided support, have been established as accessible, acceptable, and effective for use by participants with severe mental illness [29,30]. A tailor-made web-based assessment tool, the Quality Indicator for Rehabilitative Care for supported accommodation, has also been successfully used by managers of supported accommodation services, suggesting that these settings have the required resources and expertise to implement online measures [31].

The Social Inclusion Questionnaire User Experience (SInQUE) was developed as a measure of social inclusion for individuals with severe mental illness [32]. The measure has been validated across a range of mental health populations, has established reliability, is considered acceptable to service users, and has been proposed as being potentially cross-culturally suitable [32-34]. To date, the SInQUE has been used solely in offline research contexts. However, stakeholder feedback from a previous study testing the SInQUE indicated that the measure may be useful in clinical practice to assess social inclusion, facilitate important conversations with service users, and guide care and support planning [34]. Furthermore, a consistent research recommendation from the developers of the SInQUE

XSL•FO
RenderX

tool has been to investigate whether the measure has utility as a care-planning tool to promote social inclusion in routine mental health practice [13,34].

This study aimed to develop a web-based version of the SInQUE for use in mental health supported accommodation services. We sought to examine the acceptability and perceived usefulness of this tool among supported accommodation staff and service users as a means to assess their needs for greater social inclusion and promote care planning.

Aims

The study aims were as follows:

1. To develop and refine a web-based version of the SInQUE social inclusion assessment tool tailored for use in mental health supported accommodation settings.
2. To investigate the acceptability and perceived utility of the online SInQUE tool among supported accommodation staff and service users.
3. To determine the extent of uptake of the tool in supported accommodation settings with and without a locally developed implementation strategy to support its use.
4. Informed by the study findings, to develop a program theory and logic model for the online SInQUE specifying its anticipated outcomes, the mechanisms through which they may be achieved, and contextual factors affecting the use and experience of the SInQUE.

Methods

Study Design

This study comprised two stages conducted in 1 inner London borough:

1. A laboratory-testing stage to assess initial acceptability of the tool and develop it through “think-aloud” testing and semistructured interviews with supported accommodation staff.
2. A field-testing stage to assess wider acceptability, feasibility, and use of the tool. Semistructured interviews were conducted with staff and service users who had used the online SInQUE during this stage.

The 5-month field-testing stage was supported by a local implementation strategy developed in collaboration with local service leads to support the use of the online SInQUE by supported accommodation staff in the participating London borough. The online SInQUE was also made available to supported accommodation services by local service leads in 2 other areas without any accompanying implementation strategy.

Description of the SInQUE Tool

The web-based version of the SInQUE [35] can be used to assess social inclusion and inform support and care planning for people with mental health problems. It is designed to be used by staff as part of routine care planning to be completed collaboratively with service users. It can be used on a computer, tablet, or mobile device. Staff are required to register for an account on the SInQUE site using their work email address and details of their organization and can then use the tool for free. No personal data identifying service users are logged or stored on the SInQUE platform. The online SInQUE generates a unique reference number for each new service user, which is retained by the staff member completing the assessment for future reference and to link any repeat assessments.

The online tool developed for laboratory testing in our study included the 46-item version of the validated SInQUE social inclusion questionnaire, which was refined following stakeholder feedback at the end of the previous measure development study [34]. For this study, we removed 1 question from the SInQUE that asked whether the respondent was living alone as this was considered redundant for people living in residential care and supported housing. The SInQUE’s psychometric properties have been established among people with a range of mental health problems receiving input from community mental health services in previous studies [13,32,34]. Although the removal of a question from the original SInQUE questionnaire compromised its established psychometric properties, this minor adaptation is unlikely to have disrupted them substantially. We wanted to minimize changes to maintain the online SInQUE’s similarity to the validated measure and did not aim to make further significant refinements to the content of the tool. Instead, we wanted to gain feedback on its acceptability and feasibility for use in its digital format among staff and users of mental health supported accommodation services to assist in care and support planning.

The online SInQUE questionnaire yields a total score of 0 to 75, with a higher score indicating greater social inclusion. The questions and subscale scores are grouped into 9 different areas of social inclusion: leisure, social relationships, religious and cultural activities, education and employment, transport, health, crime victimization, home life and housing, and civic duties. These areas cover the 5 social inclusion domains of the validated SInQUE (social integration, productivity, consumption, access to services, and political engagement), but the aforementioned 9 areas were considered more immediately understandable for use in practice.

Using the service user’s responses, the online SInQUE generates a list of areas in which the person has said that they would like to be more socially included. It then offers a prompt for the service user and staff member to collaboratively select up to 3 priority areas that they would like to integrate into the person’s support plan. Once the assessment is completed, a summary report is generated. If the assessment is repeated with the same service user in the future, this report will also display changes in their social inclusion over time. The tool can generate management-level summary reports for each organization that is registered with it and commissioner-level summary reports of services using the tool across an entire area (such as a London borough). Multimedia Appendix 1 provides a full description of the SInQUE using the TIDieR (Template for Intervention Description and Replication) checklist [36].

Setting

This study took place in mental health supported accommodation services across 1 inner London borough. There are 21 such services in the borough run by 6 different voluntary sector organizations. They offer varying degrees of support to >270

XSL•FO
RenderX

service users who are also supported by local NHS secondary mental health services. In the borough, there are approximately 24 service users living in residential care, 159 living in supported housing, and 89 who receive floating outreach support. Supported housing services offer 24-hour support to 119 people and "9 to 5" support to 40 individuals.

Laboratory-Testing Stage: Recruitment, Data Collection, and Analysis

In total, 6 supported accommodation staff members were recruited to provide initial impressions of the online SInQUE tool. We discussed the study with service managers working in 3 different services and asked them to nominate 2 staff members each from their service who were interested in taking part. Participants were purposively sampled to include staff working in floating outreach support, 24-hour supported housing, and residential care. We asked the managers of each of the 3 services to ask for volunteers from their staff teams. Interviews were arranged with the first 2 staff members identified by the manager of each service.

Data were collected between January 2022 and February 2022. All 6 think-aloud interviews with staff members were conducted and recorded using Microsoft Teams (Microsoft Corp). The researcher first discussed the information sheet with each participant and gave them the opportunity to ask questions. Following this, participants' consent was verbally collected and audio recorded separately from the main part of the interview. Participants were then asked to fill out a short online form providing their demographic information before beginning the interview.

We conducted "think-aloud" testing of the online tool with staff using a semistructured topic guide developed by the study team (Multimedia Appendix 2). Following a process used previously in developing web-based tools [37], participants were asked to complete set tasks using the online SInQUE tool while providing a continuous commentary on their thoughts. They were asked to open the SInQUE website, register for an account, and complete an assessment as they would with a service user. At all stages, they were prompted to share their thoughts as they navigated the website and offer their initial impressions on how easy it was to understand and use and its potential suitability for their work. Once participants had completed the questionnaire, the researcher asked broader questions about their experience using the tool and any ways in which it could be improved. Throughout these interviews, participants were asked to focus on their experience using the SInQUE tool rather than offering specific feedback on individual SInQUE items. This was because we did not intend to make substantial modifications to the SInQUE questions to maintain their scope and similarity to those of the validated SInQUE measure.

Identified problems and suggestions for improvements to the online tool were collated by the researcher following the interviews. They were then reviewed by the study team, decisions about refinements to the online SInQUE were agreed upon, and the tool was revised accordingly.

Field-Testing Stage: Recruitment, Data Collection, and Analysis

Local Implementation Strategy

The revised version of the tool was made available for use in mental health supported accommodation services across the participating inner London borough. We iteratively developed and implemented a strategy to encourage and support its use by supported accommodation staff in the borough over a 5-month period beginning on May 11, 2022. This implementation strategy was informed by consultation with supported accommodation service managers and clinicians working in the Islington community mental health rehabilitation team and by individual interviews conducted with supported accommodation service users and staff.

Interviews With Field-Testing Participants

Participants and Recruitment

Individual interviews with service users (n=12) and staff (n=12) who tried out the SInQUE tool were conducted from late May 2022 to September 2022. This number was chosen to explore the views of staff and service users from a variety of supported accommodation types and service providers. Following our implementation strategy, we asked staff members to alert the study researcher once they had tried the tool in practice. Any staff member or service user in supported accommodation who tried the SInQUE tool was eligible to participate in an individual interview.

Once a staff member informed the study researcher that they had tried the tool, we asked them whether they would like to participate in an individual interview about their experience. We also invited staff to pass on information about the study to the service users with whom they had used the tool and ask them whether they would like to participate in an interview about their experience. If the service user was interested in taking part, the researcher communicated with them either directly or through the staff member they had completed the SInQUE assessment with depending on their preference. Toward the end of the recruitment stage, we recruited the final few staff members and service users purposively to ensure that participants were from a range of supported accommodation types and provider organizations.

One service user interview and 1 staff member interview were conducted online via Microsoft Teams; all other interviews were carried out in person according to participants' preferences. In-person interviews were conducted by the study researcher at the staffed supported accommodation sites, aside from 1 interview with a service user receiving floating outreach support, which was conducted at their home.

Measures and Procedures

The researcher first discussed the information sheet with the participants and gave them the opportunity to ask questions about the study. For in-person interviews, informed consent was collected via a paper consent form; for online interviews, verbal consent was audio recorded. Participants were then asked to answer brief demographic questions about themselves and their associated services. Following this, the researcher asked

XSL•FO
RenderX

each participant questions about their experience using the SInQUE; whether there were any ways in which it could be improved; the appropriateness of the online tool for use in their work; and what impacts, if any, they thought it might have on care provision and service users' experience. The interview topic guides (one for staff participants and one for service user participants) were developed by the study team as semistructured interviews—they are provided in Multimedia Appendix 3. In-person interviews were recorded using a digital voice recorder; online interviews were recorded on Microsoft Teams. Interview audio recordings were transcribed by a professional transcription company with which University College London (UCL) had a data-sharing and privacy agreement. Interview transcripts were then checked by the study researcher for accuracy. Any potentially identifiable text was anonymized. The resulting cleaned transcripts were then securely stored on the UCL university system.

Analysis

The analysis of the interviews comprised 2 stages. First, the study researcher noted any problems experienced by participants and recorded improvements to the online SInQUE they suggested. These issues and the suggested changes were reviewed by the study team, as in the previous laboratory-testing phase. Minor modifications to the online SInQUE were agreed upon, and we made adjustments to the tool in line with this.

Second, transcripts were uploaded to NVivo (version 12; QSR International) for qualitative analysis. As we aimed to develop a program theory for the online SInQUE intervention, we initially coded data into a deductively derived framework that used an intervention-context-actor-mechanism-outcome (ICAMO) configuration, with each component of this ICAMO framework representing a primary theme [38]. Within each of these 5 primary themes, we inductively derived subthemes from the data using thematic analysis. The initial coding was conducted by the lead author (SE) and was then reviewed and adjusted collaboratively by the study team. This included gaining lived-experience perspectives from a researcher with experience of mental health service use (JC) and clinical insights from a senior clinical academic working in the participating borough as a consultant rehabilitation psychiatrist supporting service users who live in supported accommodation (HK). The team brought in further perspectives from those with backgrounds in social work (BLE), clinical psychology (PM), and forensic psychiatry (GM) and from the community rehabilitation team in the borough (MD).

Data Use Monitoring

Data on the uptake and use of the online SInQUE tool were collected from the online SInQUE informatics for the 5-month field-testing period from May 11, 2022, to October 11, 2022.

At the start of this period, the study team also contacted local mental health service rehabilitation and housing leads in 2 other areas: another inner London borough and a rural county in the west of England. These service leads contacted local supported accommodation managers and invited them to use the online SInQUE in their service if they wished. The tool was made available to 7 supported accommodation services in the London borough and 10 in the rural county. No further encouragement to use the tool or implementation support was provided. This allowed us to monitor uptake and use of the tool in 2 areas without an associated implementation plan, thus making inferences about the necessity and impact of the strategy we developed.

Logic Model Development

The study team developed a preliminary logic model for the online SInQUE in planning this study. We used the findings of the aforementioned research activities to review and refine this logic model and develop an updated theory about the potential outcomes for service users and organizations from using the online SInQUE; mechanisms through which these outcomes are achieved; and factors influencing the uptake, experience, and impact of the online tool. Factors were related to (1) the intervention itself, (2) the characteristics and attitudes of staff and service users using the online SInQUE, and (3) the broader organizational and societal context. This was summarized in a logic model in the form of an "ICAMO map" [38], which was developed and refined iteratively through discussion with the study team.

Ethical Considerations

The initial laboratory-testing phase of this study (Supporting social inclusion for people with serious mental illness living in supported housing [SUSHI] phase 1) was approved by the UCL Research Ethics Committee (REC) on June 18, 2021 (REC reference 6711/002). The subsequent field-testing phase (SUSHI phase 2) was approved by the London – Camden and Kings Cross NHS REC on November 4, 2021 (REC reference 21/LO/0657). Written or audio-recorded informed consent was obtained from all participants before they took part, and they were clearly informed that they could opt out of the study at any time. All the study data were carefully deidentified. Service user participants were offered a £20 (US $25.14) shopping voucher to thank them for their time.

Results

Participants

We recruited 6 supported accommodation staff members for the "think-aloud" interviews during the laboratory-testing stage. We recruited a further 12 staff members and 12 supported accommodation service users for the individual interviews as part of the field-testing stage. Participant characteristics for both stages are summarized in Table 1.

XSL•FO
RenderX

Table 1. Laboratory testing and field testing of the online Social Inclusion Questionnaire User Experience—characteristics of participants.

Participant characteristics	Laboratory testing	Field testing	
	Staff (n=6), n (%)	Staff (n=12), n (%)	Service users (n=12), n (%)
Gender			
Male	2 (33)	5 (42)	11 (92)
Female	4 (67)	6 (50)	0 (0)
Nonbinary	0 (0)	1 (8)	1 (8)
Age group (y)			
18-30	0 (0)	8 (67)	4 (33)
31-50	2 (33)	2 (17)	3 (25)
≥51	4 (67)	1 (8)	4 (33)
Prefer not to say	0 (0)	1 (8)	1 (8)
Ethnicity			
Asian/Asian British	1 (17)	0 (0)	1 (8)
Black/Black British	3 (50)	2 (17)	4 (33)
White/White British	2 (33)	8 (67)	1 (8)
Mixed/multiple ethnic groups	0 (0)	1 (8)	6 (50)
Other ethnic background	0 (0)	1 (8)	0 (0)
Sexual orientation			
Heterosexual	N/A[a]	N/A	6 (50)
Gay/lesbian	N/A	N/A	1 (8)
Bisexual	N/A	N/A	2 (17)
Prefer not to say	N/A	N/A	3 (25)
Type of supported accommodation lived or worked in			
Floating outreach support	2 (33)	2 (17)	2 (17)
9-to-5 supported housing	2 (33)	1 (8)	2 (17)
24-h supported housing	2 (33)	7 (58)	8 (67)
Residential care	0 (0)	2 (17)	0 (0)
Length of time worked or lived in supported accommodation (y)			
<2	1 (17)	7 (58)	5 (42)
2-5	3 (50)	3 (25)	5 (42)
6-10	1 (17)	2 (17)	0 (0)
≥10	1 (17)	0 (0)	1 (8)
Prefer not to say	0 (0)	0 (0)	1 (8)

[a]N/A: not applicable; staff were not asked about their sexual orientation.

Changes Made to the SInQUE

Following phase 1 laboratory testing and phase 2 field-testing, suggestions that participants made for how the tool could be improved were collated and reviewed by the team. Accordingly, adjustments were made to the online SInQUE after each stage, an overview of which can be found in Table 2. This addressed aim 1 of this study.

XSL•FO
RenderX

Table 2. Changes made to the online Social Inclusion Questionnaire User Experience (SInQUE) following phase 1 and phase 2 testing. All changes were made following the initial laboratory-testing stage unless indicated otherwise.

Section of the SInQUE affected	Explanation of the problem	Resolution	Justification for the change
Registration and use changes			
The online SInQUE home page	Some staff members and service users suggested developing additional materials to help explain the SInQUE[a].	Developed a guidance manual for service managers and commissioners with information on using the SInQUE as well as an informational leaflet and poster aimed at service users about the SInQUE.	It is easier for managers, commissioners, staff, and service users to understand and use the SInQUE.
The initial page where staff members are asked to register for the SInQUE	Staff were asked to enter the organization they worked for in a free-text box. Some found it confusing to know which organization name they should enter.	Changed response options to a fixed-response drop-down menu with all housing providers in the borough and an "other" free-text option.	Allows for compilation of service-level data and is easier for staff to navigate.
The page where staff members enter details to set up a new SInQUE assessment	Some staff members thought that further information on the exact purpose of the questionnaire and how it should be administered would be useful in the introductory paragraph describing the assessment.	Additional guidance on how the questionnaire should be administered was added to the introduction paragraph of the SInQUE.	Important contextual information for the questionnaire was explicitly clarified.
The page where staff members enter details to set up a new SInQUE assessment	Some staff members found the wording of the following question—"Please select the type of accommodation in which the service user is living from the list below"—to be ambiguous and confusing.	Changed the wording of the question to the following: "Please select the type of housing support the service user receives from the list below."	The clarity of the question improved.
Changes to the wording of SInQUE questions			
The section covering "leisure" questions	Some staff members felt that it was unclear what the "Other" option meant in the context of question 3f: "Over the past year have you been to...Other?"	Changed the wording of this subquestion to the following: "Other leisure activity?"	The clarity of the question improved.
The section covering "leisure" questions	Some staff members and service users felt that question 7—"Do you spend time in pubs or cafés?"—was worded in a way that was potentially inappropriate for people who do not drink alcohol[a].	Changed the wording of the question to the following: "Do you go out for a coffee/drink (e.g. to a café or pub, etc) at least once a week?"	The clarity and appropriateness of the question improved.
The section covering "social" questions	Some staff members were unsure whether question 9—"How many people, outside those in your care team, could you confide in?"—related to a professional or personal care team.	Changed the wording of the question to the following: "How many people, outside the workers in your care team, could you confide in?"	The clarity of the question improved.
The section covering "home life/housing" questions	Some staff members thought that question 36—"What kind of accommodation do you live in?"—was worded ambiguously.	Changed the wording of the question to the following: "What type of housing support do you receive?"	The clarity of the question improved.
The section covering "home life/housing" questions	A statement alerting users that question 38 had been omitted from the online SInQUE, which read the following—"Not relevant for supported accommodation contexts -omitted."—was confusing.	Changed statement to the following: "Question omitted, not included in the online SInQUE."	The clarity of the question improved.
Changes to the SInQUE summary outputs			
The SInQUE summary report	Some staff members found the spider graph to be confusing to interpret as the numbers summarizing scores in each section were not standardized and, therefore, it was difficult to tell which domains scored lower than others.	Simplified the spider graph to show percentage of total score in the graph instead of frequency.	It was easier to interpret the graph as sections with different totals became standardized.
The SInQUE summary report for multiple assessments completed with the same service user	In the section summarizing scores across multiple time points, staff members thought that a visual depiction of this comparison would be useful.	Made comparative bar charts of multiple scores across time with the same service user available on the summary report.	It was easier to understand and relay the results.

[a]Changes were made following the field-testing stage.

XSL•FO
RenderX

Overall, the changes made to the online SInQUE were relatively few and minor. Following initial laboratory testing, additional information and guidance for users was added, and minor revisions to the wording of questionnaire items were made to improve clarity. Modifications to the visual representation of scores in the summary reports were also made to aid ease of interpretation. During the field-testing stage, very few suggested changes to improve the usability of the online SInQUE were made by staff or service user participants. Further changes made at this stage included a minor wording adjustment to one question to ensure its cross-cultural appropriateness. Changes were made exclusively to the web-based version of the SInQUE and did not affect the existing SInQUE measure. We also developed an additional guidance document for managers and commissioners and an informational leaflet and poster about the SInQUE.

A few participants suggested substantial modifications to the structure and wording of individual items in the SInQUE that were not implemented by the study team. These decisions were made to preserve the broad scope and logical flow of the tool. We also declined to action some suggestions that were outside the remit of the SInQUE tool, such as adding more or free-text response options to some questions. However, where these suggestions indicated important potential barriers to using the SInQUE, they were noted and integrated into the qualitative analysis and logic model development. A summary of all comments and suggestions that were proposed but not implemented after review by the team can be found in Multimedia Appendix 4.

Interview Thematic Analysis

Overview

The interviews were analyzed using thematic analysis. Primary themes were deductively imposed according to each core element of the ICAMO model: intervention, context, actors, mechanisms, and outcomes [38]. Subthemes were inductively analyzed within each of these primary themes. The resultant thematic framework considering the perceived utility and acceptability of the online SInQUE and addressing aim 2 of the study is presented in Textbox 1. The themes are summarized in the following sections with a selection of illustrative quotes. Multimedia Appendix 5 provides further illustrative quotes for each subtheme.

XSL•FO
RenderX

Textbox 1. Summary of the thematic framework (intervention-context-actor-mechanism-outcome themes and inductive subthemes).

Intervention: combination of program elements or strategies designed to produce behavior changes or improve health status among individuals

- The online Social Inclusion Questionnaire User Experience (SInQUE):
  - Promotion of positive, collaborative discussion
  - Comprehensive and novel questions
  - Ability to repeat over time
  - User-friendly design:
    - Easy-to-navigate website
    - Quick to complete
    - Fixed-response questions
    - Offers options to choose from
    - Web-based format

Context: salient conditions that are likely to enable or constrain the activation of program mechanisms

- Relevance of the SInQUE to staff role
- Inconsistency in current assessments used across services
- Absence of comparably specific assessments
- Emergence from the pandemic

Actors: the individuals, groups, and institutions who play a role in the implementation and outcomes of an intervention

- Staff:
  - Professional knowledge and skills
  - Professional boundaries
- Staff (service user views about staff):
  - Trusting relationship
  - Proactivity in offering guidance and support
- Service users:
  - Familiarity and comfort with the questions
  - Individual language and cultural differences of service users
- Service users (staff views about service users):
  - Engagement in the assessment
  - Existing mental health needs

Mechanisms: any underlying determinants or social behaviors generated in certain contexts

- Using the online SInQUE can accomplish the following:
  - Boost service user proactivity and confidence
  - Identify service users' priorities on social inclusion
  - Prompt novel, personal conversations
  - Monitor changes in social inclusion over time
  - Identify gaps in support available within the organization and local community

Outcomes: behavior changes that follow the immediate knowledge change (intermediate) and changes such as patients' health status and impact on community and health system (long term)

- Intermediate:

XSL•FO
RenderX

- Improve staff relationship with and understanding of the service user
- Help plan more relevant, targeted support for the service user

- Long term:
  - Borough-level improvements and changes in services to support social inclusion
  - Individual-level benefits for service users' recovery and social inclusion

Intervention

In general, staff and service user participants felt that the tool was user-friendly and collaborative. Many noted the ability to repeat the assessment and the web-based format as being particularly useful and felt that the website was easy to navigate. The short length of the assessment was also discussed as an important advantage, with both staff and service users commenting that it felt quick to fill out. Participants noted that, despite the short assessment length, it still offered a range of interesting, positive, and sometimes unfamiliar questions that felt comprehensive and useful to discuss:

> *I think that it wasn't too just baseline, it was a little bit more than that and I think that's good. Because it gives the option of, "Okay, you don't want something, how can we improve and what is it that you do want that could help you while you're in our service?"* [212; staff member]

Interviewees felt that the user-friendliness was aided by accessible questions that were straightforward for service users to answer and that were cross-culturally appropriate for individuals from different ethnic backgrounds. Although some participants from both groups felt that the fixed-response options for the questions were limited, certain staff members thought that this made the questionnaire more accessible to service users who may otherwise struggle with engagement.

Context

Staff members largely felt that the tool was fitting and relevant to their role in helping support service users, and most did not already use assessments that were highly similar to the SInQUE. Certain staff members highlighted a lack of continuity of support workers within their service and noted that this often made it difficult to build rapport with service users. Some also commented on an inconsistency in assessments used across different services (in the local context, where 6 different provider organizations provided supported accommodation services across the borough). They noted that individual providers currently make their own recommendations on the tools that staff should use:

> *If it was a standard central assessment that we do in all supported housing, that's similar, like this for example, it might be beneficial in the long run. But each company has their own policy around it.* [205; staff member]

One staff member noted that the tool felt particularly relevant following the COVID-19 pandemic as a means to promote engagement among service users after a period of likely sustained social isolation.

Actors

There were 2 key actors to consider in the application of the assessment: the staff members who asked the questions and the service users who responded to them.

One staff member felt that the assessment was not particularly relevant in the context of their work in a residential care service, where they had an established relationship with service users and already knew much of the queried information about them. However, this was an outlying view. Although most staff members thought that the online tool could be suitable and useful for their work, they emphasized the importance of using their professional knowledge and skills to pick when and for whom the assessment would be appropriate. They suggested that service users acutely struggling with their mental health may find it difficult to maintain concentration and engagement with the questionnaire and others may feel that the assessment is not relevant to them.

Staff members also raised the importance of maintaining professional boundaries with service users, and some expressed a concern that certain questions may feel invasive or uncomfortable for service users to answer:

> *I think there was one quite private like about if they're in a relationship or something, and that was the only question that made me feel a bit like I'm asking something very personal about a relationship. Because they might not want to say that.* [210; staff member]

However, service user participants did not express any similar concerns about intrusive questions. They generally indicated that they felt comfortable with the assessment and that they were used to answering personal questions. Both staff and service users highlighted trust between those performing the assessment as a key factor in promoting engagement with such questions.

Both respondent groups highlighted the cultural diversity of service users within supported accommodation, and many noted that the tool felt appropriate for those from a variety of religious and ethnic backgrounds. Some service users commented on the importance of the staff members being proactive and taking the time to go through the assessment in detail with them, particularly individuals for whom English was a second language, as further explanation was required for some questions. Various staff members suggested that being provided with information about the assessment and its purpose specifically aimed at service users, for instance, a guidance

XSL•FO
RenderX

leaflet, would be helpful for them to convey the essential information about the assessment.

Mechanisms

Both participant groups discussed how the tool may boost confidence and proactivity for a wide range of service users by highlighting specific, achievable ways in which an individual can improve their social inclusion. They also noted how the assessment encourages service users to open up and enables more profound conversations between them and staff members:

> *Yes I found it really interesting, so like because it's not really topics I would actually talk about. So it gave me a bit of enthusiasm to talk about some of the questions.* [409; service user]

Both groups suggested that it might be particularly useful during key working sessions as a means to get to know an individual better and identify their support preferences soon after moving into supported accommodation. Participants also noted the value of repeating assessments over time, suggesting that this could be a potentially encouraging way to demonstrate service user progress and identify gaps in available support. The most frequently suggested time between assessments was 1 to 3 months, with up to 6 months mentioned as a potential maximum gap.

Outcomes

Interviewees discussed the short- and long-term outcomes that they felt the tool could offer. They discussed how the tool enabled targeted and relevant support that prioritized the service user's interests. Both groups also mentioned the potential for the tool to improve the relationship and understanding between service users and staff members:

> *It asks questions where maybe like for your support worker to get a better understanding of you, like even though the immediate thing is highlight areas you can work on, it gives a general overview of how you are.* [410; service user]

Some staff members also discussed how prolonged use of the tool could highlight the additional borough-level support that may be needed to improve certain gaps in support and could also promote service user recovery toward the goal of more independent accommodation.

Implementation Strategy

Our implementation strategy was developed to encourage the use of the SInQUE in the supported accommodation services in the borough and was updated through consultation with clinical staff working in the borough's community rehabilitation team and supported accommodation service managers. Our strategy was further informed by feedback from staff and service user participants during both stages of the study.

Each part of the strategy was developed to target an identified potential barrier to staff using the online SInQUE with service users. Subsequently, we mapped each component of the strategy to the 3 broad domains of the Capability, Opportunity, and Motivation–Behavior framework of behavior change [39] to describe whether each element of the strategy was intended to increase the staff's capability, opportunity, or motivation to use the online SInQUE. The complete implementation strategy and the Capability, Opportunity, and Motivation–Behavior domain that each component addressed are outlined in Table 3. Strategies were related to enlisting leadership support to encourage supported accommodation staff to use the SInQUE, providing technical guidance and assistance with using the online tool, and developing bespoke summary output reports to reinforce use and increase the organizational benefits of using the SInQUE.

XSL•FO
RenderX

Table 3. Summary of the implementation strategy to support the use of the online Social Inclusion Questionnaire User Experience (SInQUE).

Activity	Implementation goal being addressed	COM-B[a] domain
Study research assistant (SE) and clinical research staff member (MD) visit all supported accommodation services to introduce the SInQUE tool to staff and offer guidance on its use.	Increase awareness of the SInQUE tool among supported accommodation staff and respond to any of their concerns or other problems.	• Motivation • Opportunity • Capability
NHS[b] community rehabilitation team manager and clinical researcher (MD) review supported accommodation caseloads to identify suitable service users for a SInQUE assessment and ask their key worker to complete an assessment with those service users.	Lack of accountability after initially asking all staff members to use the SInQUE and staff hesitancy over which service users would be suitable for an assessment	• Motivation
NHS community rehabilitation team manager contacts supported accommodation managers to ask them to support and encourage the use of the SInQUE by identified key workers within a given time frame.	Lack of supported accommodation management prioritization for staff to use the SInQUE	• Motivation • Opportunity
Local authority commissioners contact all supported accommodation managers to encourage the use of the SInQUE within services in the borough.	Lack of supported accommodation management prioritization for staff to use the SInQUE	• Motivation
Study lead (BLE) and research assistant (SE) attend local Housing Forum meetings to update on the study and encourage the use of the SInQUE among all managers and staff members present.	Increase visibility and awareness of the SInQUE among supported accommodation managers and respond to any of their concerns or other problems.	• Motivation • Opportunity
Study research assistant (SE) offers supported accommodation staff technological support with SInQUE registration and use.	Uncertainty about how to manage the technical process of using the SInQUE	• Capability
Study team develops and circulates a leaflet about the SInQUE for supported accommodation staff to give to service users to help explain the purpose of an assessment.	Uncertainty among some supported accommodation staff members about how best to explain the purpose of the SInQUE and engage service users	• Capability
Study team sends summary reports to service managers outlining use of the SInQUE and highlighting the areas of social inclusion that are most frequently prioritized and addressed in their service[c].	Increased awareness among supported accommodation managers of the value offered by the SInQUE for service planning to encourage them to prompt staff to use it.	• Motivation
Study team sends summary reports to local authority commissioners outlining which services have used the SInQUE the most and highlighting the areas of social inclusion that are most frequently prioritized and addressed across all services in the borough[c].	Increased awareness of commissioners of the value offered by the SInQUE for service planning and commissioning to encourage them to prompt services to use it.	• Motivation

[a]COM-B: Capability, Opportunity, and Motivation–Behavior.

[b]NHS: National Health Service.

[c]These actions were planned with service managers and commissioners but not carried out during the 5-month implementation period because of the small number of completed SInQUE assessments.

Usage Data

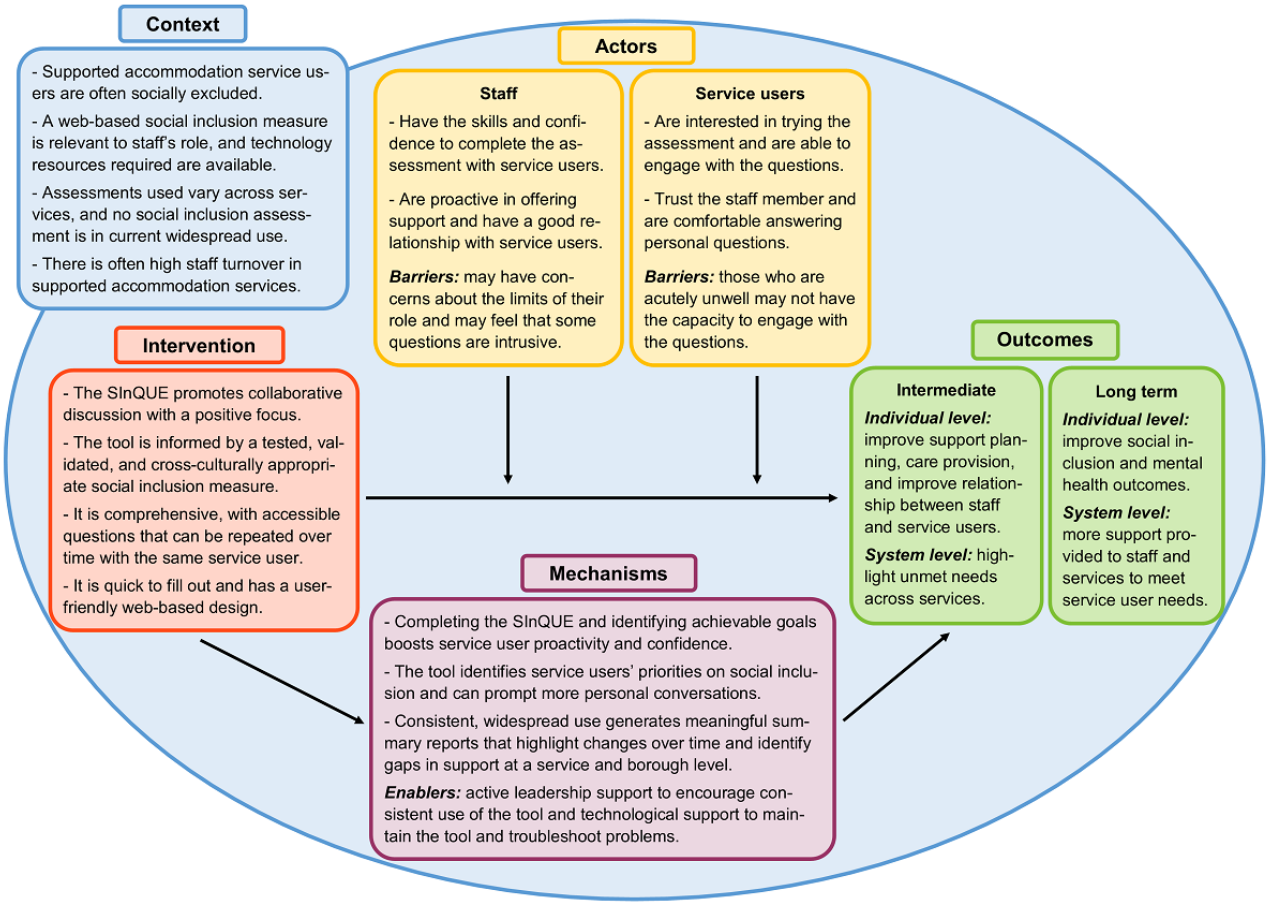
In total, 27 staff members in the inner London borough registered for an account with the online SInQUE. Of the 27 staff members who registered, 17 (63%) from 6 different supported accommodation providers started or completed a SInQUE assessment with at least 1 service user. This resulted in 30 completed SInQUE assessments with 28 service users in the borough. This represented just >10% of the estimated total number of service users living in supported accommodation in the borough. Of the 28 service users, 4 (14%) were from residential care, 19 (68%) were from 24-hour supported housing, 3 (11%) were from 9-to-5 supported housing, 1 (4%) was from floating outreach services, and 1 (4%) was registered as "other" accommodation type. One staff member from 1 of the local authority areas where there was no specific implementation strategy registered for an account with the SInQUE; however, they did not start or complete a SInQUE assessment. This addressed aim 3 of the study.

Intervention Logic Model

On the basis of the collective study findings, we developed a logic model to summarize the processes involved in using the SInQUE and address study aim 4 (Figure 1).

XSL•FO
RenderX

Figure 1. Intervention-context-actor-mechanism-outcome logic model for the online Social Inclusion Questionnaire User Experience (SInQUE).



The logic model was informed by the structure of an "ICAMO map" [38]. The model outlines the key aspects of the intervention (I), including its user-friendliness, its comprehensive nature, and the fact that it was based on a validated measure of social inclusion. It also indicates the potential outcomes (O) from using the SInQUE at both the individual and system levels, including improved support planning, better relationships, and provision of additional support for staff and services, which in turn may improve social inclusion and mental health outcomes as well as care provision more broadly. These operate within the broader societal context (C) of service users often being socially excluded and there being a high turnover of staff within these services and a high degree of variation across services in assessment tools that are recommended and in use, rendering the tool useful and relevant to the staff's role.

The key actors (A) in implementing the SInQUE are the staff and service users, who require the skills and proactivity to administer the assessment and the motivation and trust to engage with the questions, respectively. Staff may encounter barriers such as a concern that some questions are too intrusive, and service users struggling more severely with their mental health may lack the concentration or motivation to engage with the questions. The potential outcomes operate through certain mechanisms (M), which include increased service user confidence and the prompting of more in-depth personal conversations between service users and staff. The tool also identifies more relevant priorities for service users, which may or may not be chosen as an active priority for support by staff owing to individual or organizational factors. Persistent and wide-ranging use of the tool could, over time, highlight the aspects of social inclusion that are feasible to work on and those that are regularly not being prioritized.

Discussion

Principal Findings

The online SInQUE was generally perceived as acceptable and potentially useful by supported accommodation staff and service users. This is consistent with findings of previous studies that used the SInQUE with other mental health populations [32,34]. Both staff and service users generally found the tool to be user-friendly and relevant and suggested that it could promote more targeted care planning and improve the relationship between staff and service users. Owing to the lower uptake of the SInQUE in residential care and floating outreach services, findings related to the tool's utility in these settings are less conclusive than those for supported housing, where uptake was highest.

Some staff members expressed a concern that certain questions in the SInQUE could be perceived as intrusive by service users, indicating that they did not feel wholly comfortable asking what they perceived to be highly personal questions. However, this sentiment was not echoed by service users, who generally felt that the questions were appropriate and felt comfortable answering them. This finding is interesting given that supported

XSL•FO
RenderX

accommodation service users have highlighted in previous research the importance of feeling personally understood by staff in their service and have endorsed a process of familiarization with staff [25].

We found that implementation support is essential to promote the use of the tool in services, as evidenced by the lack of use of the tool in the 2 regions where the SInQUE was introduced without a concerted implementation strategy. The most effective steps in our implementation strategy were those during which the use of the tool was actively endorsed by individuals in leadership positions, particularly service managers and local service leaders. However, even with our concerted implementation strategy, uptake of the SInQUE was only achieved with approximately 10% of service users living in mental health supported accommodation in the participating borough within the 5-month study period.

Limitations

We used an established, iterative process of testing and feedback to develop the online SInQUE and determine its real-world acceptability and utility for use in mental health supported accommodation. However, it is important to acknowledge certain limitations of this study.

As mentioned previously, uptake of the SInQUE tool was highest in supported housing compared with residential care homes and floating outreach support. It is unclear whether this discrepancy reflects a greater reluctance from staff or service users in residential care and floating outreach support to use the online SInQUE. As proposed by one residential care staff member, it is possible that the staff in these services perceived the tool as being less relevant to their role. The discrepancy in part reflects the greater number of 24-hour and 9-to-5 supported housing units in the borough compared with residential care and floating outreach services, with approximately 6 times as many service users living in supported housing compared with residential care and nearly twice as many living in supported housing compared with floating outreach support. Through the local health service community mental health rehabilitation team, we also had a more direct connection with supported housing teams compared with other service types, which may have further contributed to the imbalance in services in which the SInQUE was used.

There were no female service user participants in the qualitative analysis; therefore, the findings may not be applicable to women in supported accommodation. It is unclear why it proved more difficult to recruit female participants, although it may reflect the higher proportion of male service users availing of supported accommodation in England—one review suggests that between 68% and 74% of service users are male across all supported accommodation types [40]. Furthermore, as we only tested the SInQUE in 1 London borough, the findings may not be generalizable to other regions.

As the tool was only used with approximately 10% of service users in the borough, the success of our implementation strategy was limited, and the low uptake may limit the wider generalizability of our findings. Owing to the short period and limited scope of the study, it was also not possible to assess whether use of the SInQUE in practice led to improved outcomes for service users or how useful the repeat assessments were over time. As the staff who participated volunteered to do so and they chose which service users to complete the SInQUE with, the findings may have been affected by selection bias and may not accurately reflect all supported accommodation staff members' and service users' views.

Finally, we removed 1 question from the original SInQUE questionnaire for our online version as asking people whether they lived alone was considered redundant for people living in supported housing. We also made very minor changes to the wording of 2 other questions in response to users' feedback (Table 2). We think it is unlikely that these modest changes substantially affected the SInQUE's psychometric properties. However, revalidation of the SInQUE in its web-based form is desirable in the future to confirm its validity and determine whether the minor wording changes should be retained for all versions of the SInQUE.

Implications for Practice

The SInQUE can be recommended as a potentially useful and acceptable tool for use in mental health supported accommodation settings, particularly supported housing services that offer 24-hour or 9-to-5 support, to provide a thorough assessment of social inclusion and support care planning. The tool may help meet an identified wish from service users for more discussion and support with social inclusion and relationships [41]. It was evident during the study that there is currently no universal tool in widespread use to help with social inclusion in mental health supported accommodation, highlighting the potential gap for an assessment tool such as the SInQUE. If used widely across supported accommodation services, the online SInQUE has the potential to provide benchmarking data and identify service users' most common priorities for greater social inclusion to inform service planning and evaluation.

Our findings also suggest that, for an assessment tool such as the SInQUE to be widely used, it is essential to have active leadership endorsement and support. For example, it may be required for managers or commissioners to direct staff to use the SInQUE with service users who are willing and reinforce this through team meetings, setting of use targets, or implementation of key performance indicators for its use.

Implications for Research

It is important to hear from staff and service users who chose not to use the online SInQUE to understand their reasons for not using the tool and highlight barriers to using the tool that we may have missed in this study. It would be useful to conduct further testing of the tool in residential care and floating outreach supported accommodation settings to better determine the utility of the SInQUE in these service types. It would also be useful to examine the utility of the SInQUE in other population groups within different service types to determine whether the tool may be useful in additional settings.

Future research is necessary to establish the level of uptake of the SInQUE that can be achieved in supported accommodation over a longer period and potentially establish more effective

XSL•FO
RenderX

means of implementation support. A longer-term study is also needed to establish whether the possible benefits from using the SInQUE that were mentioned by staff and service users are achievable through the use of the tool and how any potential outcomes may vary over time. A hybrid implementation-evaluation study would address these queries to determine the effectiveness of the SInQUE tool as an intervention for social inclusion and establish a precise implementation strategy for widespread uptake of the tool in supported accommodation. Further research using the SInQUE is also warranted to examine service user needs related to social inclusion and identify any additional barriers to addressing these needs in supported accommodation services. Such research could be used to inform the development of a future complex intervention to support social inclusion in supported accommodation services.

Although this study chose to examine the utility of the online SInQUE specifically in supported accommodation, the tool may also be useful in other mental health populations. Previous studies have established that the SInQUE can be used with mental health service users with a wide range of diagnoses [13,32,34]. Therefore, it is reasonable to extrapolate that the online SInQUE may be useful to assess social inclusion and inform support and care planning for other mental health service users, not just those living in mental health supported accommodation.

Acknowledgments

This paper presents independent research funded by the National Institute for Health and Care Research School for Social Care Research. The views expressed are those of the authors and not necessarily those of the National Institute for Health and Care Research or the Department of Health and Social Care.

The authors would like to thank the supported accommodation managers, staff, and service users who supported and participated in this study. They would also like to thank the members of their expert advisory group and the University College London Service User Research Forum for their advice and feedback throughout the study.

Data Availability

The qualitative data generated and analyzed during this study are not publicly available to preserve participant anonymity. The quantitative data generated during this study are available from the corresponding author (BLE) upon reasonable request.

Authors' Contributions

SE led recruitment, data collection, data analysis, and drafting of the paper. BLE led the study design and project management and supported the data analysis and drafting of the paper. HK, JC, PM, and GM co-designed the study and supervised the project. MD contributed to recruitment and data collection. GT contributed to the development of the study materials. All authors read, critically revised, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full description of the online Social Inclusion Questionnaire User Experience using the TIDieR (Template for Intervention Description and Replication) checklist for reporting interventions.
[DOCX File , 24 KB - ijmr_v13i1e45987_app1.docx]

Multimedia Appendix 2
Laboratory testing topic guide.
[DOCX File , 17 KB - ijmr_v13i1e45987_app2.docx]

Multimedia Appendix 3
Field-testing topic guides (for staff and service users).
[DOCX File , 16 KB - ijmr_v13i1e45987_app3.docx]

Multimedia Appendix 4
Explanation of suggested changes during laboratory testing and field-testing that were not made after team review.
[DOCX File , 18 KB - ijmr_v13i1e45987_app4.docx]

Multimedia Appendix 5
Interviews with staff and service users—illustrative quotes for each inductively derived subtheme.

XSL•FO
RenderX

[DOCX File , 22 KB - ijmr_v13i1e45987_app5.docx]

References

1. Burchardt T, Le Grand J, Piachaud D. Social exclusion in Britain 1991—1995. Soc Policy Adm 2002 Dec 17;33(3):227-244. [doi: 10.1111/1467-9515.00148]
2. Le Boutillier C, Croucher A. Social inclusion and mental health. Brit J Occup Ther 2010 Mar 01;73(3):136-139. [doi: 10.4276/030802210x12682330090578]
3. Morgan C, Burns T, Fitzpatrick R, Pinfold V, Priebe S. Social exclusion and mental health: conceptual and methodological review. Br J Psychiatry 2007 Dec;191:477-483. [doi: 10.1192/bjp.bp.106.034942] [Medline: 18055950]
4. Wright N, Stickley T. Concepts of social inclusion, exclusion and mental health: a review of the international literature. J Psychiatr Ment Health Nurs 2013 Feb;20(1):71-81. [doi: 10.1111/j.1365-2850.2012.01889.x] [Medline: 22369652]
5. Huxley P, Thornicroft G. Social inclusion, social quality and mental illness. Br J Psychiatry 2003 Apr;182:289-290. [doi: 10.1192/bjp.182.4.289] [Medline: 12668401]
6. Koenders JF, de Mooij LD, Dekker JM, Kikkert M. Social inclusion and relationship satisfaction of patients with a severe mental illness. Int J Soc Psychiatry 2017 Dec;63(8):773-781. [doi: 10.1177/0020764017737572] [Medline: 29067839]
7. McManus S, Meltzer H, Brugha T, Bebbington PE, Jenkins R. Adult psychiatric morbidity in England: results of a household survey. Health and Social Care Information Centre. 2009. URL: https://openaccess.city.ac.uk/id/eprint/23647/1/adul-psyc-morb-res-hou-sur-eng-2007-rep.pdf [accessed 2022-12-07]
8. Luciano A, Meara E. Employment status of people with mental illness: national survey data from 2009 and 2010. Psychiatr Serv 2014 Oct;65(10):1201-1209 [FREE Full text] [doi: 10.1176/appi.ps.201300335] [Medline: 24933361]
9. Boardman J, Grove B, Perkins R, Shepherd G. Work and employment for people with psychiatric disabilities. Br J Psychiatry 2003 Jun;182:467-468. [doi: 10.1192/bjp.182.6.467] [Medline: 12777333]
10. de Mooij LD, Kikkert M, Lommerse NM, Peen J, Meijwaard SC, Theunissen J, et al. Victimisation in adults with severe mental illness: prevalence and risk factors. Br J Psychiatry 2015 Dec;207(6):515-522. [doi: 10.1192/bjp.bp.113.143370] [Medline: 26450584]
11. Kamperman AM, Henrichs J, Bogaerts S, Lesaffre EM, Wierdsma AI, Ghauharali RR, et al. Criminal victimisation in people with severe mental illness: a multi-site prevalence and incidence survey in the Netherlands. PLoS One 2014;9(3):e91029 [FREE Full text] [doi: 10.1371/journal.pone.0091029] [Medline: 24609108]
12. Boardman J. Social exclusion and mental health – how people with mental health problems are disadvantaged: an overview. Ment Health Soc Incl 2011;15(3):112-121 [FREE Full text] [doi: 10.1108/20428301111165690]
13. Mezey G, White S, Harrison I, Bousfield J, Killaspy H, Lloyd-Evans B, et al. 'Modelling social exclusion in a diagnostically-mixed sample of people with severe mental illness'. Int J Soc Psychiatry 2022 Mar;68(2):420-428 [FREE Full text] [doi: 10.1177/00207640211001893] [Medline: 33730906]
14. Ware NC, Hopper K, Tugenberg T, Dickey B, Fisher D. Connectedness and citizenship: redefining social integration. Psychiatr Serv 2007 Apr;58(4):469-474. [doi: 10.1176/ps.2007.58.4.469] [Medline: 17412847]
15. Guidance for commissioners of rehabilitation services for people with complex mental health needs. Joint Commissioning Panel for Mental Health. 2016 Nov. URL: https://www.rcpsych.ac.uk/docs/default-source/members/faculties/rehabilitation-and-social-psychiatry/rehab-social—joint-commissioning-panel—guidance-for-commissioners-of-rehabilitation-services-for-people-with-complex-mental-health-needs—2016.pdf?sfvrsn=82cbb5e4_6 [accessed 2022-11-03]
16. Killaspy H, Priebe S. Research into mental health supported accommodation - desperately needed but challenging to deliver. Br J Psychiatry 2021 Apr;218(4):179-181. [doi: 10.1192/bjp.2020.74] [Medline: 32321596]
17. Killaspy H. Supported accommodation for people with mental health problems. World Psychiatry 2016 Feb;15(1):74-75 [FREE Full text] [doi: 10.1002/wps.20278] [Medline: 26833612]
18. Krotofil J, McPherson P, Killaspy H. Service user experiences of specialist mental health supported accommodation: a systematic review of qualitative studies and narrative synthesis. Health Soc Care Community 2018 Nov;26(6):787-800. [doi: 10.1111/hsc.12570] [Medline: 29609195]
19. Bitter NA, Roeg DP, van Nieuwenhuizen C, van Weeghel J. Identifying profiles of service users in housing services and exploring their quality of life and care needs. BMC Psychiatry 2016 Nov 23;16(1):419 [FREE Full text] [doi: 10.1186/s12888-016-1122-0] [Medline: 27881159]
20. de Heer-Wunderink C, Visser E, Caro-Nienhuis A, Sytema S, Wiersma D. Supported housing and supported independent living in the Netherlands, with a comparison with England. Community Ment Health J 2012 Jun;48(3):321-327 [FREE Full text] [doi: 10.1007/s10597-011-9381-1] [Medline: 21246274]
21. McPherson P, Krotofil J, Killaspy H. Mental health supported accommodation services: a systematic review of mental health and psychosocial outcomes. BMC Psychiatry 2018 May 15;18(1):128 [FREE Full text] [doi: 10.1186/s12888-018-1725-8] [Medline: 29764420]
22. Bitter N, Roeg D, van Nieuwenhuizen C, van Weeghel J. Recovery in supported accommodations: a scoping review and synthesis of interventions for people with severe mental illness. Community Ment Health J 2020 Aug;56(6):1053-1076 [FREE Full text] [doi: 10.1007/s10597-020-00561-3] [Medline: 32016620]

XSL•FO
RenderX

23. Piat M, Sofouli E, Sabetti J, Lambrou A, Chodos H, Briand C, et al. Protocol for a mixed studies systematic review on the implementation of the recovery approach in adult mental health services. BMJ Open 2017 Aug 30;7(8):e017080 [FREE Full text] [doi: 10.1136/bmjopen-2017-017080] [Medline: 28855202]
24. Macpherson R, Shepherd G, Thyarappa P. Supported accommodation for people with severe mental illness: an update. Adv Psychiatr Treat 2012;18(5):381-391. [doi: 10.1192/apt.bp.110.008714]
25. Sandhu S, Priebe S, Leavey G, Harrison I, Krotofil J, McPherson P, et al. Intentions and experiences of effective practice in mental health specific supported accommodation services: a qualitative interview study. BMC Health Serv Res 2017 Jul 11;17(1):471 [FREE Full text] [doi: 10.1186/s12913-017-2411-0] [Medline: 28693490]
26. Killaspy H, Priebe S, McPherson P, Zenasni Z, Greenberg L, McCrone P, et al. Predictors of moving on from mental health supported accommodation in England: national cohort study. Br J Psychiatry 2020 Jun;216(6):331-337. [doi: 10.1192/bjp.2019.101] [Medline: 31046864]
27. Roe D, Slade M, Jones N. The utility of patient-reported outcome measures in mental health. World Psychiatry 2022 Feb;21(1):56-57 [FREE Full text] [doi: 10.1002/wps.20924] [Medline: 35015343]
28. Guidance and technical packages on community mental health services: promoting person-centred and rights-based approaches. World Health Organization. 2021 May 26. URL: https://www.who.int/publications/i/item/guidance-and-technical-packages-on-community-mental-health-services [accessed 2022-09-11]
29. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
30. Berry N, Lobban F, Emsley R, Bucci S. Acceptability of interventions delivered online and through mobile phones for people who experience severe mental health problems: a systematic review. J Med Internet Res 2016 May 31;18(5):e121 [FREE Full text] [doi: 10.2196/jmir.5250] [Medline: 27245693]
31. Killaspy H, White S, Dowling S, Krotofil J, McPherson P, Sandhu S, et al. Adaptation of the Quality Indicator for Rehabilitative Care (QuIRC) for use in mental health supported accommodation services (QuIRC-SA). BMC Psychiatry 2016 Apr 14;16:101 [FREE Full text] [doi: 10.1186/s12888-016-0799-4] [Medline: 27075574]
32. Mezey G, White S, Thachil A, Berg R, Kallumparam S, Nasiruddin O, et al. Development and preliminary validation of a measure of social inclusion for use in people with mental health problems: the SInQUE. Int J Soc Psychiatry 2013 Aug;59(5):501-507. [doi: 10.1177/0020764012443752] [Medline: 22511274]
33. Baumgartner JN, Burns JK. Measuring social inclusion--a key outcome in global mental health. Int J Epidemiol 2014 Apr;43(2):354-364. [doi: 10.1093/ije/dyt224] [Medline: 24306926]
34. Mezey G, White S, Harrison I, Bousfield J, Lloyd-Evans B, Payne S, et al. Validity, reliability, acceptability, and utility of the Social Inclusion Questionnaire User Experience (SInQUE): a clinical tool to facilitate social inclusion amongst people with severe mental health problems. Soc Psychiatry Psychiatr Epidemiol 2020 Jul;55(7):953-964 [FREE Full text] [doi: 10.1007/s00127-019-01826-3] [Medline: 32016512]
35. SInQUE home page. SInQUE. URL: https://sinque.org.uk [accessed 2022-12-07]
36. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
37. Steare T, O'Hanlon P, Eskinazi M, Osborn D, Lloyd-Evans B, Jones R, et al. Smartphone-delivered self-management for first-episode psychosis: the ARIES feasibility randomised controlled trial. BMJ Open 2020 Aug 26;10(8):e034927 [FREE Full text] [doi: 10.1136/bmjopen-2019-034927] [Medline: 32847902]
38. Mukumbang FC, Marchal B, Van Belle S, van Wyk B. Using the realist interview approach to maintain theoretical awareness in realist studies. Qual Res 2020;20(4):485-515. [doi: 10.1177/1468794119881985]
39. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
40. Priebe S, Saidi M, Want A, Mangalore R, Knapp M. Housing services for people with mental disorders in England: patient characteristics, care provision and costs. Soc Psychiatry Psychiatr Epidemiol 2009 Oct;44(10):805-814. [doi: 10.1007/s00127-009-0001-0] [Medline: 19277440]
41. Pinfold V, Sweet D, Porter I, Quinn C, Byng R, Griffiths C, et al. Improving community health networks for people with severe mental illness: a case study investigation. National Institute for Health Research. 2015. URL: https://eprints.whiterose.ac.uk/84845/2/ScientificSummary_hsdr03050.pdf [accessed 2022-07-12]

Abbreviations

ICAMO: intervention-context-actor-mechanism-outcome
NHS: National Health Service
REC: Research Ethics Committee
SInQUE: Social Inclusion Questionnaire User Experience
TIDieR: Template for Intervention Description and Replication
UCL: University College London

XSL•FO
RenderX

Edited by T de Azevedo Cardoso; submitted 25.01.23; peer-reviewed by P Dabas, G Leavey, L Yang, B Hannigan; comments to author 14.11.23; revised version received 08.12.23; accepted 10.02.24; published 13.03.24.

Please cite as:
Eager S, Killaspy H, C J, Mezey G, McPherson P, Downey M, Thompson G, Lloyd-Evans B
A Web-Based Tool to Assess Social Inclusion and Support Care Planning in Mental Health Supported Accommodation: Development and Preliminary Test Study
Interact J Med Res 2024;13:e45987
URL: https://www.i-jmr.org/2024/1/e45987
doi:10.2196/45987
PMID:38477978

©Sharon Eager, Helen Killaspy, Joanna C, Gillian Mezey, Peter McPherson, Megan Downey, Georgina Thompson, Brynmor Lloyd-Evans. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 13.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Case Report

Treating Spontaneous Pneumothorax Using an Innovative Surgical Technique Called Capnodissection Pleurectomy: Case Report

Ghaith Qsous[1], MRCS, MD; Prashanth Ramaraj[1,2], MSc, MRCS; Sanjeet Singh Avtaar Singh[1], MRCS, PhD; Philip Herd[3], MBChB; Nayandra Runveer Sooraj[3], MMed, FCA (SA); Malcolm Brodie Will[1], FRCS (CTh), PhD

[1]Department of Cardiothoracic Surgery, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom

[2]Imperial College School of Medicine, Imperial College London, London, United Kingdom

[3]Department of Anaesthesia, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Prashanth Ramaraj, MSc, MRCS
Department of Cardiothoracic Surgery
Royal Infirmary of Edinburgh
Little France Crescent
Old Dalkeith Road
Edinburgh, EH16 4SA
United Kingdom
Phone: 44 01315361000 ext 21076
Email: Prashanth.ramaraj@nhs.net

Abstract

Spontaneous pneumothorax is one of the most common conditions encountered in thoracic surgery. This condition can be treated conservatively or surgically based on indications and guidelines. Traditional surgical management includes pleurodesis (mechanical or chemical) in addition to bullectomy if the bullae can be identified. Mechanical pleurodesis is usually performed by surgical pleurectomy or pleural abrasion. In this case report, we present a case of a young patient with spontaneous pneumothorax who needed a surgical intervention. We performed a new, innovative surgical technique for surgical pleurectomy where we used carbon dioxide for dissection of the parietal pleura (capnodissection). This technique may provide similar efficiency to the traditional procedure but with less risk of bleeding and complications.

(Interact J Med Res 2024;13:e54497) doi:10.2196/54497

KEYWORDS

capnodissection; pleurectomy; VATS; video-assisted thorascopic surgery; novel technique; thoracic surgery; surgical innovation; pneumothorax; spontaneous pneumothorax; pleurodesis; management; bullectomy; bullae; young patient; lung diseases; chronic obstructive pulmonary disease; COPD; surgical treatment; male; capnothorax

Introduction

Spontaneous pneumothorax (SP) is a condition in which pneumothorax occurs without trauma or iatrogenic cause. It can be classified as a primary SP if there is no obvious underlying lung disease. The most common cause is usually a small bulla or bleb in the lung [1,2]. Comparatively, secondary SP happens due to underlying lung diseases such as chronic obstructive pulmonary disease [3]. The new British Thoracic Society (BTS) guidelines advise surgical treatment for SP at initial presentation if recurrence prevention is deemed important (eg, patients presenting with tension pneumothorax or those in high-risk occupations). Elective surgery should be considered for patients with a second ipsilateral or first contralateral pneumothorax [4]. The surgical treatment that is recommended by the BTS guidelines for SP is surgical pleurodesis with or without bullectomy [4]. There are 2 common ways to perform surgical pleurodesis: the first one is surgical pleurectomy and the second one is pleural abrasion. Surgical pleurectomy is considered more efficient, but it can be associated with an increased risk of bleeding and infection [5]. The novel surgical technique that we provide here can give a similar success rate but with less risk of complications such as bleeding or infection.

Case Presentation

Our patient was a young male individual who was previously healthy. He presented with a recurrence of an SP for the first time (2 SPs in total). The previous episode was treated conservatively 7 months prior, and his computed tomography

XSL•FO
RenderX

scan for this episode showed that he had small apical bullae. The decision was made to list the patient for elective surgical treatment, and after discussion with the patient, he was listed for a pleurectomy and bullectomy.

A standard anterior video-assisted thorascopic surgery (VATS) approach was taken. A small incision was made at the sixth intercostal space, and another small port site was created for the camera, which was later converted into the drain site (Figure 1). Carbon dioxide (CO_2) insufflation at 6-8 mm Hg on high flow was used to achieve capnothorax. A small anterior VATS incision was made at the sixth intercostal space, and the dissection of the parietal pleura was performed extrapleurally using Roberts forceps with a traditional technique. The forceps were exchanged for a curved metal sucker, and the CO_2 insufflation was attached at high flow and used to mobilize the whole parietal pleura, first from apex to inferior and then from posterior to anterior (Multimedia Appendix 1). The posterior parietal pleura was then excised off the ribs using thorascopic scissors 4 cm from the sympathetic chain posteriorly, 2 cm lateral to the internal mammary vein anteriorly, and 2 cm cranially to the diaphragm. Lastly, a bullectomy was performed using a manual stapler to excise the presumed culprit apical bullae seen on the computed tomography scan. Blood loss was minimal, approximately 50 mL, predominantly from VATS entry. Operative time was approximately 40 minutes. The postoperative care was routine, and the drain was removed after 48 hours. The patient was discharged on the third postoperative day.

Figure 1. Video-assisted thorascopic surgery (VATS) ports setup and surgical instruments.

Ethical Considerations

On the day of the operation, the patient completed a written consent form. The patient kindly agreed to the recording of the procedure and the utilization of his nonidentifiable data for this case report and publication, which was further discussed between the patient and GQ. Consent was sought by GQ and given by those in the operating theater for recording of the technique and publication as a case report.

Discussion

The use of CO_2 in thoracic surgery has increased significantly with the growing use of a minimally invasive approach. Capnothorax leads to better visualization by collapsing the lung and reduces the rate of complications [6,7]. In our department, we usually use CO_2 with robot-assisted thorascopic surgery and VATS for these reasons.

Surgical pleurodesis of SP is the recommended treatment in the BTS guidelines because it gives better long-term outcomes with less risk of recurrence in the future [4]. Surgical pleurectomy, in spite of it is efficiency, carries a risk of bleeding, infection, and reoperation [8]. Surgical pleural abrasion is another method that can be used for surgical pleurodesis. Chang et al [9] published the first systematic review and meta-analysis that compared surgical abrasion against surgical apical pleurectomy. They found that there is no difference in the recurrence, but pleural abrasion has a shorter length of stay in hospital, postoperative chest tube duration, and operative time and less surgical blood loss [9]. This may cause clinicians to consider a change of practice from surgical pleurectomy to abrasion. A systematic review of randomized controlled trials found that SPs managed with a chest drain alone had recurrence rates that ranged from 26.1% to 50.1%, whereas after VATS talc pleurodesis, these ranged from 0% to 3.2%. Alternative chemical pleurodesis can be achieved with tetracycline rather than talc, although recurrence rates were reported as ranging from 13% to 33.3% [10].

Our literature search did not find any studies in which capnodissection was used for pleurectomy as a treatment of SP. However, Dai, et al [11] recently published their findings for using CO_2 for visceral pleurectomy and decortication in patients with malignant mesothelioma. They found that the positive pressure of CO_2 can facilitate dissection of the visceral pleura, making the procedure easier while achieving an acceptable postoperative air leak and chest drain output. They concluded that capnodissection can be used in pleurectomy and decortication for patients with mesothelioma [11]. It should be

XSL•FO
RenderX

taken into consideration that although the effect of capnodissection on gas exchange has not been rigorously studied, there is evidence to show that hypercarbia can result from CO_2 insufflation for capnothorax during VATS or robot-assisted thorascopic surgery procedures [12]. This must be taken into account by surgeons and anesthetists when considering compensatory ventilator strategies, especially in patients with compromised gas exchange.

Our experience with the use of capnodissection for surgical pleurectomy was successful, and after 17 months from the procedure, the patient did not have any recurrence or complications. Moreover, this technique was not time-consuming (40-minute operative time), and the patient was discharged after 48 hours, with the surgeons noticing less pain in comparison to the traditional surgical pleurectomy, although pain is subjective. After VATS talc pleurodesis, the chest drain is typically removed no sooner than the second postoperative day, with discharge later that day. There is a theoretical risk of increased recurrence, as while the relatively atraumatic nature of this technique may reduce patient pain, it may also reduce the proinflammatory process required for pleurodesis and hence recurrence prevention [13]. More cases and longer follow-up are required to investigate the noninferiority of our technique to the traditional procedure.

Conclusions

In this case, capnodissection of the parietal pleura was a novel, safe, and successful technique that may decrease the risk of bleeding and postoperative pain.

Acknowledgments

The authors would like to thank the Imperial College London Library for access to article processing charges (APCs) waiver through the Imperial-JMIR unlimited Open Access (OA) deal.

Data Availability

Data regarding this paper have been deposited in the supplemental files. Corresponding author PR will enable further data sharing upon reasonable request.

Authors' Contributions

GQ conceived the idea of the case report, assisted during this procedure, and lead the manuscript writing. PR was involved in manuscript writing, internal review, and internal editing. SSAS was involved in the internal review. PH and NRS were anesthetists for this case. MBW was the operating surgeon.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Video of capnodissection in action with surgeon narrative.
[MP4 File (MP4 Video), 17601 KB - ijmr_v13i1e54497_app1.mp4]

References

1. Louw EH, Shaw JA, Koegelenberg CFN. New insights into spontaneous pneumothorax: a review. Afr J Thorac Crit Care Med 2021 Mar 09;27(1):10.7196/AJTCCM.2021.v27i1.054. [doi: 10.7196/ajtccm.2021.v27i1.054] [Medline: 34240041]
2. Ghisalberti M, Guerrera F, de Vico A, Bertolaccini L, de Palma A, Fiorelli A, et al. Age and clinical presentation for primary spontaneous pneumothorax. Heart Lung Circ 2020 Nov;29(11):1648-1655. [doi: 10.1016/j.hlc.2020.05.107] [Medline: 32718900]
3. Nava GW, Walker SP. Management of the secondary spontaneous pneumothorax: current guidance, controversies, and recent advances. J Clin Med 2022 Feb 22;11(5):1173 [FREE Full text] [doi: 10.3390/jcm11051173] [Medline: 35268264]
4. Roberts ME, Rahman NM, Maskell NA, Bibby AC, Blyth KG, Corcoran JP, et al. British Thoracic Society guideline for pleural disease. Thorax 2023 Nov 08;78(11):1143-1156. [doi: 10.1136/thorax-2023-220304] [Medline: 37553157]
5. Ocakcioglu I, Kupeli M. Surgical treatment of spontaneous pneumothorax: pleural abrasion or pleurectomy? Surg Laparosc Endosc Percutan Tech 2019 Feb;29(1):58-63. [doi: 10.1097/SLE.0000000000000595] [Medline: 30499890]
6. Gallego-Poveda J, Guerra NC, Carvalheiro C, Ferreira H, Sena A, Junqueira N, et al. Use of CO in video assisted thoracic surgery and single-lumen endotracheal tube-a new less invasive approach. J Thorac Dis 2017 Apr;9(4):903-906 [FREE Full text] [doi: 10.21037/jtd.2017.01.53] [Medline: 28523135]
7. Brock H, Rieger R, Gabriel C, Pölz W, Moosbauer W, Necek S. Haemodynamic changes during thoracoscopic surgery the effects of one-lung ventilation compared with carbon dioxide insufflation. Anaesthesia 2000 Jan;55(1):10-16 [FREE Full text] [doi: 10.1046/j.1365-2044.2000.01123.x] [Medline: 10594427]

XSL•FO
RenderX

8. Körner H, Andersen KS, Stangeland L, Ellingsen I, Engedal H. Surgical treatment of spontaneous pneumothorax by wedge resection without pleurodesis or pleurectomy. Eur J Cardiothorac Surg 1996;10(8):656-659. [doi: 10.1016/s1010-7940(96)80381-2] [Medline: 8875174]
9. Chang J, Ratnaraj V, Fu V, Jiang M, Peri V, Nguyenhuy M, et al. Pleural abrasion versus apical pleurectomy for primary spontaneous pneumothorax: a systematic review and meta-analysis. J Cardiothorac Surg 2023 Apr 06;18(1):105 [FREE Full text] [doi: 10.1186/s13019-023-02207-3] [Medline: 37024894]
10. Hallifax RJ, Yousuf A, Jones HE, Corcoran JP, Psallidas I, Rahman NM. Effectiveness of chemical pleurodesis in spontaneous pneumothorax recurrence prevention: a systematic review. Thorax 2017 Dec 01;72(12):1121-1131 [FREE Full text] [doi: 10.1136/thoraxjnl-2015-207967] [Medline: 27803156]
11. Dai J, Liu M, Liu X, Li J, Jin K, Chen L, et al. Carbon dioxide blower facilitates visceral pleurectomy in malignant pleural mesothelioma. Ann Thorac Surg 2022 Jul;114(1):e71-e74. [doi: 10.1016/j.athoracsur.2021.09.046] [Medline: 34715083]
12. Tran DTT, Badner NH, Nicolaou G, Sischek W. Arterial pCO2 changes during thoracoscopic surgery with CO2 insufflation and one lung ventilation. HSR Proc Intensive Care Cardiovasc Anesth 2010;2(3):191-197 [FREE Full text] [Medline: 23439224]
13. van den Heuvel MM, Smit HJ, Barbierato SB, Havenith CE, Beelen RH, Postmus PE. Talc-induced inflammation in the pleural cavity. Eur Respir J 1998 Dec 01;12(6):1419-1423 [FREE Full text] [doi: 10.1183/09031936.98.12061419] [Medline: 9877502]

Abbreviations

BTS: British Thoracic Society
CO_2: carbon dioxide
SP: spontaneous pneumothorax
VATS: video-assisted thorascopic surgery

Edited by T de Azevedo Cardoso; submitted 12.11.23; peer-reviewed by UK Chalwadi, A Chambers; comments to author 16.02.24; revised version received 02.04.24; accepted 27.05.24; published 21.06.24.

Please cite as:
Qsous G, Ramaraj P, Avtaar Singh SS, Herd P, Sooraj NR, Will MB
Treating Spontaneous Pneumothorax Using an Innovative Surgical Technique Called Capnodissection Pleurectomy: Case Report
Interact J Med Res 2024;13:e54497
URL: https://www.i-jmr.org/2024/1/e54497
doi:10.2196/54497
PMID:

©Ghaith Qsous, Prashanth Ramaraj, Sanjeet Singh Avtaar Singh, Philip Herd, Nayandra Runveer Sooraj, Malcolm Brodie Will. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 21.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Case Report

Radiological Progression of Degenerative Cervical Myelopathy in a Clinically Stable Patient: Case Report

Rishi Umeria[1], MA, MB BChir; Oliver Mowforth[1], MA, MB BChir, MSt, MRCS; Munashe Veremu[1], BSc, MBBS; Benjamin Davies[1], MBChB, BSc, MRCS; Mark Kotter[1], MD, MPhil, PhD

Division of Neurosurgery, Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Mark Kotter, MD, MPhil, PhD
Division of Neurosurgery
Department of Clinical Neurosciences
University of Cambridge
Addenbrooke's Hospital
Hills Road
Cambridge, CB2 0QQ
United Kingdom
Phone: 44 1223 336946
Email: mrk25@cam.ac.uk

Abstract

Degenerative cervical myelopathy (DCM) is a common neurological condition, with disease progression that is both variable and difficult to predict. Here, we present a case of DCM in a gentleman in his late 60s with significant radiological disease progression without consequent change in clinical symptoms. The case serves as a reminder of an enduring medical aphorism that clinical history and examination should be prioritized above more complex data, such as imaging investigations. In addition, the case also highlights that guidelines should be contextualized within individual clinical circumstances.

(Interact J Med Res 2024;13:e48212) doi:10.2196/48212

KEYWORDS

degenerative cervical myelopathy; neurosurgery; radiology; magnetic resonance imaging

Introduction

Degenerative cervical myelopathy (DCM) is the most common cause of adult spinal cord dysfunction worldwide and is estimated to affect up to 2% of adults [1,2]. It arises secondary to degenerative pathology in the cervical spine, which leads to spinal cord compression. Cord compression may precipitate progressive neurological deficits including motor, sensory, and sphincter dysfunction [3].

The radiological progression of DCM does not always correlate well with its clinical progression [4-6]. Imaging findings such as the level and severity of spinal cord compression may not match the patient's experience of their symptoms over time, thereby adding complexity to the clinician's role of interpreting and reconciling clinical and radiological features of the disease [1,7].

Guidelines are clear in the event of symptom progression and advocate for surgery [8]. In contrast, the management of radiological progression without worsening symptoms is less defined and remains controversial. This report describes a case of the latter scenario, with the discussion focused on management.

Case Presentation

A retired man in his late 60s presented reporting a 2-year history of lower limb weakness, impaired balance, and calf aches. He associated this with a preceding episode of influenza.

He had a background of long-standing lower back pain secondary to lumbar stenosis, for which he had previously undergone 3 surgical decompressions and atrial fibrillation, for which he took aspirin daily. His father had undergone a cervical laminectomy many years previously. The patient was otherwise fit and well, regularly cycled for exercise, and was a nonsmoker.

His symptoms resulted in a referral to neurology and investigations including magnetic resonance imaging (MRI) of his cervical spine. The MRI revealed multilevel degenerative changes from C3 to C7 in the cervical spine, on a background of a congenitally narrow spinal canal (Figure 1). He was

XSL•FO
RenderX

therefore referred to the neurosurgery department. Consent was obtained from the patient.

On assessment in the neurosurgery clinic, the patient was diagnosed with DCM. His modified Japanese Orthopaedic Association (mJOA) score was 14 (4+6+2+2); a score of 12 to 14 indicates moderate myelopathy. Further exploration of his symptoms revealed dysesthesia in the region of the left shoulder, hypoesthesia of the third and fourth digits of the left hand, difficulties with tandem walking, and bladder issues with minor episodes of urinary incontinence. He was reluctant to undergo surgery, and hence, it was agreed to monitor his symptoms with reassessment in the clinic, with a strong recommendation to consider surgery if there was further progression.

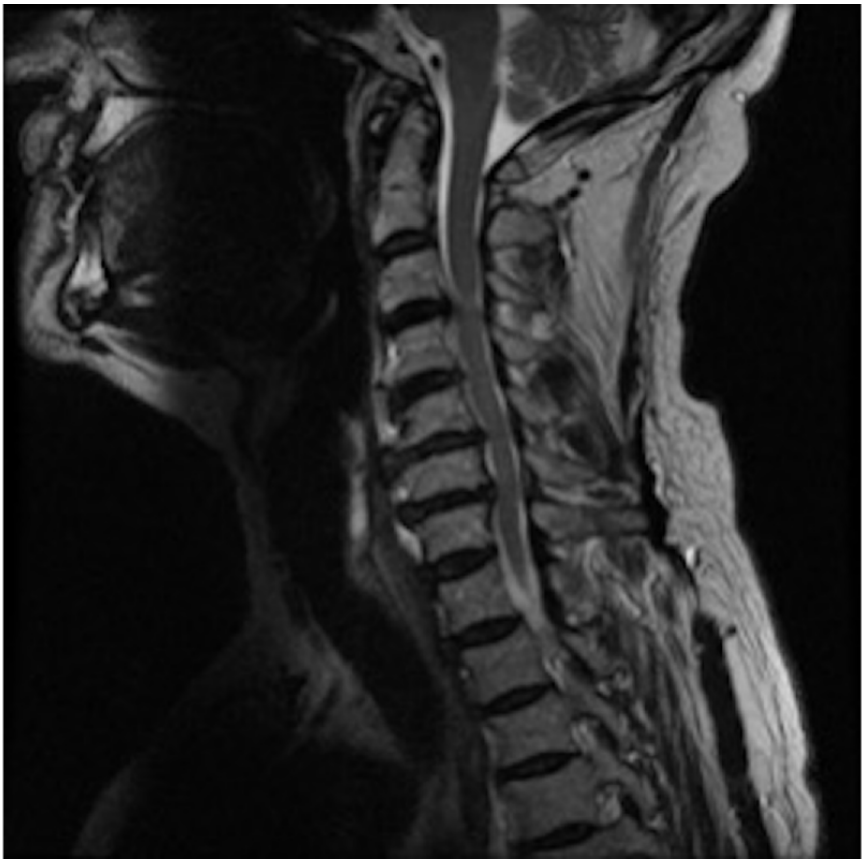
The appointment 5 months later demonstrated no further progression of his symptoms, with an unchanged mJOA score. A further 6 months later, at a third neurosurgical clinic appointment, the patient reported some deterioration in his condition; he had been finding cycling more difficult, felt lower back stiffness, had worsened pain, and reported a mild electric shock sensation on the right upper limb. His mJOA score remained unchanged. He underwent a repeat MRI of his cervical spine, with a view to considering surgical decompression. The MRI showed progression of the degenerative changes at C3 to C7 (Figure 2).

Nonetheless, at his subsequent follow-up appointment, the patient reported improvement in the symptoms and the mJOA score remained 14. He had, however, developed left L4 sciatica, which had become his main concern. Three months later at his fifth appointment, his mJOA score was 13 (4+5+2+2). He reported increasing work and concentration required to walk and climb stairs. Compared to the third appointment, there had been a slow, gradual clinical progression. At this stage, the patient was put on the waiting list for surgery, just before the national lockdown in the United Kingdom during the COVID-19 pandemic.

At 3 subsequent follow-up appointments over the succeeding 8 months, there was no further clinical progression of the patient's DCM. Therefore, it was agreed to remove him from the waiting list for surgery and continue with an expectant approach. At the ninth appointment 6 months later, symptoms were again stable. At the 10th appointment, the patient reported some worsening back and lower limb pain, as well as some deterioration in his general mobility. A further MRI was therefore requested, which demonstrated significant radiological progression of his DCM compared to his MRI 5 years earlier (Figure 3).

At the 11th appointment, 5 months later, the symptoms had improved to the patient's previous baseline, including balance, dexterity, numbness, and urinary urgency. His mJOA remained 13. The radiological progression was in stark contrast with the fact that the patient remained clinically stable. Therefore, it was agreed to continue with expectant management and continue close clinical follow-up.

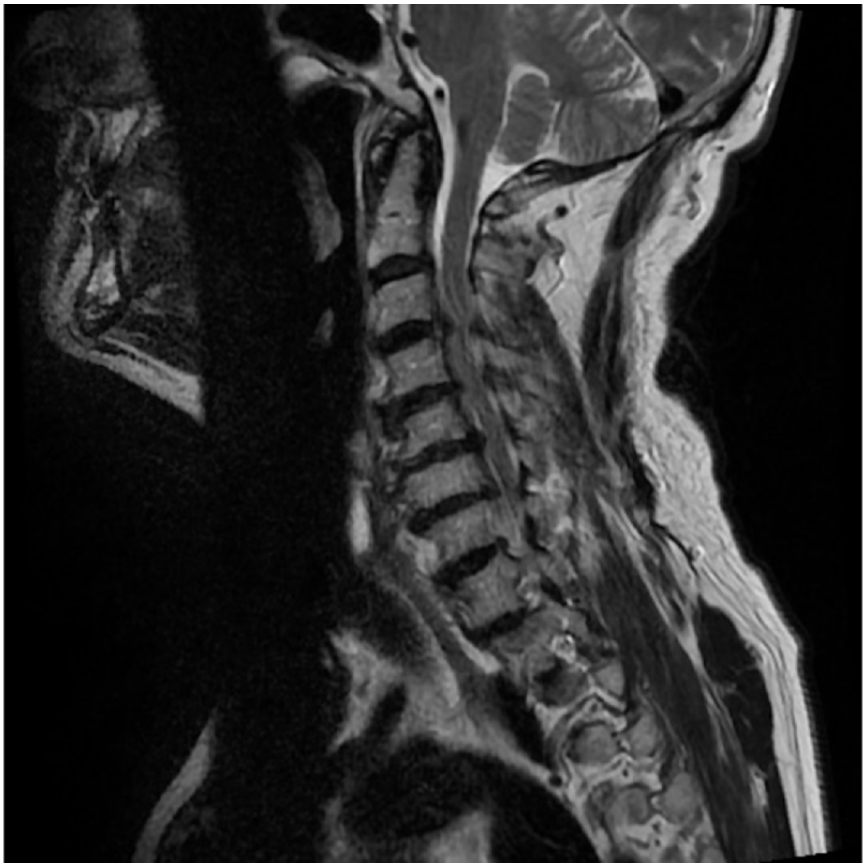
Figure 1. T2-weighted sagittal magnetic resonance imaging demonstrating multilevel degenerative cervical spondylosis and disc degeneration.



XSL•FO
RenderX

Figure 2. T2-weighted sagittal MRI demonstrating the progression of multilevel degenerative changes in the cervical spine, including disc degeneration at C3-C7.

Figure 3. T2-weighted sagittal magnetic resonance imaging showing significant radiological worsening of cervical cord compression at C3-C7 levels despite unchanged symptoms and modified Japanese Orthopaedic Association clinical severity score.



XSL•FO
RenderX

Ethical Considerations

According to the National Health Service Health Research Authority, research ethics committee approval was not required for this work. Informed patient consent was obtained for publication.

Investigations

The first MRI showed multilevel degenerative changes of the neck on the background of a congenitally narrow spinal canal: C3/4—disc osteophyte complex causing central compression of the spinal cord; C4/5—moderate stenosis; C5/6—moderate stenosis; C6/7—severe compression, with associated T2 signal changes.

The second MRI showed disc and facet joint degenerative changes between C3 and C7: C3/4—central disc protrusion causing moderate spinal canal stenosis and compression of the spinal cord, which had progressed compared to the first MRI; C4/5—broad-based disc bulge causing mild spinal canal narrowing and indenting the undersurface of the cord; C5/6—broad-based disc bulge causing moderate spinal canal narrowing and mildly compressing the spinal cord, which had progressed compared to the first MRI; C6/7—broad-based disc bulge causing mild to moderate spinal canal narrowing, indenting the anterior surface of the spinal cord.

The third MRI scan showed multilevel degenerative changes from C3 to C7, with scoliotic deformity: C3/4—severe circumferential stenosis with cord compression; C4/5—significant circumferential stenosis with cord compression; C5/6—circumferential stenosis with cord compression; C6/7—circumferential stenosis with cord compression.

Differential Diagnosis

The differential diagnosis for a presentation of lower limb weakness, imbalance, and calf ache can be divided into upper and lower motor neuron patterns of weakness. Causes for the upper motor neuron pattern include pathologies in the brain and spinal cord. In the brain, this includes demyelinating disorders such as multiple sclerosis, vascular disorders such as stroke, space-occupying lesions such as a parasagittal meningioma or abscess, and motor neuron diseases such as amyotrophic lateral sclerosis. In the spinal cord, this includes demyelinating disorders such as transverse myelitis, myelopathies such as DCM, space-occupying lesions, for example, tumor or abscess, trauma, syringomyelia, and spinal stenosis (spinal claudication).

Causes for the lower motor neuron pattern include drugs such as alcohol, metabolic disorders such as vitamin B_{12} deficiency, diabetes mellitus, inherited disorders such as Charcot-Marie-Tooth, infections such as HIV or syphilis, and autoimmune disorders such as vasculitis and chronic inflammatory demyelinating polyneuropathy. The causes of calf aches include trauma, vascular disorders such as peripheral vascular disease (intermittent claudication), and inflammatory disorders such as myositis.

Treatment

Since being diagnosed 5 years ago, an expectant approach has been taken in the management of this patient. International DCM management guidelines recommend surgical management for moderate, severe, or progressive DCM; however, for mild DCM, the optimal treatment strategy remains undefined, with a recommendation of either surgery or supervised nonsurgical management [8]. In this patient's case, there is moderate DCM, with an mJOA score of 13, with significant radiological progression but clinical and symptomatic stability across serial assessments.

Outcome and Follow-Up

The patient continues close neurosurgical follow-up, currently at 12 monthly intervals, alongside careful safety netting advice.

Discussion

Principal Findings

Degenerative changes in the cervical spine include disc herniation, osteophytosis, ligament hypertrophy, and ossification [9]. DCM is a clinical syndrome that arises when these changes result in spinal cord compression that is associated with symptoms, which may include neck pain or stiffness, limb pain or weakness, urinary incontinence, decreased manual dexterity, imbalance, or falls [1]. In this patient's case, there was also a risk factor for congenital stenosis of the cervical canal.

A challenging aspect of the management of DCM is how to deal with symptoms changing over time and correlating this with evolving imaging findings. Another challenging aspect is the important decision on the timing of any surgical management.

The guidelines advise clinicians to take a structured, consistent approach to management. Each person with DCM requires consideration of individual factors, which may mean that clinical judgment or patient preferences result in deviation from guidelines in some circumstances. The patient was diagnosed with moderate DCM; the mJOA score for the patient was initially 14 and then decreased and remained stable at 13.

Strict application of the guidelines would lead to a recommendation for surgical intervention. However, while the mJOA includes consideration of upper and lower limb motor function, upper limb sensory function, and sphincter function, it does not capture all symptoms and clinical features. For example, in this case, limb pain was not captured. Nonetheless, it is a validated scoring tool for the assessment of functional status and is responsive to changes in the severity of DCM [10].

The complexity of this case requires a nuanced approach to management. Surgical intervention within DCM is primarily aimed at halting symptom progression; however, without symptom worsening, the decision of when to operate becomes more complicated. This patient's symptoms were managed expectantly with nonsurgical interventions such as physical therapy as tolerated, oral analgesics, and neuropathic agents for any acute pain flares. Urinary symptoms were stable and were not actively managed. Waiting to operate at an older age may increase the risk of further complications, and this possibility

XSL•FO
RenderX

should be explained to the patient. This should be part of a shared decision-making approach, where patients are empowered to make decisions through collaboration with their clinicians with the understanding that, in the context of a chronic disease like DCM, this decision will likely be revisited [11].

The factors driving the disconnect between clinical and radiological severity are unknown. Nonetheless, a model proposed by Davies et al [12] postulates that DCM is a function of (1) mechanical stress, (2) duration of injury, and (3) individual vulnerability. Using this model, a scenario of limited clinical progression and significant radiological progression over time could be explained by decreased vulnerability to injury.

An individual's vulnerability to DCM comprises primary protective mechanisms such as genetics and age, in addition to adaptive protective mechanisms, such as autoregulation of spinal cord perfusion, functional reserve capacity, and nutritional status [12]. For example, certain genotypes are associated with increased regenerative capacity, such as the *HIF-1A* polymorphism rs11549467 [13]. This polymorphism is associated with susceptibility to DCM and its clinical features, including severity measured by mJOA.

Furthermore, adaptive protective mechanisms such as autoregulation of spinal cord perfusion may minimize ischemic injury. Decreased blood flow can result in blood-spinal cord barrier dysfunction, leading to microglia activation, neuroinflammation, and neuronal apoptosis [14]. In addition to the ischemia precipitating apoptosis [14], dysregulation of the autoregulatory system can occur from mechanical cord compression in DCM [15-17]. It is possible that dysregulation occurs to a lesser degree in some individuals, such as this patient. Furthermore, reserve capacity within the central nervous system [17-19] refers to resilience in the neurological system to account for the disconnect between the clinical phenotype and underlying histological pathology. In the context of DCM, cervical spinal cord compression and injury are initially asymptomatic, and the radiological changes affecting the spinal cord represent, at best, a proxy for the clinical presentation of the condition [12].

Conclusions

In summary, individuals may have different vulnerabilities and protective mechanisms that may account for the disconnect between clinical and radiological features of DCM. The message is therefore to treat the patient rather than treating the findings from the imaging. This is especially important, given that asymptomatic cervical spinal cord compression is common in the general population [2,7].

Acknowledgments

We gratefully acknowledge support from the Cambridge NIHR Brain Injury MedTech Cooperative. MK is funded by a National Institute for Health Research (NIHR) Clinician Scientist Award (CS-2015-15-023). BD is supported by an NIHR Clinical Doctoral Research Fellowship. OM is supported by an NIHR Academic Clinical Fellowship. This report is independent research arising from a Clinician Scientist Award (CS-2015-15-023) supported by the NIHR. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health. No generative AI was used in the generation of this manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

RU, OM, MV, BD, and MK contributed to the writing and review of this article.

Conflicts of Interest

Author MK is a trustee of Myelopathy.org, without any financial interests.

References

1. Davies BM, Mowforth OD, Smith EK, Kotter MR. Degenerative cervical myelopathy. BMJ 2018;360:k186 [FREE Full text] [doi: 10.1136/bmj.k186] [Medline: 29472200]
2. Smith SS, Stewart ME, Davies BM, Kotter MRN. The prevalence of asymptomatic and symptomatic spinal cord compression on magnetic resonance imaging: a systematic review and meta-analysis. Global Spine J 2021;11(4):597-607 [FREE Full text] [doi: 10.1177/2192568220934496] [Medline: 32677521]
3. Fehlings MG, Wilson JR, Yoon ST, Rhee JM, Shamji MF, Lawrence BD. Symptomatic progression of cervical myelopathy and the role of nonsurgical management: a consensus statement. Spine (Phila Pa 1976) 2013;38(22 Suppl 1):S19-S20 [FREE Full text] [doi: 10.1097/BRS.0b013e3182a7f4de] [Medline: 23963011]
4. Witiw CD, Mathieu F, Nouri A, Fehlings MG. Clinico-radiographic discordance: an evidence-based commentary on the management of degenerative cervical spinal cord compression in the absence of symptoms or with only mild symptoms of myelopathy. Global Spine J 2018;8(5):527-534 [FREE Full text] [doi: 10.1177/2192568217745519] [Medline: 30258760]

XSL•FO
RenderX

5. Harrop JS, Naroji S, Maltenfort M, Anderson DG, Albert T, Ratliff JK, et al. Cervical myelopathy: a clinical and radiographic evaluation and correlation to cervical spondylotic myelopathy. Spine (Phila Pa 1976) 2010;35(6):620-624. [doi: 10.1097/BRS.0b013e3181b723af] [Medline: 20150835]
6. Nemani VM, Kim HJ, Piyaskulkaew C, Nguyen JT, Riew KD. Correlation of cord signal change with physical examination findings in patients with cervical myelopathy. Spine (Phila Pa 1976) 2015;40(1):6-10. [doi: 10.1097/BRS.0000000000000659] [Medline: 25341986]
7. Milligan J, Ryan K, Fehlings M, Bauman C. Degenerative cervical myelopathy: diagnosis and management in primary care. Can Fam Physician 2019;65(9):619-624 [FREE Full text] [Medline: 31515310]
8. Fehlings MG, Tetreault LA, Riew KD, Middleton JW, Aarabi B, Arnold PM, et al. A clinical practice guideline for the management of patients with degenerative cervical myelopathy: recommendations for patients with mild, moderate, and severe disease and nonmyelopathic patients with evidence of cord compression. Global Spine J 2017;7(Suppl 3):70S-83S [FREE Full text] [doi: 10.1177/2192568217701914] [Medline: 29164035]
9. Akter F, Yu X, Qin X, Yao S, Nikrouz P, Syed YA, et al. The pathophysiology of degenerative cervical myelopathy and the physiology of recovery following decompression. Front Neurosci 2020;14:138 [FREE Full text] [doi: 10.3389/fnins.2020.00138] [Medline: 32425740]
10. Kopjar B, Tetreault L, Kalsi-Ryan S, Fehlings M. Psychometric properties of the modified Japanese Orthopaedic Association scale in patients with cervical spondylotic myelopathy. Spine (Phila Pa 1976) 2015;40(1):E23-E28 [FREE Full text] [doi: 10.1097/BRS.0000000000000648] [Medline: 25341993]
11. Sangeorzan I, Antonacci G, Martin A, Grodzinski B, Zipser CM, Murphy RKJ, et al. Toward shared decision-making in degenerative cervical myelopathy: protocol for a mixed methods study. JMIR Res Protoc 2023;12:e46809 [FREE Full text] [doi: 10.2196/46809] [Medline: 37812472]
12. Davies BM, Mowforth O, Gharooni AA, Tetreault L, Nouri A, Dhillon RS, et al. A new framework for investigating the biological basis of degenerative cervical myelopathy [AO spine RECODE-DCM research priority number 5]: mechanical stress, vulnerability and time. Global Spine J 2022;12(Suppl 1):78S-96S [FREE Full text] [doi: 10.1177/21925682211057546] [Medline: 35174728]
13. Wang ZC, Hou XW, Shao J, Ji YJ, Li L, Zhou Q, et al. HIF-1α polymorphism in the susceptibility of cervical spondylotic myelopathy and its outcome after anterior cervical corpectomy and fusion treatment. PLoS One 2014;9(11):e110862 [FREE Full text] [doi: 10.1371/journal.pone.0110862] [Medline: 25401740]
14. Badhiwala JH, Ahuja CS, Akbar MA, Witiw CD, Nassiri F, Furlan JC, et al. Degenerative cervical myelopathy—update and future directions. Nat Rev Neurol 2020;16(2):108-124. [doi: 10.1038/s41582-019-0303-0] [Medline: 31974455]
15. Senter HJ, Venes JL. Loss of autoregulation and posttraumatic ischemia following experimental spinal cord trauma. J Neurosurg 1979;50(2):198-206. [doi: 10.3171/jns.1979.50.2.0198] [Medline: 430132]
16. Saadoun S, Papadopoulos MC. Targeted perfusion therapy in spinal cord trauma. Neurotherapeutics 2020;17(2):511-521 [FREE Full text] [doi: 10.1007/s13311-019-00820-6] [Medline: 31916236]
17. Srihari G, Shukla D, Devi BI, Sathyaprabha TN. Subclinical autonomic nervous system dysfunction in compressive cervical myelopathy. Spine (Phila Pa 1976) 2011;36(8):654-659. [doi: 10.1097/BRS.0b013e3181dc9eb2] [Medline: 21178837]
18. Montine TJ, Cholerton BA, Corrada MM, Edland SD, Flanagan ME, Hemmy LS, et al. Concepts for brain aging: resistance, resilience, reserve, and compensation. Alzheimers Res Ther 2019;11(1):22 [FREE Full text] [doi: 10.1186/s13195-019-0479-y] [Medline: 30857563]
19. Zdunczyk A, Schwarzer V, Mikhailov M, Bagley B, Rosenstock T, Picht T, et al. The corticospinal reserve capacity: reorganization of motor area and excitability as a novel pathophysiological concept in cervical myelopathy. Neurosurgery 2018;83(4):810-818. [doi: 10.1093/neuros/nyx437] [Medline: 29165642]

Abbreviations

DCM: degenerative cervical myelopathy
mJOA: modified Japanese Orthopaedic Association
MRI: magnetic resonance imaging

Edited by T de Azevedo Cardoso; submitted 15.04.23; peer-reviewed by Y Wei, H Alsalhi; comments to author 12.12.23; revised version received 23.01.24; accepted 02.04.24; published 27.06.24.

Please cite as:
Umeria R, Mowforth O, Veremu M, Davies B, Kotter M
Radiological Progression of Degenerative Cervical Myelopathy in a Clinically Stable Patient: Case Report
Interact J Med Res 2024;13:e48212
URL: https://www.i-jmr.org/2024/1/e48212
doi:10.2196/48212
PMID:

XSL•FO
RenderX

©Rishi Umeria, Oliver Mowforth, Munashe Veremu, Benjamin Davies, Mark Kotter. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 27.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Letter

Visual "Scrollytelling": Mapping Aquatic Selfie-Related Incidents in Australia

Samuel Cornell[1], BSc, MSc; Amy E Peden[1], BA, MPP, PhD

School of Population Health, University of New South Wales, Sydney, Australia

Corresponding Author:
Samuel Cornell, BSc, MSc
School of Population Health
University of New South Wales
High St
Kensington
Sydney, 2052
Australia
Phone: 61 (2) 9065 38
Email: s.cornell@unsw.edu.au

(Interact J Med Res 2024;13:e53067) doi:10.2196/53067

KEYWORDS

selfie; map; social media; selfies; scrollama; JavaScript; scrollytelling; Mapbox; incidence; incidents; incident; fatality; fatalities; injury; injuries; retrieval; prevalence; image; images; photo; photos; photograph; photographs; prevalence; Australia; emergency; visualization; visualizations; interactive; location; geography; geographic; geographical; spatial; artificial intelligence; longitude; latitude; visual representation; visual representations

Introduction

Selfies are a modern, yet preventable cause of injury and death [1]. Medical responses and retrieval of persons, often in challenging terrain, burdens emergency medicine practitioners. To help prevent this issue, this study aimed to visualize selfie-related incidents globally by initially creating a scrollable visual story overlayed on a satellite map of the incidents in Australia. This type of visual storytelling technique using a world map helps illustrate the spatial context of this public health issue.

Methods

Overview

Incident data were acquired via publicly accessible news reports and a Wikipedia repository [2] and cleaned and prepared in Excel (Microsoft Corp). Incidents in aquatic areas (eg, coastal locations and inland waterfalls) were included; those in other settings (eg, falls from artificial structures and incidents involving trains) were excluded. Entries for each incident were created using associated media reports, incident types bring falls or drowns. Map coordinates were obtained by locating the incident using Google Maps and inputted into a coordinate finder using the Mapbox Location Helper [3]. Mapbox Studio [4] was used to create a custom map. A satellite template was chosen to best display the geographic context surrounding each selfie incident. The data set was imported into the Mapbox Studio custom map, which populated the data layer onto the map. A heat map setting was chosen for the data. To create the "scrollytelling" map story, the Mapbox storytelling template available from GitHub was used [5]. The primary input is a story comprising sections (chapters), each associated with a particular view of a map, enabling the user to "scroll" down the web page, and the resulting output is a zoomed-in view of a specific case layered on the map. The data and corresponding map visualization are hosted on GitHub and are published with GitHub pages.

Ethical Considerations

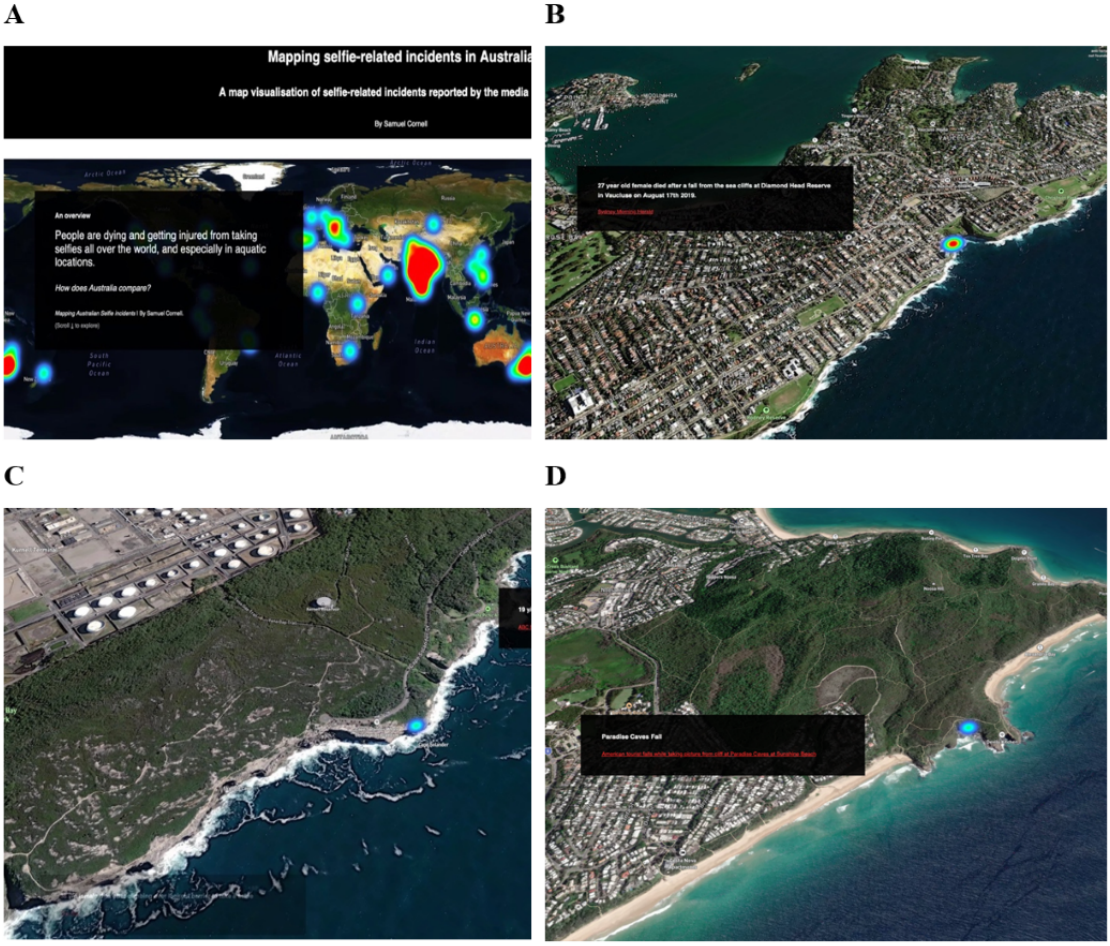
Ethics approval was not required due to the use of publicly available media reports and a Wikipedia repository to create the map.

Results

The publicly accessible map and data can be viewed on the web [6]. The heat map displays a worldwide overview of 104 cases from June 2014 through August 2023 (Figure 1 [7-9]). Once the user scrolls, the map displays 11 chapters, each showing an individual selfie-related incident in Australia including 9 deaths and 2 serious injuries. All cases reportedly involved emergency services, 3 of which occurred in the same location: Diamond Bay, Vaucluse, a suburb of Sydney, New South Wales. Two cases occurred at Gibraltar Falls, Australian Capital Territory. Further cases in Australia that are included in the map story are detailed in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. (A) Image taken from the web-based site heat map of worldwide selfie-related incidents. The image provides a worldwide overview of incidents based on the obtained media data used in this study. (B), (C), and (D) Example images acquired from the web-based site. These images illustrate the scrolling story of the heat map focusing on a location that has seen selfie-related incidents. Each incident is indicated in a "chapter," which provides a description of the incident in that location and a link to the corresponding news report. Images were acquired from Mapbox [7] and OpenStreetMap [8]. OpenStreetMap is licensed under the Open Data Commons Open Database License [9].



Discussion

Our heat map of media-based incident data provides a globally applicable visual representation of the selfie-incident phenomenon in Australia. Using a scrolling story template, overlayed on top of a map, selfie incidents can be illustrated in a geographic context. It is clear from the heat map that certain locations worldwide (nations such as India) and in Australia (cities such as Sydney) require specific and targeted prevention strategies to attenuate the incidence of selfie incidents, which is in line with their specific topographical realities.

Future research in this space should seek to further ascertain the burden on emergency and retrieval services by evaluating response times, resource allocation, retrieval or rescue methods, and health care costs associated with treating selfie-related injuries. Geographic disparities in service usage and response times, terrain, and retrieval or rescue methods should be identified and added to the visual map.

Understanding the geographic distribution and burden of selfie-related incidents is essential for designing targeted public awareness campaigns, improving safety regulations, and optimizing the allocation of resources for emergency and retrieval services.

The main limitation results from the use of media cases as the basis of these incident reports. It is not possible to acquire precise latitude and longitude coordinates. The coordinates provide the best approximation using details provided in media reports and analyzing the geography described in the report. Nevertheless, this map provides a good overview of the geographic nature of selfie-related incidents. Given the ethical limitations regarding the use of coronial data and identification of individual incidents, publicly available information of this nature remains the most appropriate data source.

In conclusion, selfie-related incidents present a significant geographic challenge for emergency services and retrievalists due to the inherent geographic context surrounding this type of event. Mapping selfie events may be a useful method of analyzing and tracking these phenomena and may be of benefit to emergency managers and land managers in collaborating to attenuate this issue.

XSL•FO
RenderX

Acknowledgments

This work was funded by a University of New South Wales (UNSW) university postgraduate award provided to SC. AP is funded by a National Health and Medical Research Council (NHMRC) Emerging Leadership Fellowship (grant APP2009306).

Data Availability

All data are publicly available and can be found in a GitHub repository [10].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table displaying cases used in the scrollytelling map.
[DOCX File , 17 KB - ijmr_v13i1e53067_app1.docx]

References

1. Cornell S, Brander RW, Peden AE. Preventing selfie-related incidents: taking a public health approach to reduce unnecessary burden on emergency medicine services. Emerg Med Australas 2023 Aug 19;35(4):691-693. [doi: 10.1111/1742-6723.14219] [Medline: 37454361]
2. List of selfie-related injuries and deaths. Wikipedia. URL: https://tinyurl.com/4xfwsfce [accessed 2023-01-11]
3. Location Helper. URL: https://labs.mapbox.com/location-helper/#3/4.48/49.2 [accessed 2023-09-06]
4. Design beautiful custom maps with Mapbox Studio. Mapbox. URL: https://www.mapbox.com/mapbox-studio [accessed 2023-09-06]
5. mapbox / storytelling. GitHub. URL: https://github.com/mapbox/storytelling [accessed 2023-09-06]
6. Cornell S. Mapping selfie and social media-related incidents in Australia. URL: https://samuelcornell.github.io/selfiemap/source/index.html [accessed 2023-09-25]
7. Mapbox. URL: https://www.mapbox.com/ [accessed 2024-05-22]
8. OpenStreetMap. URL: https://www.openstreetmap.org/ [accessed 2024-05-22]
9. Open Data Commons Open Database License. URL: https://www.openstreetmap.org/copyright [accessed 2024-05-22]
10. samuelcornell / selfiemap. GitHub. URL: https://github.com/samuelcornell/selfiemap/ [accessed 2024-05-09]

Edited by T de Azevedo Cardoso; submitted 25.09.23; peer-reviewed by N Auger, S Kang-Auger, J Luo; comments to author 06.12.23; revised version received 14.12.23; accepted 27.02.24; published 23.05.24.

Please cite as:
Cornell S, Peden AE
Visual "Scrollytelling": Mapping Aquatic Selfie-Related Incidents in Australia
Interact J Med Res 2024;13:e53067
URL: https://www.i-jmr.org/2024/1/e53067
doi:10.2196/53067
PMID:38781002

©Samuel Cornell, Amy E Peden. Originally published in the Interactive Journal of Medical Research (https://www.i-jmr.org/), 23.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.i-jmr.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX