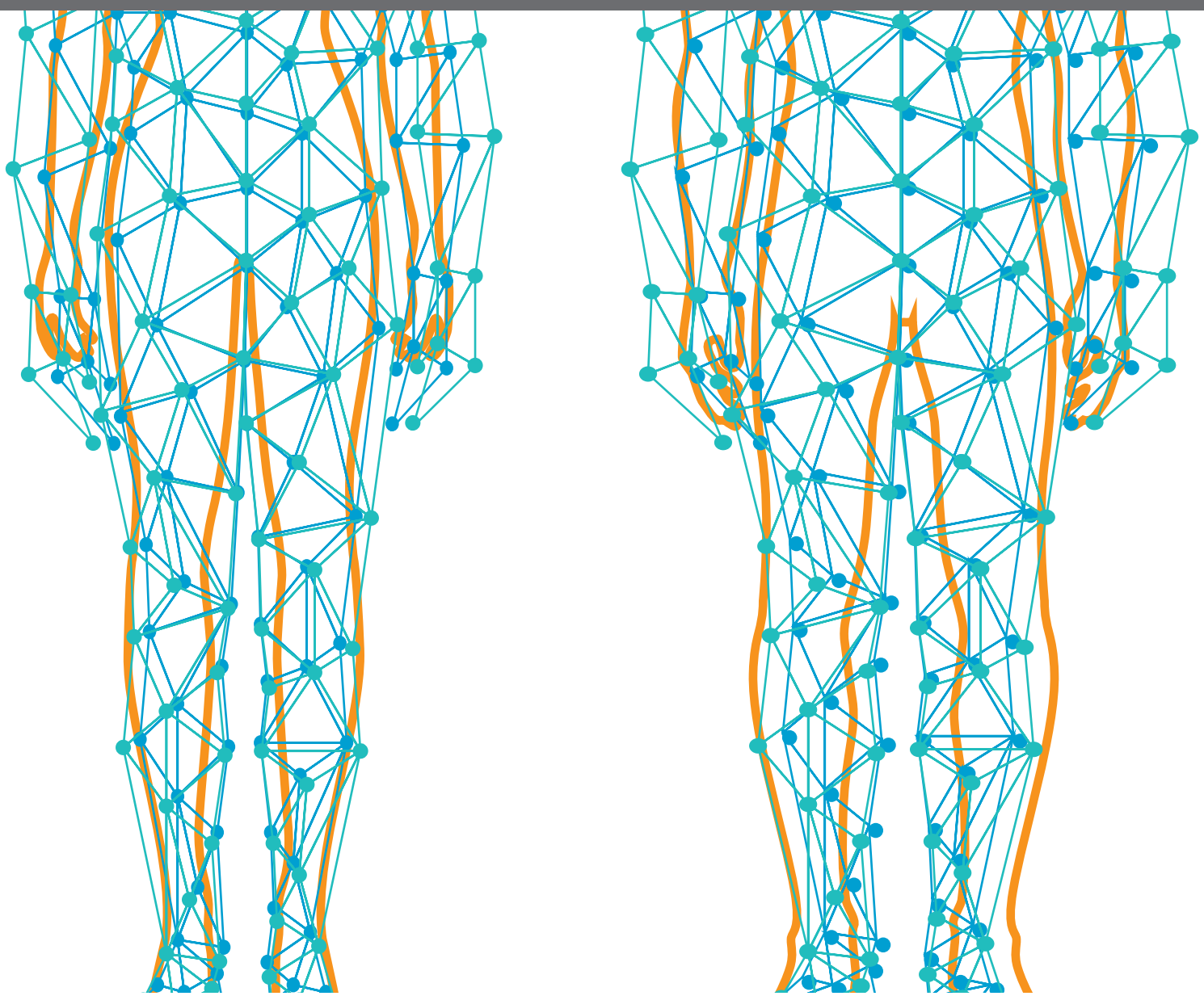
CONTEMPORARY MEDICINE: MAKING SENSE OF IMPLEMENTATION MODELS AND METHODS

EDITED BY: Michele Mario Ciulla, Vance W. Berger and Ugo Cioffi

PUBLISHED IN: Frontiers in Medicine and Frontiers in Public Health



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-298-9
DOI 10.3389/978-2-88976-298-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

CONTEMPORARY MEDICINE: MAKING SENSE OF IMPLEMENTATION MODELS AND METHODS

Topic Editors:
Michele Mario Ciulla, University of Milan, Italy
Vance W. Berger, National Cancer Institute (NIH), United States
Ugo Cioffi, University of Milan, Italy

Citation: Ciulla, M. M., Berger, V. W., Cioffi, U., eds. (2022). Contemporary Medicine: Making Sense of Implementation Models and Methods. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-298-9

Table of Contents

frontiers | Frontiers in Medicine

EDITORIAL
published: 17 May 2022
doi: 10.3389/fmed.2022.912045

Editorial: Contemporary Medicine: Making Sense of Implementation Models and Methods

Michele M. Ciulla[1,2]* *and Ugo Cioffi*[2]

[1] *Smart Laboratory of Clinical Informatics and Cardiovascular Imaging, CLO, Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy,* [2] *Department of Surgery, University of Milan, Milan, Italy*

Keywords: disease, guidelines, mathematical models, statistic, clinical medicine

Editorial on the Research Topic

Contemporary Medicine: Making Sense of Implementation Models and Methods

ARE STATISTICAL-MATHEMATICAL MODELS SUITABLE FOR STUDYING HUMAN DISEASES?

The practice of medicine, defined as "science and art of diagnosing and treating disease or injury" and, more recently, even "promoting *preventive* health care to improve patient well-being," is more and more guided by statistical-mathematical models that are, indisputably, helpful tools, nonetheless they are subjective, fallible and, thus, confutable. Given that statistics may be more or less adequate, the appropriate use of models and their output can contribute to effective preventive, diagnostic and treatment strategies, however, misuse or misinterpretation of their output can mislead decision-making. In the early'90, to summarize probabilistic data and provide practical guidance to clinical decision making, we sought the birth of the *evidence-based* model and the diffusion of *guidelines* developed accordingly (1, 2). Nowadays this system, subject to public or private funding since guidelines are drafted by invited *experts*, represents the summation of multiple instances, on the one hand the idea of rationalizing the medical intervention, based on available scientific evidences, in order to contain the costs and to build a better system of health care in the public domain and, on the other hand, to provide a legal protection that allows physicians to preserve a wide professional autonomy (Ciulla). Besides, it should also be pointed out that in clinical practice a systematic assessment of health outcomes with post-study probability tests is needed to confirm the intervention effectiveness (3, 4) and, unfortunately, this doesn't happen often.

In this Research Topic, we collected critical reviews and original papers on the use of guidelines in clinical practice to provide and evaluate research methodologies developed for the improvement of patient care, for a better understanding and guiding of clinical decisions, as well as the implementation of new models.

The *promise of digital technologies and data science* to transform healthcare (5) is discussed in the light of compelling ethical, legal, and social challenges, by Cordeiro. When processing large quantities of health data, from different sources, the Author, against the risks of *dehumanization of care*, claims a *normative framework* in order to promote fairness, inclusiveness, creativity and innovation in health.

OPEN ACCESS

Edited and reviewed by:
Arch Mainous,
University of Florida, United States

***Correspondence:**
Michele M. Ciulla
michele.ciulla@unimi.it

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *03 April 2022*
Accepted: *19 April 2022*
Published: *17 May 2022*

Citation:
Ciulla MM and Cioffi U (2022) Editorial: Contemporary Medicine: Making Sense of Implementation Models and Methods. Front. Med. 9:912045. doi: 10.3389/fmed.2022.912045

The goal of *implementation science* is to *close the gap between evidence-based practices and the extent to which research findings are integrated into real world settings and practices* (6). In such perspective, Huybrechts et al. in their review discuss the implementation of *complex interventions within primary care practices* by mapping existing theories, models, and frameworks from implementation science and combined insights across various disciplines. As a part of efforts intended to improve goal oriented care, self-management, and inter-professional collaboration, the Authors summarize three core phases to develop an overarching implementation model, development, translation and sustainment, and three main components, the intended change, the context and the implementation strategies.

A method to *close the gap between research and practice* consists in mixing design components of clinical effectiveness and implementation research in *effectiveness-implementation hybrid design* studies; unfortunately, translating these strategies into routine practice, especially in resource-constrained settings, is an arduous task. The results of a primary care-based integrated mobile health intervention for stroke management in Rural areas of China is discussed in the original paper by Gong et al. to describe the implementation indicators, related enablers and barriers, and illustrate some potential impact pathways that may influence the effectiveness of the intervention. The key factors identified to build an effective *doctor-patient alliance*, based on acceptance and fidelity on voice messages and follow-up visits, consists in supporting village doctors in *clinical decision-making* by training, financial compensation, and support from experienced physicians.

Statistical-mathematical models are often used to predict event outcomes and *nomograms*, representing geometrically the intersection of variables, are practical tools to provide rapid calculations. In the paper by Xu et al., a predictive model, consisting in a nomogram for identifying patients with acute decompensated cirrhosis at high-risk for readmission to the hospital, is proposed; the model, developed retrospectively on four independent clinical and laboratory indicators for each patient, may facilitate the development of effective interventions to reduce readmission rates.

Another predictive model is developed by Yu et al. combining data visualization and *machine learning* to predict the metabolic syndrome phenotype, especially in non-obese subjects where prevention is still challenging. As we know, *artificial intelligence* uses algorithms to learn and improve from experience and/or data without being explicitly programmed giving a perspective on the complicated relationships between metabolic components and potential risk factors.

Fragility fractures, even as pain experience, are very common in the elderly and, often, correlated; in the review by Chen et al. this relationship is analyzed supporting that frail patients with fractures were suffering from a continuous risk of pain exceeding the "typical" length of time assumed as essential for curing and resolution of pain. Thus, medical teams should develop treatment and rehabilitation protocols to prevent or reduce the pain of post-fracture, including meditation, exercises, and integrated physical treatment.

The *zero-markup policy* for essential drugs, a central point of the Chinese health reform started in 2009 (7), plays an important role in decreasing the cost of drugs for chronic diseases, such as type 2 diabetes, hypertension, metabolic syndrome, coronary heart disease, and cancer; Liu et al. in their systematic review discuss the effects of this policy on healthcare costs and utilization in China in the years 2015–2021 supporting a lower drug cost to patients but a rise in other expenditure categories such as healthcare services "induced" by physicians or response to unmet needs in the population.

High-dimensional data are unintuitive and difficult to interpret and derive insights, in the perspective article Bae et al. suggest to dimensionally reduce data by using *pattern identification* found in traditional Asian medicine, a diagnostic system using a limited amount of computation, to evaluate patient's clinical symptoms and signs and classify them. While this approach may appear biased by the underpinned *heuristic system* of belief in comparison to *pathogen-based diagnosis*, it provides an intuitive foundation for *inductive reasoning*.

Finally, when facing *how to build a better healthcare system*, a model of prevention/intervention, with a plausible design, is unavoidable, nonetheless, in the clinical setting, the application of a model requires a verification strategy, with a systematic assessment of its impact on health outcomes with post-study probability testing (Ciulla), and a continuous on-going refinement.

AUTHOR CONTRIBUTIONS

Both authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

1. Guyatt GH. Evidence-based medicine. *ACP J Club.* (1991) 114:A—16. doi: 10.7326/ACPJC-1991-114-2-A16
2. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ. The emergence of clinical practice guidelines. *Milbank Q.* (2007) 85:691–727. doi: 10.1111/j.1468-0009.2007.00505.x
3. Ioannidis JP. Why most published research findings are false. *PLoS Med.* (2005) 2:e124. doi: 10.1371/journal.pmed.0020124
4. Ciulla MM. The biological side of randomness. A starting point to rethink causation of diseases and prevention as a strategy. *PeerJ PrePrints.* (2015) 3:e1440. doi: 10.7287/peerj.preprints.1147
5. Lemmen C, Simic D, Stock S. A vision of future healthcare: potential opportunities and risks of systems medicine from a citizen and patient perspective-results of a qualitative study. *Int J Environ Res Public Health.* (2021) 18:9879. doi: 10.3390/ijerph18189879
6. Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. *BMC Med.* (2019) 17:88. doi: 10.1186/s12916-019-1322-9
7. Peng Z, Zhan C, Ma X, Yao H, Chen X, Sha X, et al. Did the universal zero-markup drug policy lower healthcare expenditures? Evidence from Changde,

China. *BMC Health Serv Res.* (2021) 21:1205. doi: 10.1186/s12913-021-07211-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Medicine

ORIGINAL RESEARCH
published: 07 April 2021
doi: 10.3389/fmed.2021.626580

Identify the Characteristics of Metabolic Syndrome and Non-obese Phenotype: Data Visualization and a Machine Learning Approach

Cheng-Sheng Yu[1,2,3,4†], *Shy-Shin Chang*[1,2†], *Chang-Hsien Lin*[1,2], *Yu-Jiun Lin*[1,2], *Jenny L. Wu*[1,2,3] *and Ray-Jade Chen*[4,5,6*]

[1] Department of Family Medicine, Taipei Medical University Hospital, Taipei, Taiwan, [2] Department of Family Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, [3] Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan, [4] Professional Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, [5] Division of General Surgery, Department of Surgery, Taipei Medical University Hospital, Taipei, Taiwan, [6] Department of Surgery, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

OPEN ACCESS

Edited by:
Michele Mario Ciulla,
University of Milan, Italy

Reviewed by:
Junjie Huang,
The Chinese University of Hong Kong, China
Jiunn-Yih Wu,
Keelung Chang Gung Memorial Hospital, Taiwan
Mengtse Lee,
National Taiwan University Hospital, Taiwan

***Correspondence:**
Ray-Jade Chen
rayjchen@tmu.edu.tw

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: 06 November 2020
Accepted: 08 March 2021
Published: 07 April 2021

Citation:
Yu C-S, Chang S-S, Lin C-H, Lin Y-J, Wu JL and Chen R-J (2021) Identify the Characteristics of Metabolic Syndrome and Non-obese Phenotype: Data Visualization and a Machine Learning Approach. Front. Med. 8:626580. doi: 10.3389/fmed.2021.626580

Introduction: A third of the world's population is classified as having Metabolic Syndrome (MetS). Traditional diagnostic criteria for MetS are based on three or more of five components. However, the outcomes of patients with different combinations of specific metabolic components are undefined. It is challenging to be discovered and introduce treatment in advance for intervention, since the related research is still insufficient.

Methods: This retrospective cohort study attempted to establish a method of visualizing metabolic components by using unsupervised machine learning and treemap technology to discover the relations between predicting factors and different metabolic components. Several supervised machine-learning models were used to explore significant predictors of MetS and to construct a powerful prediction model for preventive medicine.

Results: The random forest had the best performance with accuracy and c-statistic of 0.947 and 0.921, respectively, and found that body mass index, glycated hemoglobin, and controlled attenuation parameter (CAP) score were the optimal primary predictors of MetS. In treemap, high triglyceride level plus high fasting blood glucose or large waist circumference group had higher CAP scores (>260) than other groups. Moreover, 32.2% of patients with high CAP scores during 3 years of follow-up had metabolic diseases are observed. This reveals that the CAP score may be used for detecting MetS, especially for the non-obese MetS phenotype.

Conclusions: Machine learning and data visualization can illustrate the complicated relationships between metabolic components and potential risk factors for MetS.

Keywords: machine learning, metabolic syndrome, non-obese phenotype, data visualization, preventive medicine, artificial intelligence

INTRODUCTION

Because of the increasing prevalence of obesity, metabolic syndrome (MetS) has become a common metabolic disorder. There are several diagnostic criteria for MetS including National Cholesterol Education Program's Adult Treatment Panel III (ATP III), Modified ATP III for Asians, International Diabetes Federation (IDF) Criteria, National Heart, Lung, and Blood Institute (NHLBI) Criteria, and Joint Interim Statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity (JIS) (1–4). A comparison of the above diagnostic criteria for MetS, which is relevant for Asians can be found in **Supplementary Table 1**. In general, these different Mets criteria are very similar, all of them looks at the presence of ≥ three anthropometric characteristics or clinical factors as listed below: large waist circumference (WC), high triglyceride level (TG), high blood pressure (BP), high fasting blood glucose (FBG), and low high-density lipoprotein (HDL) cholesterol level. When evaluating Mets for Asians, the modified ATP III, JIS, and NHLBI criteria are almost identical. The IDF criteria are the most different from the above three criteria as the criteria insist that a Mets person must have abdominal obesity.

In previous studies, Beydoun et al. assessed the adiposity indices for MetS from a cohort data, the performance of detecting MetS was 0.680 and 0.733 for men using body fat mass and WC, respectively, and women (0.581 vs. 0.686) (5). Zhang et al. used a routine biomarker-based risk in Cox regression to predict MetS in an urban Han Chinese population, the performance was 0.796 and 0.897 for males and females (6). Both studies only had a better performance on females, and the selection of predictors is not objective and automated.

Non-alcoholic fatty liver disease (NAFLD) is a common comorbidity that is correlated with overweight and MetS. NAFLD is now primarily considered as a hepatic manifestation of MetS. Nevertheless, plenty of research has shown that NAFLD affects not only the liver but other chronic diseases such as chronic kidney disease (CKD), type 2 diabetes mellitus, and cardiovascular disease. Therefore, many chronic MetS-related diseases are directly caused by NAFLD, and better diagnoses and therapies of fatty liver disease are highly necessary (7–11). Currently, the detection of NAFLD has been enhanced with the capability of quantifying hepatic steatosis via measuring ultrasonic attenuation at the central frequency of the Fibroscan, termed the controlled attenuation parameter (CAP) (12–14). Previous study has found that CAP score alone can detect Mets with reasonable high accuracy of 0.79 and the combined use with machine learning can improve Mets accuracy detection to 0.904 (15, 16).

Machine learning is an artificial intelligence technique in which can the algorithm automatically learns and improves from experience or large amounts of data without being explicitly programmed. The kernel of machine learning is a statistical analysis that provides a powerful and purposeful method of observing specific patterns and correlations in health care issues by exploring undiscovered data, resulting in the establishment of data-driven prediction models (16–21). Several clinical issues—such as chronic kidney disease, postoperative sepsis, and alexithymia in fibromyalgia—have been explored using machine learning (22–24).

Data visualization is a useful technique that enhances clinicians' ability to analyze and summarize complex and large volumes of clinical data. Treemap visualization in particular is a conceivably advantageous method of visualizing clinical health care data. It enables the representation of high-dimensional hierarchical data in one diagram (21, 25, 26).

In this study, we will like to combine the use of data visualization and machine learning to find out if different levels of Mets will have different prediction accuracies. This is because the non-obese MetS population is difficult to discover, and this population is always the most challenging target in preventive medicine. In addition, we will like to find out if the CAP score alone can detect non-obese patients, as currently there are limited tools to detect non-obese patients without the invasive blood draw and inconvenient starvation. Use of CAP score for screening offers the clinical advantage of non-invasiveness, and no requirement for overnight starvation.

METHODS

Setting and Study Design

This retrospective cohort study was executed at Taipei Medical University Hospital (TMUH), a private teaching hospital with 800 beds in Taiwan. The electronic health care records of all participants were analyzed. The ability of treemap visualization and supervised machine learning to cluster different combinations of five metabolic components was assessed using patients who took a self-paid health examination at the Healthcare Center (HC) of TMUH, which has approximately 60 visits per day. This study was approved by the Institutional Review Board of TMUH (TMU-JIRB No.: N202003088).

Data Collection and Criteria

Patients had to meet the following inclusion criteria: older than 17 years, underwent a self-paid health examination at the HC of TMUH between March 2015 and May 2019, and underwent abdominal transient electrography inspection using the FibroScan 502 Touch (Echosens, Paris, France).

All patients underwent the regular processes of the HC (**Supplementary Table 2**). The blood samples required were collected from laboratory tests, and other anthropometric characteristics were also recorded (**Supplementary Table 3**). The definitions of measurement cut-offs and calculations are presented in **Supplementary Tables 1**, **4**. The included patients were than follow-up for 3 years at Taipei Medical University Hospital (**Figure 1**), and it was found that ∼60% of patients do not have follow-up data.

Statistical Analysis

The chi-square test and Kruskal–Wallis rank sum test were used to compare the groups of various participants with different

numbers of metabolic components. Descriptive characteristics were also analyzed and are presented as discrete or continuous variables with frequencies or percentages and medians or interquartile ranges, respectively. A box plot was drawn for presenting data distributions and comparing groups. Multinomial stepwise logistic regression was used to determine

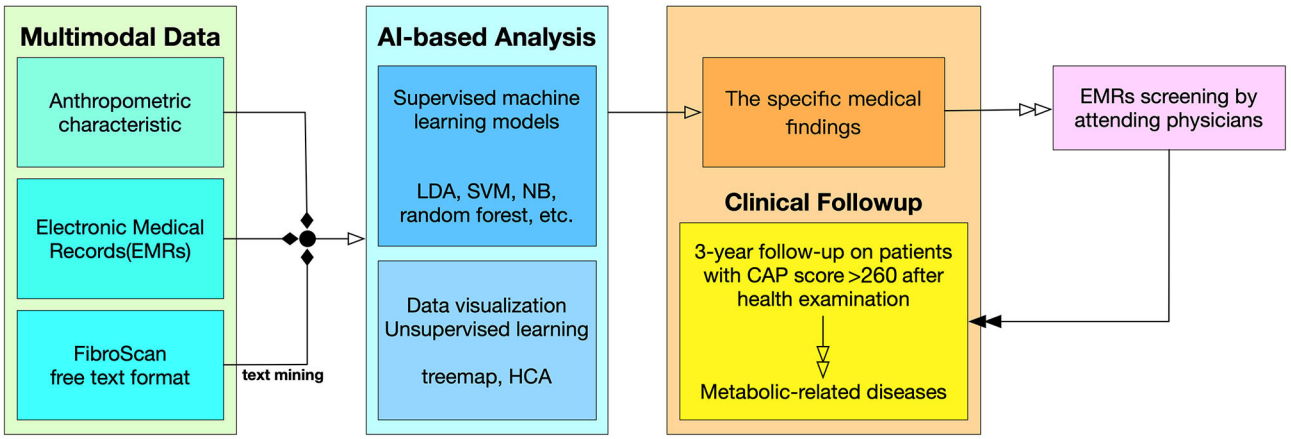


FIGURE 1 | Flowchart from data collection to clinical follow-up. Multimodal Data. The input variables are divided into three types of sources in the clinical data collection. The detail list of inputs can be found in **Supplementary Table 3**. AI-Based Analysis. This study used both supervised machine learning model and unsupervised machine learning along with data visualization techniques. Clinical Follow-up. The result of the AI-based analysis feedback to physicians, and the physicians investigate the outcome of the patients through EMRs screening.

TABLE 1 | Descriptive statistics and non-parametric multinomial test for multiple levels of metabolic syndrome in health examination data.

Factors	Health (0/5)	Met (1/5)	Met (2/5)	MetS (3/5)	MetS (4/5)	MetS (5/5)	*p*-value
	$n_0 = 477$ No. (%)	$n_1 = 295$ No. (%)	$n_2 = 200$ No. (%)	$n_3 = 102$ No. (%)	$n_4 = 42$ No. (%)	$n_5 = 13$ No. (%)	
CKD							
No	279 (58.5%)	148 (50.2%)	84 (42%)	39 (38.2%)	16 (38.1%)	4 (30.8%)	<0.001
Yes	198 (41.5%)	147 (49.8%)	116 (58%)	63 (61.8%)	26 (61.9%)	9 (69.2%)	
Sex							
Female	296 (62.1%)	126 (42.7%)	64 (32%)	26 (25.5%)	6 (14.3%)	1 (7.7%)	<0.001
Male	181 (37.9%)	169 (57.3%)	136 (68%)	76 (74.5%)	36 (85.7%)	12 (92.3%)	
MEDIAN (IQR)							
Age	42 (36–48)	45 (37–51)	45 (40–52)	45 (40–52)	45 (39–51)	44 (40–50)	<0.001
BMI	21.5 (19.9–23.2)	23.9 (22.3–25.9)	25 (23.3–27.5)	26.8 (24.9–29.8)	28.2 (26.6–30.9)	28.8 (25.8–31.7)	<0.001
Cholesterol	182 (163–202)	193 (170–213)	195 (172–219)	195 (166–214)	190 (158–215)	190 (139–241)	<0.001
LDL	112 (95–132)	128 (107–149)	134 (114–155)	134 (107–155)	124 (90–158)	135 (86–173)	<0.001
HbA1C	5.3 (5.1–5.4)	5.4 (5.2–5.6)	5.5 (5.3–5.7)	5.6 (5.3–5.9)	5.7 (5.4–6.0)	6.5 (6.0–7.3)	<0.001
GOT	19 (16–23)	20 (17–25)	21 (18–26)	23 (18–30)	26 (21–35)	26 (22–52)	<0.001
GPT	16 (12–22)	21 (15–31)	25 (17–35)	30 (20–47)	40.5 (24–58)	43 (23–99)	<0.001
γGT	13 (10–19)	18 (13–27)	22 (17–36)	25 (18–42)	35 (26–55)	37 (23–74)	<0.001
T_bilirubin	0.6 (0.4–0.8)	0.6 (0.5–0.8)	0.7 (0.4–0.9)	0.65 (0.5–1.0)	0.6 (0.4–0.9)	0.8 (0.5–1.2)	0.221
ALKp	55 (46–65)	58 (49–69)	62 (53–74)	61 (53–71)	67 (58–79)	59 (52–70)	<0.001
AFP	2.21 (1.56–3.02)	2.15 (1.57–3.2)	2.36 (1.72–3.24)	2.31 (1.62–3.07)	2.31 (1.66–3.15)	2.83 (2.28–4.70)	0.068
E score	3.9(3.3–4.6)	4.0 (3.4–4.7)	4.3 (3.5–5.1)	4.9 (4.0–5.5)	5.1 (4.4–6.8)	6.1 (4.6–6.8)	<0.001
CAP score	221(197–248)	250 (217–281)	272 (242–310)	298 (251–331)	327 (296.5–359)	323 (276–370)	<0.001
Albumin	4.6 (4.4–4.7)	4.6 (4.4–4.8)	4.6 (4.4–4.8)	4.6 (4.5–4.8)	4.6 (4.4–4.9)	4.8 (4.5–5.0)	0.007
BUN	12 (10–14)	12 (10–14)	12 (10–15)	12 (10–15)	13 (11–16)	12 (11–15)	0.009
Creatinine	0.7 (0.6–0.9)	0.8 (0.6–0.9)	0.8 (0.7–1.0)	0.9 (0.7–1.0)	0.9 (0.8–1.0)	1.0 (0.8–1.1)	<0.001
UA	4.8 (4.1–5.9)	5.5 (4.6–6.7)	6.0 (5.2–7.1)	6.3 (5.4–7.1)	6.9 (5.8–7.8)	7.1 (6.7–7.9)	<0.001
TSH	1.87 (1.24–2.61)	1.83 (1.30–2.48)	1.91 (1.24–2.54)	1.74 (1.25–2.45)	1.82 (1.12–2.75)	2.12 (1.39–2.65)	0.971

which variables had significant differences and the odds ratios among the groups of patients with different numbers of metabolic components. Receiver operating characteristic (ROC) curves were plotted to demonstrate the diagnostic ability of machine-learning prediction models for MetS. Model performance was measured using c-statistic, sensitivity (recall), and specificity in ROC plots (27, 28).

Figure 1 describes the procedure of this study from data collection to clinical outcomes. In data preprocessing, multimodal data were summarized; a series of machine learning models were then constructed, and statistical analyses were performed. A feedback mechanism was working clinically as a prospective survey when remarkable findings were obtained by the machine learning models. A recommended threshold of risk factor was targeted before clinical physician scrutinized the potential MetS patients' follow up (16).

Machine Learning

Several supervised machine-learning models—k-nearest neighbor classification (KNN), linear discriminant analysis (LDA), logistic regression for classification, ensemble learning, support vector machine (SVM), naive Bayes classification (NB), and hierarchical clustering analysis (HCA)—were also executed using R (version 3.6.3). The factors used as input to each machine learning models were listed in **Supplementary Table 3**. And a series of data preprocessing, including structured query language command, database merging and text mining, were applied to integrate these databases in the study.

KNN has relatively simple implementation and is robust because the classes do not have to be linearly separable in the searching space. This advantage was the reason it was applied to missing value mutation in our study (29, 30). Variables will be excluded if the number of missing values is more than 10% of the sample size in this study.

LDA is a statistical method in which a linear combination of features separating two or more classes of objects is located. It can handle multivariate problems because its linear combination is more commonly used for dimensionality reduction before classification (31, 32).

Logistic regression is usually used in machine learning for classification because the probability of some obtained event is represented as a linear function of a combination of predictor variables. The technique is used when the response variable is categorical in nature, for instance, when it has the value yes/no or true/false. In contrast to linear regression, a linear relationship between dependent and independent variables is not required (33).

The main advantage of ensemble models in machine learning is that decisions from multiple models are combined to improve overall performance (34, 35). Random forest is a parallel ensemble method used for classification, regression, or other applications and is based on the structure of a decision tree. It eliminates the possibility of bias that a decision tree model may induce in the system. Moreover, it improves the predictive power considerably (36). Adaptive boosting (AdaBoost) is a sequential ensemble method in which the base learners are generated in series. The underlying purpose of sequential learning is to use the dependence between the base learners, and overall performance can be improved by giving previously mislabeled samples higher weights in the sequential training processes (37, 38).

SVM model constructs a hyperplane or set of hyperplanes in a high-dimensional space, which is used for classification, regression, or outlier detection. Although SVM performs relatively favorably when a clear margin of separation exists between classes, and it is effective in high-dimensional spaces (39).

NB classifiers are probabilistic classifiers based on the use of Bayes' theorem with naive assumptions of independence between features. They are simple and easy to implement and do not require as much training data as other methods. The leading advantage of NB classification is that it is highly scalable with the number of predictors and data points (40).

The machine learning algorithms were executed in R program, the library, package and function using in this study are listed in **Supplementary Table 5**.

TABLE 2 | Multinomial stepwise logistic regression analysis of risk factors related to metabolic syndrome.

Factor	Met (1/5) $n_1 = 295$		Met (2/5) $n_2 = 200$		MetS (3/5) $n_3 = 102$		MetS (4/5) $n_4 = 42$		MetS (5/5) $n_5 = 13$		Likelihood Ratio Test
	OR	*p*-value	OR	*p*-value	OR	*p*-value	OR	*p*-value	OR	*p*-value	*p*-value
Age	1.011	0.295	1.037	0.003	1.038	0.016	1.047	0.037	1.055	0.132	**0.032**
BMI	**1.392**	<0.001	**1.525**	<0.001	**1.825**	<0.001	**1.795**	<0.001	**1.877**	<0.001	**<0.001**
γGT	1.025	<0.001	1.033	<0.001	1.027	0.001	1.035	<0.001	1.039	<0.001	**<0.001**
CAPscore	1.003	0.230	1.005	0.051	1.009	0.008	1.024	<0.001	1.017	0.027	**<0.001**
UA	0.930	0.277	1.064	0.430	1.035	0.734	**1.282**	0.067	**1.772**	0.005	**0.014**
Cholesterol	0.985	0.028	0.962	<0.001	0.983	0.154	0.990	0.490	0.975	0.304	**0.002**
LDL	1.029	<0.001	1.056	<0.001	1.030	0.017	1.016	0.301	1.034	0.189	**<0.001**
HbA1C	**1.559**	0.087	**3.264**	<0.001	**4.717**	<0.001	**4.403**	<0.001	**7.447**	<0.001	**<0.001**

The baseline of multinomial logistic regression for the health group is (0/5) without any metabolic syndrome disorders. After stepwise regression, only eight factors were retained. High odds ratios are in bold.

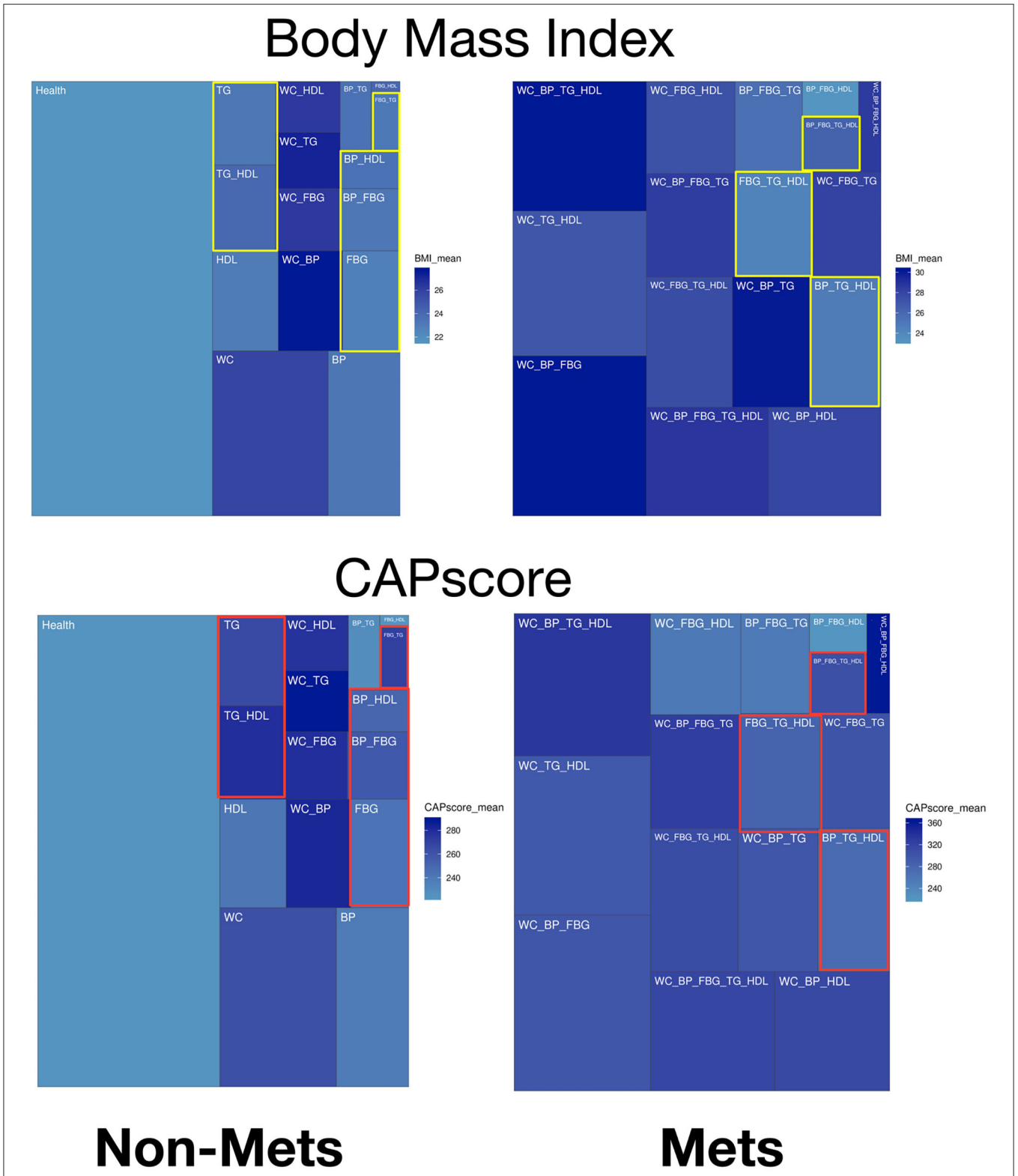


FIGURE 2 | Treemaps of significant predictors within different combinations of metabolic components for non-MetS and MetS subjects. Body mass index (BMI) is the upper panel and CAP score is the lower panel.

Data Visualization

In data analysis, visualization is always the most intuitive and sufficient method of exploring a specific pattern in data reflecting unknown or complicated issues. In this study, we used an unsupervised learning model called HCA in heatmap and a large and complex data-mapping technique called treemap to depict the characterization of metabolic components, because these approaches clearly enable recognition of special patterns in high-dimensional data through the use of gradient colors and grids of different areas (26, 41).

RESULT

The statistical distribution and differences between patient groups with different numbers of metabolic components are shown in **Table 1**. The combinations of metabolic components are listed in **Supplementary Table 6**. Stepwise multinomial logistic regression reveals the odds ratios, compared with the healthy group without any metabolic components, among the significant variables in **Table 2**. When the number of metabolic components increases, a significant difference was observed in several predictors, such as age, body mass index (BMI), gamma-glutamyl transferase (γGT), CAP score, serum uric acid (UA), cholesterol, low density lipoprotein (LDL), and glycosylated hemoglobin (HbA1C) ($p < 0.01$).

In the treemaps presented in **Figure 2** and **Supplementary Figure A**, gradient colors display specific patterns of significant predictors comparing groups with different numbers of metabolic components. The non-obese potential MetS populations are highlighted with color rectangles as comparison in treemaps. In **Figure 2**, the upper panel on BMI shows there is general positive correlation between BMI

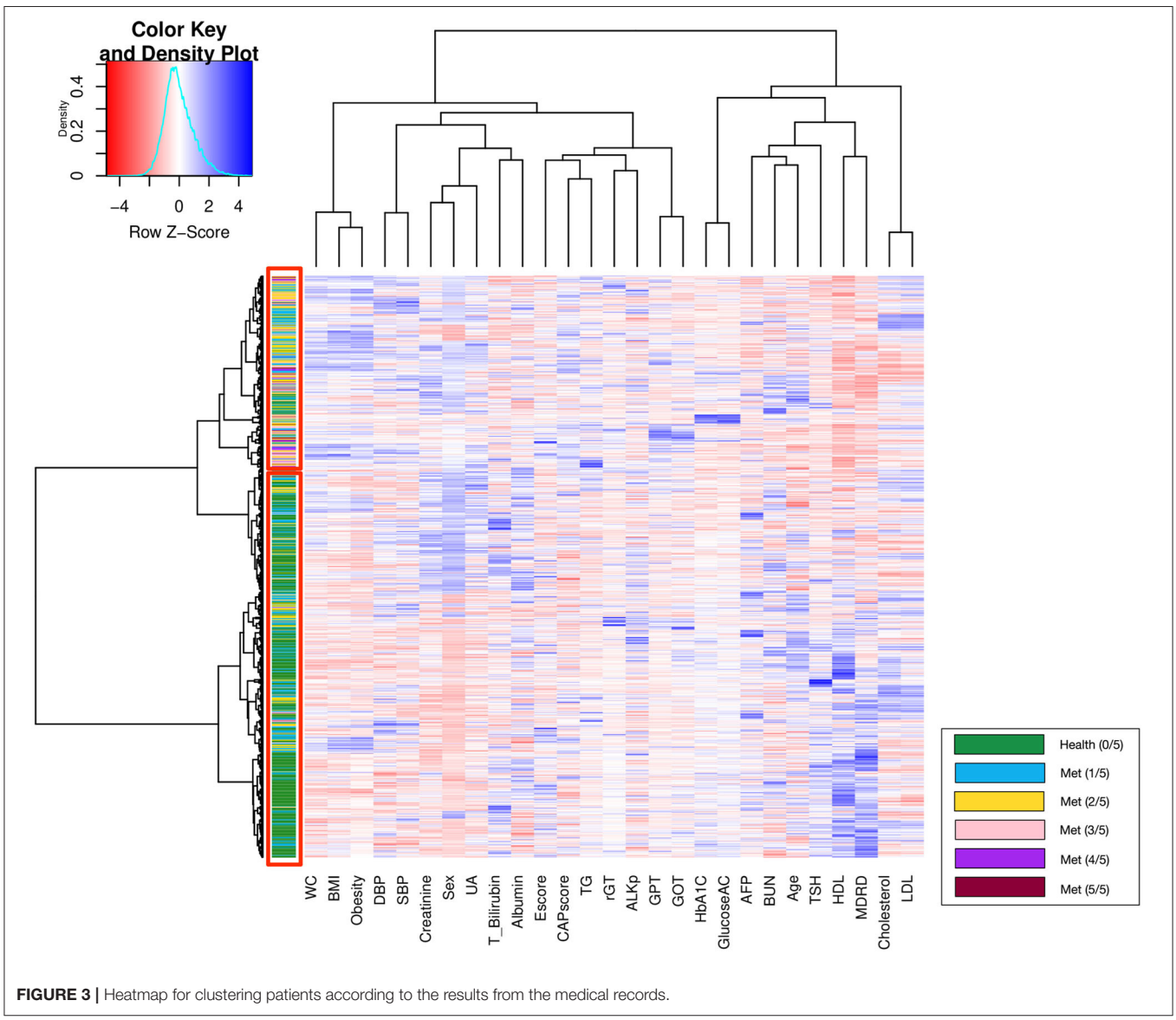


FIGURE 3 | Heatmap for clustering patients according to the results from the medical records.

and waist circumference (WC). However, the highlighted yellow rectangles show some patients without elevated WC/ has low BMI, and yet many of these subjects have Mets. CAP score. **Figure 2**'s lower panel on CAP score, shows the distribution of CAP score for different types of subjects. The highlighted red rectangles show the non-obese subjects, where the mean of CAP score is ∼260.

Unsupervised hierarchical clustering determined the similarity and classification between groups with different numbers of metabolic components; the corresponding heatmap is displayed in **Figure 3**. Patients with similar physiological records were clustered into the same group via hierarchical clustering analysis. In general, the upper red rectangle contains subjects with increased numbers of metabolic components, and the lower red rectangle contains healthy subjects (green), which do not have any Mets components. However, occasionally a few of the subjects do not follow the above described pattern.

Several supervised learning models were used to predict MetS according to both ATP III and International Diabetes Federation (IDF) criteria as the ground truth, and the performance of these models is illustrated in **Figures 4A,B** and **Table 3**. The rank of variable importance for ensemble learning summarization of multiple classifiers is represented in **Figure 4C**.

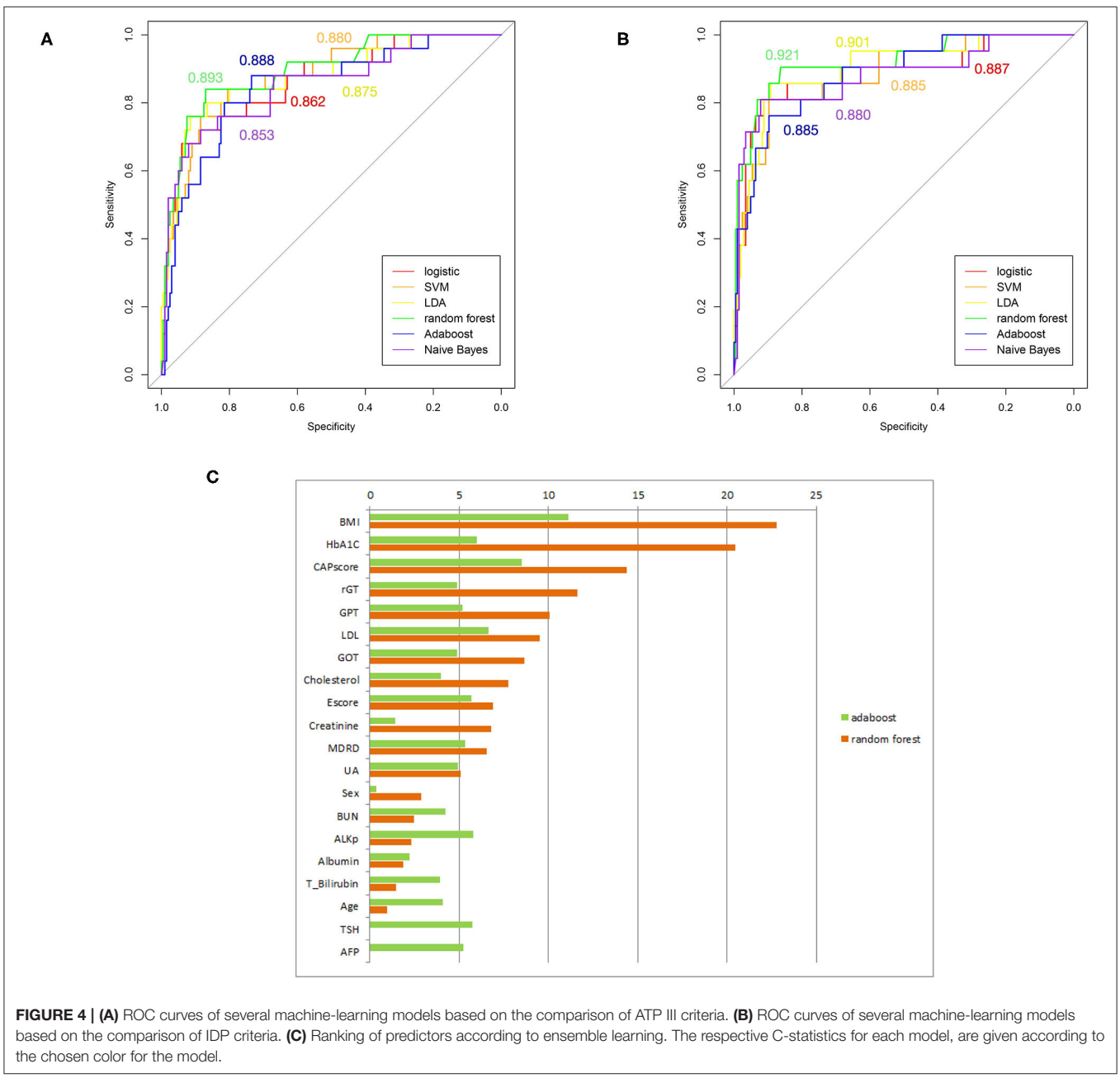


FIGURE 4 | (A) ROC curves of several machine-learning models based on the comparison of ATP III criteria. **(B)** ROC curves of several machine-learning models based on the comparison of IDP criteria. **(C)** Ranking of predictors according to ensemble learning. The respective C-statistics for each model, are given according to the chosen color for the model.

TABLE 3 | Performance of different machine-learning models on predicting metabolic syndrome using ATP III, JIS, NHLBI, and IDF criteria.

Model	Criteria	Accuracy	Sensitivity	Specificity	c–statistic
Logistic	ATPIII	0.902	0.520	0.950	0.862
LDA	&	0.898	0.545	0.936	0.875
SVM	JIS	0.902	0.400	0.965	0.880
Random Forest	&	0.922	0.440	0.980	0.893
Adaboost	NHLBI	0.893	0.440	0.950	0.888
Naïve Bayes		0.853	0.720	0.870	0.853
Logistic	IDF	0.929	0.619	0.961	0.887
LDA		0.916	0.545	0.956	0.901
SVM		0.916	0.476	0.961	0.885
Random Forest		0.947	0.571	0.985	0.921
Adaboost		0.911	0.429	0.961	0.885
Naïve Bayes		0.893	0.810	0.902	0.880

AUC, area under curve.

The relationship between CAP score and obesity, as well as MetS, is shown in **Figures 5A,B**. The box plots presented in **Figure 5B** show that CAP score was positively related to MetS. Four attending physicians conduct an approximately 3 year follow-up of the patients with a CAP score higher than 260, and recorded metabolically associated diseases—including diabetes, cardiovascular disease, stroke, CKD, and dyslipidemia. The follow-up results are presented in **Figure 5C**, which shows that 32.2%, 22.4%, 18.6%, and 16.4% of the patients had metabolic diseases, liver-related diseases, kidney diseases, and cardiovascular diseases, respectively.

DISCUSSION

In statistical analysis, significant differences between groups with different numbers of metabolic components were discovered for several predictors. Because patients who have the same number of metabolic components may nonetheless have different combinations of the five components, their physical characteristics are diverse.

Furthermore, the classification of patients with different numbers of metabolic components that was visualized using clustering and a heat map revealed an overlapping representation at the left cluster label, although unsupervised machine learning made a strong contribution to the separation of the group with severe MetS (more than three components) and group with mild MetS (fewer than two components). Most patients with MetS were clustered in the upper portion, whereas healthier patients were clustered in the lower portion. Therefore, we applied several supervised learning models to predict MetS and found some representative predictors—such as CAP score, BMI, HbA1C, and γGT—that resulted in high accuracy and performance without any of the five criteria being involved in the models. Ensemble learning of random forest had highest performance in both ATP III and IDF criteria as ground truth with respective accuracy of 0.922 and 0.947 and c-statistic of 0.893 and 0.921; BMI with obesity, HbA1C, and CAP score were observed to be the best primary predictors for MetS (**Figure 4C**).

CAP score represents the severity of MetS because it reveals the extent of NAFLD (15, 42, 43). In previous research of several decision tree algorithms for MetS prediction, the threshold range of CAP score is also found to be approximately 290–300 (16). Similar to previous study, we found that if the goal is to identify both obese and non-obese patients, the cut off is ∼290 (average of the 320 obese cut-off and 260 non-obese cut-off). Using the 260 non-obese CAP cut-off, we found that ∼60% (43/72) of non-obese patients can be identified. This is likely because CAP is detecting NAFLD. In liver cells, NAFLD is caused by a considerable accumulation of triglycerides (44). Many evidence supports the connection between MetS and NAFLD. NAFLD is actually considered as the hepatic manifestation of MetS. Insulin resistance is the failure of cell to normally respond to insulin to reduce blood glucose level and is the key pathogenic feature of MetS. Insulin resistance is now identified as the most common risk factor for development and progression of NAFLD (45–47). In clinical laboratory examination, TG and FBG measurements are easily disturbed by many factors including incomplete fasting and medication. Therefore, CAP score measurements are more convenient and may be an alternative tool for detecting MetS, especially for the hard-to-detect non-obese patients.

The patients in the WC plus TG and WC plus BP metabolic component groups had higher BMI than those in the other groups (**Figure 2**). This implies that obesity is one of the leading risk factors for MetS (16). Moreover, multiple machine-learning models had high accuracy and performance for both the ATP III and IDF criteria. In particular, CAP score is also one of the primary variables in ensemble learning, giving machine-learning models high prediction ability (**Figure 4C** and **Table 3**). In addition, **Figure 5A** reveals that CAP score was proportional to degree of obesity. Fibroscan, a non-invasive method of screening for liver disease, is widely applied in detecting and treating NAFLD patients with MetS may be taken into consideration by experts and physicians.

Numerous cross-sectional and prospective studies have investigated the relationship between baseline γGT and the development of MetS (48–51). According our study, γGT is a valuable predictor of MetS because patients with TG and FBG metabolic components have elevated γGT (**Supplementary Figure A**). The non-obese metabolic health patients can be detected early to prevent progress of metabolic disorders to MetS. Moreover, the more glycemic level increases, the higher prevalence of NAFLD is (52). Several methods can evaluate the ranges of glycemia, containing HbA1c and FBG. HbA1c reflects the mean of glycaemia over the past 8–12 weeks and is applied to assess chronic glycemic level (53). Insulin resistance is a primary factor of NAFLD, and HbA1c correlates more strongly with insulin resistance than does FBG (54, 55).

A prominent relationship was illustrated between serum UA level and the risk for metabolic disorders in a meta-analysis of prospective studies. A linear relationship was speculated to exist between elevated UA and MetS/NAFLD incidence (56). Hyperuricemia is associated with histologically severe NAFLD (57). Furthermore, several research has identified UA as an

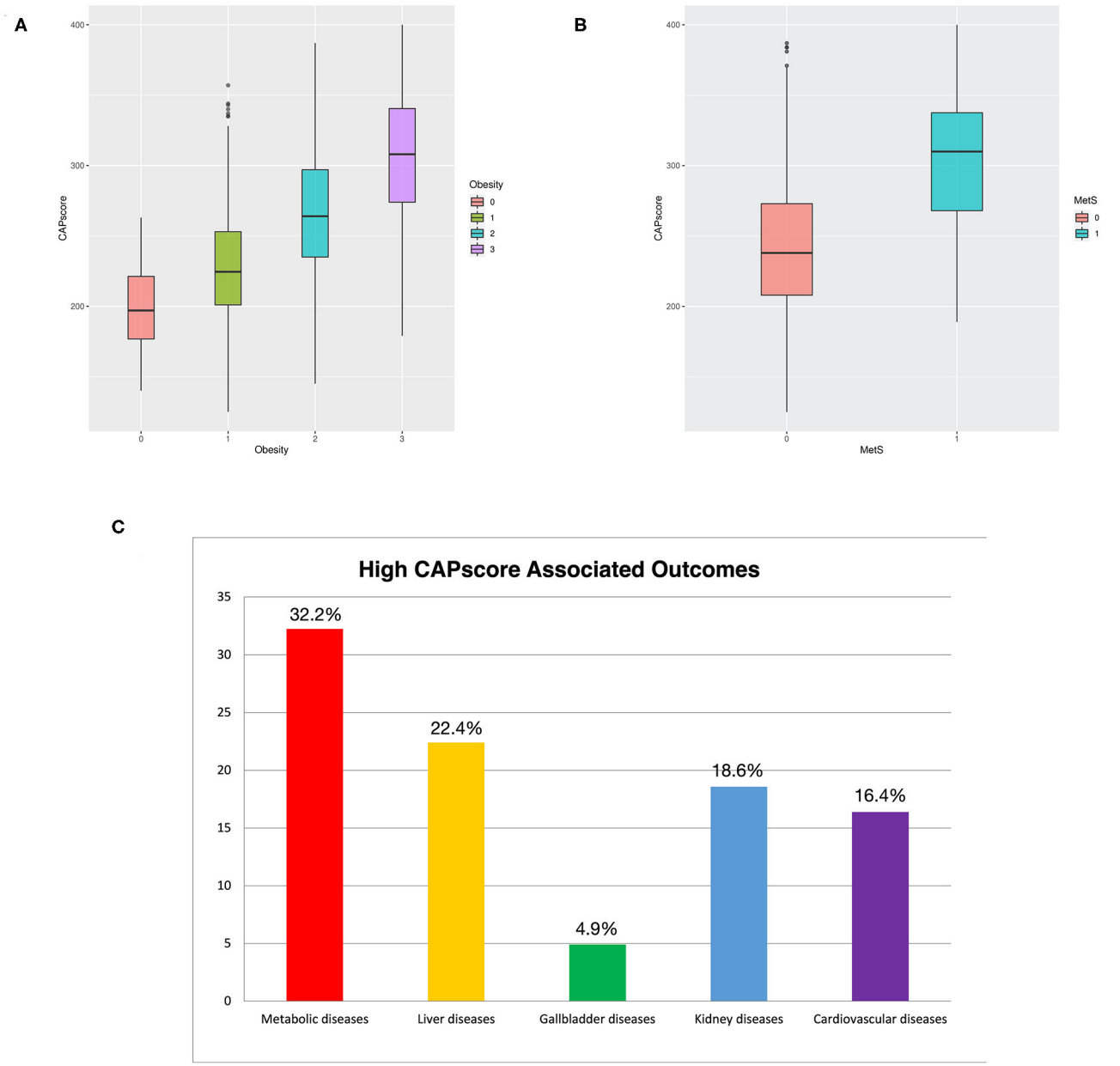


FIGURE 5 | (A) Boxplot of CAP score comparing patients with and without obesity. **(B)** Boxplot of CAP score comparing patients with and without MetS. **(C)** Percentage of outcome diseases among patients with CAP score of higher more than 260. Patients that were loss during follow-up were deleted in the calculation of percentage.

independent risk factor for cardiometabolic diseases, indicating that UA can be regarded as a essential therapeutic target for patients with these diseases and particularly those with hyperuricemia (58).

This study has some limitations. First, the data only represent an Asian population; the CAP score cut-off at which fatty liver disease increases metabolic risks may vary for different races. Second, the data are collected from one HC and reflected the information of healthier population. Therefore, the bias in data distribution cannot be avoided. The more the information included on patients with severe MetS, the more robust is the distribution represented. Because of this limitation, this study focused on early intervention for patients to prevent the occurrence of MetS. Third, this is a single-center study involving self-paid health examination subjects that were prospectively follow-up in the same hospital, and only 40.4% of patients with CAP score >260 were successfully tracked in our hospital. A large number of patients with loss of follow-up implies that the metabolic-related risks may have been underestimated; therefore, the significance of fatty liver disease, measured using FibroScan,

in MetS is probably higher than that determined in our study. In the future, it will be interesting to follow-up the medical record of these patients at other hospitals and apply machine learning in improving the prediction for cardiometabolic events for different types of Mets patients.

CONCLUSION

Machine learning and big data visualization can depict the complicated relationships between metabolic components and potential risk factors. The potential MetS patients can be captured by machine learning for prevention especially for those non-obese population. In the future, more data on CAP scores from the healthy population and those with severe MetS should be collected to establish a more robust investigation. Moreover, analyzing data of different races could enhance the achievement of data visualization to describe the association between CAP score cut-off and MetS for different particular populations.

DATA AVAILABILITY STATEMENT

The datasets generated for this article are not publicly available due to the confidentiality concerns/ethical restriction but are available from the author on reasonable request. Requests to access the datasets should be directed to Shy-Shin Chang, sschang0529@gmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Taipei Medical University-Joint Institutional Review Board (TMU-JIRB No.: N202003088). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

C-SY and S-SC: study conception and design, analysis and interpretation of data, and acquisition of funding. C-SY, S-SC, R-JC, and JLW: acquisition of data. R-JC, S-SC, C-HL, and Y-JL: medical insight consultation. C-SY and R-JC: intelligence insight consultation. R-JC, C-SY, and S-SC: drafting of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Ministry of Science and Technology Grant (MOST109-2314-B-038-080 and MOST 110-2314-B-038-025) and Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. No funding bodies had any role in study design, data collection, or analysis; decision to publish; or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.626580/full#supplementary-material

REFERENCES

1. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation*. (2009) 120:1640–5. doi: 10.1161/CIRCULATIONAHA.109.192644
2. Thomas GN, Ho SY, Janus ED, Lam KS, Hedley AJ, Lam TH. The US National Cholesterol Education Programme Adult Treatment Panel III (NCEP ATP III) prevalence of the metabolic syndrome in a Chinese population. *Diabetes Res Clin Pract*. (2005) 67:251–7. doi: 10.1016/j.diabres.2004.07.022
3. Alberti KG, Zimmet P, Shaw J. The metabolic syndrome–a new worldwide definition. *Lancet*. (2005) 366:1059–62. doi: 10.1016/S0140-6736(05)67402-8
4. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome. *Circulation*. (2005) 112:2735–52. doi: 10.1161/CIRCULATIONAHA.105.169404
5. Beydoun MA, Kuczmarski MT, Wang Y, Mason MA, Evans MK, Zonderman AB. Receiver-operating characteristics of adiposity for metabolic syndrome: the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study. *Public Health Nutr*. (2011) 14:77–92. doi: 10.1017/S1368980010002648
6. Zhang W, Chen Q, Yuan Z, Liu J, Du Z, Tang F, et al. A routine biomarker-based risk prediction model for metabolic syndrome in urban Han Chinese population. *BMC Public Health*. (2015) 15:64. doi: 10.1186/s12889-015-1424-z
7. Bang KB, Cho YK. Comorbidities and metabolic derangement of NAFLD. *J Lifestyle Med*. (2015) 5:7–13. doi: 10.15280/jlm.2015.5.1.7
8. Mikolasevic I, Milic S, Turk Wensveen T, Grgic I, Jakopcic I, Stimac D, et al. Nonalcoholic fatty liver disease - a multisystem disease? *World J Gastroenterol*. (2016) 22:9488–505. doi: 10.3748/wjg.v22.i43.9488
9. Byrne CD, Targher G. NAFLD: a multisystem disease. *J Hepatol*. (2015) 62:S47–64. doi: 10.1016/j.jhep.2014.12.012
10. Targher G, Chonchol MB, Byrne CD. CKD and nonalcoholic fatty liver disease. *Am J Kidney Dis*. (2014) 64:638–52. doi: 10.1053/j.ajkd.2014.05.019
11. Armstrong MJ, Adams LA, Canbay A, Syn WK. Extrahepatic complications of nonalcoholic fatty liver disease. *Hepatology*. (2014) 59:1174–97. doi: 10.1002/hep.26717
12. Sasso M, Audière S, Kemgang A, Gaouar F, Corpechot C, Chazouillères O, et al. Liver steatosis assessed by Controlled Attenuation Parameter (CAP) measured with the XL probe of the fibroscan: a pilot study assessing diagnostic accuracy. *Ultrasound Med Biol*. (2016) 42:92–103. doi: 10.1016/j.ultrasmedbio.2015.08.008
13. Sasso M, Beaugrand M, De Ledinghen V, Douvin C, Marcellin P, Poupon R, et al. Controlled attenuation parameter (CAP): a novel VCTE^{TM} guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. *Ultrasound Med Biol*. (2010) 36:1825–35. doi: 10.1016/j.ultrasmedbio.2010.07.005

14. Eddowes PJ, Sasso M, Allison M, Tsochatzis E, Anstee QM, Sheridan D, et al. Accuracy of fibroscan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology*. (2019) 156:1717–30. doi: 10.1053/j.gastro.2019.01.042
15. Lin YJ, Lin CH, Wang ST, Lin SY, Chang SS. Noninvasive and convenient screening of metabolic syndrome using the controlled attenuation parameter technology: an evaluation based on self-paid health examination participants. *J Clin Med*. (2019) 8:1775. doi: 10.3390/jcm8111775
16. Yu CS, Lin YJ, Lin CH, Wang ST, Lin SY, Lin SH, et al. Predicting metabolic syndrome with machine learning models using a decision tree algorithm: retrospective cohort study. *JMIR Med Inform*. (2020) 8:e17110. doi: 10.2196/17110
17. Rojas E, Munoz-Gama J, Sepúlveda M, Capurro D. Process mining in healthcare: a literature review. *J Biomed Inform*. (2016) 61:224–36. doi: 10.1016/j.jbi.2016.04.007
18. Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Netw*. (2002) 15:11–39. doi: 10.1016/S0893-6080(01)00111-3
19. Joshi S, Shenoy D, Simha GGV, Rrashmi PL, Venugopal KR, Patnaik LM. Classification of Alzheimer's disease and Parkinson's disease by using machine learning and neural network methods. In: *2010 Second International Conference on Machine Learning and Computing*. (2010). p. 218–22. doi: 10.1109/ICMLC.2010.45
20. Moradi M, Ghadiri N. Different approaches for identifying important concepts in probabilistic biomedical text summarization. *Artif Intell Med*. (2018) 84:101–16. doi: 10.1016/j.artmed.2017.11.004
21. Yu CS, Lin YJ, Lin CH, Lin SY, Wu JL, Chang SS. Development of an online health care assessment for preventive medicine: a machine learning approach. *J Med Internet Res*. (2020) 22:e18585. doi: 10.2196/18585
22. Yu CS, Lin CH, Lin YJ, Lin SY, Wang ST, Wu JL, et al. Clustering heatmap for visualizing and exploring complex and high-dimensional data related to chronic kidney disease. *J Clin Med*. (2020) 9:403. doi: 10.3390/jcm9020403
23. Yao R-Q, Jin X, Wang G-W, Yu Y, Wu G-S, Zhu Y-B, et al. A machine learning-based prediction of hospital mortality in patients with postoperative sepsis. *Front Med (Lausanne)*. (2020) 7:445. doi: 10.21203/rs.2.24188/v1
24. Orrù G, Gemignani A, Ciacchini R, Bazzichi L, Conversano C. Machine learning increases diagnosticity in psychometric evaluation of alexithymia in fibromyalgia. *Front Med (Lausanne)*. (2020) 6:319. doi: 10.3389/fmed.2019.00319
25. Shneiderman B. Tree visualization with tree-maps: 2-d space-filling approach. *J ACM Trans Graph*. (1992) 11:92–9. doi: 10.1145/102377.115768
26. Hugine AL, Guerlain SA, Turrentine FE. Visualizing surgical quality data with treemaps. *J Surg Res*. (2014) 191:74–83. doi: 10.1016/j.jss.2014.03.046
27. Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, et al. *Package 'caret'* (2020). Available online at: https://cran.r-project.org/web/packages/caret/caret.pdf (accessed June 20, 2020).
28. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer (2016). doi: 10.1007/978-3-319-24277-4
29. Minakshi, Vohra R, Gimpy. Missing value imputation in multi attribute data set. *Int J Comput Sci Inf Technol*. (2014) 5:1–7. Available online at: http://www.ijcsit.com/docs/Volume%205/vol5issue04/ijcsit20140504108.pdf
30. Torgo, L. *Data Mining With R, Learning With Case Studies Chapman and Hall/CRC*. Boca Raton, FL (2010). Available online at: http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR
31. Balakrishnama S, Ganapathiraju A. Linear discriminant analysis-a brief tutorial In: *Institute for Signal and information Processing*. Mississippi (1998). p. 1–8.
32. Izenman A. *Linear Discriminant Analysis. Modern Multivariate Statistical Techniques*. New York, NY: Springer (2013). doi: 10.1007/978-0-387-78189-1_8
33. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*. New York, NY: John Wiley & Sons (2013). doi: 10.1002/9781118548387
34. Zhang C, Ma Y. *Ensemble Machine Learning: Methods and Applications*. New York, NY: Springer (2012). doi: 10.1007/978-1-4419-9326-7
35. Rokach L. Ensemble-based classifiers. *Artif Intell Rev*. (2010) 33:1–39. doi: 10.1007/s10462-009-9124-7
36. Breiman L. Random forests. *J Mach Learn*. (2001) 45:5–32. doi: 10.1023/A:1010933404324
37. Freund Y, Schapire RE. A desicion-theoretic generalization of on-line learning and an application to boosting. In: *European conference on computational learning theory*. New York, NY: Springer (1995). p. 23–37. doi: 10.1007/3-540-59119-2_166
38. Alfaro E, Gámez M, Garcia NJJOS. adabag: an R package for classification with boosting and bagging. *J Stat Softw*. (2013) 54:1–35. doi: 10.18637/jss.v054.i02
39. Karatzoglou A, Meyer D, Hornik KJJOS. Support vector machines in R. *J Stat Softw*. (2006) 15:1–28. doi: 10.18637/jss.v015.i09
40. Rish I. An empirical study of the naive Bayes classifier. In: *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*. Seattle, WA (2001). p. 41–6.
41. Perrot A, Bourqui R, Hanusse N, Lalanne F, Auber D. Large interactive visualization of density functions on big data infrastructure. In: *2015 IEEE 5th Symposium on Large Data Analysis and Visualization (LDAV)*. Chicago, IL (2015). p. 99–106. doi: 10.1109/LDAV.2015.7348077
42. De Lédinghen V, Vergniol J, Capdepont M, Chermak F, Hiriart JB, Cassinotto C, et al. Controlled attenuation parameter (CAP) for the diagnosis of steatosis: a prospective study of 5323 examinations. *J Hepatol*. (2014) 60:1026–103. doi: 10.1016/j.jhep.2013.12.018
43. Cardoso AC, Cravo C, Calçado FL, Rezende G, Campos CFF, Neto JMA, et al. The performance of M and XL probes of FibroScan for the diagnosis of steatosis and fibrosis on a Brazilian nonalcoholic fatty liver disease cohort. *Eur J Gastroenterol Hepatol*. (2020) 32:231–8. doi: 10.1097/MEG.0000000000001496
44. Sherlock S, Dooley J. *Diseases of the Liver and Biliary System*. London: Oxford: Blackwell Science (2002). doi: 10.1002/9780470986820
45. Pagano G, Pacini G, Musso G, Gambino R, Mecca F, Depetris N, et al. Nonalcoholic steatohepatitis, insulin resistance, and metabolic syndrome: further evidence for an etiologic association. *Hepatology*. (2002) 35:367–72. doi: 10.1053/jhep.2002.30690
46. Hamaguchi M, Kojima T, Takeda N, Nakagawa T, Taniguchi H, Fujii K, et al. The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. *Ann Intern Med*. (2005) 143:722–8. doi: 10.7326/0003-4819-143-10-200511150-00009
47. Lomonaco R, Ortiz-Lopez C, Orsak B, Webb A, Hardies J, Darland C, et al. Effect of adipose tissue insulin resistance on metabolic parameters and liver histology in obese patients with nonalcoholic fatty liver disease. *Hepatology*. (2012) 55:1389–97. doi: 10.1002/hep.25539
48. Chang MW, Chang SS, Lee CC, Sheu BF, Young YR. Hypokalemia and hypothermia are associated with 30-day mortality in patients with acute paraquat poisoning. *Am J Med Sci*. (2008) 335:451–6. doi: 10.1097/MAJ.0b013e318157cb6d
49. Kim JY, Ahn SV, Yoon JH, Koh SB, Yoon J, Yoo BS, et al. Prospective study of serum adiponectin and incident metabolic syndrome: the ARIRANG study. *Diabetes Care*. (2013) 36:1547–53. doi: 10.2337/dc12-0223
50. Kawamoto R, Kohara K, Tabara Y, Miki T, Otsuka N. Serum gamma-glutamyl transferase levels are associated with metabolic syndrome in community-dwelling individuals. *J Atheroscler Thromb*. (2009) 16:355–62. doi: 10.5551/jat.No414
51. Liu CF, Zhou WN, Fang NY. Gamma-glutamyltransferase levels and risk of metabolic syndrome: a meta-analysis of prospective cohort studies. *Int J Clin Pract*. (2012) 66:692–8. doi: 10.1111/j.1742-1241.2012.02959.x
52. Jimba S, Nakagami T, Takahashi M, Wakamatsu T, Hirota Y, Iwamoto Y, et al. Prevalence of non-alcoholic fatty liver disease and its association with impaired glucose metabolism in Japanese adults. *Diabet Med*. (2005) 22:1141–5. doi: 10.1111/j.1464-5491.2005.01582.x
53. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. (2011) 34:S62-9. doi: 10.2337/dc11-S062
54. Ziemer DC, Kolm P, Weintraub WS, Vaccarino V, Rhee MK, Twombly JG, et al. Glucose-independent, black-white differences in hemoglobin A1c levels: a cross-sectional analysis of 2 studies. *Ann Intern Med*. (2010) 152:770–7. doi: 10.7326/0003-4819-152-12-201006150-00004

55. Yki-Järvinen H. Liver fat in the pathogenesis of insulin resistance and type 2 diabetes. *Dig Dis*. (2010) 28:203–9. doi: 10.1159/000282087
56. Liu Z, Que S, Zhou L, Zheng S. Dose-response relationship of serum uric acid with metabolic syndrome and non-alcoholic fatty liver disease incidence: a meta-analysis of prospective studies. *Sci Rep*. (2015) 5:14325. doi: 10.1038/srep14325
57. Huang Q, Yu J, Zhang X, Liu S, Ge Y. Association of the serum uric acid level with liver histology in biopsy-proven non-alcoholic fatty liver disease. *Biomed Rep*. (2016) 5:188–92. doi: 10.3892/br.2016.698
58. Lee SJ, Oh BK, Sung K-C. Uric acid and cardiometabolic diseases. *Clin Hypertens*. (2020) 26:13. doi: 10.1186/s40885-020-00146-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

SYSTEMATIC REVIEW
published: 24 May 2021
doi: 10.3389/fmed.2021.609318

The Relationship Between Fragility Fractures and Pain Experience: A Systematic Review

Pei-En Chen[1,2†], *Ching-Wen Chien*[3†] *and Tao-Hsin Tung*[4*]

[1] *Institute of Health Policy and Management, College of Public Health, National Taiwan University, Taipei, Taiwan,* [2] *Taiwan Association of Health Industry Management and Development, Taipei, Taiwan,* [3] *Institute for Hospital Management, Tsing Hua University, Shenzhen, China,* [4] *Evidence-Based Medicine Center, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, China*

Purpose: This systematic review is conducted to explore the relationship between fragility fractures and pain experience.

Methods: We searched for relevant studies on Pubmed, Embase, Web of Science, and the Cochrane library without restrictions on language from inception until February 4th, 2021. The risk of bias and methodological quality was evaluated using the Newcastle-Ottawa Scale and ROBINS-I tool.

Results: Twenty-one studies were included in this systematic review. The so-called study reported participants with continuous post-fracture pain. The included studies showed that post- fractured pain can decrease with time, however, the continual pain can last at least 1 year even longer, and some participants would need to self-manage pain. Moreover, the limited range of motion was considered as a factor that might distress the normal development of daily activities.

Conclusions: The current evidence could not fully support that pain continues to influence patients' lives after a fragility fracture. However, it still showed the pain might come with fracture. The findings also could be useful to help health care providers better recognize and manage this clinical consequence of fractures. Nonetheless, future large-scale longitudinal studies will be required to evaluate the long-term effects of pain in fragility fractures.

Keywords: fragility fracture, pain, systematic review, fracture, discomfort

OPEN ACCESS

Edited by:
Ugo Cioffi,
University of Milan, Italy

Reviewed by:
Giovanni Iolascon,
University of Campania Luigi Vanvitelli, Italy
Marwan Bukhari,
University Hospitals of Morecambe Bay NHS Foundation Trust, United Kingdom

***Correspondence:**
Tao-Hsin Tung
ch2876@yeah.net

†These authors share first authorship

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *23 September 2020*
Accepted: *26 April 2021*
Published: *24 May 2021*

Citation:
Chen P-E, Chien C-W and Tung T-H (2021) The Relationship Between Fragility Fractures and Pain Experience: A Systematic Review. Front. Med. 8:609318. doi: 10.3389/fmed.2021.609318

INTRODUCTION

The World Health Organization located osteoporosis at the primary health care level, reporting that *"a fracture caused by injury that would be insufficient to fracture a normal bone...the result of reduced compressive and/or torsional strength of bone"* (1). From a clinical perspective, fragility fractures are considered as skeletal complications, leading to substantial morbidity, longer hospitalization period, higher health care costs, poorest quality of life, more severe disabilities, and death (2). Different fracture locations may as well involve diverse symptoms across time. Hip fractures are conceived as the most serious kind; with a 1-year mortality rate of 21% for women and 31% for men (3). Other kind of fragility-related fractures at other anatomical locations has been related with lower quality conditions of life, although most studies tend to focus on the impacts of the hip or the vertebral fractures (4).

While significant improvements have been achieved— both in surgical procedures and treatment tools—in this area, current information on incidences, risk factors, and medical costs of pain appears to be highly restricted. Currently, pain assessment and relief for patients with fragility fractures before and after surgery are placed as crucial topics for research. Besides, as surgical indications tend to be a direct procedure for most patients suffering from this kind of fragility fracture and adequate anesthesia is the basis for these successful so-called surgical procedures, there is still room for improvement in various anesthesia and sedation techniques. A part of the patients with fragility fractures, however, show surgical contraindications or prefer conservative treatment; in this kind of situation it becomes, indeed, highly significant to use methods to help to relieve this related discomfort, reduce the risk of adverse effects and improve the overall quality of life, accurate diagnosis and efficient pain eradication.

The purpose of this study was to assess the relationship between fragility fractures and pain experience. The results from this systematic review could further understand the fragility of fractures related to pain and guide health care to address the issues which matter to such patients.

MATERIALS AND METHODS

The methodology was written based on several studies published (5–7).

Literature Review

The Pubmed, Embase, Web of Science, and the Cochrane library for relevant studies without language limitations were used from inception until February 04, 2021. These databases included most of the academic research articles on this topic. The searched eligible studies were identified by scanning electronic databases using various combinations of Medical Subject Headings (MeSH) and non-MeSH terms.

Data Sources and Search Methods

The search process was extended by (1) perusing the reference section of all relevant studies, and (2) manually searching through the abstracts of key journals and articles published at major annual meetings. The review's population, intervention, comparison, outcome (PICO) items defined the search strategy: Population: all population, Intervention: fracture, Comparison: not applicable; Outcome: Pain. The search terms included all field and the following: (fragility OR fracture OR low traumatic) AND (fracture OR break OR split OR crack) AND (Pain OR Long-term pain OR Painful OR suffer OR discomfort OR hurt OR irritation OR tenderness OR soreness OR Fracture-related limitations OR disability Or disabled). **Table 1** shows the search strategy of databases.

Data Extraction and Quality Assessment

A data extraction form was used to obtain the following data from the included studies: first author (publication year), country, study duration, study subjects, age of study subjects, sex, assigned

TABLE 1 | Search strategy in PubMed up till Febuary 4th, 2021.

Pubmed		***N***
#1	fragility [all field]	19,914
#2	fragile [all field]	22,727
#3	low traumatic [all field]	12,625
#4	#1 OR #2 O #3	53,368
#5	fracture [all field]	322,337
#6	break [all field]	119,498
#7	split [all field]	102,207
#8	crack [all field]	20,423
#9	#5 OR #6 OR #7 OR #8	554,762
#10	Pain [all field]	869,082
#11	Long-term pain [all field]	45,413
#12	Painful [all field]	900,316
#13	Suffer [all field]	1,307,321
#14	discomfort [all field]	48,254
#15	hurt [all field]	4,855
#16	irritation [all field]	52,785
#17	tenderness [all field]	24,481
#18	soreness [all field]	3,526
#19	Fracture-related limitations [all field]	85
#20	disability [all field]	374,533
#21	disabled [all field]	374,533
#22	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	1,740,707
#23	#4 AND #9 AND #22	1,479
Cochrane		***N***
#1	fragility [all field]	1,454
#2	fragile [all field]	736
#3	low traumatic [all field]	2,223
#4	#1 OR #2 O #3	4,313
#5	fracture [all field]	18,713
#6	break [all field]	4,066
#7	split [all field]	10,282
#8	crack [all field]	390
#9	#5 OR #6 OR #7 OR #8	32,832
#10	Pain [all field]	191,678
#11	Long-term pain [all field]	13,367
#12	Painful [all field]	12,200
#13	Suffer [all field]	7,294
#14	discomfort [all field]	16,901
#15	hurt [all field]	519
#16	irritation [all field]	5,867
#17	tenderness [all field]	3,618
#18	soreness [all field]	1,995
#19	Fracture-related limitations [all field]	10
#20	disability [all field]	37,247
#21	disabled [all field]	3,175
#22	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	234,933
#23	#4 AND #9 AND #22	745

(Continued)

TABLE 1 | Continued

Embase		*N*
#1	fragility [all field]	41,107
#2	fragile [all field]	29,980
#3	low traumatic [all field]	144
#4	#1 OR #2 O #3	69,324
#5	fracture [all field]	396,777
#6	break [all field]	72,142
#7	split [all field]	91,222
#8	crack [all field]	9,824
#9	#5 OR #6 OR #7 OR #8	563,937
#10	Pain [all field]	1,378,467
#11	Long-term pain [all field]	1,777
#12	Painful [all field]	90,626
#13	Suffer [all field]	98,951
#14	discomfort [all field]	84,393
#15	hurt [all field]	6,051
#16	irritation [all field]	48,794
#17	tenderness [all field]	37,726
#18	soreness [all field]	4,539
#19	Fracture-related limitations [all field]	1
#20	disability [all field]	334,924
#21	disabled [all field]	60,695
#22	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	1,916,555
#23	#4 AND #9 AND #22	4,262
Web of Science		***N***
#1	fragility [all field]	27,390
#2	fragile [all field]	40,871
#3	low traumatic [all field]	24,839
#4	#1 OR #2 O #3	91,094
#5	fracture [all field]	509,008
#6	break [all field]	407,261
#7	split [all field]	287,946
#8	crack [all field]	261,946
#9	#5 OR #6 OR #7 OR #8	1,351,490
#10	Pain [all field]	708,471
#11	Long-term pain [all field]	42,587
#12	Painful [all field]	53,923
#13	Suffer [all field]	349,061
#14	discomfort [all field]	44,770
#15	hurt [all field]	12,487
#16	irritation [all field]	21,493
#17	tenderness [all field]	16,422
#18	soreness [all field]	3,980
#19	Fracture-related limitations [all field]	27
#20	disability [all field]	285,912
#21	disabled [all field]	57,642
#22	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	1,399,557
#23	#4 AND #9 AND #22	1,764

groups, and outcomes. The abstract and full-text screening was undertaken by Pei-En Chen and Tao-Hsin Tung. An assessment of methodological quality was performed independently by the authors (Pei-En Chen and Tao-Hsin Tung). The Newcastle-Ottawa Scale (NOS) was applied independently by two authors to determine the consistency of the selected studies (6). Any disagreement was discussed with a third senior author (Ching-Wen Chien). The NOS applies three domains (selection of study groups, comparability, and outcome assessment) to assess the quality of studies. A study could be awarded up to one star for each item within the selection and outcome domains and up to two stars for comparability. We viewed it as a study of high quality if seven or more stars were awarded. Besides, to increase the reproducibility and comparability of this systematic review to future reviews on a similar topic, we also evaluated risk of bias assessment using Risk of Bias in Non-randomized studies of Interventions (ROBINS-I) due to since it is the newest and most robust method of identifying the risk of bias in systematic reviews and meta-analyses (7).

Data Synthesis

The outcomes of the selected studies were assessed focusing on various measurements.

Follow by characteristics of outcome measurement:

1. Von Kroff questionnaire (which points both pain intensity score and disability score): It was developed in order to grade the severity of chronic pain (8).
2. EQ-5D (pain/discomfort): It is a standardized tool used to assess general health problems, which covers 5 main domains such as mobility, self-care, daily activities, pain/discomfort, and anxiety/depression (9). Thoughout this present research, only domains related to pain/discomfort will be discussed.
3. Numeric rating scale (NRS): Which displays results verabally reported by patients. The scale ranges from 0 (no pain) to 10 (worst imaginable pain) (10).
4. Visual Analog Scale (VAS, 0–10): Scale used in order to quantify a trait or attitude that is assumed to extend across a spectrum of values and cannot be directly measured easily. It is also used to measure the severity or frequency of different symptoms through epidemiological and clinical research (10).
5. SF-36 (pain/discomfort), which consists of an eight scaled score, containing weighted sums of questions (0–100). The eight sections are: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, mental health (10).
6. The Quality of Life Questionnaire (QUALEFFO-41; pain domain): To investigate about the improvements in the quality of life associated to day-to-day living, general well-being, and specific well-being of patients who have had any kind of the vertebral fractures named by International Osteoporosis Foundation (IOF) (11).
7. Geriatric Pain Scale (GPS, 0–100): Used to classify pain in patients and to assess physical, mental, cognitive, and behavioral responses to pain (12).

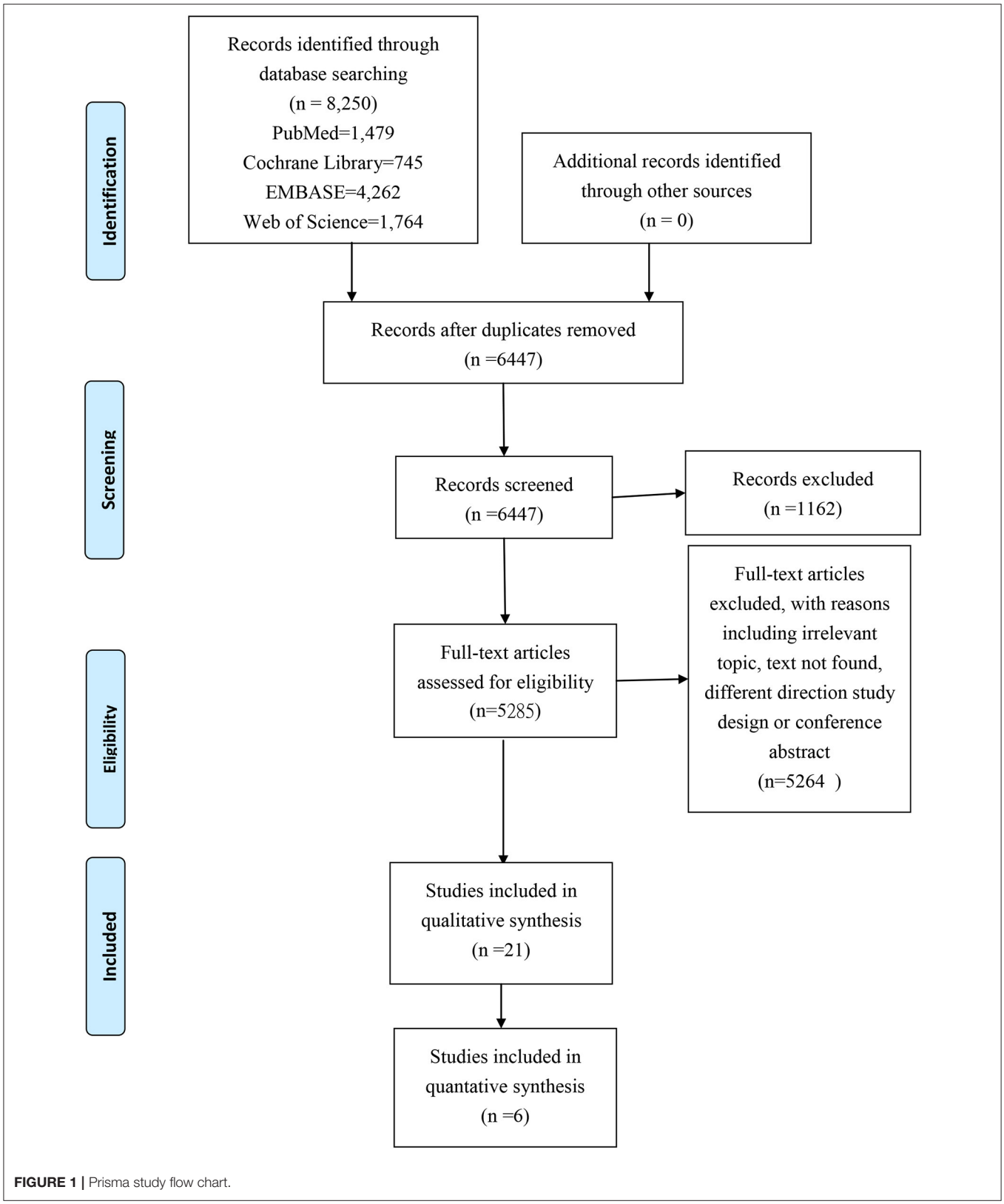


FIGURE 1 | Prisma study flow chart.

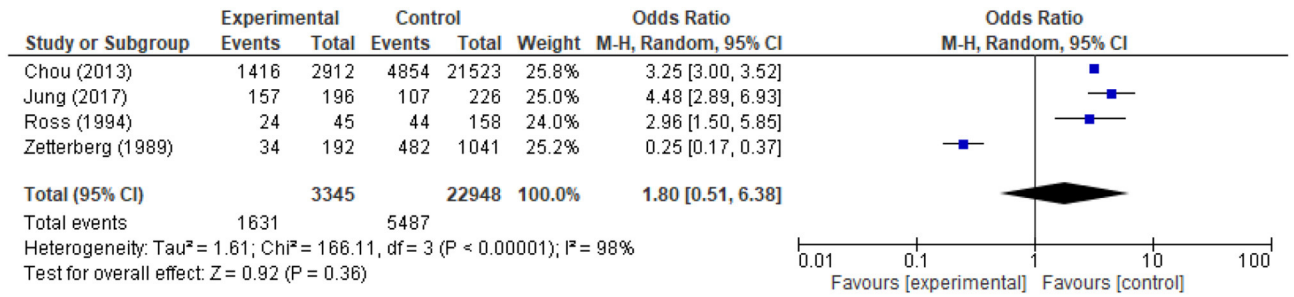


FIGURE 2 | Odds of pain in patients with fracture.

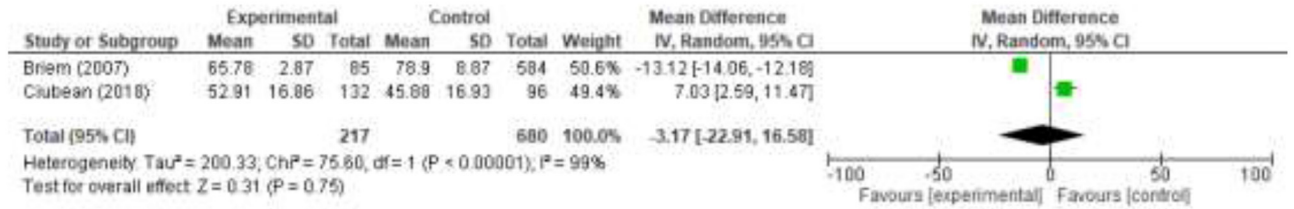


FIGURE 3 | Mean difference of pain in patients with fracture: SF-36 (pain/discomfort domain).

8. Pain Regulation Questionnaire (PRQ) which includes competences, intensity, anxiety, depression, avoidance, withdrawal and distraction of pain (13).
9. The amounts of individuals reporting pain.

The ROBIS tool was applied to assess the risk of bias in this systematic review. This tool consists of three phases; and this systematic review more specifically evaluated phase 2 and phase 3. During phase 2 there were four aspects evaluated: study eligibility criteria, identification and selection of studies, data collection and study appraisal, and synthesis and findings. Moreover, phase 3 integrated the overall risk from phase 2 (14).

Statistical Analysis

The tool Review Manager 5.3 (The Nordic Cochrane Center, The Cochrane Collaboration, 2014) was used in this study. We presented the risk of pain as OR with 95% CI and assessed heterogeneity by using the I^2 statistic. The I^2 statistic is used to evaluate the degree of variation across studies due to heterogeneity rather than by chance alone (15).

RESULTS

Study Characteristics

Figure 1 illustrates the results of this systematic review. From all the databases we searched in, with 8,250 records collected and after removing 1,803 duplicate articles, there was a total of 1,162 records which were excluded due to is protocol or other conference abstracts. Five thousand two hundred sixty-four full-text articles were also excluded for reasons such as irrelevancy of the topic, incapability to find the related text, differences on the purposes on the study design, conferences abstracts, or fracture-pain related articles with drugs/treatment. Finally, following a thorough review of all candidate papers, we identified a total of 21 studies that addressed the relationship between fragility fractures and pain experience.

Fragility Fracture and Pain Experience

Assessment of Pain and Disability

In Chou et al. (16), Jung et al. (17), and Ross et al. (18) fracture group showed a higher frequency on reported pain. However, Zetterberg et al. (19) indicated a different tendency from the other 3 studies. **Figure 2** displays that pain was not significantly associated with fracture than the control group. **Figure 3** specifies that pain was not related to the fracture group ($p = 0.75$). Among the 21 included studies, there were 8 which investigated the risk for fragility fracture and pain in postmenopausal females and 3 in the elderly, due to its higher vulnerability.

Apart from the above mentioned studies, other studies that were also utilized used different measurements to illustrate pain and disability. **Table 2** demonstrated some characteristics of the included studies. Firstly, Jin et al. (27), Hallal (25), and Kapucu and Unver (29) classified similar pain grades as *Slight, Mild, Moderate, Severe, Extreme*. However, we were not able to synthesize the data together, owing to the fact that the data classification was not standardized. Secondly, the duration of the presence of pain was delved in researches of authors as Jin et al. (27), Hallal (25), and Ozdemir et al. (31). Jin et al. (27) research indicated that pain would continue for <2, 2–8, ≥8 week, respectively, 183 (51.1%), 116 (32.4%), 59 (16.5%). In Hallal (25), pain could be sever hour: 40 (47.6%), 1 day:11 (13.1%), several days: 11 (13.1%), several weeks: 4 (4.8%), constantly:

TABLE 2 | Characteristics of included studies.

References	Country	Duration	Participants	Female/male	Age(year)	Outcome (20, 21)	NOS score[a]
Briem et al. (13)	Germany	Average follow-up of 5.3 ± 1.7 years after injury Range 3–8	Thoracolumbar fracture group = 85; control = 584	Thoracolumbar fracture group: 41/44 control: N/A	Thoracolumbar fracture group- Mean ± SD: 47.8 ± 12.8 Range 25–65 control: N/A	Thoracolumbar fracture vs. control group *(*Mean ± SD) *SF-36 questionnaire -pain domain 65.78 ± 2.87 vs. 78.90 ± 8.87 ($p < 0.05$) *Pain Regulation Questionnaire (PRQ) Pain competence: 41.40 ± 1.12 vs. 36.98 ± 8.41 ($p < 0.001$) Pain intensity: 23.84 ± 1.30 vs. 29.55 ± 8.62 ($p < 0.001$) Pain anxiety:26.63 ± 11.59 vs. 31.90 ± 8.58 ($p < 0.001$) Pain depression: 21.15 ± 1.27 vs. 25.33 ± 9.60 ($p < 0.001$) Pain avoidance: 25.72 ± 0.94 vs. 25.38 ± 8.10 Pain withdrawal: 24.93 ± 1.23 vs. 28.90 ± 10.90 ($p < 0.001$) Pain distraction: 33.29 ± 1.03 vs. 32.37 ± 8.21	S*** C** O***
Chou et al. (16)	Taiwan	Oct. 2002–Mar. 2003	24,435	11,937/12,498	Over 20 years	1. 2,912 participants with osteoporosis, 1,416 reported low back pain ($p < 0.001$). 2. Osteoporosis vs. non- osteoporosis, with low back pain OR = 2.55 (95% CI = 2.33–2.78); with frequent low back pain OR = 4.15 (95% CI = 3.66–4.70) 3. Adjusted sociodemographic factors, ORs of associated osteoporosis or not for frequent low back pain in females and males were 3.49 (95% CI = 2.99–4.07) and 5.77 (95% CI = 4.66–7.15), respectively.	S**** C** O***
Ciubean et al. (22)	Romania	Jun. 2016–Aug. 2017	364 postmenopausal women	364/0	Mean ± SD: Osteoporosis ($n = 228$): 65.5 ± 7.39 Control ($n = 136$): 63.45 ± 8.16 Range: 46–85	*SF-36 questionnaire -pain domain [median (IQR)] 1. Osteoporosis vs. Control:45 (45; 67.5) vs. 72.5 (55; 77.5) ($p < 0.001$) 2. Osteoporosis patients with fracture ($n = 132$) vs. without fracture ($n = 96$): 45 (45; 67.5) vs. 45 (35; 57.5) ($p = 0.035$) *QUALEFFO-41 -pain domain [median (IQR)] 1. Osteoporosis patients with fracture ($n = 132$) vs. without fracture ($n = 96$): 55 (30; 65) vs. 50 (30; 65) ($p = 0.446$)	S** C* O***
Fechtenbaum et al. (23)	France	–	588 have osteoporosis	588/0	Mean ± SD vertebral fracture ($n = 548$) vs. control group: 71.61 ± 5.01 vs. 71.00 ± 5.13 ($p = 0.43$)	QUALEFFO scores- pain domain(0-100) patients with no fracture ($n = 40$): 60 patients with sum of grade of fracture is 1 or 2 ($n = 133$): 51 patients with sum of grade of fracture is 3 or 4 ($n = 189$): 58 patients with sum of grade of fracture is 5–9 ($n = 146$): 58 patients with sum of grade of fracture is ≥10 ($n = 80$): 55	S** C* O**
Finsen (24)	Norway	–	307 subjects age of 50 years	222/85	Over 50 years	Patients self-reported pain (Some gave more than one answer and horizontal aggregates of percentages are therefore >loo) None (no infirmity): 31 (10.1%); foot (foot/leg/knee pain): 135 (44%); back (back pain): 96 (31.3%); hip (hip pain): 53 (17.3%)	S** C* O**
Gheorghita et al. (4)	Canada	At least 1 year	67	55/12	Range: 47–89	34 participants reported pain (30 female,4 male).	S*** C* O***
Hallal (25)	USA	–	101 women with diagnosed postmenopausal osteoporosis	101/0	Mean: 62.6	1. 84 participants reported the presence of back pain. 2. Frequency of back pain (daily: 33, weekly: 6, monthly: 20, less than once per month: 15) 3. Duration of back pain (sever hour: 40, 1 day: 11, several days:11, several weeks:4, constantly:18) 4.severity of back pain (very: 14, moderately: 45, mildly: 25)	S** C* O*

(Continued)

TABLE 2 | Continued

References	Country	Duration	Participants	Female/male	Age(year)	Outcome (20, 21)	NOS score[a]
Jahelka et al. (26)	Austria	Jun 2007–Jun. 2008	222	173/49	Mean ± SD: total: 79.3 ± 8.5	Visual analog scale (0–10) Osteopenic patients: 3.2 ± 2.6 Osteoporotic patients without fracture history: 3.2 ± 2.5 Osteoporotic patients with fracture history: 3.9 ± 2.7 ($p > 0.05$)	S*** C** O***
Jin et al. (27)	China	Nov. 1, 2016–Sep. 30, 2018	358 with vertebral fractures	284/74	Mean ± SD: 72.3 ± 9.4	1. Pain duration, weeks (< 2:183, 2–8:116, ≥8:59) 2. Spinal palpation tenderness: 197 3. Axial spinal percussion pain: 83 4. Radiating pain: 76 5. Pain grades (mild: 17, moderate: 121, severe: 220)	S*** C** O***
Jung et al. (17)	Korea	At least 6 month	196 with an osteoporotic vertebral compression fracture Reference population (28) = 600	Fracture group:165/31 Reference population:303/297	Mean ± SD: 72.7 ± 7.9	*EQ-5D (pain/discomfort domain) 1. No problem-39 (19.9%); 1 some problems: 139 (70.9%); serious problems: 18 (9.2%) 2. Age 50–59 ($n = 13$) vs. reference population 84.6 vs. 30.6% ($P < 0.001$) 3. Age ≥ 60 ($n = 183$) vs. Reference population 79.8 vs. 62.7% ($P < 0.001$)	S*** C* O**
Kapucu and Ünver (29)	Turkey	–	105 females with osteoporosis	105/0	Mean: 74.3 ± 7.5	Geriatric pain scale (0–100) 1. Mean: 57.6 ± 17.5; Min = 16.6; Max = 92.8 2. Pain level ($n = 104$) Slight (0–30) = 7 (6.7%); Mild (31–69) = 70 (67.3%); Severe (70–100) = 27 (26.0%)	S*** C* O**
Miyakoshi et al. (30)	JAPAN	–	174 consecutive women with postmenopausal osteoporosis	174/0	Mean ± SD back pain ($n = 159$) vs. Non-back pain ($n = 15$): 67.8 ± 6.5 vs. 65.5 ± 7.0 ($p = 0.18$)	1. 159 patients (91.4%) complained of back pain.	S** C* O**
Qzdemir et al. (31)	Turkey	–	909 patients		Mean: 60 Range: 33–89	1. 695 patients (76.45%) reported experiencing pain 2. The duration of the presence of pain was 8.7 ± 5.27 year [Min:1, Max: 26]	S** C* O**
Ramírez-Pérez et al. (32)	Mexico	6 month	136 with hip fracture	95/41	Mean ± SD: 77 ± 10	EQ-5D(pain/discomfort domain) 1. 1st, 3rd, and 6th month patients report pain, respectively, 122 (89.7%), 92 (68%), 72 (52.9%) 2. patients report pain, respectively, in level 1, 2, 3 1st month:148,735; 6th month: 646,210 (level 1: indicating no problem; level 2: indicating some problems; level 3: indicating extreme problems)	S** C** O***
Ribom et al. (33)	Sweden	–	36 women with osteoporosis and verified with vertebral fracture	36/0	Mean ± SD: 74.6 ± 8.3 Median: 76.6 Range: 57–87	Numeric rating scale (NRS) 1. Maximum pain: Mean ± SD: 5.9 ± 1.8; median: 6; range: 2–8 2. Minimum pain. Mean ± SD: 1.9 ± 2.5; median: 2; range: 0–8 3. Average pain: Mean ± SD: 4.8 ± 2.1; median: 5; range: 0–8	S** C* O***
Ross et al. (18)	USA	Each of ∼1.5 years duration	1,098 Japanese ancestry	1,098/0	Mean: 63.3 Range: 43–80	*The original population ($n = 1,098$) 1. 200 of these women had responded to questions about back pain, the number who reported increased frequency of back pain after the fracture was 16 (46%) of 35 subjects with new vertebral fractures, 1 (10%) of 10 subjects with prevalent fractures only, and 21 (14%) of 155 subjects without vertebral fractures.	S*** C*

(Continued)

TABLE 2 | Continued

References	Country	Duration	Participants	Female/male	Age(year)	Outcome (20, 21)	NOS score[a]
						2. Incidence of increased frequency of back pain With vertebral fractures vs. without vertebral fractures Incident fractures: OR = 6.4 (95% CI: 2.6–15.6); $p < 0.05$ Prevalent fractures OR = 1.7 (95% CI: 0.5–5.6): $p > 0.05$ *The most examination ($n = 203$) 1. 28.1% reported some frequency of back pain since their previous visit. 2. Among the subjects with and without incident vertebral fractures ($n = 45$ and 158), the proportions reporting some frequency of back pain were 53 and 21%, respectively.	O**
Sale et al. (34)	Canada	6 month	21 who had sustained fractures	16/5	Range: 51–87	11 participants reported persistent pain	S*** C* O***
Scaturro et al. (35)	Italy	Jan. 2016–Jan. 2018	513 post-menopausal women over 50, having back pain for at least 3 months, not responding to conservative treatment, with NRS between 2 and 4 (mild pain) and SF 36 between 60 and 100.	513/0	Mean: 72 Range: 50–89	Numeric rating scale (NRS) 1. 77.5 % ($n = 165$) of patients referred an NRS rate between 2 and 3 (first group) and 22.5% ($n = 48$) a rate of 4 (second group). 2. The correlation between the pain (NRS) and the number of vertebral fragility fractures ($P < 0.001$).	S** C* O**
Suzuki et al. (36)	Sweden	A year (Dec. 2003–Nov. 2006)	107	72/35	Mean ± SD: 75.5 ± 11.9 Range: 42–96	*Von Korff Pain Intensity score(0-100) 70.9 ± 19.3 (3 weeks), 61.5 ± 21.4 (3 months), 60.7 ± 21.6 (6 months), 60.5 ± 23.0 (12 months); *Von Korff Disability score (0–100) disability means scored 68.9 ± 23.6 (3 weeks), 56.4 ± 25.5 (3 months), 51.0 ± 27.5 (6 months), 53.9 ± 27.8 (12 months) ($P < 0.001$). *EQ-5D 1.Total score: 0.37 ± 0.37 (3 week), 0.52 ± 0.35 (3 months), 0.54 ± 0.36 (6 months), 0.52 ± 0.38 (12 month) ($p < 0.001$). 2. The number of patients reporting moderate or severe problems in pain/discomfort domain 97% (3 week), 89% (3 months), 87% (6 months), 89%(12 month) ($p < 0.001$).	S** C* O**
Tulay et al. (37)	Turkey	Jan.-Dec. 2016	172 with rib fracture	66/106	Medican: 47 Range: 18–85	Numeric rating scale (NRS) (0–10) 1. At 15th days, 3rd month, 6th month, the pain level of <65 yr participants were significant lower than ≥65 yrs group. 2. At 15th days, 3rd month, 6th month, the pain level of < female were significant higher than ≥ male. 3. Patients have 2 rib fractures with significant higher pain level than who has only one fracture.	S** C* O***

(Continued)

TABLE 2 | Continued

References	Country	Duration	Participants	Female/male	Age(year)	Outcome (20, 21)	NOS score[a]
Zetterberg et al. (19)	Sweden	1 year	Hip fracture patients:868 (final was 840 patients) Control group: 2,251	Hip fracture patients: 623/245 Control group: 1,333/918	Hip fracture patients mean Female: 79.0 Male: 73.9	Back pain during last 10 years 1. Female ($P < 0.001$) Hip fracture patients: 23% ($n = 143$) vs. Control group: 45% ($n = 600$) 2. Male ($P < 0.001$) Hip fracture patients:20% ($n = 49$) vs. Control group: 48% ($n = 441$)	S** C* O**

[a]*Scale domains: S selection of study groups, C comparability, O outcome assessment. Each "*" counts one point in different domain.*

18 (21.4%). Ozdemir et al. (31) demonstrated 8.7 ± 5.27 year. Thirdly, NRS was used in Tulay et al. (37), Ribom et al. (33), and Scaturro et al. (35). Tulay et al. (37) focused on illustrating the pain duration. Ribom et al. (33) did not specify how low is the participants with fracture, which encountered difficulties in the comparison process. Scaturro et al. (35) study only mentioned that the pain (NRS) was significant if it was directly related to the number of vertebral fractures. Fourthly, EQ-5D was applied both by Ramírez-Pérez et al. (32) and Jung et al. (17). Ramírez-Pérez et al. (32) did not have the control group, resulting on the comparison not being successfully completed. Fifthly, Qualleffo-41 was used throughout the study of Fechtenbaum et al. (23) and Ciubean et al. (22) and provided also a different illustration, which we could not compare. Sixthly, there was a certain pain assessment used in the study: Von Korff Pain Intensity and Disability questionnaires (36), Geriatric Pain Scale (0–100) (29), Pain Regulation Questionnaire (PRQ) (13), Visual Analog Scale (VAS) (26). From the study of Jahelka et al. (26), which used the SF-36 and QUALEFFO, we could not access the pain domain solely and, thus. we did not include the data in **Figure 3**.

Self-Reported Pain

In 10 studies have patients self-reported pain (4, 16, 18, 23–25, 27, 30, 31, 34). Chou et al. (16) study had 2,912 participants from which 1,416 reported some kind of low back pain (48.6%, $p < 0.001$). Fechtenbaum et al. (23) 548 of 588 reported pain problems. Finsen (24) could report 276 cases of pain out of 307 subjects who had suffered any kind of fracture, among the 307 participants who reported pain, the different body parts affected would be seen as followed, foot (foot/leg/knee pain):135 (44%); back (back pain): 96 (31.3%); hip (hip pain):53 (17.3%). Gheorghita et al. (4) enrolled 67 applicants who had suffered from any kind of fracture, from which 34 of them reported fracture-related pain (5.7%) ith fragility fracture. During Hallal's (25) research, 84 participants (83.1% out of the total) reported some presence of back pain. Jin et al. (27) reported 197 (55.0%) spinal palpation tenderness, 82 (23.2%) axial spinal percussion pain, and 76 (21.2%) radiating pain. Miyakoshi et al. (30) reported 159 patients (91.4%) who would show some kind of discomfort related to back pain. Ozdemir et al. (31) showed 695 patients (76.45%) who had reported experiencing pain. Ross et al. (18), 28.1% ($n = 203$) reported frequent back pain in the studied patients. Sale et al. (34) recruited 21 participants whose ages ranged from 51 to 87, from which 11 individuals self-reported constant pain after a fracture. They also reported movement-related limitations including difficulties related to a range of motion, lifting capacity, or insufficient strength. On the other hand, the other 10 participants reported not suffering pain at the site of fracture. Similarly, there were four of them who reported a limited range of motion.

Risk of Bias Assessments

Twenty-one of the included studies were rated on the NOS scale, where the higher the score rated the better quality it would proof (see **Table 2**), a total of 7 were rated as "high quality," and none of the included studies was rated as "low quality." Also, ROBINS-I was used to evaluate the risk of bias throughout the

TABLE 3 | Risk of bias assessment using ROBINS-I.

References	Type of research	Pre-intervention		At intervention bias in classification of intervention	Post-intervention				Total
		Bias due to confounding	Bias in selection participants into study		Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported outcomes	Total bias
Briem et al. (13)	Retrospective	Low	Low	Low	Low	Low	Low	Low	Low
Chou et al. (16)	Cross-sectional	Low	Low	Low	Low	Low	Moderate	Moderate	Moderate
Ciubean et al. (22)	Cross-sectional	Low	Low	Low	Low	Low	Low	Low	Low
Fechtenbaum et al. (23)	Prospective	Low	Low	Low	Low	Low	Low	Moderate	Low
Finsen (24)	Quantitative cohort	Low	Low	Low	Low	Moderate	Moderate	Moderate	High
Gheorghita et al. (4)	Qualitative cohort	Low	Low	Unclear	Low	Low	Moderate	Moderate	Moderate
Hallal (25)	Prospective	Low	Low	Low	Low	Low	Moderate	Moderate	Moderate
Jahelka et al. (26)	Prospective	Low	Low	Low	Low	Low	Low	Low	Low
Jin et al. (27)	Prospective	Low	Low	Low	Low	Low	Low	Moderate	Low
Jung et al. (17)	Ambispective	Low	Low	Low	Low	Low	Low	Low	Low
Kapucu and Ünver (29)	Descriptive	Low	Low	Low	Low	Low	Low	Low	Low
Miyakoshi et al. (30)	Observational	Low	Low	Low	Low	Low	Low	Moderate	Low
Qzdemir et al. (31)	Retrospective	Low	Low	Low	Low	Low	Low	Moderate	Low
Ramírez-Pérez et al. (32)	Prospective	Low	Low	Low	Low	Moderate	Low	Low	Low
Ribom et al. (33)	Prospective	Low	Low	Low	Low	Moderate	Low	Moderate	Moderate
Ross et al. (18)	Cross-sectional	Low	Low	Low	Low	Low	Low	Moderate	Low
Sale et al. (34)	Qualitative cohort	Low	Low	Unclear	Low	Low	Moderate	Moderate	Moderate
Scaturro et al. (35)	Observational	Low	Low	Low	Low	Low	Low	Low	Low
Suzuki et al. (36)	Prospective cohort	Low	Low	Low	Low	Low	Low	Moderate	Low
Tulay et al. (37)	Prospective	Low	Low	Low	Low	Low	Low	Low	Low
Zetterberg et al. (19)	Prospective	Low	Low	Low	Low	Moderate	Moderate	Low	Moderate

TABLE 4 | Risk of Bias in Systematic Reviews (ROBIS tool) of the study.

Phase 2				Phase 3
Study eligibility criteria	**Identification and selection of studies**	**Data collection and study appraisal**	**Synthesis and findings**	**Risk of bias in review**
Low risk	Low risk	Low risk	High risk	Low risk

**Possible risk of bias levels: low, high, unclear.*

study (**Table 3**). Six articles were evaluated as "moderate risk" and 1 as "high risk" to have bias. Due to the small number of papers and the degree of heterogeneity in study designs, interventions, and outcome indices, the meta-analysis was not considered fully appropriate.

The Result Summary of ROBIS

A summary of the findings and the ROBIS assessment for each domain can be seen in **Table 4**. In phase 2, study eligibility criteria, identification and selection of studies, data collection, and study appraisal were rated as "low risk." Due to the fact that the used studies included different assessments, synthesis, and findings, the domain was rated as "high risk." Throughout phase 3, the overall risk of bias was rated "low risk."

DISCUSSION

Clinical Implications

To the best of our knowledge, we comprehend this study to be the first systematic review and meta-analysis to examine the impacts of pain among the fragility fracture population. Our study results support the hypothesis that frail patients with fractures were suffering from a continuous risk of pain, and as this further exceeded the typical length of time assumed essential for curing and resolution of pain. Also, our results also provide more clear evidence related to patients undergoing fragility fracture may experience significant long-term pain effects.

The trajectories of frailty degrees in the elder population could vary substantially, particularly when estimating the short-term or long-term treatment effect, personal lifestyle change, related comorbidity development, and severe disease progression (38). Exploring the frailty transition chronologically would help clinicians to obtain a further knowledge related to the effects modifications of frailty and a better estimate on future fracture risk. Previous studies indicated that due to changes in the spine shape and height loss, patients may experience uninterrupted back pain even after the acutely painful episode subsides after a vertebral fracture (4, 39). Wrist fractures also have been projected to be able to recover after 6 weeks' post-fracture (40). However, another study showed one physiological cycle of bone which would remodel in healthy adults, lasting from 4 to 6 months (41). Risk fracture is a most frequent feature on the Complex regional pain syndrome(CRPS) (42). Patients with hip fractures often presented comorbidities and cognitive impairment that frequently prevented their recovery (43).

To examine the relationship between fragility fracture and pain, our study included only observational studies. Ordinary meta-analyses on the efficacy of interventions merely obtained high-quality evidence from randomized controlled trials (6). However, randomized controlled trials are often not the best source of evidence for harm due to the study duration is often too short to detect long-term or rare adverse events (44, 45). In addition, it is not possible to randomize patients into the categories "with fragility fracture" or "without fragility fracture." Including observational studies in this systematic review was a strong point, as these studies could indicate the effect of short-term and long-term pain in the fragility fracture population.

Clinical Practice

This systematic review found that there is an influence of pain following fragility fracture. Based on the results, medical teams should develop the treatment and rehabilitation protocol to prevent or reduce the pain of post-fracture, and the protocol should include meditation, exercises, and integrated physical treatment. Some consideration about the role of pain killers as well as anti-osteoporosis drugs for pain relief in fragility fracture patients should be provided (46–48). For better pain improvement, the program should be continuous, progressive, and combined alternative strategies.

Methodological Considerations

From the methodological viewpoint, our study included several limitations. Firstly, based on the current information, we could not assess the fluctuating frailty status concerning the risk of fragility fracture. Secondly, due to the number of selected studies that could be quantified, it was not sufficient. Due to the various measurements, it was difficult to conduct a meta-analysis with enough sample sizes. Finally, when we used the ROBIS approach to assess the quality of the evidence for the systematic review, the evidence from all the included observational studies was initially rated as relatively low quality because of imprecision. The addition of more studies in the future may increase the quality of evidence.

Conclusions

The current evidence could not fully support that pain continues to influence patients' lives after a fragility fracture. However, it still exposed the pain might come with fracture. The findings also could be useful to help health care providers to better recognize and manage this clinical consequence of fractures. We

recommend research on a wider range of populations to provide more comprehensive and accurate findings.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

P-EC: conceptualization, methodology, software, data curation, and writing-original draft preparation. T-HT: writing-reviewing and editing. C-WC: supervision. All authors contributed to the article and approved the submitted version.

REFERENCES

1. World Health Organization. *Guidelines for Preclinical Evaluation and Clinical Trials in Osteoporosis*. Geneva: World Health Organization (1998).
2. Li G, Prior JC, Leslie WD, Thabane L, Papaioannou A, Josse RG, et al. Frailty and risk of fractures in patients with Type 2 diabetes. *Diabetes Care.* (2019) 42:507–13. doi: 10.2337/dc18-1965
3. von Friesendorff M, McGuigan FE, Wizert A, Rogmark C, Holmberg AH, Woolf AD, et al. Hip fracture, mortality risk, and cause of death over two decades. *Osteoporos Int.* (2016) 27:2945–53. doi: 10.1007/s00198-016-3616-5
4. Gheorghita A, Webster F, Thielke S, Sale JEM. Long-term pain experience after a fragility fracture. *Osteoporos Int.* (2018) 29:1093–104. doi: 10.1007/s00198-018-4399-7
5. Chen PE, Liu CY, Chien WH, Chien CW, Tung TH. Effectiveness of cherries in reducing uric acid and gout: a systematic review. *Evid Based Complement Alternat Med.* (2019) 2019:9896757. doi: 10.1155/2019/9896757
6. Chi CC, Chen TH, Wang SH, Tung TH. Risk of suicidality in people with psoriasis: a systematic review and meta-analysis of cohort studies. *Am J Clin Dermatol.* (2017) 18:621–7. doi: 10.1007/s40257-017-0281-1
7. Jiesisibieke ZL, Tung TH, Xu QY, Chen PE, Hsu SY, Liu Y, et al. Association of acute kidney injury with frailty in elderly population: a systematic review and meta-analysis. *Ren Fail.* (2019) 41:1021–7. doi: 10.1080/0886022X.2019.1679644
8. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain.* (1992) 50:133–49. doi: 10.1016/0304-3959(92)90154-4
9. Kim SH, Hwang JS, Kim TW, Hong YS, Jo MW. Validity and reliability of the EQ-5D for cancer patients in Korea. *Support Care Cancer.* (2012) 20:3155–60. doi: 10.1007/s00520-012-1457-0
10. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res.* (2011). 63(Suppl. 11):S240–52. doi: 10.1002/acr.20543
11. Cooper C, Harvey N, Dennison E. Worldwide epidemiology of osteoporotic fractures. In: Reginster J, Rizzoli R, edsitors. *Innovation in Skeletal Medicine*. Hampshire: Elsevier (2008).
12. Ferrell BA, Stein WM, Beck JC. The Geriatric Pain Measure: validity, reliability and factor analysis. *J Am Geriatr Soc.* (2000) 48:1669–73. doi: 10.1111/j.1532-5415.2000.tb03881.x
13. Briem D, Behechtnejad A, Ouchmaev A, Morfeld M, Schermelleh-Engel K, Amling M, et al. Pain regulation and health-related quality of life after thoracolumbar fractures of the spine. *Eur Spine J.* (2007) 16:1925–33. doi: 10.1007/s00586-007-0395-x
14. Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* (2016) 69:225–34. doi: 10.1016/j.jclinepi.2015.06.005
15. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* (2003) 327:557–60. doi: 10.1136/bmj.327.7414.557
16. Chou YC, Shih CC, Lin JG, Chen TL, Liao CC. Low back pain associated with sociodemographic factors, lifestyle and osteoporosis: a population-based study. *J Rehabil Med.* (2013) 45:76–80. doi: 10.2340/16501977-1070
17. Jung HJ, Park YS, Seo HY, Lee JC, An KC, Kim JH, et al. Quality of life in patients with osteoporotic vertebral compression fractures. *J Bone Metab.* (2017) 24:187–96. doi: 10.11005/jbm.2017.24.3.187
18. Ross PD, Davis JW, Epstein RS, Wasnich RD. Pain and disability associated with new vertebral fractures and other spinal conditions. *J Clin Epidemiol.* (1994) 47:231–9. doi: 10.1016/0895-4356(94)90004-3
19. Zetterberg C, Mannius S, Mellstrom D, Rundgren A, Astrand K. Osteoporosis and back pain in the elderly. A controlled epidemiologic and radiographic study. *Spine.* (1990) 15:783–6. doi: 10.1097/00007632-199008010-00009
20. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol.* (2014) 14:135. doi: 10.1186/1471-2288-14-135
21. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat Methods Med Res.* (2018) 27:1785–805. doi: 10.1177/0962280216669183
22. Ciubean AD, Ungur RA, Irsay L, Ciortea VM, Borda IM, Onac I, et al. Health-related quality of life in Romanian postmenopausal women with osteoporosis and fragility fractures. *Clin Interv Aging.* (2018) 13:2465–72. doi: 10.2147/CIA.S190440
23. Fechtenbaum J, Cropet C, Kolta S, Horlait S, Orcel P, Roux C. The severity of vertebral fractures and health-related quality of life in osteoporotic postmenopausal women. *Osteoporos Int.* (2005) 16:2175–9. doi: 10.1007/s00198-005-2023-0
24. Finsen V. Osteoporosis and back pain among the elderly. *Acta Med Scand.* (1988) 223:443–9. doi: 10.1111/j.0954-6820.1988.tb15896.x
25. Hallal JC. Back Pain with Postmenopausal osteoporosis and vertebral fractures. *Geriatr Nurs.* (1991) 12:285–7. doi: 10.1016/S0197-4572(05)80266-6
26. Jahelka B, Dorner T, Terkula R, Quittan M, Bröll H, Erlacher L. Health-related quality of life in patients with osteopenia or osteoporosis with and without fractures in a geriatric rehabilitation department. *Wien Med Wochenschr.* (2009) 159:235–40. doi: 10.1007/s10354-009-0655-y
27. Jin H, Ma X, Liu Y, Liu M, Yin X, Fan W, et al. Back pain from painful osteoporotic vertebral fractures: discrepancy between the actual fracture location and the location suggested by patient-reported pain or physical examination findings. *Osteoporos Int.* (2020) 31:1721–32. doi: 10.1007/s00198-020-05434-9
28. Lee SI. *Validity and Reliability Evaluation for EQ-5D in Korea*. Osong: Korea Centers for Disease Control and Prevention (2011).
29. Kapucu S, Ünver G. Pain and frailty in elderly females with osteoporosis. *Turk J Geriatr.* (2017) 20:306–14.
30. Miyakoshi N, Hongo M, Shimada Y. *Prevalence of Back Pain in Postmenopausal Osteoporosis and Associations With Multiple Spinal Factors, Osteoporosis, Yannis Dionyssiotis*. Akita: IntechOpen (2012).
31. Ozdemir F, Rodoplu M, Zateri C. Pain, body mass index, and bone mineral density in patients with postmenopausal and senile osteoporosis. *J Musculoskelet Pain.* (2009) 17:59–66. doi: 10.1080/10582450802672362
32. Ramírez-Pérez E, Clark P, Carlos F, Camacho A, Franco-Marina F. Health-related quality of life after surgery for hip fracture: a multicentric study in Mexican population. *Medwave.* (2014) 14:e5972. doi: 10.5867/medwave.2014.05.5972
33. Ribom EL, Kindmark A, Ljunggren Ö. Hyperkyphosis and back pain are not associated with prevalent vertebral fractures in women with osteoporosis. *Physiother Theory Pract.* (2015) 31:182–5. doi: 10.3109/09593985.2014.989295
34. Sale JEM, Frankel L, Thielke S, Funnell L. Pain and fracture-related limitations persist 6 months after a fragility fracture. *Rheumatol Int.* (2017) 37:1317–22. doi: 10.1007/s00296-017-3761-y
35. Scaturro D, Lauricella L, Tumnoinelli LG, Tomasello S, Mauro GI. Is there a relationship between mild-moderate back pain and fragility fractures? Original investigation. *Acta Med.* (2020) 36:2149–53. doi: 10.19193/0393-6384_2020_3_336

36. Suzuki N, Ogikubo O, Hansson T. The course of the acute vertebral body fragility fracture: its effect on pain, disability and quality of life during 12 months. *Eur Spine J.* (2008) 17:1380–90. doi: 10.1007/s00586-008-0753-3
37. Tulay CM, Yaldiz S, Bilge A. Do we really know the duration of pain after rib fracture? *Kardiochir Torakochirurgia Pol.* (2018) 15:147–50. doi: 10.5114/kitp.2018.78437
38. Li G, Papaioannou A, Thabane L, Cheng J, Adachi JD. Frailty change and major osteoporotic fracture in the elderly: data from the global longitudinal study of osteoporosis in women 3-year hamilton cohort. *J Bone Miner Res.* (2016) 31:718–24. doi: 10.1002/jbmr.2739
39. Nevitt MC, Ettinger B, Black DM, Stone K, Jamal SA, Ensrud K, et al. The association of radiographically detected vertebral fractures with back pain and function: a prospective study. *Ann Intern Med.* (1998) 128:793–800. doi: 10.7326/0003-4819-128-10-199805150-00001
40. Hollevoet N, Verdonk R. Outcome of distal radius fractures in relation to bone mineral density. *Acta Orthop Belg.* (2003) 69:510–4.
41. Das S, Crockett JC. Osteoporosis—a current view of pharmacological prevention and treatment. *Drug Des Dev Ther.* (2013) 7:435–48. doi: 10.2147/DDDT.S31504
42. de Sire A, Paoletta M, Moretti A, Brandi ML, Iolascon G. Complex regional pain syndrome: facts on causes, diagnosis and therapy. *Clin Cases Miner Bone Metab.* (2018) 15:166–72. doi: 10.11138/ccmbm/2015.12.3s.004
43. Beaupre LA, Binder EF, Cameron ID, Jones CA, Orwig D, Sherrington C, et al. Maximising functional recovery following hip fracture in frail seniors. *Best Pract Res Clin Rheumatol.* (2013) 27:771–88. doi: 10.1016/j.berh.2014.01.001
44. Chi CC, Wang SH, Wojnarowska F, Kirtschig G, Davies E, Bennett C. Safety of topical corticosteroids in pregnancy. *Cochrane Database Syst Rev.* (2015) 26:CD007346. doi: 10.1002/14651858.CD007346.pub3
45. Chi CC, Wang SH, Kirtschig G. Safety of topical corticosteroids in pregnancy. *JAMA Dermatol.* (2016) 152:934–5. doi: 10.1001/jamadermatol.2016.1009
46. Iolascon G, Cisari C, Moretti A, Frizzi L, Gimigliano R, Gimigliano F. NSAIDs and opioids in management of fragility fractures. *Aging Clin Exp Res.* (2013) 25(Suppl. 1):S97–100. doi: 10.1007/s40520-013-0100-3
47. Iolascon G, Moretti A, Toro G, Gimigliano F, Liguori S, Paoletta M. Pharmacological therapy of osteoporosis: what's new? *Clin Interv Aging.* (2020) 15:485–91. doi: 10.2147/CIA.S242038
48. Moretti A, de Sire A, Curci C, Toro G, Gimigliano F, Iolascon G. Effectiveness of denosumab on back pain-related disability and quality-of-life in patients with vertebral fragility fractures. *Curr Med Res Opin.* (2019) 35:151–5. doi: 10.1080/03007995.2018.1545636

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

ORIGINAL RESEARCH
published: 31 May 2021
doi: 10.3389/fmed.2021.646875

Risk Stratification Score to Predict Readmission of Patients With Acute Decompensated Cirrhosis Within 90 Days

Xiaomei Xu[1,2†], *Juntao Tan*[3†], *Haolin Wang*[3], *Wenlong Zhao*[3,4] *and Bo Qin*[1*]

[1] Department of Infectious Diseases, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [2] Department of Gastroenterology, The Fifth People's Hospital of Chengdu, Chengdu, China, [3] College of Medical Informatics, Chongqing Medical University, Chongqing, China, [4] Medical Data Science Academy, Chongqing Medical University, Chongqing, China

OPEN ACCESS

Edited by:
Ugo Cioffi,
University of Milan, Italy

Reviewed by:
Jun Lyu,
First Affiliated Hospital of Jinan University, China
Huiguo Ding,
Capital Medical University, China
Keqing Shi,
First Affiliated Hospital of Wenzhou Medical University, China

***Correspondence:**
Bo Qin
cqqinbo@126.com

[†] *These authors have contributed equally to this work and share first authorship*

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *28 December 2020*
Accepted: *10 May 2021*
Published: *31 May 2021*

Citation:
Xu X, Tan J, Wang H, Zhao W and Qin B (2021) Risk Stratification Score to Predict Readmission of Patients With Acute Decompensated Cirrhosis Within 90 Days. Front. Med. 8:646875. doi: 10.3389/fmed.2021.646875

Background and Aims: Patients with acute decompensated (AD) cirrhosis are frequently readmitted to the hospital. An accurate predictive model for identifying high-risk patients may facilitate the development of effective interventions to reduce readmission rates.

Methods: This cohort study of patients with AD cirrhosis was conducted at six tertiary hospitals in China between September 2012 and December 2016 (with 705 patients in the derivation cohort) and between January 2017 and April 2020 (with 251 patients in the temporal validation cohort). Least absolute shrinkage and selection operator Cox regression was used to identify the prognostic factors and construct a nomogram. The discriminative ability, calibration, and clinical net benefit were evaluated based on the C-index, area under the curve, calibration curve, and decision curve analysis. Kaplan–Meier curves were constructed for stratified risk groups, and log-rank tests were used to determine significant differences between the curves.

Results: Among 956 patients, readmission rates were 24.58, 42.99, and 51.78%, at 30, 60, and 90 days, respectively. Bacterial infection was the main reason for index hospitalization and readmission. Independent factors in the nomogram included gastrointestinal bleeding [hazard rate (HR): 2.787; 95% confidence interval (CI): 2.221–3.499], serum sodium (HR: 0.955; 95% CI: 0.933–0.978), total bilirubin (HR: 1.004; 95% CI: 1.003–1.005), and international normalized ratio (HR: 1.398; 95% CI: 1.126–1.734). For the convenience of clinicians, we provided a web-based calculator tool (https://cqykdx1111.shinyapps.io/dynnomapp/). The nomogram exhibited good discrimination ability, both in the derivation and validation cohorts. The predicted and observed readmission probabilities were calibrated with reliable agreement. The nomogram demonstrated superior net benefits over other score models. The high-risk group (nomogram score >56.8) was significantly likely to have higher rates of readmission than the low-risk group (nomogram score ≤56.8; $p < 0.0001$).

Conclusions: The nomogram is useful for assessing the probability of short-term readmission in patients with AD cirrhosis and to guide clinicians to develop individualized treatments based on risk stratification.

Keywords: acute decompensated cirrhosis, readmission, independent predictors, nomogram, risk stratification

INTRODUCTION

Acute decompensated (AD) cirrhosis, defined as the acute development of ascites, hepatic encephalopathy, gastrointestinal hemorrhage, bacterial infections, or a combination of these factors, is the main cause of hospitalization and mortality in patients with cirrhosis (1–3). Owing to these complications, patients with decompensated cirrhosis are more frequently hospitalized and rapidly readmitted shortly after discharge. Hospital readmission is considered as a surrogate marker of the quality of healthcare delivery systems. Moreover, readmissions are associated with negative outcomes in patients and their families, and they have a significant impact on the overall costs of health care (4).

An estimated 27.1% of readmissions may be avoidable (5). To date, several studies have focused on assessing the predictors of readmission in decompensated cirrhosis (6–18). However, hospital readmission rates of patients with cirrhosis remain high, ranging from 10 to 50%, with a pooled estimate of 26% at 30 days and 21–71% at 90 days (19). The reason for these findings may be that the effective implementation of interventions requires not only understanding risk factors, but also identifying high-risk patients on the basis of highly accurate individualized risk predictive models, given that misleading risk estimates often lead to inappropriate treatment choices (20, 21). Therefore, the establishment of a model that can effectively predict and distinguish the individual risk of readmission remains an urgent medical requirement.

The current study was conducted to determine the readmission risk factors for patients with AD cirrhosis-related complications, to develop and temporally validate a nomogram to estimate the individual probability of readmission within 90 days, and to guide clinicians to develop individualized counseling programs and treatments based on risk stratification.

Abbreviations: AD, acute decompensated; TRIPOD, transparent reporting of a multivariable prediction model for individual prognosis or diagnosis; ACLF, acute-on-chronic liver failure; TIPS, transjugular intrahepatic portosystemic shunt; GI, gastrointestinal; HBV, hepatitis B virus; HCV, hepatitis C virus; INR, international normalized ratio; MELDs, model for end-stage liver disease score; CTPs, Child-Turcotte-Pugh score; CLIF-ADs, chronic liver failure-consortium acute decompensation scores; MELD-Nas, MELD-Na score; SD, standard deviation; LASSO, least absolute shrinkage and selection operator; λ, lambda; AUC, area under the curve; ROC, receiver operating characteristic; C-index, concordance index; DCA, decision curve analyses; TB, total bilirubin; CI, confidence interval; EVB, esophageal variceal bleeding.

PATIENTS AND METHODS

Population and Study Design

We conducted a multicenter retrospective prognostic study of inpatients with AD cirrhosis at six tertiary hospitals in Chongqing, China, including the Second Affiliated Hospital of Chongqing Medical University, Yong Chuan Hospital of Chongqing Medical University, Third Affiliated Hospital of Chongqing Medical University, University-Town Hospital of Chongqing Medical University, People's Hospital of Tong Liang District, and Southeast Hospital of Chongqing. We followed the transparent reporting of a multivariable predictive model for individual prognosis or diagnosis (TRIPOD) guidelines for model development and validation (22). Clinical data were collected using electronic medical record systems. Consecutive patients with AD cirrhosis admitted to the above hospitals from September 2012 to December 2016 were enrolled as the derivation cohort, and those with the same clinical characteristics hospitalized between January 2017 and April 2020 were enrolled as the temporal validation cohort. To determine whether patients admitted in April 2020 were readmitted to the hospital, we conducted follow-up until August 2020. The end points were cirrhosis-related readmissions within 90 days from the date of hospital discharge. In cases of multiple admissions, only the first readmission was considered.

The study protocol was reviewed and approved by the ethics committee of the First Affiliated Hospital of Chongqing Medical University. Due to its retrospective nature, this study required no conformed consent.

Inclusion and Exclusion Criteria

The inclusion criteria were as follows: (1) patients aged $\geq$ 18 years and (2) hospital admission for AD cirrhosis.

The exclusion criteria were as follows: (1) patients with acute-on-chronic liver failure (ACLF): diagnosis of ACLF was based on the criteria from the consensus recommendation of the Asian Pacific Association for the Study of the Liver (23); (2) liver cancer or other active malignancies; (3) evidence of congestive heart failure, chronic kidney disease, or other significant chronic extrahepatic diseases; (4) hospital stay $\leq$1 day; (5) endoscopic ligation of esophageal varices or transjugular intrahepatic portosystemic shunt (TIPS) in the initial hospitalization or elective hospital admission; (6) discharge against medical advice; (7) patients lost to follow-up or death during index hospitalization; and (8) patients with >30% of data missing. Details of readmission in other hospitals or planned procedures, surgery, and therapy were compiled from medical history.

AD cirrhosis was defined as the rapid development of one or more major complications of liver disease, such as ascites,

encephalopathy, gastrointestinal hemorrhage, bacterial infection, or a combination of these factors, requiring hospitalization (1, 24–27).

Ascites was recorded as the primary reason for admission if this was the sole criterion for admission, and infection was absent. Hepatic encephalopathy was characterized by altered mental status or neuropsychiatric abnormalities in the presence of liver cirrhosis after exclusion of other causes. Gastrointestinal (GI) bleeding was defined as the development of an upper and/or lower gastrointestinal hemorrhage of any etiology (27). Bacterial infection was defined in cases of spontaneous bacterial peritonitis, pneumonia, cellulitis, biliary tract infection, urinary system infection, and spontaneous bacteremia (26).

In the presence of more than one contributory factor, the main cause of admission was defined as follows: (1) in patients admitted with GI bleeding in the presence of ascites, bacterial infection, or hepatic encephalopathy, GI bleeding was considered the main cause; (2) in the absence of bleeding at admission, bacterial infection was the main cause of hospitalization; and (3) in patients with hepatic encephalopathy and ascites, the main cause was the former (11). The principal cause of hospitalization was subsequently assessed independently by two subspecialist physicians.

Treatment

Medical therapies were used for all patients during hospitalization and after discharge, such as antiviral therapy, diuretics, lactulose, non-selective beta-blockers, antibiotics, symptomatic and supportive therapies. Prophylactic antibiotics were not routinely administered after discharge.

Data Collection

Demographic, etiological, clinical, and laboratory data were recorded within 24 h of the first hospital visit. Demographic characteristics included age and sex. The etiological characteristics, including hepatitis B virus (HBV)/hepatitis C virus (HCV) infection, autoimmunity, and alcohol consumption, were assessed from medical history. Clinical data included length of hospital stay, complications related to liver cirrhosis, comorbidities, smoking history, alcohol consumption, and family history of liver disease. Laboratory analyses included liver function test, routine blood test, creatinine, blood urea nitrogen, serum sodium, serum potassium, and international normalized ratio (INR). End-stage Liver Disease score (MELDs), Child–Turcotte–Pugh score (CTPs), chronic liver failure-consortium acute decompensation scores (CLIF-C ADs), and MELD-Na score (MELD-Nas) were calculated at admission according to previously published criteria (28–31).

Statistical Analyses

Continuous variables were expressed as mean ± standard deviation (SD) or median (interquartile range), according to the distribution of normality. Categorical variables were reported as counts with percentages. Group comparisons of continuous variables were analyzed with the Mann–Whitney *U* test and categorical variables with χ^2 or Fisher's exact test. For variables with omission rates <30%, multiple imputation was used.

To avoid overfitting, we performed two steps of variable selection. First, we evaluated the association between readmission within 90 days and a set of potential predictors by using univariate analyses (Mann–Whitney *U* test for quantitative predictors and χ^2 or Fisher's exact test for binary predictors). Predictors with $p < 0.05$ were subsequently considered in an automated variable selection procedure within the least absolute shrinkage and selection operator (LASSO) framework to select the best predictor subset (32). The complexity of LASSO regression was controlled by a tuning parameter lambda (λ) with the rule that the penalty for each variable coefficient increases with λ value, and the relevant features with non-zero coefficients were selected that contributed to the final LASSO regression (33). The number of variables involved in the final model was considered based on the optimal λ value to balance the accuracy and simplicity of the model. Then, the retained variables were used to construct the nomogram using multivariate Cox regression. The nomogram was based on the fitted predictive model using R version 4.0.2 with the rms package (34). To streamline the power calculation estimation, we produced PowerTools, an interactive open-source web application, written in R code by using the Shiny framework (http://www.shinyapps.io/).

The discriminatory value of the models was assessed based on the concordance index (C-index). The area under the curve (AUC) of the receiver operating characteristic (ROC) curve was also used to evaluate the prognostic accuracy of the nomogram. To demonstrate the stability of the model, we applied bootstraps with 200 resamples to correct the C-index to overcome overfitting. Calibration curves were additionally drawn to evaluate the concordance between the predicted and observed probabilities. The nomogram model was validated with a temporal validation cohort using the same process of capability assessment. Decision curve analyses (DCA) were applied to compare the benefits and improved performance of different models (35).

According to the nomogram score, the patients were classified into two groups representing low and high risk. The optimal cut-off values for the total points of the nomogram were determined by maximizing the Youden index (sensitivity + specificity – 1). Kaplan–Meier curves were constructed for stratified risk groups, and log-rank tests were used to determine significant differences between the curves.

All tests were two-sided, and data were considered statistically significant at $p < 0.05$. Statistical analyses were performed using R software (version 4.0.2, Vienna, Austria).

RESULTS

Characteristics of the Study Cohort

A total of 8,402 patients met the inclusion criteria. Following the application of the exclusion criteria, 956 patients were finally included in the study, specifically 705 patients in the derivation cohort (from September 2012 to December 2016) and 251 patients in the temporal validation cohort (from January 2017 to April 2020). The study selection process is depicted as a flow chart (see **Supplementary Figure 1** for details).

TABLE 1 | Cirrhosis-related index hospitalizations and readmissions.

	GI bleeding	Bacterial infection	HE	Ascites	others
Index, No. (%)	335 (35.0)	519 (54.3)	47 (4.9)	55 (5.8)	N/A
30-day, No. (%)	53 (22.6)	105 (44.7)	21 (8.9)	9 (3.8)	47 (20.0)
60-day, No. (%)	89 (21.7)	198 (48.2)	36 (8.8)	15 (3.7)	73 (17.8)
90-day, No. (%)	104 (21.0)	242 (48.9)	40 (8.1)	19 (3.8)	90 (18.2)

GI, gastrointestinal; HE, hepatic encephalopathy; NA, not applicable.

TABLE 2 | The HR values of the independent risk factors for prediction of 90-day readmission in patients with acute decompensated cirrhosis.

	HR	95%CI	*P*-value
GI bleeding	2.787451	(2.2207,3.4988)	<0.001
Serum sodium	0.9552653	(0.9333,0.9777)	<0.001
Total bilirubin	1.0037945	(1.0026,1.0050)	<0.001
INR	1.3975204	(1.1264,1.7339)	0.002

GI, gastrointestinal; INR, international normalized ratio; HR, hazard ratio; CI, confidence interval.

The mean (SD) age of all patients was 58.8 (12.6) years, and 68.31% were male. The etiologies of cirrhosis were chronic hepatitis B (50.4%), alcoholic (10.7%), autoimmune liver disease (9.9%), chronic hepatitis C (5.0%), and other/cryptogenic factors (19.6%). The overall readmission rates for patients at 30, 60, and 90 days were 24.6, 43.0, and 51.8%, respectively. As shown in **Table 1**, bacterial infection was the main reason for index admission (54.3%), followed by GI bleeding (35.0%), hepatic encephalopathy (5.8%), and ascites (4.9%). Regarding the main reason for readmission, bacterial infection was the most common cause, followed by GI bleeding, other cirrhosis-related diseases, hepatic encephalopathy, and ascites at the 30-, 60-, and 90-day time points.

Based on the baseline characteristics of the two cohorts of patients as listed in **Supplementary Table 1**, patients in the derivation set were older and had lower neutrophil percentages and blood urea nitrogen levels; higher total protein, total bilirubin (TB), direct bilirubin, hemoglobin, aspartate aminotransferase, and alkaline phosphatase levels; lower rates of gastrointestinal hemorrhage; higher rates of bacterial infection; higher lengths of stay at initial admission; and higher CTPs and CLIF-C ADs ($p < 0.05$). The 30-, 60-, and 90-day risk of readmission were higher for the temporal validation set than for the derivation set ($p < 0.05$). The remaining clinical and laboratory parameters at initial admission as well as MELDs and MELD-Nas were similar between the derivation and temporal validation sets.

Development of a Nomogram

Supplementary Table 2 provides the results of the univariate analyses for all 36 factors considered as potential predictors in our scoring system. Fifteen candidate predictor variables with $p < 0.05$ were used as the input data in the LASSO regression. When the lambda value was collected as 1 standard error [log (λ1se) $= -2.10$], four variables were selected (see **Supplementary Figure 2** for details). Then, the four retained variables were used to construct the nomogram using multivariate Cox regression. As shown in **Table 2**, GI bleeding, serum sodium, total bilirubin (TB), and INR composed a panel of significant predictors of readmission in patients with AD cirrhosis.

A nomogram was constructed based on the four aforementioned independent prognosticators (**Figure 1**). The values of each risk factor were assigned a score on the point scale axis. By adding each single score and using that value in the total point scale axis, the total score could be easily calculated to assign the probability of readmission for individual patients at 30, 60, and 90 days. For the convenience of clinicians, we have provided the nomogram as a web-based calculator tool (https://cqykdx1111.shinyapps.io/dynnomapp/). Doctors can enter the indicators for each patient to automatically calculate the patient's probability of readmission within 90 days.

The model exhibited good discrimination ability. The C-index values of the nomogram for 30-, 60-, and 90-day readmission were 0.770 [95% confidence interval (CI): 0.728–0.812], 0.754 (95% CI: 0.716–0.791), and 0.731 (95% CI: 0.693–0.768), respectively. Furthermore, the predictive performance of nomogram was calculated by AUC of ROC curve. The AUC values were 0.775 (95% CI: 0.733–0.817), 0.753 (95% CI: 0.715–0.791), and 0.733 (95% CI: 0.695–0.770) for 30-, 60-, and 90-day, respectively (**Figure 2A**). Bootstraps with 200 resamples were performed to correct the predictive model. The adjusted C-index values of the nomogram for 30-, 60-, and 90-day readmission were 0.770, 0.750, and 0.732, respectively.

The calibration curves showed good agreement between nomogram predictions and observed probabilities for 30-, 60-, and 90-day outcomes in the derivation cohort (**Figure 3A**).

Temporal Validation of the Nomogram

The validation set was estimated using the established nomogram, and the C-index values obtained were 0.703 (95% CI 0.638–0.768), 0.694 (95% CI: 0.627–0.762), and 0.707 (95% CI: 0.636–0.777), respectively. The AUC values were 0.714 (95% CI: 0.649–0.778), 0.682 (95% CI: 0.614–0.751), and 0.712 (95% CI: 0.644–0.780), respectively (**Figure 2B**), supporting the suitability of the nomogram for estimating 30-, 60-, and 90-day readmission.

The calibration curves revealed good agreement between the nomogram predictions and observed probabilities for 30-, 60-, and 90-day outcomes in the temporal validation cohort (**Figure 3B**).

Comparison of Predictive Accuracy for Readmission Among Nomogram, MELDs, CLIF-C ADs, CTPs, and MELD-Nas

Predictive power for readmission was compared for the nomogram, MELDs, CLIF-C ADs, CTPs, and MELD-Nas based on C-indexes. DCA was performed to determine

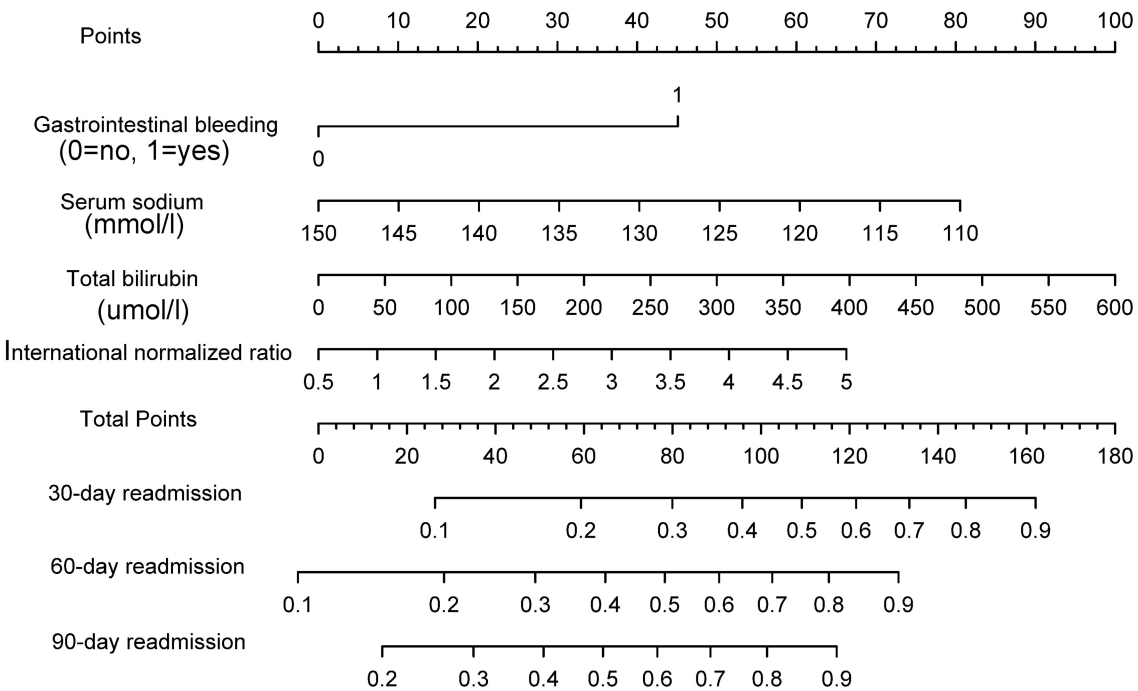


FIGURE 1 | The nomogram to predict the risk of readmission in patients with acute decompensated cirrhosis.

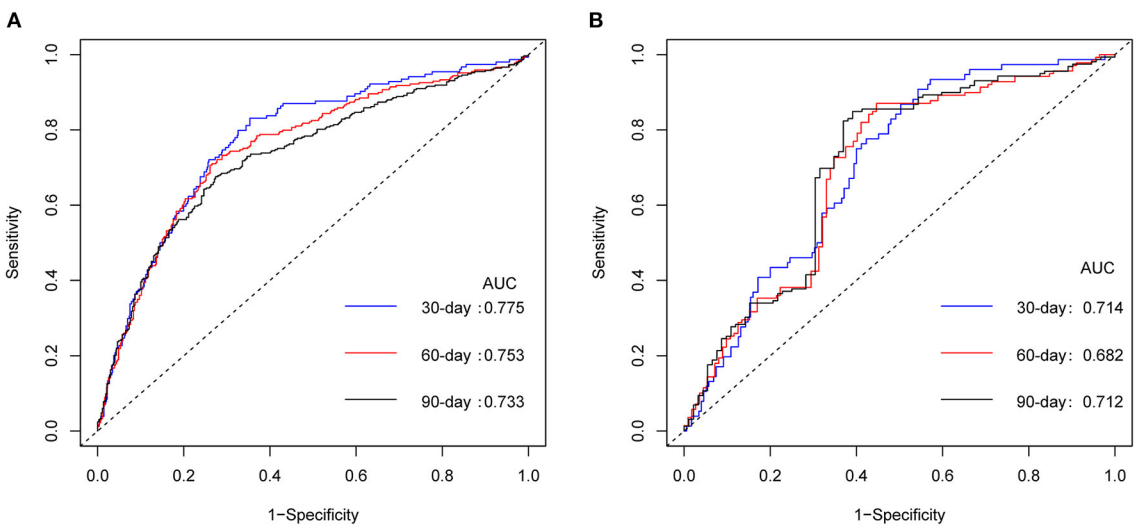


FIGURE 2 | ROC curves for nomogram. ROC curves of nomogram in derivation cohort **(A)** and in temporal validation cohort **(B)**. ROC, receiver operating characteristic; AUC, area under the curve.

the clinical utility of the nomogram by calculating the net benefits at different threshold probabilities. The C-indexes for 30-, 60-, and 90-day readmission predicted with MELDs, CLIF-C ADs, CTPs, and MELD-Nas were significantly lower than those with the nomogram, in both the derivation and temporal validation cohorts (**Table 3**).

Using DCA, our nomogram provided superior net benefit and displayed improved performance in prognostic evaluation over the 30-, 60-, and 90-day periods, both in the

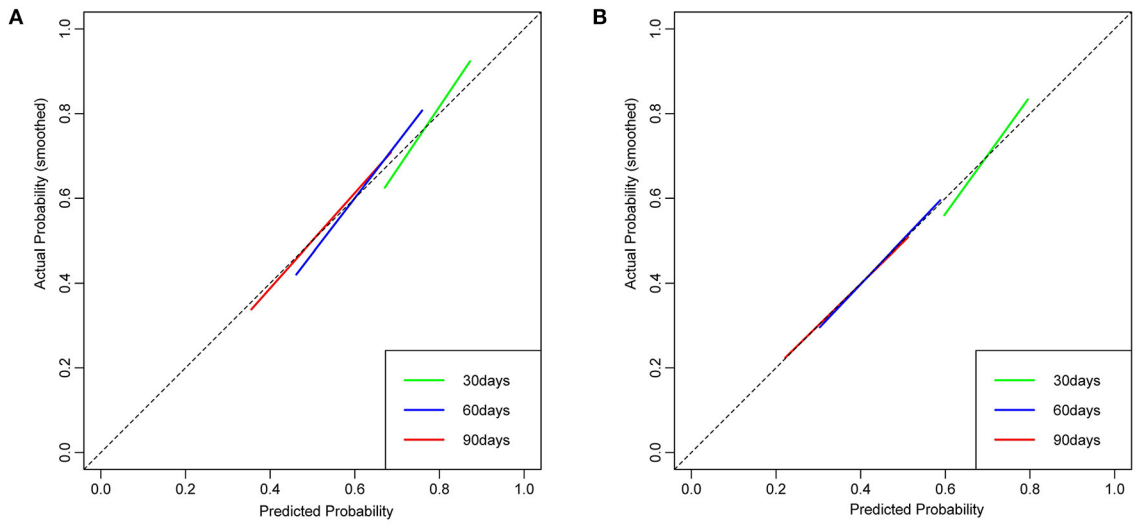


FIGURE 3 | The calibration curve of nomogram at 30, 60, and 90 days for the derivation cohort **(A)** and the temporal validation cohort **(B)**. Dashed lines along the 45-degree line through the point of origin represent the perfect calibration models in which the predicted probabilities are identical to the actual probabilities.

derivation (**Figures 4A–C**) and validation (**Figures 4D–F**) cohorts, compared to MELDs, CLIF-C ADs, CTPs, and MELD-Nas models.

Performance of the Nomogram in Stratifying Patient Risk

When patients were stratified according to the optimal cut-off value by the total nomogram points (high risk: >56.8 and low risk: ≤56.8), each group represented a distinct prognosis. The high-risk group was more likely to have readmission than the low-risk group, with statistical significance in both the derivation cohort and temporal validation cohort ($p < 0.0001$, **Figures 5A,B**).

DISCUSSION

In the present study, we generated an easy-to-perform nomogram consisting of clinical complications and laboratory indicators for the first time that could be effectively used to prognosticate the readmission probability of AD cirrhotic patients receiving drug-based therapy at different time points within a 90-day period. To avoid the effects of other coexisting medical conditions, we only included patients with cirrhosis-related complications as the reason for initial hospitalization or readmission. The nomogram model performed well, as determined from C-indexes, AUC values, and the calibration curves both in the derivation cohort and temporal validation cohort at 30, 60, and 90 days, respectively. According to DCA, our nomogram demonstrated better net benefit and improved performance in 30-, 60-, and 90-day prognostic evaluations in both the derivation and validation cohorts, compared with MELDs, CLIF-C ADs, CTPs, and MELD-Nas. Furthermore, the model was able to stratify patients into groups with high and low risk of readmission within 90 days. Finally, our nomogram is accessible to medical staff via a link to the algorithm to automatically calculate a patient's probability of readmission within 90 days (https://cqykdx1111.shinyapps.io/dynnomapp/). This scoring system can facilitate early identification of high-risk patients, thus allowing implementation of interventions during hospitalization to reduce readmission.

Readmission among patients with AD cirrhosis in the current study was common, with incidence rates of 24.6, 43.0, and 51.8% at 30, 60, and 90 days, respectively. This finding was similar to that obtained from earlier studies based in India and North America (10, 16, 18). The etiology of cirrhosis in India was mainly hepatitis B virus (50.4%), while in North America and Europe, the main causes of cirrhosis were alcohol consumption (29.4%) and HCV (39.3%) (10, 11). We identified bacterial infection as the main reason for index admission and readmission (54.3% at initial admission, 44.7% at 30 days, 48.2% at 60 days, and 48.9% at 90 days), distinct from findings from India, North America, and Europe, where hepatic encephalopathy and ascites were identified as the main contributory factors (10, 11, 16, 18, 36). This difference may be associated with the distinct inclusion/exclusion criteria and medical conditions in different regions. The admission of cirrhosis patients with bacteremia to the intensive care unit was associated with an increase in the severity of the disease and an increase in the need for extrahepatic organ support. Bacteremia was an independent predictor of mortality in patients with ACLF (37, 38). Recent novel perspectives in the management of decompensated cirrhosis suggest that the systemic inflammatory response is one of the upstream events underlying the development of complications of liver cirrhosis (39). A PREDICT study showed that the most severe course of acute decompensation occurs in patients with pre-ACLF who display rapid progression of systemic inflammation leading to the development of ACLF

TABLE 3 | Predictive discrimination ability of nomogram as compared to MELDs, CLIF-C ADs, CTPs, and MELD-Nas in the derivation and temporal validation cohort.

	Nomogram C-index 95%CI	MELDs C-index 95%CI	CLIF-C ADs C-index 95%CI	CTPs C-index 95%CI	MELD-Nas C-index 95%CI
Derivation cohort *N* = 705					
30-day	0.77 (0.728,0.812)	0.659 (0.610,0.709)	0.678 (0.630,0.727)	0.638 (0.587,0.690)	0.674 (0.624,0.723)
P-value vs. readmission		<0.001	<0.001	<0.001	<0.001
60-day	0.754 (0.716,0.791)	0.621 (0.578,0.663)	0.655 (0.614,0.697)	0.598 (0.555,0.641)	0.649 (0.608,0.691)
P-value vs. readmission		<0.001	<0.001	<0.001	<0.001
90-day	0.731 (0.693,0.768)	0.609 (0.567,0.650)	0.662 (0.622,0.702)	0.604 (0.563,0.645)	0.647 (0.606,0.687)
P-value vs. readmission		<0.001	<0.001	<0.001	<0.001
Validation cohort *N* = 251					
30-day	0.703 (0.638,0.768)	0.574 (0.495,0.652)	0.619 (0.543,0.696)	0.590 (0.514,0.666)	0.606 (0.528,0.683)
P-value vs. readmission		<0.001	<0.001	0.029	<0.001
60-day	0.694 (0.627,0.762)	0.569 (0.498,0.640)	0.558 (0.487,0.629)	0.627 (0.559,0.696)	0.520 (0.449,0.592)
P-value vs. readmission		<0.001	<0.001	0.029	<0.001
90-day	0.707 (0.636,0.777)	0.607 (0.537,0.677)	0.599 (0.528,0.669)	0.649 (0.579,0.718)	0.556 (0.484,0.628)
P-value vs. readmission		<0.001	<0.001	0.029	<0.001

95% CI, 95% confidence interval; MELDs, model for end-stage liver disease score; CLIF-C ADs, chronic liver failure-consortium acute decompensation scores; CTPs, child-turcotte-pugh score; MELD-Nas, MELD-Na score.

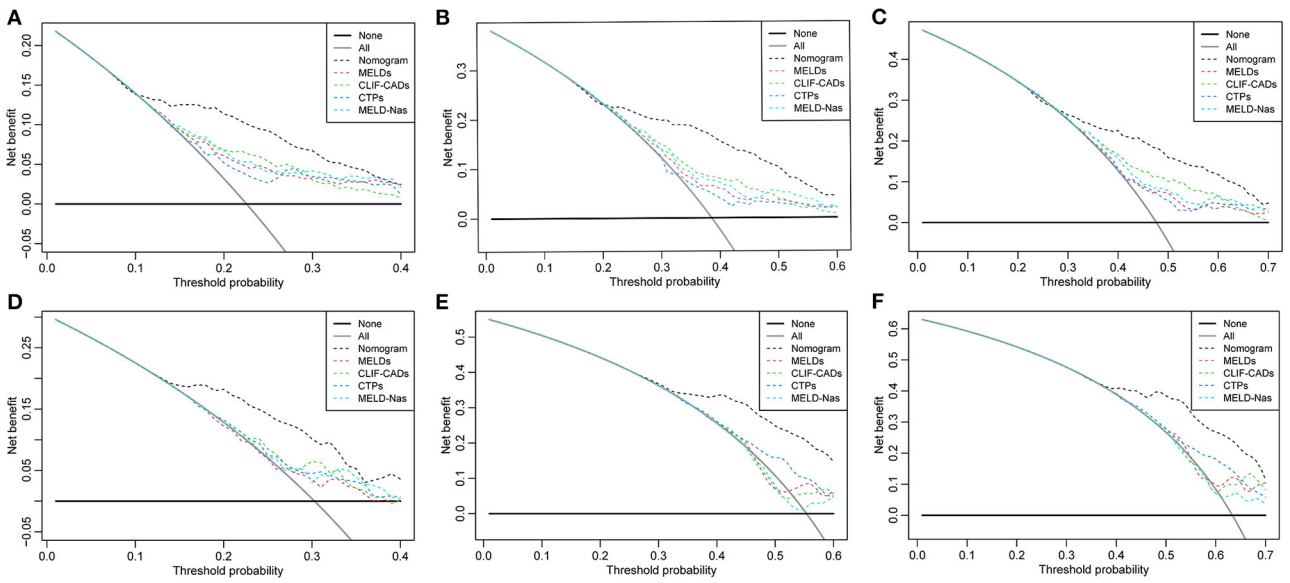


FIGURE 4 | Decision curve analysis at 30, 60, and 90 days for the derivation cohort **(A–C)** and the temporal validation cohort **(D–F)**. The horizontal solid black line represents the assumption that no patients will experience the event, and the solid gray line represents the assumption that all patients will relapse. On decision curve analysis, the readmission nomogram showed superior net benefit compared with other models across a range of threshold probabilities.

and death within 90 days (40). These findings clearly suggest that systemic inflammatory response is predictive of poor prognosis. Therefore, clinicians should pay significant attention to the prevention of infections, which could avoid downstream complications (further decompensation, repeat infections, ACLF, or death) of cirrhosis (41).

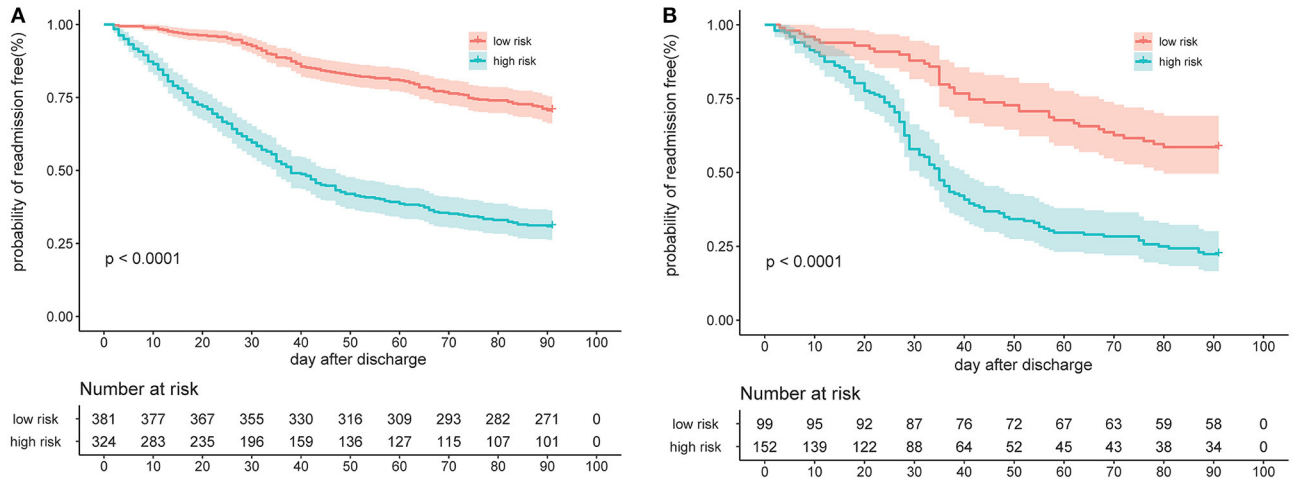


FIGURE 5 | Risk group stratification according to bisection of the nomogram predicted readmission in the derivation cohort **(A)** and the temporal validation cohort **(B)**.

Our nomogram includes laboratory and clinical indicators, which is better compared with other models, reflecting the severity of disease. Hyponatremia has been associated with hepatorenal syndrome and ascites and is an important predictor of readmission and mortality among patients with decompensated cirrhosis (30, 42). INR and TB are critical markers of liver protein synthesis function and the extent of hepatocellular necrosis (23, 43). These three indicators are all or partly involved in the construction of CLIF-C ADs, MELD-Nas, MELDs, and CTPs, reported to be significantly associated with the prognosis of AD cirrhosis or readmission (6, 11, 16, 18, 31, 36). However, owing to the involvement of the logarithm in calculations of MELDs, MELD-Nas, and CLIF-C ADs, clinicians have to use calculators, making it impractical in busy clinical practice. Although the calculation of CTPs is relatively simple, there are still some limitations, such as the narrow range of disease severity and subjective criteria, including hepatic encephalopathy and ascites (44). Notably, our model was more accurate than CLIF-C ADs, MELD-Nas, MELDs, and CTPs in predicting the risk of AD cirrhosis readmission. GI bleeding is a frequent and serious complication of cirrhosis. Mortality rates associated with acute esophageal variceal bleeding (EVB) are 12–20% and as high as 50% with EVB rebleeding (45). We found that although GI bleeding was a risk factor, the main cause for readmission in our study was bacterial infection. This finding underscores that patients with decompensated cirrhosis with gastrointestinal bleeding may be more likely to develop community-acquired infections after being discharged. Further studies are needed to confirm this possibility.

Our study showed that the readmission rate of AD cirrhosis patients in the low-risk group with total points $\leq$56.8 was significantly lower than that of high-risk patients with total points >56.8. In situations of limited medical resources and to improve cost effectiveness, low-risk patients could be considered for early discharge, whereas high-risk patients, especially those with GI bleeding, might need intensive management to prevent short-term readmission.

Our study has several strengths. First, we included only patients readmitted to the hospital for the first time after the initial discharge, thus avoiding the effects of multiple admissions. Second, our study was based on the Cox proportional hazard model, which predicts the probability of readmission at different time points within 90 days, rather than a logistic model that predicts readmission risk at a single time point, as used in most previous studies. In addition, we used temporal validation to demonstrate that our model can be generalized to further applications.

LIMITATIONS

This study has several limitations. First, selection bias may exist due to the retrospective nature of the investigation. However, we used a relatively large training cohort to construct the model, which was further subjected to temporal validation. Second, we excluded planned readmissions. Most of these patients underwent endoscopic variceal ligation during the initial hospitalization. Therefore, we did not evaluate the effects of this intervention on readmission. However, because the main reason for readmission is bacterial infection, the treatment target for patients with GI bleeding should not be limited to only preventing rebleeding. Third, data on social support, level of education, and socioeconomic status were not available. Further research is warranted to explore the impact of these important indicators.

CONCLUSIONS

The rates of short-term readmission related to cirrhosis were high in our patients. Bacterial infection was the main cause of

index admission and readmission. We developed and temporally validated a prognostic model that accurately predicts the incidence of cirrhosis-related readmissions in patients with AD cirrhosis receiving drug-based therapy. The readmission probability can be obtained with the nomogram scoring system, which is based on four independent variables for each patient (https://cqykdx1111.shinyapps.io/dynnomapp/). The present nomogram can assist in clinical decision-making, counseling for treatment, and, most importantly, risk stratification of patients to help differentiate patients who need intensive management to prevent short-term readmission from those who could discharge earlier.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethics Committee of the First Affiliated Hospital of Chongqing Medical University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

XX, JT, and BQ: concept and design. XX and JT: drafting of the manuscript. XX, JT, HW, and WZ: statistical analysis. HW, WZ, and BQ: administrative, technical, and material support. XX and BQ: had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis and supervision. All authors critical revision of the manuscript for important intellectual content, acquisition, analysis, or interpretation of data, and read and approved the final manuscript.

FUNDING

This research was funded by Chongqing Science and Health Joint Project (Grant no. 2020MSXM026) and Intelligence Medicine Project of Chongqing Medical University (Grant no. YJSZHYX202001).

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.com) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.646875/full#supplementary-material

REFERENCES

1. Arvaniti V, D'Amico G, Fede G, Manousou P, Tsochatzis E, Pleguezuelo M, et al. Infections in patients with cirrhosis increase mortality four-fold and should be used in determining prognosis. *Gastroenterology.* (2010) 139:1246–56, 56.e1–5. doi: 10.1053/j.gastro.2010.06.019
2. Arroyo V, Moreau R, Kamath PS, Jalan R, Ginès P, Nevens F, et al. Acute-on-chronic liver failure in cirrhosis. *Nat Rev Dis Primers.* (2016) 2:16041. doi: 10.1038/nrdp.2016.41
3. Fernández J, Acevedo J, Wiest R, Gustot T, Amoros A, Deulofeu C, et al. Bacterial and fungal infections in acute-on-chronic liver failure: prevalence, characteristics and impact on prognosis. *Gut.* (2018) 67:1870–80. doi: 10.1136/gutjnl-2017-314240
4. Chirapongsathorn S, Talwalkar JA, Kamath PS. Readmission in cirrhosis: a growing problem. *Curr Treat Options Gastroenterol.* (2016) 14:236–46. doi: 10.1007/s11938-016-0091-1
5. van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ.* (2011) 183: E391–402. doi: 10.1503/cmaj.101860
6. Berman K, Tandra S, Forssell K, Vuppalanchi R, Burton JR Jr, Nguyen J, et al. Incidence and predictors of 30-day readmission among patients hospitalized for advanced liver disease. *Clin Gastroenterol Hepatol.* (2011) 9:254–9. doi: 10.1016/j.cgh.2010.10.035
7. Volk ML, Tocco RS, Bazick J, Rakoski MO, Lok AS. Hospital readmissions among patients with decompensated cirrhosis. *Am J Gastroenterol.* (2012) 107:247–52. doi: 10.1038/ajg.2011.314
8. Singal AG, Rahimi RS, Clark C, Ma Y, Cuthbert JA, Rockey DC, et al. An automated model using electronic medical record data identifies patients with cirrhosis at high risk for readmission. *Clin Gastroenterol Hepatol.* (2013) 11:1335–41.e1. doi: 10.1016/j.cgh.2013.03.022
9. Tapper EB, Finkelstein D, Mittleman MA, Piatkowski G, Lai M. Standard assessments of frailty are validated predictors of mortality in hospitalized patients with cirrhosis. *Hepatology.* (2015) 62:584–90. doi: 10.1002/hep.27830
10. Koola JD, Ho SB, Cao A, Chen G, Perkins AM, Davis SE, et al. Predicting 30-day hospital readmission risk in a national cohort of patients with cirrhosis. *Dig Dis Sci.* (2020) 65:1003–31. doi: 10.1007/s10620-019-05826-w
11. Morales BP, Planas R, Bartoli R, Morillas RM, Sala M, Cabré E, et al. Early hospital readmission in decompensated cirrhosis: incidence, impact on mortality, and predictive factors. *Dig Liver Dis.* (2017) 49:903–9. doi: 10.1016/j.dld.2017.03.005
12. Chirapongsathorn S, Krittanawong C, Enders FT, Pendegraft R, Mara KC, Borah BJ, et al. Incidence and cost analysis of hospital admission and 30-day readmission among patients with cirrhosis. *Hepatol Commun.* (2018) 2:188–98. doi: 10.1002/hep4.1137
13. Gaspar R, Rodrigues S, Silva M, Costa-Moreira P, Morais R, Andrade P, et al. Predictive models of mortality and hospital readmission of patients with decompensated liver cirrhosis. *Dig Liver Dis.* (2019) 51:1423–9. doi: 10.1016/j.dld.2019.03.016
14. Chirapongsathorn S, Poovorawan K, Soonthornworasiri N, Pan-Ngum W, Phaosawasdi K, Treeprasertsuk S. Thirty-day readmission and cost analysis in patients with cirrhosis: a nationwide population-based data. *Hepatol Commun.* (2020) 4:453–60. doi: 10.1002/hep4.1472
15. Graupera I, Solà E, Fabrellas N, Moreira R, Solé C, Huelin P, et al. Urine monocyte chemoattractant protein-1 is an independent predictive factor of hospital readmission and survival in cirrhosis. *PLoS ONE.* (2016) 11:e0157371. doi: 10.1371/journal.pone.0157371
16. Bajaj JS, Reddy KR, Tandon P, Wong F, Kamath PS, Garcia-Tsao G, et al. The 3-month readmission rate remains unacceptably high in a large North American cohort of patients with cirrhosis. *Hepatology.* (2016) 64:200–8. doi: 10.1002/hep.28414

17. Shaheen AA, Nguyen HH, Congly SE, Kaplan GG, Swain MG. Nationwide estimates and risk factors of hospital readmission in patients with cirrhosis in the United States. *Liver Int.* (2019) 39:878–84. doi: 10.1111/liv.14054
18. Patel R, Poddar P, Choksi D, Pandey V, Ingle M, Khairnar H, et al. Predictors of 1-month and 3-months hospital readmissions in decompensated cirrhosis: a prospective study in a large Asian cohort. *Ann Hepatol.* (2019) 18:30–9. doi: 10.5604/01.3001.0012.7859
19. Orman ES, Ghabril M, Emmett TW, Chalasani N. Hospital readmissions in patients with Cirrhosis: a systematic review. *J Hosp Med.* (2018) 13:490–5. doi: 10.12788/jhm.2967
20. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ.* (2009) 338:b606. doi: 10.1136/bmj.b606
21. Amarasingham R, Patzer RE, Huesch M, Nguyen NQ, Xie B. Implementing electronic health care predictive analytics: considerations and challenges. *Health Aff.* (2014) 33:1148–54. doi: 10.1377/hlthaff.2014.0352
22. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* (2015) 350:g7594. doi: 10.1136/bmj.g7594
23. Sarin SK, Choudhury A, Sharma MK, Maiwall R, Al Mahtab M, Rahman S, et al. Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific association for the study of the liver (APASL): an update. *Hepatol Int.* (2019) 13:353–90. doi: 10.1007/s12072-019-09946-3
24. Blei AT, Cordoba J. Hepatic encephalopathy. *Am J Gastroenterol.* (2001) 96:1968–76. doi: 10.1111/j.1572-0241.2001.03964.x
25. Moore KP, Wong F, Gines P, Bernardi M, Ochs A, Salerno F, et al. The management of ascites in cirrhosis: report on the consensus conference of the International Ascites Club. *Hepatology.* (2003) 38:258–66. doi: 10.1053/jhep.2003.50315
26. Gustot T, Durand F, Lebrec D, Vincent JL, Moreau R. Severe sepsis in cirrhosis. *Hepatology.* (2009) 50:2022–33. doi: 10.1002/hep.23264
27. Garcia-Tsao G, Bosch J. Management of varices and variceal hemorrhage in cirrhosis. *N Engl J Med.* (2010) 362:823–32. doi: 10.1056/NEJMra0901512
28. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg.* (1973) 60:646–9. doi: 10.1002/bjs.1800600817
29. Kamath PS, Kim WR. The model for end-stage liver disease (MELD). *Hepatology.* (2007) 45:797–805. doi: 10.1002/hep.21563
30. Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med.* (2008) 359:1018–26. doi: 10.1056/NEJMoa0801209
31. Jalan R, Pavesi M, Saliba F, Amoros A, Fernandez J, Holland-Fischer P, et al. The CLIF Consortium Acute Decompensation score (CLIF-C ADs) for prognosis of hospitalised cirrhotic patients without acute-on-chronic liver failure. *J Hepatol.* (2015) 62:831–40. doi: 10.1016/j.jhep.2014.11.012
32. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* (2010) 33:1–22. doi: 10.18637/jss.v033.i01
33. Vasquez MM, Hu C, Roe DJ, Chen Z, Halonen M, Guerra S. Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. *BMC Med Res Methodol.* (2016) 16:154. doi: 10.1186/s12874-016-0254-8
34. Frank E, Harrell J. Rms: Regression Modeling Strategies. R Package version 3.4–0. Available online at: https://CRAN.R-project.org/package=Hmisc (accessed April 21, 2017).
35. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* (2006) 26:565–74. doi: 10.1177/0272989X06295361
36. Louissaint J, Foster C, Harding-Theobald E, Lok AS, Tapper EB. Social support does not modify the risk of readmission for patients with decompensated cirrhosis. *Dig Dis Sci.* (2020). doi: 10.1007/s10620-020-06421-0. [Epub ahead of print].
37. Feltracco P, Brezzi M, Barbieri S, Milevoj M, Galligioni H, Cillo U, et al. Intensive care unit admission of decompensated cirrhotic patients: prognostic scoring systems. *Transplant Proc.* (2011) 43:1079–84. doi: 10.1016/j.transproceed.2011.01.153
38. Karvellas CJ, Pink F, McPhail M, Austin M, Auzinger G, Bernal W, et al. Bacteremia, acute physiology and chronic health evaluation II and modified end stage liver disease are independent predictors of mortality in critically ill nontransplanted patients with acute on chronic liver failure. *Crit Care Med.* (2010) 38:121–6. doi: 10.1097/CCM.0b013e3181b42a1c
39. Bernardi M, Caraceni P. Novel perspectives in the management of decompensated cirrhosis. *Nat Rev Gastroenterol Hepatol.* (2018) 15:753–64. doi: 10.1038/s41575-018-0045-2
40. Trebicka J, Fernandez J, Papp M, Caraceni P, Laleman W, Gambino C, et al. The PREDICT study uncovers three clinical courses of acutely decompensated cirrhosis that have distinct pathophysiology. *J Hepatol.* (2020) 73:842–54. doi: 10.1016/j.jhep.2020.06.013
41. Garcia-Tsao G. Prophylactic antibiotics in cirrhosis: are they promoting or preventing infections? *Clin Liver Dis.* (2019) 14:98–102. doi: 10.1002/cld.819
42. Deitelzweig S, Amin A, Christian R, Friend K, Lin J, Lowe TJ. Hyponatremia-associated healthcare burden among US patients hospitalized for cirrhosis. *Adv Ther.* (2013) 30:71–80. doi: 10.1007/s12325-012-0073-1
43. Moreau R, Jalan R, Gines P, Pavesi M, Angeli P, Cordoba J, et al. Acute-on-chronic liver failure is a distinct syndrome that develops in patients with acute decompensation of cirrhosis. *Gastroenterology.* (2013) 144:1426–37, 37.e1–9. doi: 10.1053/j.gastro.2013.02.042
44. Cai YJ, Dong JJ, Dong JZ, Chen Y, Lin Z, Song M, et al. A nomogram for predicting prognostic value of inflammatory response biomarkers in decompensated cirrhotic patients without acute-on-chronic liver failure. *Aliment Pharmacol Ther.* (2017) 45:1413–26. doi: 10.1111/apt.14046
45. Rush B, Deol N, Teriyaki A, Sey M, Jairath V, Walley KR, et al. Lower 90-day hospital readmission rates for esophageal variceal bleeding after TIPS: a nationwide linked analysis. *J Clin Gastroenterol.* (2020) 54:90–5. doi: 10.1097/MCG.0000000000001199

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

PERSPECTIVE
published: 16 June 2021
doi: 10.3389/fmed.2021.510421

Predictability in Contemporary Medicine

Michele M. Ciulla [1,2,3]*

[1] Laboratory of Clinical Informatics and Cardiovascular Imaging, University of Milan, Milan, Italy, [2] Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy, [3] Cardiovascular Diseases Unit, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy

Medical practice is increasingly coming under the guidance of statistical-mathematical models that are, undoubtedly, valuable tools but are also only a partial representation of reality. Indeed, given that statistics may be more or less adequate, a model is still a subjective interpretation of the researcher and is also influenced by the historical context in which it operates. From this opinion, I will provide a short historical excursus that retraces the advent of probabilistic medicine as a long process that has a beginning that should be sought in the discovery of the complexity of disease. By supporting the belonging of this evolution to the scientific domain it is also acknowledged that the underlying model can be imperfect or fallible and, therefore, confutable as any product of science. Indeed, it seems non-trivial here to recover these concepts, especially today where clinical decisions are entrusted to practical guidelines, which are a hybrid product resulting from the aggregation of multiple perspectives, including the probabilistic approach, to disease. Finally, before the advent of precision medicine, by limiting the use of guidelines to the original consultative context, an *aged* approach is supported, namely, a relationship with the individual patient.

Keywords: diseases, Koch's postulates, multifactorial, risk factors, COVID-19, susceptibility, mathematical-statistical models, guidelines

OPEN ACCESS

Edited by:
Tao-Hsin Tung,
Cheng Hsin General Hospital, Taiwan

Reviewed by:
Harm Peters,
Charité – Universitätsmedizin Berlin, Germany
Maurice Mars,
University of KwaZulu-Natal, South Africa

***Correspondence:**
Michele M. Ciulla
michele.ciulla@unimi.it

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *03 December 2019*
Accepted: *26 April 2021*
Published: *16 June 2021*

Citation:
Ciulla MM (2021) Predictability in Contemporary Medicine. Front. Med. 8:510421. doi: 10.3389/fmed.2021.510421

"Medicine is the most humane of all sciences, it is practiced by humans for human health".

INTRODUCTION

The aim of this paper is to present in a historical perspective the process that led to the affirmation of the *mathematical-statistical models* in medicine as a surrogate system for making clinical decisions in order to bring it back to its original theoretical and rebuttable domain. The intent is not to give a sterile criticism but to support the proper use of any statistical-mathematical interpretative model, that is, subjective, fallible, and, therefore, confutable as any product of science. Since any constructive criticism must be proactive, in this article I will therefore support the recovery of the relationship with the individual patient that cannot be reduced to a mathematical average. The discussion will be articulated in a few points: the first will outline the process that led to the discovery of the complexity of disease and the advent of probabilistic medicine; the second, starting from the definition of illness as an unfavorable interaction between genes and environment, will support the general unpredictability of the disease as the result of an excess of variables whose control is very unlikely; the third point will try to demystify statistical-mathematical models by bringing them back to the main watercourse of science that should be understood as

TABLE 1 | Prevalence, causation, and complexity of diseases.

A. Factors affecting the prevalence and severity of infectious diseases.

- Climatic factors
- Social conditions
- Molecular variations in pathogens
- Dynamic of the vector agent
- Genetics of the host population
- Interactions with concurrent diseases

B. Cause vs. causal association.

→ Cause

- any factor that produces an effect
- in medicine, also reported as etiology, pathogenesis, mechanism.

→ Causal association

- any factor that reveals an increased frequency of association in exposed vs. non-exposed
- in medicine, a factor associated is a risk factor.

C. Complexity of individual adaptation.

There are many degrees of freedom:

- the initial condition is uncertain
- the sensitivity/susceptibility is individual
- the future exposure to the environment depends on space-time
- different systems are dynamically involved
- the results may be of opposite sign, unfavorable, neutral, or even favorable.

a continuous process of knowledge made of successes and failures; in this regard, in the fourth point, mathematics at the bedside will be supported as a tool that, till now, is the most profitable one in developing interpretative models in medicine; and the fifth point will criticize the formal application of guidelines as a hybrid product resulting from the aggregation of multiple instances, often incompatible, by supporting their original consultative role.

DISCUSSION

The Question of Identifying the Causes of Diseases

In clinical medicine, to identify the causes of injury and/or illness is essential for the progress of knowledge and to prevent the progress of disease and/or to develop appropriate care; unfortunately, the simplified causative approach, well-summarized in *Koch's Postulates*, is no longer appropriate (1). This is because the outcome of illness, either healing or death, is never predictable with certainty but is, still, a causal inference based on the *study of probability*. Formulated to identify the *pathogen* responsible for a specific disease, the postulates contributed to the spread of an etiological approach based on a mono-factorial cause-effect relationship that has been replaced, since the Framingham Heart Study (2), in favor of a *multifactorial* one based on the concept of *risk factors*. This model has also expanded in the area of infectious diseases as even the Black Death of the 14th century, caused by Yersinia pestis, has been claimed to be a multifactorial pandemic (3) by postulating a series of causative factors that are reported in **Table 1A**.

The very recent emergence of SARS-CoV-2 and the COVID-19 pandemic, while pointing out the dynamic of spreading-over-species of infectious diseases (4), has provided an unexpected opportunity to link the puzzle of variable clinical manifestations and outcomes with host genomic factors (5).

Thus, all diseases, including the infectious ones are, possibly, multifactorial with an inherent increase in the complexity of the model since a cluster of interacting causative factors are, usually, associated and none of them are sufficient (6). From such a perspective, the association between events affecting biological life is not simply deterministic but more elusive as it follows a *non-linear dynamic* (7) making their effects, or outcomes, almost unpredictable in the individual subject. Thus, leaving behind the mono-factorial approach and focusing on the multi-factorial one, the main problem is to determine the probabilities of an event, such as illness or death, or the success of an intervention/treatment, whose degrees of freedom are innumerable. In **Table 1B** the main differences between cause and causal association are listed. Finally, in a clinical setting, when observing two events-diseases apparently connected with each other, with a possible causal relationship, it must be considered that there are, also, *recurring events*, sometimes cyclical, whose consequentiality may be casual and/or influenced by the direction, forward or backward, of the observation. Thus it may be difficult to reconstruct the time-line (8) like in the *"which came first: the chicken or the egg?"* dilemma.

Why It Is Rather Impossible to Predict the Future for the Individual Subject

At present, a disease is the result of an *unfavorable* interaction between genes and environment, thus we must shift our focus to the *pattern of interaction* that has to do more with *individual adaptation*. Indeed, changes in the environment are handled by the same strategy that drives development and evolution by using biological resources involved in tissue maintenance and repair of damage (9). This kind of *susceptibility* is linked to a genetic risk as the disease manifests itself in a certain environmental context making the interaction unfavorable, and this is only known afterwards. Thus, we are facing *uncertainty* in the initial condition, with *many degrees of freedom* affecting susceptibility and future exposures depending on space-time and whether the environmental context is neutral or, sometimes, even favorable (10). This also explains the issue of s*elective advantages,* in terms of probability of survival, with the emergence of different *phenotypes* from a single *genotype* that are the result of the *epigenetic* machinery as proposed by Waddington in 1957. Having a unique genotype, more or less fixed, and an epigenome to provide dynamic and flexible responses to environmental changes does not simplify the matter, but rather makes it more complex as the individual process of adaptation is tricky (**Table 1C**). We can, therefore, say that no one is healthy and everyone is sick since the *"boundary between health and disease is, at least, fuzzy as it moves according to the reciprocal interaction between phenotype and environment and each individual is a different phenotype"* (11).

Science and the Ability to Prevent and Treat Diseases Through the Forecasts

Science can be viewed as a process, carried out by scientists, aimed at increasing the knowledge of the internal/external human environment, through the formulation of relevant questions and using appropriate methodology for getting answers. The scientific method is based on the claim and refutation of *evidence* that is collected to support the answers, thus, having obtained an answer, the desirable next goal is to use this experience, which is now part of the knowledge. In the clinical setting, as we have already said, we can use knowledge to prevent evolution toward disease, better identify the causal process of injury and/or illness, and/or to develop/use appropriate care. As we are in an *open thermodynamic system* where life, by using the *individual adaptation pattern*, tries, temporarily, to curb the drift toward chaos (9), any clinical decision, taken in the patient's interest, follows this general *attempt* to gain persistence and maintain order by capitalizing on the previous experience. This issue is a *generalization* of results obtained from population averages to individuals based on the similarity of the clinical profile. Indeed, in clinical epidemiology, it is generally assumed, for example, that *"the risk of a disease at equal levels of known risk factors is similar in any individual belonging to the studied population"* (12), but this is not the truth since individual susceptibility is different from the collective one. Furthermore, the *predictability* of (any) previous clinical experiences is a more complex question since it is based on less/more evidence and weak/strong algorithms applied to set-up an *interpretative model*. It must be said that a model is, always, a subjective interpretation of the researcher, thus a certain bias must be taken into account. Indeed, modeling is subjected to the influence of the *dominant scientific culture* which, inevitably, creates *consensus* on, more or less, conventional interpretative models that are, often, too dogmatic. This is a general question affecting science as a *collective process* that makes it difficult to *falsify the evidence* since consensus is maintained by sociological processes which are explained with organizational and institutional influences (13, 14) (arguments that go beyond this discussion).

In summary, making *forecasts* about the future state of a *thermodynamic system* is at the basis of scientific knowledge, with theoretical and practical implications of the utmost importance that are listed in **Table 2A**.

TABLE 2 | Predicting health and disease.

A. Why it is difficult to predict the (clinical) future for the individual: the boundary between health and disease is fuzzy.

→ Predictability relies on:
- formulation of relevant questions
- using appropriate methodology
- having more evidence
- strong algorithms
- interpretative model
- generalizability of population averages to individuals is problematic:

→ Interpretative models are:
- subjective interpretation of the researcher
- influenced by the dominant scientific culture

B. Limitations in the use of statistical models in medicine.
- A model is a subjective interpretation of the researcher.
- A model lacks a complete -structural-systemic- understanding.
- The predictability of a model relies on algorithms.
- In statistics, any correlation found does not imply causation.
- Reproducibility is a standard, mainly, for science.
- Future exposures may not be predictable precisely.
- The generalization of results is impossible.
- The verification of the prevention effectiveness is a complex issue.

C. Critical approach to the guidelines system.

Multiple instances, difficult to meet at the same time, *built a (better) system of (public) health care*
- rationalizing the medical intervention
- reduce costs
- ensure legal protection of physicians
- preserve professional autonomy
- fall within the public/private funding of research
- ensure an appropriate statistical-mathematical standard
- check the prevention effectiveness

D. Different approaches to the patient if viewed as an individual or as an average.

Individual patient	Average patient
Requires an empathic relationship	Relationship should be limited
Benefit from the consultation of the guidelines	Use of guidelines mandatory
Takes more time	Takes less time
Decision making assumes a high level of responsibility	Responsibilities are shared

Mathematics at the Bedside and Complexity of Biology

While approaching the complexity of biological life and death, we have no profitable approach except the scientific one that has been applied to the study of biological networks and their relationships. Among the different approaches, *mathematics* is undoubtedly the most profitable in developing *interpretative models* (12) since it has an axiomatic structure, uses logic, and has a method, scientific accuracy, and flexibility.

Indeed, when clinicians use *statistics* to test/develop a model by applying mathematical formulas, they often use a conventional logic that does not fit the *complexity of biology* that goes beyond, but instead follows non-deterministic rules and is characterized by a *non-linear* and, often, *chaotic* dynamics. But this is not a problem for mathematics that is, indeed, flexible; in this regard, one of the applications in biology of a non-linear mathematical model based on attractors of chaos - an attractor is a geometric place to which a dynamic system evolves after a long enough time - is the study of heart rate (15), a biological phenomenon known for its high variability (16). Nonetheless, assuming that the chaotic dynamics are appropriate to describe some human phenomena such as diseases, we could only predict

events knowing perfectly the *initial conditions* of the system, and this is not always the case. The lack of knowledge of the initial conditions, obviously, does not rule out that they exist, but we must consider that even at the extreme opposite of *determinism* we can find *non-determinism* or truly *random patterns* that physicists, with the scattering of protons, have suggested to describe the intimate behavior of physical matter (17). Again, is this model suitable for human diseases? The question presents different facets: one is to consider events related to adaptive mechanisms such as *pseudo-random* or phenomena waiting for a new algorithm to formulate a suitable predictive model (11). As humans have a mind oriented to capture causal links and to reject the randomness of natural events, this kind of confidence in the advancement of knowledge allows, temporarily, to avoid the hardness of having *chaos* in the *hospital*. The issue of having a model, however, does not seem to solve the problem as it is, still, a subjective interpretation of the events made by the researcher; thus, *random* or *pseudo-random* are just models. It may be necessary to rethink the science of certainty and that of uncertainty (18) bearing in mind that mathematics is just a tool, it uses logic and creates knowledge by following hypotheses that, primarily, are inspired by *intuition*, a peculiar type of ability that does not use inference or reason (19). Finally, it seems non-trivial to question even the mechanistic approach since it is limited because of the lack of a complete understanding that can result, only, from a broader vision that, till now, seems to have been associated with philosophy. In **Table 2B**, the main limitations in the use of statistical models in medicine are listed.

Evidence-Based Guidelines System: From Sources of Bias to Inherent Limitations When Approaching the Individual Patient

The consequences of statistical models at the bedside are *evidence-based guidelines* that are, ideally, useful tools produced to summarize probabilistic data and provide practical guidance. The cultural background which has led to the widespread diffusion of evidence-based guidelines was the setting of the etiological model in favor of a multifactorial one and the progressive adoption of mathematical-statistical models to estimate the risk of an adverse event and to assess possible intervention strategies in order to prevent it (20). This system represents the summation of multiple instances: on the one hand the idea of rationalizing the medical intervention, based on available scientific evidence, in order to contain the costs and to build a better system of health care in the public domain and, on the other hand, to provide a legal protection that allows physicians to preserve a wide professional autonomy. To combine multiple instances is never an easy task, but there are, as we have seen, several drawbacks inherent to the statistical model used in published research since its predictability is, still, limited and, in any case, a research finding provides, only, a partial representation of reality obtained from a *finite* number of subjects. Furthermore, it should also be remembered that in the clinical setting *verification* of the prevention/intervention effectiveness (11) is a complex issue that requires a systematic assessment of its impact on health outcomes with post-study probability testing (21). Another source of bias depends on the fact that scientific research is not a free domain but is subject to public or private funding, according to a complex and questionable interference pattern. Even guidelines, which generally summarize probabilistic data from research studies, are *possibly* biased since they are drafted, mostly, by those who declare a *conflict of interest*; in this regard, several papers are available discussing the issue of financial conflict of interest.

Historically, the need to declare a conflict of interest began in 2003 when the *pharmaceutical industry* established guidelines on *Good Publication Practice* to make the publication of *industry-sponsored trials* more transparent (22). This followed the well-known law suit against *Pfizer* that produced "fraudulent scientific evidence" by "suppressing unfavorable study results to promote off-label uses of gabapentin" (23). This position was followed by the International Committee of Medical Journal Editors requiring, starting from 2004, a registration in a *public trial's registry* as a "condition of consideration for publication" of clinical trials (24). At that time, a review published by *Jama* showed a significant association between industry sponsorship and pro-industry conclusions (25), but the question remains on whether *clinical studies/trials*, even when summarized into practical guidelines, are still now influenced by financial relationship that, "disclosed or undisclosed, relevant or not relevant" have been demonstrated to impact on "whether studies report findings favorable to industry sponsors," which is supported by a recent survey dated 2019 (26). This potential bias is more evident if we look at the most worldwide prescribed therapies; an 18 year retrospective study supported a significant association between industry-funded *randomized controlled trials* and statistically significant outcomes for antidepressants (27).

Thus, it seems evident that declaring a conflict of interest does not guarantee a lack of bias. Furthermore, a very recent study supported that, years after this obligation, a low percentage of primary studies, such as *randomized controlled trials*, include a declaration of conflict of interest (28). If we look at *research validity,* it can vary considerably; in a meta-analysis of survey data published by PLoS in 2009, "misconducting research" seems to be a fairly common practice that has been *self-admitted,* regardless of the reasons why, by up to 34% of authors explicitly asked about "questionable research practices," including having *fabricated/falsified* research data or *altered/modified* results to improve the outcome, and 29% of the cases of misconduct known by respondents were never discovered (29). A recent update by the same author of the previous survey supported that *self-admission* rates for both fabrication/falsification, including *plagiarism*, seems to have declined over the years, but *non-self-admission* rates have not changed (30).

In summary, *clinical studies/trials* may not only be *methodologically* incorrect, underpowered, and even misinterpreted (31) but, at the same time, biased by a disclosed or undisclosed conflict of interest and this is, also, for their summarization in *evidence-based guidelines.*

If we assume that available *clinical guidelines* are reliable, as we ideally do expect (32), their widespread diffusion support obvious potential benefits to rationalize medical intervention. Nonetheless, as we stated before, using *heterogeneous*

information, made up of more-or-less *valid evidence*, needs a *systematic strategy* for the evaluation/verification of the prevention/intervention effectiveness and the "epistemological responsibility of doctors" when drafting/using guidelines (33).

Finally, a simple question that is scarcely taken into consideration, is that what appears beneficial for our patients *as a group* may not always be suitable for the *individual* and, certainly, not for patients with *comorbidities* (34) and the routine use of guidelines is possibly conflicting with the emerging concept of *personalized medicine* and the model of *shared decision-making* (35).

That is to say, even when using the best on *average* up-to-date evidence, because of the *heterogeneous* response to *any* intervention, we are not able to predict how this strategy may work in a specific patient, even when subgroup analysis is available; thus, it is possible that physicians, informed by last trials' results, but with their direct *clinical experience*, do better than others at prescribing the same evidence-based best option to everyone, failing to profile patients who may not benefit (32). A critical approach to the guidelines system is reported in **Table 2C**.

CONCLUSIONS AS A STARTING POINT

The complexity of life and death is demonstrated by their unpredictability, thus making predictions on the *clinical future* of an individual seem an arduous task. The individual adaptation profile to environmental changes seems to have part of the answer, but predictability uses statistical modeling that, to increase the rate of probability, needs more subjects, thus moving far away from the individual subject in favor of the collective. This means that the individual subject/patient is, still, missing from research papers and guidelines, appearing only in case reports. This does not mean that forecasts in medicine are banned but, simply, that they have to be brought back to the main watercourse of science that should be understood as a continuous process and, even if mathematics at the bedside could be extremely profitable in developing interpretative models, these models are still subjective, fallible, and, therefore, confutable. Before the advent of *precision medicine,* doctors pursued a relationship with the individual patient, who cannot be reduced to a mathematical average, and to recollect Osler's thoughts when he wrote in his most famous essay, Aequanimitas, delivered to new doctors in 1889 at Pennsylvania School of Medicine: "The practice of medicine is an art, based on science." That is to say, medicine is not an art like painting but, neither is it a science like physics; it needs humanity, empathy, respect, communication, and *fact checking ability* when using evidence- based medicine algorithms for the individual patient to plan a strategy and reach the best outcome.

In **Table 2D** the different approaches to the patient if viewed as an individual or as an average is summarized.

LIMITATION OF THE STUDY

Perspectives are intrinsically limited but, sometimes, useful.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. All datasets generated for this study are included in the article/ supplementary materials.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

1. Fredericks DN, Relman DA. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin Microbiol Rev.* (1996) 9:18–33. doi: 10.1128/CMR.9.1.18-33.1996
2. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J III. Factors of risk in the development of coronary heart disease-six year follow-up experience. The Framingham Study. *Ann Intern Med.* (1961) 55:33–50. doi: 10.7326/0003-4819-55-1-33
3. Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, et al. A draft genome of Yersinia pestis from victims of the black death. *Nature.* (2011) 478:506–10. doi: 10.1038/nature 10549
4. Ciulla MM. The puzzle of sharing bio-molecular targets between coronaviruses and mediators of the cardiovascular system in humans: looking for plausible hypotheses. *Med Hypotheses.* (2020) 141:109793. doi: 10.1016/j.mehy.2020.109793
5. Murray MF, Kenny EE, Ritchie MD, Rader DJ, Bale AE, Giovanni MA, et al. COVID-19 outcomes and the human genome. *Genet Med.* (2020) 22:1175–7. doi: 10.1038/s41436-020-0832-3
6. Rothman KJ. Synergy and antagonism in cause-effect relationships. *Am J Epidemiol.* (1974) 99:385–8. doi: 10.1093/oxfordjournals.aje.a121626
7. Higgins JP. Nonlinear systems in medicine. *Yale J Biol Med.* (2002) 75:247–60.
8. Ciulla MM. Atrial fibrillation and gastroesophageal reflux disease: what about vice-versa? The arrow of time in medicine and biology. Letter to the Editor. *Medicine.* (2015). Available online at: http://journals.lww.com/md-journal/Blog/MedicineCorrespondenceBlog/pages/post.aspx?PostID=11
9. Ciulla MM. Epistemology of natural strategies for cardiac tissue repair. *Front Cardiovasc Med.* (2017) 4:61. doi: 10.3389/fcvm.2017.00061
10. Ciulla MM. Darwin and the patent foramen ovale phenotype. *Int J Cardiol.* (2014) 172:e124. doi: 10.1016/j.ijcard.2013.12.116
11. Ciulla MM. The biological side of randomness. A starting point to rethink causation of diseases and prevention as a strategy. *PeerJ PrePrints.* (2015) 3:e1440. doi: 10.7287/peerj.preprints.1147
12. Aalen OO, Valberg M, Grotmol T, Tretli S. Understanding variation in disease risk: the elusive concept of frailty. *Int J Epidemiol.* (2015) 44:1408–21. doi: 10.1093/ije/dyu192
13. Heinze T, Shapira P, c. Rogers JD, Senker JM. Organizational and institutional influences on creativity in scientific research. *Res Policy.* (2009) 38:610–23. doi: 10.1016/j.respol.2009.01.014
14. Zola IK. Medicine as an institution of social control. *Sociol Rev.* (1972) 20:487–504. doi: 10.1111/j.1467-954X.1972.tb00220.x
15. Goldberger AL. Cardiac chaos. *Science.* (1989) 243:1419. doi: 10.1126/science.2928773
16. Nicolini P, Ciulla MM, De Asmundis C, Magrini F, Brugada P. The prognostic value of heart rate variability in the elderly, changing the perspective: from

sympathovagal balance to chaos theory. *Pacing Clin Electrophysiol.* (2012) 35:622–38. doi: 10.1111/j.1540-8159.2012.03335.x
17. James F. Chaos and randomness. *Chaos Solit Fractals.* (1995) 6:221–6. doi: 10.1016/0960-0779(95)80028-F
18. Colani EV. Defining and modeling uncertainty. *J Uncertain Syst.* (2008) 2:202–11.
19. Beveridge WIB. *The Art of Scientific Investigation.* New York, NY: W W Norton & Company Inc. (1957).
20. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ. The emergence of clinical practice guidelines. *Milbank Q.* (2007) 85:691–727. doi: 10.1111/j.1468-0009.2007.00505.x
21. Ioannidis JP. Why most published research findings are false. *PLoS Med.* (2005) 2:e124. doi: 10.1371/journal.pmed.0020124
22. Wager E, Field EA, Grossman L. Good publication practice for pharmaceutical companies. *Curr Med Res Opin.* (2003) 19:149–54. doi: 10.1185/030079903125001767
23. Lenzer J. Pfizer pleads guilty, but drug sales continue to soar. *BMJ.* (2004) 328:1217. doi: 10.1136/bmj.328.7450.1217
24. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Jour-nal Editors. *Lancet.* (2004) 364:911–2. doi: 10.1016/S0140-6736(04)17034-7
25. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA.* (2003) 289:454–65. doi: 10.1001/jama.289.4.454
26. Cherla DV, Viso CP, Holihan JL, Bernardi K, Moses ML, Mueck KM, et al. The effect of financial conflict of interest, disclosure status, and relevance on medical research from the United States. *J Gen Intern Med.* (2019) 34:429–34. doi: 10.1007/s11606-018-4784-0
27. Waqas A, Baig AA, Khalid MA, Aedma KK, Naveed S. Conflicts of interest and outcomes of clinical trials of antidepressants: an 18-year retrospective study. *J Psychiatr Res.* (2019) 116:83–7. doi: 10.1016/j.jpsychires.2019.05.029
28. Hakoum MB, Bou-Karroum L, Al-Gibbawi M, Khamis AM, Raslan AS, Badour S, et al. Reporting of conflicts of interest by authors of primary studies on health policy and systems research: a cross-sectional survey. *BMJ Open.* (2020) 10:e032425. doi: 10.1136/bmjopen-2019-032425
29. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE.* (2009) 4:e5738. doi: 10.1371/journal.pone.0005738
30. Pupovac V, Fanelli D. Scientists admitting to plagiarism: a meta-analysis of surveys. *Sci Eng Ethics.* (2015) 21:1331–52. doi: 10.1007/s11948-014-9600-6
31. Brown AW, Kaiser KA, Allison DB. Issues with data and analyses. *Proc Natl Acad Sci USA.* (2018) 115:2563–70. doi: 10.1073/pnas.1708279115
32. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E. *Clinical Practice Guidelines We Can Trust, Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines.* Washington, DC: National Academies Press (2011).
33. van Baalen S, Boon M. An epistemological shift: from evidence-based medicine to epistemological responsibility. *J Eval Clin Pract.* (2015) 21:433–9. doi: 10.1111/jep.12282
34. Afshari A, De Hert S. Pitfalls of clinical practice guidelines in the era of broken science: let's raise the standards. *Eur J Anaesthesiol.* (2018) 35:903–6. doi: 10.1097/EJA.0000000000000892
35. Tinetti ME, Bogardus ST Jr, Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *N Engl J Med.* (2004) 351:2870–4. doi: 10.1056/NEJMsb042458

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

MINI REVIEW
published: 08 July 2021
doi: 10.3389/fmed.2021.647897

Digital Technologies and Data Science as Health Enablers: An Outline of Appealing Promises and Compelling Ethical, Legal, and Social Challenges

João V. Cordeiro [1,2,3]*

[1] Public Health Research Centre, NOVA National School of Public Health, Universidade NOVA de Lisboa, Lisboa, Portugal, [2] Comprehensive Health Research Center, Universidade NOVA de Lisboa, Lisboa, Portugal, [3] Centro Interdisciplinar de Ciências Sociais, Lisboa, Portugal

Digital technologies and data science have laid down the promise to revolutionize healthcare by transforming the way health and disease are analyzed and managed in the future. Digital health applications in healthcare include telemedicine, electronic health records, wearable, implantable, injectable and ingestible digital medical devices, health mobile apps as well as the application of artificial intelligence and machine learning algorithms to medical and public health prognosis and decision-making. As is often the case with technological advancement, progress in digital health raises compelling ethical, legal, and social implications (ELSI). This article aims to succinctly map relevant ELSI of the digital health field. The issues of patient autonomy; assessment, value attribution, and validation of health innovation; equity and trustworthiness in healthcare; professional roles and skills and data protection and security are highlighted against the backdrop of the risks of dehumanization of care, the limitations of machine learning-based decision-making and, ultimately, the future contours of human interaction in medicine and public health. The running theme to this article is the underlying tension between the promises of digital health and its many challenges, which is heightened by the contrasting pace of scientific progress and the timed responses provided by law and ethics. Digital applications can prove to be valuable allies for human skills in medicine and public health. Similarly, ethics and the law can be interpreted and perceived as more than obstacles, but also promoters of fairness, inclusiveness, creativity and innovation in health.

Keywords: digital health, ethics, law, artificial intelligence, telemedicine, big data, patient–doctor relationship

OPEN ACCESS

Edited by:
Michele Mario Ciulla, University of Milan, Italy

Reviewed by:
Maurice Mars, University of KwaZulu-Natal, South Africa
Bernard Kamsu Foguem, Université de Toulouse, France
Mara Almeida, University of Lisbon, Portugal

***Correspondence:**
João V. Cordeiro
joao.cordeiro@ensp.unl.pt

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *30 December 2020*
Accepted: *10 June 2021*
Published: *08 July 2021*

Citation:
Cordeiro JV (2021) Digital Technologies and Data Science as Health Enablers: An Outline of Appealing Promises and Compelling Ethical, Legal, and Social Challenges. Front. Med. 8:647897. doi: 10.3389/fmed.2021.647897

INTRODUCTION

Innovative solutions to both classic and emergent medical problems have resulted from the impact of the digital revolution on healthcare (1, 2). Prominent examples include telemedicine, electronic health records, wearable, implantable, injectable and ingestible medical devices, health mobile apps, and the application of artificial intelligence (AI) algorithms to health settings (3). Correspondingly, computer power, interconnectivity and storage capacity, have potentiated the collection, analysis,

and sharing of health data. These advancements, coupled with the expansion of data generation capabilities have led to an era of big data in healthcare, which promises to facilitate timely and precise healthcare interventions (4, 5). In order to achieve this aim, the extraction of knowledge from big data through the interdisciplinary work of data science is fundamental (6).

Better healthcare quality as a result of digital health applications and data science methods is an appealing promise, which also elicits significant ethical, legal, and social challenges (**Table 1**). This article aims to outline this dichotomy, with a limited focus on the examples of telemedicine and AI, which are two specific and interconnected areas in expansion.

In broad terms, telemedicine and telehealth consist in the practice of healthcare through information and telecommunication systems (7). This branch of digital health has had notorious growth in the last years and particularly during the COVID-19 pandemic (8–10). Its applications include, but are not restricted to, real-time health consultations from a distance, remote health data collection, analysis, interpretation, and monitoring, and digital interactions with health assistants, including virtual ones (7). Accordingly, these subjects have deserved particular ethical, legal and scholar attention in recent years (7, 11–13).

In parallel, AI applications in healthcare have gathered significant interest (14, 15). These include, but are also not restricted to, analysis of health data to predict health events and outcomes, check symptoms and improve diagnosis, suggest preventive strategies, design and develop new medicines, improve the organization and conduction of clinical trials, enhance patient experiences, and advance the structure and intelligibility of electronic health records (5, 14, 16–19). Consequently, the influence of AI and machine learning in the health sector is projected to expand and affect the work of healthcare professionals, the efficiency of health systems, and the capacity of patients to interpret their own health data (18, 20, 21). Similarly, awareness about the ethical, legal and social dimensions of AI broadened (22–25), which will hopefully translate into better regulation (26–28).

With a balanced view between promises and challenges and telemedicine and AI as examples, this article aims to provide a succinct review of the main ethical, legal and social implications (ELSI) of the adoption of digital technologies and the processing of health data in medicine and public health.

METHODS

The main topics highlighted in this article resulted from an initial literature search using electronic platforms *PubMed* and *Google Scholar* and general search terms (corresponding to the article keywords in combination with "ethics," "law," and "ELSI"). Additional references were identified by backward and forward citation chaining. Finally, articles from scholars known to the author were also analyzed to complement the analysis.

ELSI OF THE ADOPTION OF DIGITAL TECHNOLOGIES IN HEALTHCARE

Trust, Quality, and the Doctor–Patient Relationship

Trust is a fundamental and reciprocal value in healthcare. To obtain guidance and care, patients trust health professionals in a context of asymmetrical information (29). Conversely, health professionals trust patients to describe their individual experience and their medical history, as well as adhering to recommended behaviors and treatments (30). Therefore, the doctor–patient relationship is based on mutual trust, which is fundamental to ensure quality of care (31, 32). Admittedly, the uptake of digital technologies in healthcare can affect the doctor–patient relationship by reducing human contact and proximity (33). This effect has been substantially debated in the context of telemedicine (9, 12, 34, 35). In particular, the possible devaluation of the importance of continuous face-to-face interactions between patients and doctors, of non-verbal cues and of established ways of building empathy and rapport for economical or efficiency reasons have been highlighted (9). Notably, the impact of digital technologies (and telemedicine in particular) on the doctor-relationship might be very different depending on medical specialty. For example, it might be low for some interactions in dermatology, and completely reshape relations in the context of mental health specialties (35). Likewise, specific functions (image analysis/evaluation of health parameters vs. communication of diagnosis, for example) might also be impacted differently. These considerations reinforce the need to value patient context and preferences in order to improve quality in health, as superficial relationships (even if quantitatively informed) might lead to superficial care.

In a context of increased reliance on digital technologies, the trustworthiness of digital services and goods is fundamental to preserve the value of trust, to strengthen the doctor–patient relationship and increase healthcare quality (36). To achieve these goals, it is essential to rigorously assess the analytical validity and clinical validity and utility of digital technologies (37, 38). Particularly, validity assessment must consider scientific standards for market clearance and authorization, licensing and periodical evaluation, definition and enforceability of quality control norms, and professional requirements for usage and operation (including training, registration and authentication rules). However, despite ongoing efforts, audit and certification procedures are often variable, opaque and incompatible with the rhythm of technological progress (37, 39–41). Furthermore, clinical utility must be critically estimated and, subsequently, communicated to users, which involves the capacity to properly perceive and transmit technology's benefits and risks, as well as uncertain notions such as probability and variance. These issues, together with legal liability clarification and definition of malpractice norms across different geographies and jurisdictions, have been clearly identified as fundamental to guarantee that quality of care and patient safety are protected and hopefully improved during

a sharp uptake of telemedicine and telehealth (8, 9, 34, 35).

Different health stakeholders will deliberate and define relations of trust differently (42). Nonetheless, in broad terms, the recommendation or adoption of subpar digital health services and products risks fostering mistrust in healthcare professionals, institutions and systems, which might ultimately affect the whole digital health field and the broader scientific endeavor (43, 44). As an example, paradigm cases like *Theranos*, in which a hyped promise to revolutionize the blood tests industry turned out to be fraudulent, illustrate the damage that can result from the lack of adequate scrutiny (45). Digital health has lessons to learn from this and other similar cases. Ultimately, the implementation of rigorous, updated and intelligible assessment models is a key component of the digital health promise to promote and advance ethics, evidence, and value-based healthcare (46).

In particular, the introduction of AI and machine learning algorithms in healthcare provides a relevant illustration of the growing need for dedicated assessment and validation. Although progress in the area of deep neural networks allows for assessment of algorithmic capacity using synthetic data, including medical images (47), peer-reviewed real-world clinical data must not be abandoned if poor accuracy or undetermined clinical utility are to be avoided (18). Cases like the IBM Watson-mediated promotion of unsafe and incorrect treatment recommendations to hospitals and medical doctors globally, illustrate how risky it is to adopt AI-based approaches without proper validation, particularly as these flaws can potentially affect large numbers of patients in a short timespan, severely affecting the elements of trust and quality in healthcare (48).

Finally, the extreme case of superficiality in contemporary and future medicine, some argue, is machinal healthcare (33). In fact, it might be precisely in healthcare that dehumanization might prove more costly as human vulnerability, hope, suffering, dependence and, ultimately, death are at stake (49). In accordance, healthcare practice extends beyond technical analysis to include ethics and morality. Interestingly, the competence of machines for moral reasoning, judgement and decision-making is a developing discussion (50–52). Either way, it would reinforce trust, promote quality and strengthen the doctor patient-relationship if digital health tools, and AI in particular, provide stronger incentives for healthcare professionals to focus on caring, compassion, and communication. These are fundamental skills, which are perceived by patients to be in decline (33, 53).

TABLE 1 | Summary of relevant ethical, legal, and social issues (ELSI) raised by digital technologies and health data processing in healthcare.

Digital health ELSI
Ethical
• Promotion of patient autonomy and empowerment
• Design, obtainment, and interpretation informed consent
• Identity confirmation and authentication
• Achieving fair distribution of risks, benefits, and costs
• Guaranteeing quality of care
• Strengthening the doctor–patient relationship
• Assuring continuity of care
• Defining professional duties and responsibilities
• Maintaining confidentiality
• Patient-generated health data
• Direct to consumer telemedicine and unsolicited requests for diagnosis and individual health management
• Dehumanization of care
• Moral status and ethical judgement of machines
• Human nature, *quantified-self* and *technological singularity*
Legal
• Appropriateness, coherence, and accessibility of regulation, including inconsistencies in interpretation by oversight bodies
• Assessment of validity, utility and quality of products, services, strategies, and interventions
• Data protection rights (including privacy by default, privacy by design, data destruction policies, the right to know and the right not to know)
• Data access, return of information and non-discrimination
• Data ownership rights, fair, transparent, and harmonized data sharing rules
• Compliance standards, oversight, and sanctions
• Broader data and device security issues
• Jurisdiction and licensure for telemedicine
Social
• Level of public participation and awareness
• Digital literacy levels of patients and health professionals
• Academic curricula adequacy (upgrades and updates)
• Limits to privacy and confidentiality in health
• Inequality and social stigma
• Lifestyle changes and adoption of healthy behaviors
• Impact on health access (economic, geographical, and informational)

Despite categorization, issues are mainly hybrid in nature.

Transparency, Bias, and Exclusion

In order to achieve its highest aims while preserving trust, the uptake of digital technologies in healthcare must be transparent (40). In a world where data can be artificially created, it is increasingly important that the non-human dimensions of healthcare are disclosed, including the usage of models and algorithms, both in and outside the context of telehealth (9, 14, 15, 54). Therefore, medical decisions supported by digital technologies should be more transparent and understandable, in order to simultaneously guarantee accountability and avoid patient disenfranchizement and exclusion. For example, the incapacity to understand and scrutinize algorithm decision-making leading to less transparency, a challenge known as the black box problem, is currently subject to intense debate including in the healthcare context (55, 56). Notably, the "right to explanation" of algorithm decisions and requirements for human intervention are legally established in different jurisdictions (57). In contrast, some authors view a degree of uncertainty as inevitable and perceive these conditions as obstacles to progress (18, 58). Nonetheless, common ground can be found in the need for better integration of scientific disciplines to surpass hyper technical discussions (and limited understanding and explainability) in contemporary medicine in the context of digital

health in general, and AI in particular. Consequently, adaptation of academic curricula to digital health developments should be prioritized (59). In parallel, the emergence of new health skills or professions that oversee the development of common languages, intersect different disciplines, assist in science implementation and facilitate interactions between different stakeholders is likely (60).

On a separate yet related note, flawed algorithms can feed into human bias and potentiate discrimination (61, 62). Moreover, as the widespread use of facial recognition technology (FRT) edges closer, questionable studies portraying facial traits as proxies for different characteristics (including economic condition, emotional status, and sexual orientation) multiply, raising justified concerns about bias (63–65). The same is true for extrapolating conclusions from online digital behavior (66, 67). Healthcare is not foreign to this debate as FRT can be used to diagnose medical and genetic conditions, for example (68). Notoriously, the issue of AI bias is still open and evolving (69, 70). Nonetheless, it is unreasonable to expect that simply dehumanizing the flawed dimensions of healthcare will *per se* facilitate fairer health outcomes. Expectedly, feeding AI with biased data will lead to biased and unjust decisions (71, 72). Hence, AI implementation in healthcare demands great responsibility (73, 74). Additionally, data is context-dependent and biased context can result in biased conclusions, as studies have shown (75). Conversely, decontextualization of data can also result in algorithm bias, flawed decisions, and discrimination (75, 76). These are renewed arguments to keep fairness and justice at the center of the healthcare debate.

ELSI OF DIGITAL HEALTH DATA PROCESSING

Autonomy, Consent, and Patient Participation

Grounded on the principle of autonomy, informed consent is a cornerstone of medical ethics. Valid informed consent requires a clear and precise acknowledgement of the situation, freedom from coercion (physical or psychological), and competence for decision making (or representation, in the case of minors and incompetent adults) (77). Notably, guaranteeing informed consent for heath research or care purposes, faces singular challenges in the digital era, including identity confirmation, remote evaluation of voluntariness, assessment of understanding levels and competence determination (78, 79).

Defining the scope of consent for health data processing is especially difficult. On one hand, single purpose consent is problematic as secondary uses are often necessary for research and care purposes and re-consent is impracticable (80). On the other hand, in an increasingly fluid ecosystem with expanding interactions between different stakeholders and infrastructures (hospitals, clinics, biobanks, research institutes, biotechnology, and pharma) and possible cycling of health data between health research and healthcare contexts, significant challenges are posed to classical informed consent models. Consequently, in alternative to otherwise open consent options, dynamic consent models have been proposed and justify continuous efforts of implementation (81, 82). Additionally, as health data anonymity is a commodity in an interconnected digital context, legal compliance, management of expectations and risk assessment and communication add extra pressure and complexity to the informed consent process (83–85). Furthermore, some types of health data (for example genetic data), might be shared by more than one person, blurring the limits of individual consent and rights while urging extra care in defining norms and interpreting the law (86).

Along with these challenges to informed consent, the issues of patient autonomy, participation and the doctor–patient relationship converge on other equally challenging digital health data trends. For example, health data can now be generated by patients themselves (*via* apps, wearables, and other digital means) (87). In parallel, healthcare interactions can be patient-initiated (requests of diagnosis and treatment, for example) (88). Thirdly, direct-to-consumer health services, including telemedicine and AI, are expanding (17, 89). These trends highlight the need to extract meaning and knowledge from large quantities of data, while protecting patients from misinformation, misjudgement and disenfranchizement (90). For example, individual and collective risks such as unjustified anxiety, false reassurance and overconsumption of scarce health resources are magnified by digital health illiteracy, neglect, or abandonment to excessive technicality (15, 17, 87, 91). Equally, confirming the accuracy of patient-generated health data and streamlining its integration with electronic health records should be harmonized if healthcare quality is to be promoted (92).

It is well-established that health illiteracy and the digital divide affect patient participation, possibly compromising access to healthcare (93–96). Therefore, to respect autonomy and promote patient participation, the most vulnerable (due to isolation, disability, age, illiteracy, and other factors) deserve special attention and protection. This implies rejecting a *one-size-fits-all* approach and tailoring digital healthcare encounters to individual needs and histories, which are told by different types of data. Among different digital health services and products, which can promote health data sharing and healthcare access, telemedicine has understandably gathered special recognition due to its proven capacity to extend healthcare access to isolated communities (35, 97). Nonetheless, challenges remain as some studies have also indicated that telemedicine services might contribute to medicalize the home context and end up worsening isolation and dependence in some cases (35). This fact further underlines the relevance of context-dependent assessment models and implementation efforts. Furthermore, as telehealth services grew quickly to meet demand in a context of reduced physical interactions such as the COVID-19 pandemic, one must not use the lack of immediate alternatives as an excuse for neglecting fundamental aspects of healthcare ethics. Particularly, the delimitation of professional responsibilities (clinical, administrative, and other) related with health data accessibility and sharing [including the clarification of End User License Agreements (EULAs)] and improved risk communication and cultural respect in digital

interactions must be focused on in order to attribute real meaning to health data and achieve the highest hopes for this technology (9).

Finally, it is the human pondering of different alternatives that enriches the consent process and furthers patient participation and autonomy in healthcare. As machine learning algorithms gather more influence and prove to be increasingly autonomous, new challenges are posed to these ethical principles in the context of health data processing (98). Therefore, progress in healthcare should not consist in the promotion of machine autonomy at the expense of human autonomy. On the contrary, well-established human values in healthcare, such as integrity, conscientiousness and compassion must guide health data processing in a digital context and work as allies of digital health innovation.

Privacy, Confidentiality, and Security

The issues of privacy, confidentiality and data protection are recognized as fundamental rights in most jurisdictions and are especially challenging for digital health (99–103). The old debate surrounding the erosion of privacy and confidentiality in health settings endures (104, 105). For example, the protection of electronic health records has been widely recognized as insufficient (106, 107). Furthermore, privacy protection, data access, interoperability, and quality of recorded data are recurrently reported as ELSI of digital health, including of telemedicine (9, 12, 34, 35). Undoubtedly, health data processing is essential for medical and scientific progress and should follow transparent, balanced and fair rules (108–112). In order to strengthen fundamental rights in the digital age and regulate the free movement of personal data, including health data, the EU has adopted the General Data Protection Regulation (GDPR) (113). This broad ranging legal document reinforced mechanisms of data protection, including more transparency and accountability, mandatory impact assessments, pan-European validation of codes of conduct, certification procedures, and more severe sanctions (114). Furthermore, rights of data subjects were enhanced, including a right of access and rights to information, explanation, rectification, erasure, restriction of processing, data portability, object and not to be subject to automated individual decision-making (113). Due to its broad scope and recent nature, it is still early to judge the impact of GDPR in the health area. Additionally, harmonization of health data flows across jurisdictions is still problematic. For example, the successive European Court of Justice *Schrems* cases (115), leading to successive annulments of legal agreements regulating EU-US data flows, have impacted health research and healthcare (116).

In parallel, health data security has become a serious concern as poor protection measures combined with high transactional value, exacerbate the risk of violations and damages (117, 118). Particular cybersecurity concerns emerge from the expansion of digital health, including telemedicine and the multiplication of interconnected sensors and medical devices (119), which has rightly deserved regulatory attention (120–123). Ultimately, healthcare institutions should see their data security infrastructure strengthened and recent technological progress can provide the tools for risk mitigation (124).

In complement to data protection measures, responsible use of data is key. Especially, projects with public notoriety demand greater responsibility if public trust is to be preserved. For example, cases such as the Google Deepmind collaboration with the UK NHS (125, 126), NHS England's care.data programme (127), or "Project Nightingale" in the US, in which patient data was accessed by commercial companies without informed consent, emphasize the need for transparency, communication and responsibility in order to guarantee the positive impact of data sharing on healthcare. In contrast, opacity extends power imbalances and unprotects citizens (128). Therefore, clear and fair health data ownership rules, beyond the traditional property approaches, should continue to be developed and harmonized. These should guarantee patient access to their data (129) while limiting access and usage by third parties without a compelling interest. Moreover, different strategies can be adopted to encourage responsible use of health data. For example, investing in healthcare ethics literacy programs, implementing validated codes of conduct (institutional, national, and international), refining deontology rules (for health professionals and data scientists), protecting whistleblowers of data mistreatment practices, setting fair procedures, and imposing dissuasive sanctions (disciplinary, legal, and social) for confirmed misconduct. Obviously, such strategies must include electronic health records and health data shared in telemedicine settings, processed by mobile apps and medical devices, as well as AI algorithms (130, 131).

DISCUSSION

The digital revolution has impacted different areas of society, including healthcare (19). At the core of this transformation is an increased capacity to process large quantities of health data using digital means (15, 16). Big data in healthcare originates from different sources, including biological and social determinants, health records, environmental signals, habits, and behaviors (19, 132, 133). Against this backdrop, telemedicine and telehealth services expanded significantly in recent years, particularly during the COVID-19 pandemic (10). Furthermore, AI applications are gaining ground in complementing even the most knowledgeable or skilled professionals (134, 135). Concomitantly, a new era of precision healthcare is promised, where the right individual and public intervention is available for the right patient or population at the right time (136–141).

In this context, the optimization of health data processing using digital means and the general uptake of digital technologies can rightly be perceived as health enablers. In parallel, compelling related ELSI must be considered and dealt with.

Expectedly, healthier activities and wiser health choices should result from better data science. In this sense, digitally-mediated health data processing can promote individual autonomy and patient empowerment (142–144). However, the adoption of healthy behaviors does not emerge linearly from better health information (which must be extracted from health data) as human decision making is complex and affected by context and cognitive biases, combining emotion and reason.

Therefore, data-assisted decision making in healthcare, justifies a closer collaboration between healthcare professionals, decision, and data scientists and ethicists (15, 33). In fact, the links between health data and statistics literacy and healthcare quality is a classical debate, which is expected to intensify (145). Indeed, the adoption of digital technologies has the potential to improve the volume and quality of health data processing in order to expand knowledge to professionals and patients alike. However, a lack of common platforms and cross-disciplinary languages to deal with increasing technical complexity are significant challenges (146). Also, can technology, data and analytical models alone capture human vulnerability, suffering, fears, hopes and potential? Evidently not. Nonetheless, these can elevate the standard of care by providing healthcare professionals with invaluable (an otherwise inaccessible) information and knowledge, while alleviating the burden of repetitive and laborious tasks to focus on compassion and emotional connections, which are associated with the highest quality (15, 33). To this end, patient stories, particularly those of the most vulnerable, must be heard and understood and one must be mindful that health data misuse can contribute to misinformation, poorer care, reinforced exclusion or stigmatization. Notably, such risks are exacerbated if human health and disease are looked at from a purely quantitative lens (147–150). There is, however, cause for optimism as digital technologies can also be used to promote scientific robustness and tackle the very risks it potentially generates (151–154).

Balanced health data processing and usage of digital technologies to improve healthcare quality is a matter of public interest. Presently, there is significant data access asymmetries between citizens, corporations and governments (128). Therefore, urgent efforts are necessary to reach an inclusive and democratic deliberation leading to the simultaneous advancement of science and human rights. Importantly, digital technologies can advance the fulfillment of the human right to health. Specifically, they can improve the availability of health facilities, services and goods; increase the acceptability of practices (by incorporating medical ethics and approximating cultures); raise the quality of scientific and medical services, goods and professional skills; and promote access without discrimination (155, 156). The specific issue of fairness regarding access to digital technologies has been a key topic in the context of the accelerated uptake of telemedicine. This is particularly relevant as those who are more likely to benefit from this technology and its applications (isolated communities, including in rural areas) are also those who, predictably, are less capable of affording or using them (35). Therefore, public and individual interests must be properly balanced in order to maximize the potential of this technology while respecting human rights. Additionally, big data, telehealth, machine learning algorithms and the era of individual profiling might be ingredients for deeper discrimination and stigmatization (142). In summary, the positive application of digital technologies and data science in medicine and public health should promote, not defer progress in social justice.

In conclusion, the ELSI of the digital health field (**Table 1**) are compelling and proportional to the positive impact of digital technologies in healthcare. Consequently, normative orders such as law and ethics should act as beneficial limit-setters and promoters of just, creative and innovative realities. Accordingly, digital health ELSI convoke ethicists, legal scholars, patients, scientists, health professionals, health providers and payers, regulators, managers, and other decision makers to play a role in this fascinating field, which promises to decisively shape the way health and disease are perceived, assessed and managed in the future (157, 158).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This present publication was funded by Fundação Ciência e Tecnologia, IP national support through CHRC (UIDP/04923/2020).

REFERENCES

1. Elenko E, Underwood L, Zohar D. Defining digital medicine. *Nat Biotechnol.* (2015) 33:456–61. doi: 10.1038/nbt.3222
2. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol.* (2015) 33:462–3. doi: 10.1038/nbt.3223
3. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med.* (2020) 3:118. doi: 10.1038/s41746-020-00324-0
4. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA.* (2013) 309:1351. doi: 10.1001/jama.2013.393
5. Topol EJ. A decade of digital medicine innovation. *Sci Transl Med.* (2019) 11:eaaw7610. doi: 10.1126/scitranslmed.aaw7610
6. EMC Education Services. *Data Science & Big Data Analytics* (2015).
7. American Medical Association. *Telehealth Implementation Playbook.* Digital Health Implementation Playbook Series (2020).
8. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. *J Med Internet Res.* (2020) 22:e22280. doi: 10.2196/preprints.22280
9. Kaplan B. REVISITING HEALTH INFORMATION TECHNOLOGY ETHICAL, LEGAL, and SOCIAL ISSUES and EVALUATION: TELEHEALTH/TELEMEDICINE and COVID-19. *Int J Med Inform.* (2020) 143:104239. doi: 10.1016/j.ijmedinf.2020.104239
10. Brody JE. A pandemic benefit: the expansion of telemedicine. *The New York Times* (2020). Available online at: https://www.nytimes.com/2020/05/11/well/live/coronavirus-telemedicine-telehealth.html (accessed December 30, 2020).
11. World Medical Association. *WMA Statement on the Ethics of Telemedicine* (2018).
12. Botrugno C. Towards an ethics for telehealth. *Nurs Ethics.* (2019) 26:357–67. doi: 10.1177/0969733017705004
13. Raposo VL. Telemedicine: the legal framework (or the lack of it) in Europe. *GMS Health Technol Assess.* (2016) 12:Doc03. doi: 10.3205/hta000126
14. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* (2019) 6:94–8. doi: 10.7861/futurehosp.6-2-94
15. Topol EJ. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York, NY: Basic Books (2019).
16. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med.* (2016) 375:1216–9. doi: 10.1056/NEJMp1606181

17. Babic B, Gerke S, Evgeniou T, Cohen IG. Direct-to-consumer medical machine learning and artificial intelligence applications. *Nat Mach Intell.* (2021) 3:283–7. doi: 10.1038/s42256-021-00331-0
18. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* (2019) 25:44–59. doi: 10.1038/s41591-018-0300-7
19. Topol EJ. Individualized medicine from prewomb to tomb. *Cell.* (2014) 157:241–53. doi: 10.1016/j.cell.2014.02.012
20. Anthes E. The shape of work to come. *Nature.* (2017) 550:316–9. doi: 10.1038/550316a
21. Nedelkoska L, Glenda Q. Automation, skills use and training. In: *OECD Social, Employment and Migration Working Papers*. Paris: OECD Publishing (2018).
22. Etzioni A, Etzioni O. Incorporating ethics into artificial intelligence. *J Ethics.* (2017) 21:403–18. doi: 10.1007/s10892-017-9252-2
23. High-Level Expert Group on AI. *Ethics Guidelines for Trustworthy AI*. European Commission (2019).
24. High-Level Expert Group on AI. *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment*. European Commission (2020).
25. Courtland R. Bias detectives: The researchers striving to make algorithms fair news-feature. *Nature.* (2018) 558:357–60. doi: 10.1038/d41586-018-05469-3
26. Digital Economy Task Force. Trustworthy AI in health. In: *Background Paper for the G20 AI Dialogue*. OECD (2020). Available online at: https://www.oecd.org/health/trustworthy-artificial-intelligence-in-health.pdf
27. Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. Practical guidance on artificial intelligence for health-care data. *Lancet Digit Health.* (2019) 1:e157–e9. doi: 10.1016/S2589-7500(19)30084-6
28. Cohen IG, Evgeniou T, Gerke S, Minssen T. The European artificial intelligence strategy: implications and challenges for digital health. *Lancet Digit Health.* (2020) 2:e376–e9. doi: 10.1016/S2589-7500(20)30112-6
29. Zaner RM. The phenomenon of trust and the patient-physician relationship. In: Pellegrino ED, Veatch RM, Langan JP, editors. *ETHICS, TRUST, AND THE PROFESSIONS Philosophical and Cultural Aspects*. Washington, DC: Georgetown University Press (1991). p. 45–68.
30. Sousa-Duarte F, Brown P, Mendes AM. Healthcare professionals' trust in patients: a review of the empirical and theoretical literatures. *Sociol Compass.* (2020) 14:1–15. doi: 10.1111/soc4.12828
31. Donabedian A. Evaluating the quality of medical care. *Milbank Q.* (2005) 83:691–725. doi: 10.1111/j.1468-0009.2005.00397.x
32. Rasiah S, Jaafar S, Yusof S, Ponnudurai G, Chung KPY, Amirthalingam SD. A study of the nature and level of trust between patients and healthcare providers, its dimensions and determinants: a scoping review protocol. *BMJ Open.* (2020) 10:e028061. doi: 10.1136/bmjopen-2018-028061
33. Verghese A. How tech can turn doctors into clerical workers. *The New York Times Magazine* (2018). Available online at: https://www.nytimes.com/interactive/2018/05/16/magazine/health-issue-what-we-lose-with-data-driven-medicine.html?mtrref=www.google.com&gwh=50127696B49AB16FC5EA028724BB6740&gwt=regi&assetType=REGIWALL (accessed December 30, 2020).
34. Nittari G, Khuman R, Baldoni S, Pallotta G, Battineni G, Sirignano A, et al. Telemedicine practice: review of the current ethical and legal challenges. *Telemed eHealth.* (2020) 26:1427–37. doi: 10.1089/tmj.2019.0158
35. Keenan AJ, Cert G, Tsourtos G, Tieman J. The value of applying ethical principles in telehealth practices: systematic review. *J Med Internet Res.* (2021) 23:e25698. doi: 10.2196/25698
36. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. *J Med Internet Res.* (2018) 20:e11254. doi: 10.2196/11254
37. Garell C, Svedberg P, Nygren JM. A legal framework to support development and assessment of digital health services. *JMIR Med Inform.* (2016) 4:e17. doi: 10.2196/medinform.5401
38. Perakslis E, Ginsburg GS. Digital health-the need to assess benefits, risks, and value. *JAMA*. (2021) 325:127–8. doi: 10.1001/jama.2020.22919
39. Schulke DF. The regulatory arms race: mobile health applications and agency posturing. *Boston Univ Law Rev.* (2013) 93:1699–752.
40. Rodriguez-Villa E, Torous J. Regulating digital health technologies with transparency: the case for dynamic and multi-stakeholder evaluation. *BMC Med.* (2019) 17:226. doi: 10.1186/s12916-019-1447-x
41. World Health Organisation. *Monitoring and Evaluating Digital Health Interventions: A Practical Guide to Conducting Research and Assessment*. WHO (2016).
42. Pilgrim D, Tomasini F, Vassilev I. *Examining Trust in Healthcare: A Multidisciplinary Perspective.* London: Palgrave Macmillan (2011).
43. Kabat GC. Taking distrust of science seriously. *EMBO Rep*. (2017) 18:1052–5. doi: 10.15252/embr.201744294
44. Baron RJ, Berinsky AJ. Mistrust in science - a threat to the patient-physician relationship. *N Engl J Med.* (2019) 381:182–5. doi: 10.1056/NEJMms1813043
45. Carreyrou J. *Bad Blood: Secrets and Lies in a Silicon Valley Startup*. New York, NY: Vintage Books, Pengun Random House (2018).
46. Rigby M, Ammenwerth E. The need for evidence in health informatics. In: Ammenwerth E, Rigby M, editors. *Evidence-Based Health Informatics: Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*. Amsterdam: IOS Press (2016). p. 3–14.
47. Guibas JT, Virdi TS, Li PS. Synthetic medical images from dual generative adversarial networks. *arXiv:1709.01872 [Preprint]*. (2018).
48. Ross C, Swetlitz I. *IBM's Watson Supercomputer Recommended 'Unsafe and Incorrect' Cancer Treatments, Internal Documents Show*. Stat+ (2018).
49. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. *Algorithmic Bias In Health Care: A Path Forward.* Health Affairs (2019).
50. Anderson M, Susan LA. Robot be good. *Sci Am.* (2010) 303:72–7. doi: 10.1038/scientificamerican1010-72
51. Several authors. Morals and the machine. *The Economist*. (2012). Available online at: https://www.economist.com/leaders/2012/06/02/morals-and-the-machine (accessed December 30, 2020).
52. Cervantes JA, López S, Rodríguez LF, Cervantes S, Cervantes F, Ramos F. Artificial moral agents: a survey of the current status. *Sci Eng Ethics.* (2020) 26:501–32. doi: 10.1007/s11948-019-00151-x
53. Singletary B, Patel N, Heslin M. Patient perceptions about their physician in 2 words. *JAMA Surg.* (2017) 152:1169–70. doi: 10.1001/jamasurg.2017.3851
54. Lima MR, Wairagkar M, Natarajan N, Vaitheswaran S, Vaidyanathan R. Robotic telemedicine for mental health: a multimodal approach to improve human-robot engagement. *Front Robotics AI.* (2021) 8:618866. doi: 10.3389/frobt.2021.618866
55. Knight W. The dark secret at the heart of AI. *Technol Rev.* (2017) 120:54–61.
56. Bjerring JC, Busch J. Artificial intelligence and patient-centered decision-making. *Philos Technol.* (2021) 34:349–71. doi: 10.1007/s13347-019-00391-6
57. Atkinson K, Bench-Capon T, Bollegala D. Explanation in AI and law: past, present and future. *Artif Intell.* (2020) 289:103387. doi: 10.1016/j.artint.2020.103387
58. Brouillette M. *Deep Learning Is a Black Box, But Health Care Won't Mind - MIT Technology Review*. MIT Technology Review. *(*2017).
59. Aungst TD, Patel R. Integrating digital health into the curriculum-considerations on the current landscape and future developments. *J Med Educ Curric Dev.* (2020) 7:1–7. doi: 10.1177/2382120519901275
60. Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. *BMC Med.* (2019) 17:88. doi: 10.1186/s12916-019-1322-9
61. Angwin J, Larson J, Mattu S, Kirchner L. *Machine Bias*. ProPublica (2016).
62. Simonite T. When it comes to gorillas, google photos remains blind. *Wired* (2018). Available online at: https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/ (accessed December 30, 2020).
63. Burdick A. The A.I. "Gaydar" study and the real dangers of big data. *The New Yorker* (2017, September 15). Available online at: https://www.newyorker.com/news/daily-comment/the-ai-gaydar-study-and-the-real-dangers-of-big-data
64. van Noorden R. The ethical questions that haunt facial-recognition research. *Nature.* (2020) 587:354–8. doi: 10.1038/d41586-020-03187-3
65. Castelvecchi D. Is facial recognition too biased to be let loose? The technology is improving - but the bigger issue is how it's used. *Nature.* (2020) 587:347–9. doi: 10.1038/d41586-020-03186-4
66. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci U S A*. (2014) 111:8788–90. doi: 10.1073/pnas.1320040111

67. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci U S A*. (2013) 110:5802–5. doi: 10.1073/pnas.1218772110
68. Martinez-Martin N. What are important ethical implications of using facial recognition technology in health care? *AMA J Ethics*. (2019) 21:E180–E7. doi: 10.1001/amajethics.2019.180
69. Miller AP. Want Less-Biased Decisions? Use Algorithms. *Harvard Business Review Digital Articles* (2018, July 26). Available online at: https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms
70. Thomas R. *What HBR Gets Wrong About Algorithms and Bias*. fast.ai. (2018).
71. Zou J, Schiebinger L. AI can be sexist and racist - it's time to make it fair. *Nature*. (2018) 559:324–6. doi: 10.1038/d41586-018-05707-8
72. Adamson AS, Smith A. Machine learning and health care disparities in dermatology. *JAMA Dermatol*. (2018) 154:1247–8. doi: 10.1001/jamadermatol.2018.2348
73. Castelvecchi D. Prestigious AI meeting takes steps to improve ethics of research. *Nature*. (2021) 589:12–3. doi: 10.1038/d41586-020-03611-8
74. Gibney E. The battle for ethical AI at the world's biggest machine-learning conference. *Nature*. (2020) 577:609. doi: 10.1038/d41586-020-00160-y
75. Borgesius F. *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*. Council of Europe (2018).
76. Lorentz A. With big data, context is a big issue. *Wired* (2018). Available online at: https://www.wired.com/insights/2013/04/with-big-data-context-is-a-big-issue/ (accessed December 30, 2020).
77. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics, 8th Edn*. Oxford: Oxford University Press (2019).
78. Grady C. Enduring and emerging challenges of informed consent. *N Engl J Med*. (2015) 372:855–62. doi: 10.1056/NEJMra1411250
79. Grady C, Cummings SR, Rowbotham MC, McConnell Mv, Ashley EA, Kang G. Informed consent. *N Engl J Med*. (2017) 376:856–67. doi: 10.1056/NEJMra1603773
80. Aicardi C, del Savio L, Dove ES, Lucivero F, Tempini N, Prainsack B. Emerging ethical issues regarding digital health data. On the World Medical Association Draft Declaration on Ethical Considerations Regarding Health Databases and Biobanks. *Croatian Med J*. (2016) 57:207–13. doi: 10.3325/cmj.2016.57.207
81. Schneble CO, Elger BS, Shaw DM. All our data will be health data one day: the need for universal data protection and comprehensive consent. *J Med Internet Res*. (2020) 22:e16879. doi: 10.2196/16879
82. Vayena E, Blasimme A. Biomedical big data: new models of control over access, use and governance. *J Bioethical Inquiry*. (2017) 14:501–13. doi: 10.1007/s11673-017-9809-6
83. Rocher L, Hendrickx JM, de Montjoye Y-A. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun*. (2019) 10:3069. doi: 10.1038/s41467-019-10933-3
84. Editorial. Time to discuss consent in digital-data studies. *Nature*. (2019) 572:5. doi: 10.1038/d41586-019-02322-z
85. Mann SP, Savulescu J, Sahakian BJ. Facilitating the ethical use of health data for the benefit of society: electronic health records, consent and the duty of easy rescue. *Philos Trans A Math Phys Eng Sci*. (2016) 374:20160130. doi: 10.1098/rsta.2016.0130
86. Annas GJ. Bioidentifiers. In: *Worst Case Bioethics*. New York, NY: Oxford University Press (2010). p. 235–50.
87. Lordon RJ, Mikles SP, Kneale L, Evans HL, Munson SA, Backonja U, et al. How patient-generated health data and patient-reported outcomes affect patient-clinician relationships: a systematic review. *Health Inform J*. (2020) 26:2689–706. doi: 10.1177/1460458220928184
88. Whear R, Thompson-Coon J, Rogers M, Abbott RA, Anderson L, Ukoumunne O, et al. Patient-initiated appointment systems for adults with chronic conditions in secondary care. *Cochrane Database Syst Rev*. (2020) 4:CD010763. doi: 10.1002/14651858.CD010763.pub2
89. Elliott T, Shih J. Direct to consumer telemedicine. *Curr Allergy Asthma Rep*. (2019) 19:1. doi: 10.1007/s11882-019-0837-7
90. Ho A, Quick O. Leaving patients to their own devices? Smart technology, safety and therapeutic relationships. *BMC Med Ethics*. (2018) 19:18. doi: 10.1186/s12910-018-0255-8
91. Bollmeier SG, Stevenson E, Finnegan P, Griggs SK. Direct to consumer telemedicine: is healthcare from home best? *Missouri Med*. (2020) 117:303–9. doi: 10.1007/978-3-030-53879-8_11
92. Abdolkhani R, Gray K, Borda A, DeSouza R. Patient-generated health data management and quality challenges in remote patient monitoring. *JAMIA Open*. (2019) 2:417–8. doi: 10.1093/jamiaopen/ooz036
93. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: Literacy in underserved populations in the United States. *Inform Health Soc Care*. (2015) 41:1–19. doi: 10.3109/17538157.2014.948171
94. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: Measuring a broad spectrum of health 1.0 and health 2.0 skills. *J Med Internet Res*. (2017) 19:e27. doi: 10.2196/jmir.6709
95. Brall C, Schröder-Bäck P, Maeckelberghe E. Ethical aspects of digital health from a justice point of view. *Eur J Public Health*. (2019) 29(Suppl. 3):18–22. doi: 10.1093/eurpub/ckz167
96. McAuley A. Digital health interventions: widening access or widening inequalities? *Public Health*. (2014) 128:1118–20. doi: 10.1016/j.puhe.2014.10.008
97. Rubeis G, Schochow M, Steger F. Patient autonomy and quality of care in telehealthcare. *Sci Eng Ethics*. (2018) 24:93–107. doi: 10.1007/s11948-017-9885-3
98. Bitterman DS, Aerts HJWL, Mak RH. Approaching autonomy in medical artificial intelligence. *Lancet Digit Health*. (2020) 2:e447–e9. doi: 10.1016/S2589-7500(20)30187-4
99. Buchner B. Privacy. In: den Exter A, editor. *European Health Law*. Apeldoorn: Maklu-Uitgevers nv (2017). p. 273–90.
100. Rothstein MA. Privacy and confidentiality. In: Joly Y, Knoppers BM, editors. *Routledge Handbook of Medical Law and Ethics*. New York, NY: Routledge (2015). p. 52–66.
101. González Fuster G. *The Emergence of Personal Data Protection as a Fundamental Right of the EU*. Cham: Springer International Publishing (2014).
102. McDermott Y. Conceptualising the right to data protection in an era of Big Data. *Big Data Soc*. (2017) 4:205395171668699. doi: 10.1177/2053951716686994
103. High Commissioner for Human Rights. *The Right to Privacy in the Digital Age*. (2018).
104. Siegler M. Confidentiality in medicine - a decrepit concept. *N Engl J Med*. (1982) 307:1518–21. doi: 10.1056/NEJM198212093072411
105. de Faria PL, Cordeiro JV. Health data privacy and confidentiality rights: Crisis or redemption? *Rev Portuguesa Saude Publ*. (2014) 32:123–33. doi: 10.1016/j.rpsp.2014.10.001
106. Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. *J Med Syst*. (2016) 40:252. doi: 10.1007/s10916-016-0628-9
107. Rezaeibagha F, Win KT, Susilo W. A systematic literature review on security and privacy of electronic health record systems: technical perspectives. *Health Inform Manag J*. (2015) 44:23–38. doi: 10.1177/183335831504400304
108. Kosseim P, Dove ES, Baggaley C, Meslin EM, Cate FH, Kaye J, et al. Building a data sharing model for global genomic research. *Genome Biol*. (2014) 15:430. doi: 10.1186/s13059-014-0430-2
109. Cook-Deegan R, Ankeny RA, Maxson Jones K. Sharing data to build a medical information commons: from bermuda to the global alliance. *Annu Rev Genomics Human Genet*. (2017) 18:389–415. doi: 10.1146/annurev-genom-083115-022515
110. Baker DB, Kaye J, Terry SF. Privacy, fairness, and respect for individuals. *eGEMs*. (2016) 4:7. doi: 10.13063/2327-9214.1207
111. Woolley JP, McGowan ML, Teare HJA, Coathup V, Fishman JR, Settersten RA, et al. Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives Donna Dickenson, Sandra Soo-Jin Lee, and Michael Morrison. *BMC Medical Ethics*. (2016) 17:33. doi: 10.1186/s12910-016-0117-1
112. Atasoy H, Greenwood BN, McCullough JS. The digitization of patient care: a review of the effects of electronic health records on health care quality and utilization. *Annu Rev Public Health*. (2019) 40:487–500. doi: 10.1146/annurev-publhealth-040218-044206
113. European Commission. *Data Protection in the EU* (2020).
114. Buttarelli G. The EU GDPR as a clarion call for a new global digital gold standard. *Int Data Privacy Law*. (2016) 6:77–8. doi: 10.1093/idpl/ipw006
115. Costello RA. Róisín Áine. Schrems II: everything is illuminated? *Eur Pap J Law Integr*. (2020) 2020:1045–59.

116. Global Alliance for Genomics and Health. *GA4GH GDPR Brief: Transferring Genomic and Health-Related Data Following Schrems II* (2020).
117. Gillum J, Kao J, Larson J. *Millions of Americans' Medical Images and Data Are Available on the Internet. Anyone Can Take a Peek*. ProPublica. (2019).
118. CBS News. Hackers are stealing millions of medical records - and selling them on the dark web. *CBS News* (2019) (accessed December 30, 2020).
119. Thierer AD. The internet of things & wearable technology: addressing privacy & security concerns without derailing innovation. *SSRN Electr J.* (2015). doi: 10.2139/ssrn.2494382
120. Williams PAH, Woodward AJ. Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Med Dev Evid Res.* (2015) 8:305–16. doi: 10.2147/MDER.S50048
121. Food and Drug Administration (FDA). *Medical Device Cybersecurity: What You Need to Know* (2020).
122. European Commission. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS on Enabling the Digital Transformation of Health and Care in the Digital Single Market; Empowering Citizens and Building a Healthier Society*. COM/2018/233 Final (2018).
123. Schwartz S, Ross A, Carmody S, Chase P, Coley SC, Connolly J, et al. The evolving state of medical device cybersecurity. *Biomed Instrum Technol.* (2018) 52:103–11. doi: 10.2345/0899-8205-52.2.103
124. Shi S, He D, Li L, Kumar N, Khan MK, Choo K-KR. Applications of blockchain in ensuring the security and privacy of electronic health record systems: a survey. *Comput Secur.* (2020) 97:101966. doi: 10.1016/j.cose.2020.101966
125. Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol.* (2017) 7:351–67. doi: 10.1007/s12553-017-0179-1
126. Vaughan A. Google gets green light to access five years of NHS patient data. *New Scientist* (2019). Available online at: https://www.newscientist.com/article/2220344-google-gets-green-light-to-access-five-years-of-nhs-patient-data/ (accessed December 30, 2020).
127. Godlee F. What can we salvage from care.data? *BMJ.* (2016) 354:i3907. doi: 10.1136/bmj.i3907
128. Véliz C. *Privacy is Power: Why and How You Should Take Back Control of Your Data*. London: Penguin Random House UK (2020).
129. Neves AL, Freise L, Laranjo L, Carter AW, Darzi A, Mayer E. Impact of providing patients access to electronic health records on quality and safety of care: a systematic review and meta-analysis. *BMJ Qual Saf.* (2020) 29:1019–32. doi: 10.1136/bmjqs-2019-010581
130. Kish LJ, Topol EJ. Unpatients-why patients should own their medical data. *Nat Biotechnol.* (2015) 33:921–4. doi: 10.1038/nbt.3340
131. Wilbanks JT, Topol EJ. Stop the privatization of health data. *Nature*. (2016) 535:345–8. doi: 10.1038/535345a
132. Cordeiro JV. Ethical and legal challenges of personalized medicine: paradigmatic examples of research, prevention, diagnosis and treatment. *Rev Portuguesa Saude Publ.* (2014) 32:164–80. doi: 10.1016/j.rpsp.2014.10.002
133. Ho D, Quake SR, McCabe ERB, Chng WJ, Chow EK, Ding X, et al. Enabling technologies for personalized and precision medicine. *Trends Biotechnol.* (2020) 38:497–518. doi: 10.1016/j.tibtech.2019.12.021
134. Schmietow B, Marckmann G. Mobile health ethics and the expanding role of autonomy. *Med Health Care Philos.* (2019) 22:623–30. doi: 10.1007/s11019-019-09900-y
135. Lamanna C, Byrne L. Should artificial intelligence augment medical decision making? The Case for an Autonomy Algorithm. *AMA J Ethics.* (2018) 20:E902–10. doi: 10.1001/amajethics.2018.902
136. Sunstein CR. The ethics of nudging. *Yale J Regul.* (2015) 32:413–50. doi: 10.2139/ssrn.2526341
137. Gornick MC, Zikmund-Fisher BJ. What clinical ethics can learn from decision science. *AMA J Ethics.* (2019) 21:E906–E12. doi: 10.1001/amajethics.2019.906
138. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. *Lancet Digit Health.* (2020) 2:e425–34. doi: 10.1016/S2589-7500(20)30137-0
139. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med.* (2016) 50:398–401. doi: 10.1016/j.amepre.2015.08.031
140. Rasmussen SA, Khoury MJ, del Rio C. Precision public health as a key tool in the COVID-19 response. *JAMA.* (2020) 324:933–4. doi: 10.1001/jama.2020.14992
141. Khoury MJ, Bowen MS, Clyne M, Dotson WD, Gwinn ML, Green RF, et al. From public health genomics to precision public health: A 20-year journey. *Genet Med.* (2018) 20:574–82. doi: 10.1038/gim.2017.211
142. Lupton D. *Digital Health: Critical and Cross-Disciplinary Perspectives.* Abingdon, Oxfordshire: Routledge (2017).
143. White RW, Horvitz E. Experiences with web search on medical concerns and self diagnosis. *AMIA Annu Symp Proc*. (2009) 2009:696–700.
144. de Faria PL, Cordeiro JV. Public health: current and emergent legal and ethical issues in a nutshell. In: Joly Y, Knoppers BM, editors. *Routledge Handbook of Medical Law and Ethics.* New York, NY: Routledge (2014). p. 381–401.
145. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest.* (2007) 8:53–96. doi: 10.1111/j.1539-6053.2008.00033.x
146. Land MK, Aronson JD. Human rights and technology: new challenges for justice and accountability. *Annu Rev Law Soc Sci.* (2020) 16:223–40. doi: 10.1146/annurev-lawsocsci-060220-081955
147. Ajana B. Digital health and the biopolitics of the Quantified Self. *Digit Health.* (2017) 3:205520761668950. doi: 10.1177/2055207616689509
148. Ruckenstein M, Schüll ND. The datafication of health. *Annu Rev Anthropol.* (2017) 46:261–78. doi: 10.1146/annurev-anthro-102116-041244
149. Sharon T. Self-tracking for health and the quantified self: re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philos Technol.* (2017) 30:93–121. doi: 10.1007/s13347-016-0215-5
150. Abnousi F, Rumsfeld JS, Krumholz HM. Social determinants of health in the digital age. *JAMA.* (2019) 321:247–8. doi: 10.1001/jama.2018.19763
151. Nissim K, Wood A. Is privacy privacy? *Philos Trans R Soc A Math Phys Eng Sci.* (2018) 376:1–17. doi: 10.1098/rsta.2017.0358
152. Papernot N, Abadi M, Erlingsson Ú, Goodfellow I, Talwar K. Semi-supervised knowledge transfer for deep learning from private training data. *arXiv:1610.05755 [Preprint]*. (2017).
153. Yang Y, Youyou W, Uzzi B. Estimating the deep replicability of scientific findings using human and artificial intelligence. *Proc Natl Acad Sci U S A.* (2020) 117:10762–8. doi: 10.1073/pnas.1909046117
154. Volpp KG, Mohta NS. *Social Networks to Improve Patient Health Advisor Analysis. NEJM Catalyst.* (2017).
155. UN. *CESCR General Comment no. 14: The Right to the Highest Attainable Standard of Health (Article 12 of the International Covenant on Economic, Social and Cultural Rights)*. E/C.12/2000/4 (2000).
156. Faria PL, Cordeiro JV. Managing a difficult ethical and legal equilibrium in healthcare: assuring access to the basics while keeping up with innovation. *Rev Portuguesa Saude Publ.* (2014) 32:121–2. doi: 10.1016/j.rpsp.2014.11.001
157. Topol E. *Preparing the Healthcare Workforce to Deliver the Digital Future. The Topol Review*. An Independent Report on Behalf of the Secretary of State for Health and Social Care (2019).
158. Intelligence Unit. The future of healthcare. *The Economist* (2017). Available online at: https://eiuperspectives.economist.com/healthcare/future-healthcare (accessed December 30, 2020).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

SYSTEMATIC REVIEW
published: 22 July 2021
doi: 10.3389/fmed.2021.618046

Systematic Review of the Effect of a Zero-Markup Policy for Essential Drugs on Healthcare Costs and Utilization in China, 2015–2021

Wen-Yi Liu[1,2,3], *Chia-Hsien Hsu*[4], *Ting-Jun Liu*[5], *Pei-En Chen*[6], *Boyuan Zheng*[2], *Ching-Wen Chien*[2*] *and Tao-Hsin Tung*[7*]

[1] Department of Health Policy Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States, [2] Institute for Hospital Management, Tsing Hua University, Shenzhen, China, [3] Shanghai Bluecross Medical Science Institute, Shanghai, China, [4] Department of Public Health, College of Health Science, Kaohsiung Medical University, Kaohsiung, China, [5] Tai Kang Institute of Healthcare Management, Beijing, China, [6] Institute of Health Policy and Management, College of Public Health, National Taiwan University, Taipei, Taiwan, [7] Evidence-Based Medicine Center, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, China

OPEN ACCESS

Edited by:
Ugo Cioffi,
University of Milan, Italy

Reviewed by:
Ryoung Choi,
Dongshin University, South Korea
Yanfang Su,
University of Washington,
United States

***Correspondence:**
Tao-Hsin Tung
ch2876@yeah.net
Ching-Wen Chien
ihhca@sz.tsinghua.edu.tw

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: 16 October 2020
Accepted: 18 June 2021
Published: 22 July 2021

Citation:
Liu W-Y, Hsu C-H, Liu T-J, Chen P-E, Zheng B, Chien C-W and Tung T-H (2021) Systematic Review of the Effect of a Zero-Markup Policy for Essential Drugs on Healthcare Costs and Utilization in China, 2015–2021. Front. Med. 8:618046. doi: 10.3389/fmed.2021.618046

Objective: This systematic review aimed to discuss the effects of a zero-markup policy for essential drugs (ZPED) on healthcare costs and utilization in China in the years 2015–2021.

Methods: We searched the PubMed, Embase, Scopus, and CINAHL databases for all associated studies carried out from January 1, 2015, to May 31, 2021, without any limitations regarding the language the studies were written in. To prevent selection bias, gray documents were tackled by other means. The methodological approaches were assessed by applying the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the Newcastle-Ottawa Scale (NOS) collaboration tool.

Results: Forty studies were selected at first and then 15 studies that met the inclusion criterion. Most of the studies showed a considerable decrease in total medical spending and drug spending in both outpatient and inpatient services. After the implementation of ZPED, studies showed that the medical services increased and total hospital income sustained, despite a decrease in drug revenue. Minimal or no government subsidy is required from a financial perspective.

Conclusions: Although, the government could implement ZEPD with lower medical cost and drug cost to patients, and sustained income for health facilities, we have limited understanding of whether the increase in medical services was induced by the provider or was a response to unmet needs in the population. Further, studies using rigorous and advanced methods to study health policy, patient behaviors, provider behaviors, and government decisions are warranted.

Keywords: zero-markup policy, medical expenditure, systematic review, China, healthcare

INTRODUCTION

The Policy on Drug Markups (PoDM) implemented by the government in 1954 allowed medical institutions in China to increase drug prices by a maximum of 15% (1). From the 1980's, it became evident that patients were struggling with obtaining medical care due to the insufficient fiscal allowance and increases in drug costs (2). The Chinese government launched a campaign of clinical health modernization to mitigate these problems and implemented the zero-markup policy for essential drugs (ZPED) in 2009. The ZPED mandated that the markup from medication bills could no longer be retained and that 10% of the original 15% markup under the PoDM would be substituted by fiscal allowance. The answer for the remainder of the discharge was believed to be a preposterous method, and it was thought that diagnostic expenses would be increased to obtain 80% of the former markup. The others were the main parts of the hospitals themselves (3).

Even though the implementation of ZPED was essential to restrict the increase in drug prices, it could not prevent patients from experiencing financial difficulties, which are now caused by paying for other treatments rather than expensive drugs. For instance, when the policy was initially implemented (2009–2011), the pharmaceutical fees of patients per visit were decreased, particularly in rural areas and counties. An evidence-based study showed that the substitution effect of medical consumables offsets the decrease in total expenditure in the long term (4). Hospitals that rely on earnings by drugs to cover their costs could be mitigated by ZPED (5). From 2011 to 2015, this strategy was no longer a pilot scheme and came to be used in all county-level hospitals in China. However, medical expenses were still rising, in spite of efforts to control the increase (2). This meant that hospitals were able to increase their spending on other medical areas covered by ZPED, such as discharge diagnosis fees, nursing payments, surgery expenses, and treatment fees (2).

Most previous studies on this topic focused on the enduring effects of ZPED on expenses of patients per visit, in particular, drug costs, but neglected the overall cost of the process of therapy. Therefore, this systematic review aimed to discuss the effects of ZPED on healthcare costs and utilization in China from 2015 to 2021.

MATERIALS AND METHODS

Data Sources and Selection

We searched the PubMed, Embase, Scopus, and CINAHL databases for all associated studies carried out from January 1, 2015, to May 31, 2021, without any limitations regarding the language the studies were written in. To prevent selection bias, gray documents—for example, OpenGrey and Open Access Theses and Dissertations—were tackled by other means. We performed a search of the aforementioned electronic databases, applying keywords included in the title and/or abstract as follows: ("pharmaceutical*" OR "drug*" OR "medicine*") AND [("zero") AND ("markup" OR "mark-up") AND ("China")]. The selection sheets of the associated studies were evaluated manually to identify comparable works (**Table 1**). Due to the fact that no study patients were enrolled, as we only used published studies, it was not necessary for us to obtain approval from the institutional review board (IRB) for this systematic review. Two reviewers evaluated regular studies that assessed the effects of ZEPD on the annual medical expenditure per subject and the expense of courses of therapy. Disagreements were resolved *via* conversation with a well-trained third reviewer. The studies we selected were original articles rather than letters to the editor, editorials, commentaries, or congress documents. The results of these included investigations should be related to the financial indicators of medical institutions. For instance, we aimed to find data on the total expense per inpatient or outpatient visit, the costs of drugs per visit, the number of visits a patient required, etc.,

TABLE 1 | Search strategy in PubMed up until 31th May 2021 (similar search run in other databases).

1	"Pharmaceutical*" [Title/Abstract]
2	"Drug*" [Title/Abstract]
3	"Medicine*" [Title/Abstract]
4	1 OR 2 OR 3
5	"Zero" [Title/Abstract]
6	"Markup" [Title/Abstract]
7	"Mark-up" [Title/Abstract]
8	6 OR 7
9	"China" [Title/Abstract]
10	4 AND 5 AND 8 AND 9

Data Extraction and Quality Assessment

The Newcastle-Ottawa Scale (NOS) was used for quality assessment (6). This is an approach that has been proven to be effective for appraising methodological quality in non-randomized controlled trials. The two reviewers also summarized the relevant features of the selected studies using a standardized data collection form. **Table 2** indicates the results of the quality ratings. Each asterisk means one star, and the total scale of NOS is the summation of the stars (nine is the maximum), which are allocated for selection (four stars), comparability (two stars), and outcome (three stars).

Data Synthesis

Four extensive outcomes were considered: (1) medical cost; (2) drug cost; (3) healthcare utilization; and (4) others (facility revenue, drug revenue, and government subsidy). The baseline and intervention for the outcome variables were evaluated.

RESULTS

Characteristics of the Included Studies

As shown in **Figure 1**, our investigation started with 40 records after ruling out repeats. We discarded 25 records that could not fulfill the selection criteria. Fifteen studies (including

TABLE 2 | Quality assessment of included studies using the Newcastle-Ottawa Scale (NOS).

Source	Selection				Comparability	Exposure			Total NOS score
	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	
Zhou et al. 2015 (5)	★	★	★	★	★★	★	★		8
Zhou et al. 2015 (7)	★	★	★	★	★★	★	★		8
Tian et al. 2016 (8)			★			★			2
Wei et al. 2017 (9)	★	★	★		★★	★			6
Yang et al. 2017 (2)	★		★			★			3
Fu et al. 2018 (10)	★		★		★★	★			5
He et al. 2018 (3)	★		★			★			3
Tang et al. 2018 (1)			★			★			2
Yin et al. 2018 (11)	★		★			★			3
Mao et al. 2019 (12)	★		★			★			3
Shi et al. 2019 (13)	★		★		★★	★			5
Zeng et al. 2019 (4)	★	★	★	★	★	★			6
Jiang et al. 2020 (14)	★		★		★★	★	★		6
Li et al., 2021 (15)	★		★		★★	★			5
Du et al., 2021 (16)	★		★		★★	★			5

Selection
(1) Representativeness of the exposed cohort.
(2) Selection of the non-exposed cohort.
(3) Ascertainment of exposure.
(4) Demonstration that outcome of interest was not present at start of study.
Comparability
(1) Comparability of cohorts on the basis of the design or analysis.
Exposure
(1) Assessment of outcome.
(2) Was follow-up long enough for outcomes to occur.
(3) Non-response rate.

retrospective cohort studies, time series studies, and quasi-experimental studies) were eventually included (1–5, 7–16). **Table 3** outlines the features of the selected studies. All of the studies included were from China. Not all studies were classified as having more than seven stars based on the NOSs, and some were believed to be of lower quality. **Table 3** also shows the various estimated results.

Medical Cost

The total expenses decreased by 3.12 and 65.6 U.S. dollars for outpatients and inpatients, respectively, per visit according to the quasi-experimental study design. The expense of each visit was predicted to show a decrease of 11% for both outpatient and inpatient medical services (5). One retrospective follow-up study showed that there was an increase in the annual number of patient visits (8). Another study revealed that there was a significant reduction in the expense of hospitalization per month (2). In addition, an increase in expenditure on clinical services was noted, while no alteration in overall healthcare costs was observed (10). The medication reform did not manage to longitudinally decrease the overall health expenditure of the patients (3). The great growth in the investment of the government in hospitals offsets the revenue from the sales of Western medicine, meaning that the zero-markup drug policy (ZMDP) can be considered a success; however, whether the investment of the government can last for a long duration is a factor that needs to be considered (13). A positive relationship was also found between the compensation for the situation of services and the contents of the annual income of clinical services (1). However, no meaningful modification of the average medical expenditure per treatment was noticed (12).

Drug Cost

One selected study revealed that not only the costs of drugs per patient visit declined by 4.47 U.S. dollars (outpatient services) or 45.75 U.S. dollars (inpatient services) but also the proportion of money spent on drugs out of the overall medical expenditure per patient visit decreased by 11.73% for outpatient visits and 3.92% for inpatient visits (5). Another study showed that the implementation of ZPED was associated with a reduction in the use of antibiotics (11). ZPED succeeded in removing the profits from Western medicines from the revenues of county hospitals without significantly disrupting their normal operation (17). In addition, the change in policy caused a reduction in drug expenditure (10). The medication reform was only able to decrease drug expenses in the short term (3). One study showed that the number of medicines prescribed per patient and the use of antibiotics, intramuscular injections, and intravenous injections reduced, while the use of hormones increased (12). The total drug expenditure decreased by 14.4%, while the drug expenditure of outpatients with chronic diseases in tertiary

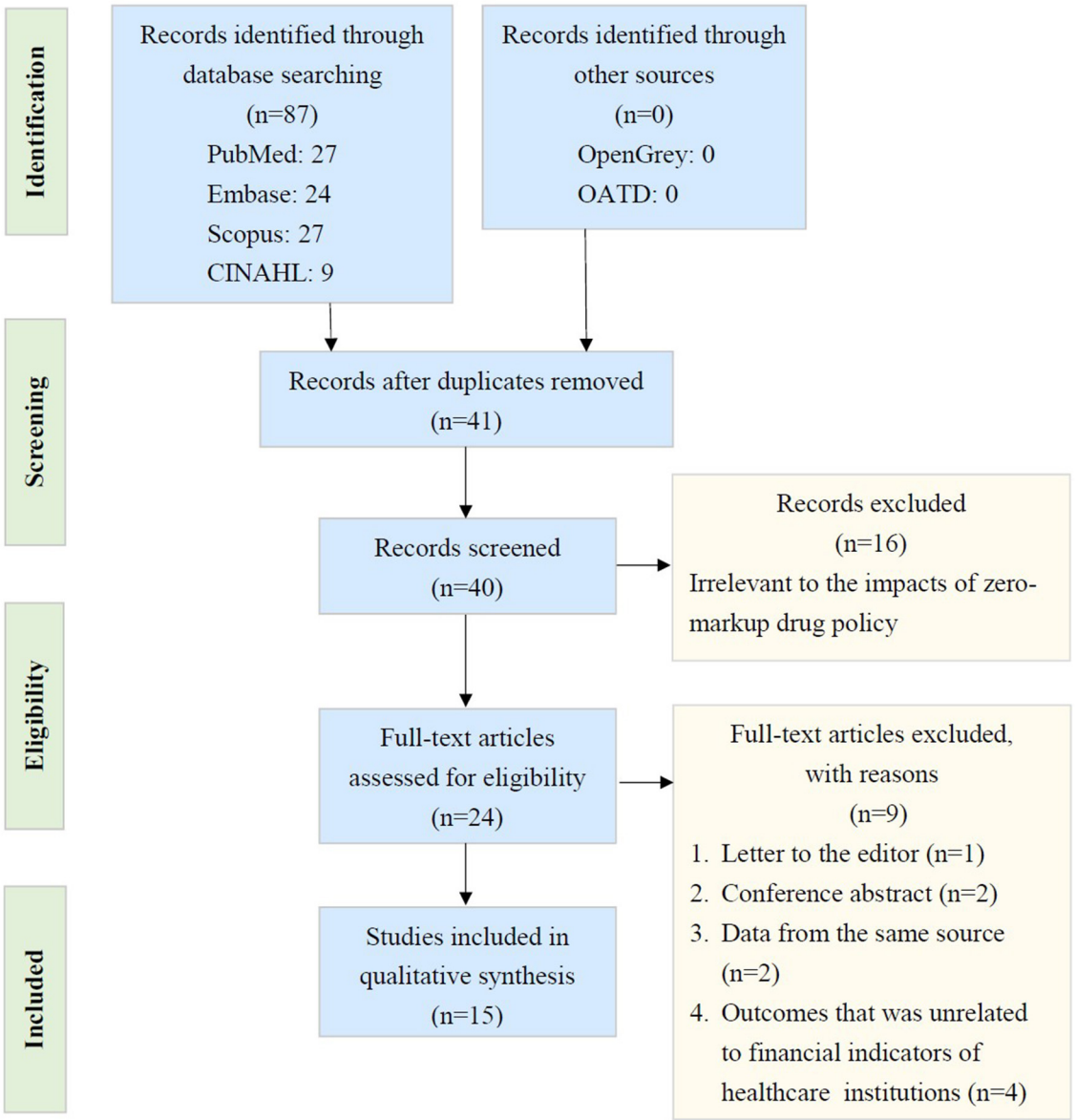


FIGURE 1 | PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) flow diagram. (CINAHL: Cumulative Index to Nursing and Allied Health Literature; QATD: Open Access Theses and Dissertations).

hospitals was reduced due to the implementation of the ZPED (14, 16).

Healthcare Utilization

With regard to the reduction in the ratio of medicine to healthcare charges in services to patients, both outpatient and inpatient services showed increases in the annual number of patient visits (8). ZPED may be associated with decreases in antibiotic prescriptions and intravenous infusions for outpatient visits (9).

The provision of outpatient services increased in the treatment group by 28.55%, while the provision of inpatient services increased by 16.17%. The provision of outpatient services was similar to that of the control group, and the provision of inpatient services only increased by 1.31%. Following the implementation of a zero-markup policy, the provision of outpatient and inpatient services in the treatment group increased by 9,697 and 398 visits, respectively [27].

Others

The great reduction in expenditure on drugs caused more physicians to induce patients/healthcare service needs. The separation of hospital revenue from drug sales (SHRDS) policy is not an effective means of controlling healthcare expenditure

(16). In addition, the increase in inpatient physician workdays decreased the mortality rate of inpatients. Workloads and inflation-adjusted per visit medical care charges of physicians increased in the outpatient services (8).

DISCUSSION

Clinical Implications of a Zero-Markup Policy

Few systematic reviews have been conducted to explore the effects of ZPED for essential drugs on healthcare costs and utilization in China. Based on the included studies, we showed that there were considerable decreases in both the drug cost and the total expenditure per patient visit. Medical services also revealed increasing levels of inpatient visits annually. In China, the economic benefit of prescribing medicines was most regularly referred to as a reason potentially causing the illogical use of medicines in a previous systematic review (18). However, the improvement of the reasonable use of medicines still has many unpredictable deficiencies. ZPED covers national-level medicines, while lower-level governments develop the list according to local requirements. The main effect of the medical policy is to avoid hospitals not only being seldomly sufficiently reimbursed but also from having to deal with financial embarrassments due to the complicated medical circumstances involved (19).

ZPED has been proven to decrease the medical costs of patients, leading to the reform of the inpatient and outpatient structures (20). Hospitalization does not depend on the cost of outpatient care (20); however, both outpatient medical care and inpatient medical care are determined by medical factors, such as the health status of patients. Outpatient care is a short-term medical service that does not require an overnight stay in a hospital or a medical facility. Meanwhile, inpatient care involves continuous processes between patients and medical staff, in which the perception of the interaction of inpatients between the environment and service process is valued (21). Further studies should be conducted to explore the improvement of the medical care service system, the public health system, and the drug supply system. In addition, the results from the behavioral economic studies have indicated that people often make decisions according to not only absolute but also relative changes in price (5, 22). This manner is at the polytechnic of relative thinking theory, which shows that people are affected more by relative changes than absolute changes in a given initial stage (5).

In addition, the results of the selected studies concern the cost containment policy of healthcare, which changes prices for drugs and medical services. Changing prices is a broadly applied policy instrument in most Western countries (10). Rules and regulations of price alone would not yield a successful decrease in expenses because healthcare providers could avoid regulations by message merit. It is a fact that China experiences pressure regarding healthcare. Increasing the provisions of other medical services by increasing prices may compensate for the loss of revenue in most public hospitals in China (10).

Clinical Practice of Zero-Markup Policy

This systemic review showed that there has been a major improvement in drug costs and the number of patient visits in spite of there being some heterogeneity in terms of total expenses. According to these findings, medical teams should set up a customized agreement to manage or decrease medical expenses using ZPED due to the growth or reduction in healthcare costs primarily being dependent on hospital practices (23). For example, both outpatient and inpatient payments were reduced in health centers in towns but not in those of county status or above (8). The monitoring system should include a longitudinal evaluation with advanced and alternative approaches for the improvement of expenditure.

The zero-markup policy for essential drugs plays an important role in decreasing the cost of drugs for chronic non-communicable diseases, such as type 2 diabetes, hypertension, metabolic syndrome, coronary heart disease, and cancer (21). In China, urban employees with basic medical insurance cannot afford outpatient expenses. Outpatient expenses could be deducted from the amount in the personal account, such as a partial return of the paid amount (16). A previous study also demonstrated that the abolishment of the drug markup fee at public hospitals has more predictable benefits for patients if they are urban employees with basic medical insurance (24). In addition, due to the limited reimbursement rate for patients who are urban residents with basic medical insurance, such patients may be more inclined to purchase drugs independently outside of the hospital, meaning that the execution of the ZPED in tertiary hospitals is less effective on medical costs. In contrast, urban employees with basic medical insurance may have a relatively higher awareness of reasonable drug use or may be more willing to make the decision to purchase drugs at hospitals (16).

Methodological Considerations

Several methodological perspectives should be addressed when applying the findings of this systematic review. First, the relatively small number of selected studies limits the power of our conclusions. Second, the included studies vary in terms of methodological quality, which may have introduced some risk of bias. Third, from a statistical viewpoint, it is worth using either qualitatively or statistically presented aggregated evidence; however, it would be difficult to conduct a meta-analysis of the selected studies because the included studies do not provide consistent information. Finally, the findings might not be able to be generalized to other medical institutions, as the studies we included were conducted in only a few areas of China. The external validity of our outcomes should also be further explored.

Conclusions

Although the government could implement ZEPD with the lower medical cost and drug cost to patients, and sustained income for health facilities, we have limited understanding of whether the increase in medical services was induced by the provider or was a response to unmet needs in the population. Further, studies using rigorous and advanced methods to study health policy, patient behaviors, provider behaviors, and government decisions are warranted.

TABLE 3 | Characteristics of included studies in China.

Author	Publication year	Study design	Area	Hospital accreditation	Outcomes	Statistical methods	Conclusions
Zhou et al. (5)	2015	Quasi-experimental study	Ningshan County, Zhenping County, and Shaanxi Province	Secondary	• Total expense per visit (inpatient/outpatient service). • Drug expense per visit (inpatient/outpatient service).	• Hospital-data difference-in-differences. • Individual-data regressions.	• The absolute monetary reduction of the per-visit inpatient expense is 20 times of that in outpatient care. • The relative reductions are the same for outpatient and inpatient visits. • The incentive to utilize outpatient or inpatient care attributed to Zero-markup Policy for Essential Drugs is equivalent, regardless of the price difference in absolute terms.
Zhou et al. (7)	2015	Quasi-experimental study	Ningshan County Hospital in Ankang city, Shaanxi province		• The effects of zero-markup on medical expense per visit. • The effects of zero-markup policy on medical service provision. • The effects of zero-markup policy on the health care revenue for hospitals. • Estimates of government subsidy.	• A difference-in-difference model to measure the difference in several indicators between two hospitals.	With minimal or no subsidy, the government can catalyze the zero-markup policy and potentially generate positive outcomes for county hospitals.
Tian et al. (8)	2016	Descriptive study (retrospective longitudinal study)	Beijing	Tertiary	• Annual patient-visits. • Ratios of medicine-to-healthcare-charges (RMOH). • Physician work-days (inpatient/outpatient service). • Physician-workload (inpatient/outpatient service). • Inflation-adjusted per-visit healthcare charges (inpatient/outpatient service). • Mortality-rate (inpatient/outpatient service).	• Rank-sum tests. • Join-point regression analyses.	Implementation of Universal Zero-Markup Drug Policy: • Increase annual patient-visits • Reduce RMOH • Have different impacts on outpatient and inpatient services

(Continued)

TABLE 3 | Continued

Author	Publication year	Study design	Area	Hospital accreditation	Outcomes	Statistical methods	Conclusions
Wei et al. (9)	2017	Natural experiment	Guangxi	N/A	• Antibiotic prescribing rate (outpatients with a primary diagnosis of upper respiratory tract infection).	• Difference-in-difference analyses.	The national essential medicines scheme and zero-mark-up policy may be associated with reductions in outpatient antibiotic prescribing and intravenous infusions.
Yang et al. (2)	2017	Time series study	Shaanxi Province	Primary	• Monthly average hospitalization expenditure (AHE). • Monthly average hospitalization expenditure after reimbursement (AHER).	• Segmented regression analysis of interrupted time series data.	A statistically significant absolute decrease in the level or trend of monthly AHE and AHER was detected after the introduction of the zero-markup drug policy in western China. However, hospitalization expenditure and hospitalization expenditure after reimbursement were still increasing. More effective policies are needed to prevent these costs from continuing to rise.
Fu et al. (10)	2018	Penal study	1,880 counties	N/A	**Outpatient care** • Total expenditures per visit. • Drug expenditures per visit. • Expenditures for diagnostic tests/medical consumables per visit. • Expenditures for medical services per visit. • Outpatient visits. **Inpatient care** • Total expenditures per admission. • Drug expenditures per admission. • Expenditures for diagnostic tests/medical consumables per admission. • Expenditures for medical services per admission. • Inpatient admissions. • Average length of inpatient stay.	• Pre-trend test based on linear regressions.	The policy change led to a reduction in drug expenditures, a rise in expenditures for medical services, and no measurable changes in total health expenditures. However, this study also found an increase in expenditures for diagnostic tests/medical consumables at hospitals that had a greater reliance on drug revenues before the reform, which is unintended by policymakers. Further, these results were more likely to be driven by the supply side, suggesting that hospitals offset the reductions in drug revenues by increasing the provision of services and products with higher price-cost margins.

(Continued)

TABLE 3 | Continued

Author	Publication year	Study design	Area	Hospital accreditation	Outcomes	Statistical methods	Conclusions
He et al. (3)	2018	Time series study	Sanming City, Fujian Province	Secondary ($n = 4$) and tertiary ($n = 21$)	• Outpatient drug expenditure. • Outpatient total health expenditure. • Inpatient drug expenditure. • Inpatient total health expenditure.	• Interrupted time series analysis with three segments divided by two intervention points.	Although, the pharmaceutical reform could control or reduced drug expenditure and total health expenditure in short term, expenditures gradually resumed growing again and reached or even exceeded their baseline levels of pre-reform period, indicating the effect became weakened or even faded out in long term.
Tang et al. (1)	2018	N/A	Nanjing City, Jiangsu Province	Secondary and tertiary	• The markup ratio of drug sales. • The growth rate of medical service revenue.	• Simple linear interrupted time series regressions.	Nanjing's pricing and compensation reform has basically achieved the policy targets of eliminating the drug markups, promoting the growth of medical services revenue, and adjusting the structure of medical revenue. However, the growth rate of service revenue of hospitals varied significantly from one another.
Yin et al. (11)	2018	Time series study	Shandong Province	Secondary tertiary, and urban/rural primary healthcare centers (PHCs)	• Total annual antibiotic expenditure. • Antibiotic expenditure per person per year.	• Descriptive statistics.	The overall antibiotic expenditure increased over time in Shandong, China. However, the increase rate of expenditure began to decline in 2016, possibly related to the implementation of antibiotic stewardship initiatives.

(Continued)

TABLE 3 | Continued

Author	Publication year	Study design	Area	Hospital accreditation	Outcomes	Statistical methods	Conclusions
Mao et al. (12)	2019	Penal study	Hangzhou City, Zhejiang Province	Primary ($n = 6$), secondary ($n = 2$), and tertiary ($n = 9$)	• Average number of medicines. • Average number of antibiotics. • Average expenditure per prescription.	• *T*-test. • Pearson Chi-square test or Fisher exact test.	The average number of medicines per prescription, use of antibiotics, intramuscular (IM) injections and intravenous (IV) injections decreased while the use of hormones increased. No significant change of the average medicine expenditure per prescription was observed. The problems of poly-pharmacy, overuse of antibiotics, intramuscular (IM) injections and intravenous (IV) injections and hormones still existed, however mitigated after the implementation of The National Essential Medicine Policy and the Zero Mark-up Policy.
Shi et al. (13)	2019	Penal study	All TCM county hospitals		• Revenue from government subsidies. • Share of revenue from drug sales (divide the amount of revenue from TCM and Western medicine by total revenue except government subsidies). • Revenue from Chinese medicine sales. • Revenue from western medicine sales. • Revenue from medical care services. • Gross revenue. • The number of annual outpatient visits and the number of inpatient admissions.	• Difference-in-difference.	• ZMDP achieved its stated goal through reducing the share of revenue from drug sales without disrupting the availability of healthcare services at TCM county hospitals no matter in the short term or long term. • The success of ZMDP was mainly due to the huge growth in the government's financial investment in TCM hospitals, which offset western medicine sales revenue, while maintaining current hospital service levels. However, whether government financial investment can continue such long-term growth remains an open question.
Zeng et al. (4)	2019		Beijing City	Urban employee basic medical insurance	• The total expenditure and other expenditure components of the pilot hospitals.	• 1:1 propensity score-matched analysis (Propensity scores were calculated by logistic regression).	After the zero markup drug policy, expenditure on drugs revealed a continuous decline. However, the decline in total expenditure was weakened by the substitution effect of medical consumables in the long term.

(Continued)

TABLE 3 | Continued

Author	Publication year	Study design	Area	Hospital accreditation	Outcomes	Statistical methods	Conclusions
Jiang et al. (14)	2020	Penal study	Shandong Province	Secondary and tertiary	• Revenue from medicine sales. • The share of revenue from medicine sales. • Evenue from medical care services. • Government subsidies. • Revenue and expenditure surplus. • Gross revenue. • The number of annual outpatient and inpatient visits.	• Difference-in-difference analyses.	• The ZMDP achieved its some initial goals of removing the profits from western medicines in county hospitals' revenue without disrupting the normal operation, and had different impacts between county general andTCM hospitals. • Meanwhile, some unintended consequences were also recognized through the analysis, such as the decline of the utilization of the TCM.
Li et al. (15)	2021	Penal study	Chengdu City	Urban employee basic medical insurance	• A series of expenditure variables (actual reimbursement expenditure, reimbursement ratio, total healthcare expenditures, drug expenditure, examinations expenditure, material expenditure, nursing expenditure, etc.,).	• Difference-in-difference analyses.	• After implementing the SHRDS policy, the significant reduction in drug expenditure led to more physicians inducing patients' healthcare service needs, and the increased social healthcare burden was partially transferred to the patients' personal economic burden through the decline in the reimbursement ratio. • The SHRDS policy is not an effective way to control healthcare expenditure.
Du et al. (16)	2021	Interrupted time series study	Chongqing City	Tertiary	• Average drug cost 11 per month for NCDs' outpatients analyzed overall.	• Interrupted time series analysis.	• The ITS analysis is an effective method of health policy evaluation. • The implementation of the ZMDP can reduce the drug cost for chronic disease outpatients in the tertiary hospital and their economic burden. • Follow-up policies still require targeted price adjustments in the health service system to adjust the drug cost effectively.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

W-YL, C-HH, T-JL, P-EC, T-HT, and C-WC conducted the study and drafted the manuscript. W-YL, C-HH, BZ, and T-JL participated in the design of the study and performed data synthesis. P-EC, BZ, T-HT, and C-WC conceived the study and participated in its design and coordination. All of the authors read and approved the final manuscript.

FUNDING

This study was supported by the Taizhou Hospital of Zhejiang Province, Linhai, China (No 27).

REFERENCES

1. Tang W, Xie J, Lu Y, Liu Q, Mallone D, Ma A. Effects on the medical revenue of comprehensive pricing reform in Chinese urban public hospitals after removing drug markups: case of Nanjing. *J Med Econ.* (2018) 21:326–39. doi: 10.1080/13696998.2017.1405817
2. Yang C, Shen Q, Cai W, Zhu W, Li Z, Wu L, et al. Impact of the zero-markup drug policy on hospitalisation expenditure in western rural China: an interrupted time series analysis. *Trop Med Int Health.* (2017) 22:180–6. doi: 10.1111/tmi.12817
3. He Y, Dou G, Huang Q, Zhang X, Ye Y, Qian M, et al. Does the leading pharmaceutical reform in China really solve the issue of overly expensive healthcare services? Evidence from an empirical study. *PLoS ONE.* (2018) 16 13:e0190320. doi: 10.1371/journal.pone.0190320
4. Zeng J, Chen X, Fu H, Lu M, Jian W. Short-term and long-term unintended impacts of a pilot reform on Beijing's zero markup drug policy: a propensity score-matched study. *BMC Health Serv Res.* (2019) 19:916. doi: 10.1186/s12913-019-4764-z
5. Zhou Z, Su Y, Campbell B, Zhou Z, Gao J, Yu Q, et al. The financial impact of the 'zero-markup policy for essential drugs' on patients in county hospitals in western rural China. *PLoS ONE.* (2015) 10:e0121630. doi: 10.1371/journal.pone.0121630
6. Lo CK, Mertz D, Loeb M. Newcastle-Ottawa scale: comparing reviewers' to authors' assessments. *BMC Med Res Methodol.* (2014) 14:45. doi: 10.1186/1471-2288-14-45
7. Zhou Z, Su Y, Campbell B, Zhou Z, Gao J, Yu Q, et al. The impact of China's Zero-Markup Drug Policy on county hospital revenue and government subsidy levels. *J Asian Public Policy*. (2015) 8:102–16. doi: 10.1080/17516234.2015.1005561
8. Tian W, Yuan J, Yang D, Zhang L. Descriptive analysis on the impacts of universal zero-markup drug policy on a Chinese Urban Tertiary Hospital. *PLoS ONE.* (2016) 11:e0162795. doi: 10.1371/journal.pone.0162795
9. Wei X, Yin J, Walley JD, Zhang Z, Hicks JP, Zhou Y, et al. Impact of China's essential medicines scheme and zero-mark-up policy on antibiotic prescriptions in county hospitals: a mixed methods study. *Trop Med Int Health.* (2017) 22:1166–74. doi: 10.1111/tmi.12922
10. Fu H, Li L, Yip W. Intended and unintended impacts of price changes for drugs and medical services: evidence from China. *Soc Sci Med.* (2018) 211:114–22. doi: 10.1016/j.socscimed.2018.06.007
11. Yin J, Li Q, Sun Q. Antibiotic consumption in Shandong Province, China: an analysis of provincial pharmaceutical centralized bidding procurement data at public healthcare institutions, 2012-16. *J Antimicrob Chemother.* (2018) 73:814–20. doi: 10.1093/jac/dkx469
12. Mao W, Huang Y, Chen W. An analysis on rational use and affordability of medicine after the implementation of National Essential Medicines Policy and Zero Mark-up Policy in Hangzhou, China. *PLoS ONE.* (2019) 14:e0213638. doi: 10.1371/journal.pone.0213638
13. Shi X, Zhu D, Man X, Wang W, Zhu K, Nicholas S, et al. "The biggest reform to China's health system": did the zero-markup drug policy achieve its goal at traditional Chinese medicines county hospitals? *Health Policy Plan.* (2019) 34:483–91. doi: 10.1093/heapol/czz053
14. Jiang X, He P, Zhu D, Shi X, Meng Q. Different impacts of the zero-markup drug policy on county general and traditional Chinese medicine hospitals: evidence from Shandong province, China. *Int J Equity Health.* (2020) 19:219. doi: 10.1186/s12939-020-01326-w
15. Li L, Yu Q. Does the separating of hospital revenue from drug sales reduce the burden on patients? Evidence from China. *Int J Equity Health.* (2021) 20:12. doi: 10.1186/s12939-020-01363-5
16. Du J, Xie J, Qian Y, Wu M, Huang W, Yin J, et al. Impact of China's zero mark-up drug policy on drug cost of NCDs' outpatients: an interrupted time series analysis. *BMC Health Serv Res.* (2021) 21:404. doi: 10.1186/s12913-021-06414-3
17. Lv Y, Xue C, Ge Y, Ye F, Liu X, Liu Y, et al. Analysis of factors influencing inpatient and outpatient satisfaction with the Chinese military health service. *PLoS ONE.* (2016) 11:e0151234. doi: 10.1371/journal.pone.0151234
18. Mao W, Vu H, Xie Z, Chen W, Tang S. Systematic review on irrational use of medicines in China and Vietnam. *PLoS ONE.* (2015) 10:e0117710. doi: 10.1371/journal.pone.0117710
19. Mao W, Chen W. The Zero Mark-up Policy for essential medicines at primary level facilities, China case study. In: Xu K, editor. *Improving Health System Efficiency, Health Systems Governance & Financing*. Geneva: World Health Organization (2015).
20. Zhou Z, Su Y, Gao J, Xu L, Zhang Y. New estimates of elasticity of demand for healthcare in rural China. *Health Policy.* (2011) 103:255–65. doi: 10.1016/j.healthpol.2011.09.005
21. Arabyat RM, Nusair MB, Al-Azzam SI, Alzoubi KH. Analysis of prevalence, risk factors, and potential costs of unnecessary drug therapy in patients with chronic diseases at the outpatient setting. *Expert Rev Pharmacoecon Outcomes Res.* (2020) 20:125–32. doi: 10.1080/14737167.2019.1612243
22. Azar OH. *Relative Thinking Theory. Microeconomics 0504002.* Munich: University Library of Munich (2005). doi: 10.2139/ssrn.701581
23. Wang J, Zhen J, Fan W, Zhang Q, Zhang G. Investigation and analysis of the effects of national essential drug system on medical expenses in different levels of medical Institutions. *Chin Pharm.* (2013) 23:2982–4. doi: 10.6039/j.issn.1001-0408.2012.32.03
24. Wang J. *Research on the Influence of Zero-Profit Drugs Sales Policy to the Patient Personal Burden–in the TianJin BinHai New Area as an Example.* Master: Hebei United University (2015).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

REVIEW
published: 03 August 2021
doi: 10.3389/fpubh.2021.675171

The Building Blocks of Implementation Frameworks and Models in Primary Care: A Narrative Review

Ine Huybrechts[1,2]*, *Anja Declercq*[3], *Emily Verté*[1,2], *Peter Raeymaeckers*[1†] and *Sibyl Anthierens*[1†] *On behalf of the Primary Care Academy*

[1] Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium, [2] Department of Family Medicine and Chronic Care, Free University of Brussels, Brussels, Belgium, [3] LUCAS – Centre for Care Research and Consultancy & CESO – Centre for Sociological Research, Catholic University of Leuven, Leuven, Belgium

OPEN ACCESS

Edited by:
Michele Mario Ciulla, University of Milan, Italy

Reviewed by:
Innocent Besigye, Makerere University, Uganda
Christopher M. Westgard, University of North Carolina at Chapel Hill, United States

***Correspondence:**
Ine Huybrechts
ine.huybrechts@uantwerpen.be

[†] *These authors have contributed equally to this work and share last and senior authorship*

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Public Health

Received: *02 March 2021*
Accepted: *08 July 2021*
Published: *03 August 2021*

Citation:
Huybrechts I, Declercq A, Verté E, Raeymaeckers P and Anthierens S (2021) The Building Blocks of Implementation Frameworks and Models in Primary Care: A Narrative Review. Front. Public Health 9:675171. doi: 10.3389/fpubh.2021.675171

Background: Our aim is to identify the core building blocks of existing implementation frameworks and models, which can be used as a basis to further develop a framework for the implementation of complex interventions within primary care practices. Within the field of implementation science, various frameworks, and models exist to support the uptake of research findings and evidence-based practices. However, these frameworks and models often are not sufficiently actionable or targeted for use by intervention designers. The objective of this research is to map the similarities and differences of various frameworks and models, in order to find key constructs that form the foundation of an implementation framework or model that is to be developed.

Methods: A narrative review was conducted, searching for papers that describe a framework or model for implementation by means of various search terms, and a snowball approach. The core phases, components, or other elements of each framework or model are extracted and listed. We analyze the similarities and differences between the frameworks and models and elaborate on their core building blocks. These core building blocks form the basis of an overarching model that we will develop based upon this review and put into practice.

Results: A total of 28 implementation frameworks and models are included in our analysis. Throughout 15 process models, a total of 67 phases, steps or requirements are extracted and throughout 17 determinant frameworks a total of 90 components, constructs, or elements are extracted and listed into an Excel file. They are bundled and categorized using NVivo 12© and synthesized into three core phases and three core components of an implementation process as common elements of most implementation frameworks or models. The core phases are a development phase, a translation phase, and a sustainment phase. The core components are the intended change, the context, and implementation strategies.

Discussion: We have identified the core building blocks of an implementation framework or model, which can be synthesized in three core phases and three core components. These will be the foundation for further research that aims to develop a new model that will guide and support intervention designers to develop and implement complex interventions, while taking account contextual factors.

Keywords: primary care interventions, implementation, implementation frameworks, implementation models, implementation process, implementation science

INTRODUCTION

Initiating and sustaining change within primary care is challenging (1). Most change that is introduced in primary care takes the form of a complex intervention, meaning that it involves concepts that are rather difficult to measure and its components are often interconnected (2, 3). Nowadays, there are increased efforts to shift toward a more patient-centered approach (1), as this proves to improve disease outcomes and quality of life (4). However, such a shift highly challenges current primary care practices and there is therefore no consensus on how to best implement it (5). This indicates a gap between scientific evidence and actual practice: an evidence-to-practice gap (3). This can also be referred to as *"the black box of knowledge translation"* (6), meaning that much uncertainty exists about understanding why evidence-based practices do not find their way into real world settings and investigating how such complexities can best be approached.

Concrete initiatives and strategies for implementation often do not match with targeted problems (7). In the end, too much is expected from practitioners' ability and goodwill to consult, interpret, and adapt their practices in line with best evidence of research findings (8). The World Health Report 2008[1] stated that "*providing a sense of direction to health systems requires a set of specific and context-sensitive reforms that respond to the health challenges of today and prepare for those of tomorrow*." It is thus key to carefully define specific interventions that aim to transform current practices, while at the same time tailoring them to local circumstances (9, 10). To do this, it is essential to gain insight in the process of implementation as well as in potential barriers and facilitators that might hinder or support the implementation process. This is studied in the field of implementation science, which is "*the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice (11).*" The goal of implementation science is to close the gap between evidence-based practices and the extent to which research findings are integrated into real world settings and practices (3, 12).

Within the field of implementation science, many theories, models and frameworks have been created by various disciplines. Moreover, there is a variety of guidelines and tools aimed at facilitating the integration of knowledge of implementation science into either the development or the initiation of interventions and how to document this process. Examples are the ImpRes tool (13), NCEC Implementation Guide & Toolkit for National Clinical Guidelines (14), RNAO Toolkit: Implementation of Best Practice Guidelines (2nd ed). (15), STaRi Standards for Reporting Implementation Studies (16) and Implementation Research Logic Model (17). However, the landscape of implementation science is rather difficult to navigate, as there is a lack of guidance for selecting theories, frameworks, models, or tools that best fit specific implementation objectives (18). A first step toward a better comprehension of such guidance on implementation efforts and to focus on concepts that are more meaningful to the actors in the field, is to gain better understanding in the common thread throughout the wide variety of models and frameworks that form the basis of such tools.

Current approaches to guide the implementation process are mainly characterized by a single-discipline, medical perspective in which a limited number and types of barriers are taken into account (19). This is insufficient to provide a deeper understanding of implementation success or failure or to increase the chance of success of the implementation (20). Existing frameworks and models tend to incorporate a selection of barriers, but do not allow to give more guidance about their validity or relative importance in specific contexts (20). Moreover, many frameworks and models remain very abstract and fall short in giving concrete guidance for intervention designers on how to navigate the implementation process (21). As many of such frameworks or models remain untested, this again questions their operability (7).

Therefore, an overarching framework is needed that provides both an explanatory approach (3), but also allows to prioritize those variables that are essential to achieve implementation success (22). This means that such a framework should provide a pathway that clarifies the core phases and steps throughout an implementation process and that highlight the core constructs that, within each phase, need to be defined, acted upon, and reflected upon. These phases and constructs should be accessible and meaningful to actors that will conduct implementation efforts. It is key for such framework to transcend disciplines and to bundle insights from different approaches (7).

This research is a first step in the development of a generic framework that incorporates such an approach. We therefore looked into existing theories, models, and frameworks from implementation science and combined insights across various disciplines. The similarities and differences between various frameworks informed us about the main building blocks of such

[1]World Health Organization. (2008). Primary care, now more than ever. https://www.who.int/whr/2008/whr08_en.pdf (accessed May 15, 2020).

frameworks and about how and why they differ. In doing this, we were guided by a rather broad research question: "*What are the main components of implementation frameworks and models in order to structure and guide implementation processes?*" This resulted into the identification of core building blocks that form a common thread throughout implementation models and frameworks. Such synthesis will in future research help to develop an overarching model that puts forward clear and meaningful constructs for intervention designers, and that provides both a pathway as well as an explanatory structure to define, act, and reflect upon each component of a complex intervention.

METHODS

To determine the building blocks of an overarching implementation framework, we conducted a literature review. Various disciplines were represented in the included literature, for which the initial search had been conducted by a multi-disciplinary team of medical researchers, sociologists, social work, and agogic sciences. We opted for a narrative review, which can be defined as "*comprehensive narrative syntheses of previously published information* (23)" and which helps to "*pull many pieces of information together into a readable format* (23)." This reviewing technique is particularly helpful for grasping a broad perspective on a topic; it enables us to transcend a purely medical view on primary care and incorporate other perspectives such as social welfare. Moreover, since the field of implementation science is rather fragmented and consists of a wide range of sources, it requires a wider scoping (24). Instead of focusing on a more rigor methodology to answer a very specific, narrowly-focused research question (24), a narrative review allows for interpretation and critique, aiming to deepen the overall understanding of the subject specifically targeted at our problem (24). This corresponds to our goal to identify and possibly simplify the complexities of implementing an intervention by extracting the core phases and components that are common in most models. According to Green, Johnson and Adams (23), a successful narrative review synthesizes available evidence in relation to a topic and present it in a structured way, conveying a clear message. Our aim is thus to provide an overview of existing implementation frameworks and models and to analyze how they are structured and build.

Our initial search started with articles that were key in identifying other models and frameworks: Nilsen (25) which categorized many frameworks and models and Damschroder et al. (26) which provided a list of references on which the consolidated framework for advancing implementation science was based. Our search continued with consulting the three databases PubMed, Web of Science, and Google Scholar, which are most commonly used in this type of literature. The key words that were used are listed in **Table 1**. Article titles and abstracts were screened for references about a specific framework, model, or theory for implementation, followed by an additional search for theoretical papers on these frameworks, models, or theories. Subsequently, the search terms were adapted and redefined based upon our findings, thus creating an iterative process that ensures covering literature in a comprehensive way (27). Also, a snowball approach was used and additional literature was found in the references of the papers.

TABLE 1 | Overview of the process of searching articles.

1) References in key articles	
Nilsen (25)	Provides a categorization of frameworks and models and gives many examples of each type.
Damschroder et al. (26)	Provides a list of references on which the consolidated framework for advancing implementation science was based.
2) Database search	
Databases (Between 2000 and May 2020)	**List of search terms**
PubMed Web of Science Google Scholar	"Primary care" or "primary care interventions" or "health services" AND "implementation framework" or "implementation model" or "implementation science"
3) Adaptation of search terms based on findings	

Articles were searched for and consulted between October 2019 and May 2020. They were mostly published between the years 2000 and 2020, but we did include some older source material if a model or framework was considered to be relevant (e.g., the paper was often referred to by other relevant articles). All articles were available as full text in English. We looked for articles which primarily consisted of a theoretical elaboration (and/or application) of a specific framework or model. Frameworks and models that were highly targeted toward a single case or strategy were excluded, as they were difficult to generalize for overall primary care settings.

To compare and analyze the frameworks and models, they were listed and classified according to Nilsen's (25) categorization (see: **Table 2**). We built our analysis upon process models and determinant frameworks, as they allowed to extract clear steps, actions, barriers, and facilitators that can be transformed into guidance for intervention designers, which was the main aim of our research. For additional understanding of the component evaluation that came up in several models and frameworks, we also looked into three evaluation frameworks. Several classic theories [e.g., Theory of Diffusion (28)] and implementation theories [e.g., Normalization Process Theory (29)] were initially identified, but were not included in our analysis as their approach and structure did not match with our goal to extract clear building blocks of an implementation process that could be used to reconstruct a generic framework.

To analyze, all relevant frameworks and models were listed in an Excel file, with an overview of how they were constructed. For process models, their main phases (steps, stages) were listed, together with relevant details or components within the process they described. For determinant frameworks, the main components (constructs, elements) were listed, together with any details or further clarification about each of the components

TABLE 2 | Five categories of theories, models and frameworks used in implementation science.

Category	Description
Process models	Specify steps (stages, phases) in the process of translating research into practice, including the implementation and use of research. The aim of process models is to describe and/or guide the process of translating research into practice. An action model is a type of process model that provides practical guidance in the planning and execution of implementation endeavors and/or implementation strategies to facilitate implementation.
Determinant frameworks	Specify types (also known as classes or domains) of determinants and individual determinants, which act as barriers and enablers (independent variables) that influence implementation outcomes (dependent variables). Some frameworks also specify relationships between some types of determinants. The overarching aim is to understand and/or explain influences on implementation outcomes, e.g., predicting outcomes or interpreting outcomes retrospectively
Classic theories	Theories that originate from fields external to implementation science, e.g., psychology, sociology, and organizational theory, which can be applied to provide understanding and/or explanation of aspects of implementation
Implementation theories	Theories that have been developed by implementation researchers (from scratch or by adapting existing theories and concepts) to provide understanding and/or explanation of aspects of implementation
Evaluation frameworks	Specify aspects of implementation that could be evaluated to determine implementation success

Categorization and definitions by Nilsen (25).

described. The first step to analyze was to bundle each of the phases or components that had a similar approach or meaning. This was done by the main researcher and validated by the three senior researchers. An overarching concept was appointed to each group of concepts. Then, NVivo 12© was used to structure the main themes and concepts and to analyze their similarities and differences. The overarching concepts were entered as the main nodes in NVivo 12©, whereby details or explanation about each concept from the different models and frameworks were again coded when we noticed overlap with approaches from different frameworks or models. By structuring the phases and components this way and by analyzing the details that were given for each component, we could synthesize it into core building blocks.

RESULTS

Fifteen process models and 17 determinant frameworks were identified. Four models had characteristics of both a process model as well as a determinant framework: the Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors (22), the Consolidated Framework for Implementation Research (26), The Ottawa Model of Health Care Research (30) and the Generic Implementation Framework (19). The frameworks or models focus on various domains. They were either developed specifically to apply within a certain research domain or development was based upon a single discipline. **Table 3** gives an overview of the process models and determinant frameworks that were incorporated in our analysis per research domain. As we have only included English literature, this is largely represented in the geographical distribution of the included literature: 18 articles derive from authors affiliated with institutions located in the United States of America, 5 in the United Kingdom, 2 in Canada, 1 in Australia (in collaboration with a Spanish and Portuguese institution), 1 in Ireland, and 1 in Sweden.

Through analysis of both process models and determinant frameworks, we were able to grasp (1) a logical pathway in which different actions need to be taken in order to successfully implement a complex intervention, and (2) the main building blocks of which the intervention consists.

Table 4 gives an overview of the 15 process models with the main phases, steps, or requirements we could detract in each model (67 in total) and **Table 5** gives an overview of the 17 determinant frameworks and the main components, constructs, or elements that were put forward in these frameworks (90 in total). This served as a basis on which we detracted the common thread in each of these models and frameworks. We identified three main phases which most models have in common: a development phase, a translation phase, and a sustainment phase. Throughout all process models, 54 phases, steps, or requirements could directly be linked to these three phases. We also identified three main components: the intended change, the context, and the implementation strategies. A total of 67 components, constructs, or elements from all determinant frameworks could be directly linked to these three main components (see: **Table 5**). Thirteen components from 10 different process models could also be linked to these three main components (see: **Table 4**). Additionally, 17 components from 10 different determinant frameworks could indirectly be linked to the three main components as either outcomes, actors or processes (see: **Table 5**), leaving only 6 components that were not linked to the core phases and components we identified.

The three core phases we identified simplify the implementation process and are relevant to distinguish between different actions that need to be taken at different points in the process. The three components we identified are the core building blocks of the intervention: the way these components are approached and interact with each other will determine implementation success. Therefore, intervention designers need to reflect on how to approach each of the components within each of the phases.

Phases of an Implementation Process

To examine different phases of an implementation process, we look at process models, as defined by Nilsen (25). Such models are built to make sense of the different phases or steps of the implementation process of an intervention (25). The goal is to construct and clarify a "logical pathway" that can give concrete guidance for intervention designers. Many models were designed with the objective of translating research evidence into real world practice (33, 39, 44, 46) or the so called shift from knowledge to action [cfr. Wilson et al. (46)]. They tend to depart from

TABLE 3 | Overview of process models and determinant frameworks per domain.

Domain	Process models	Determinant frameworks
Implementation science or interdisciplinary	Consolidated Framework for Implementation Research (CFIR) (26), Advancing Understanding of Mechanism of Change in Implementation Science (31), Quality Implementation Framework (32), Ottawa Model of Health Care Research (30), Generic Implementation Framework (GIF) (19)	Consolidated Framework for Implementation Research (CFIR) (26), Integrated Promoting Action Research in Health Services Framework (i-PARiHS) (33), Understanding User Context Framework for Knowledge Translation (34), Interdisciplinary Conceptual Framework of Clinicians' Compliance with Evidence-based Guidelines (35), A Practical, Robust Implementation and Sustainability Model (PRISM) (36), Determinants and Consequences of Implementation Effectiveness (37), Conceptual Framework (3), Generic Implementation Framework (GIF) (19)
Medical sciences	Medical Research Council guidance (38), A Model for Large Scale Knowledge Translation (39)	Four levels of change for improving quality (40), Translating Research into Practice (41), Barrier Assessment (20)
Nursing	IOWA Model (42), Stetler Model of Research Utilization (43), ACE Star Model of Knowledge Transformation (44)	
Pharmacy	Active Implementation Frameworks (45)	
Public health or prevention research	The NCCDPHP Knowledge to Action Framework for Public Health (46), Research Utilization Model (modified from Rogers) (47)	Ecological Framework—Interactive Systems Framework for Dissemination and Implementation (48)
Organization research or service innovations	Organizational model for transformational change in health care systems (49)	Conceptual Model for Considering the Determinants of Diffusion, Dissemination, and Implementation of Health Service Delivery and Organization (50)
Social and behavioral sciences		Theoretical Domains Framework (V2.0) (51)
Social work	Conceptual Model of Evidence-Based Practice Implementation in Public Services Sectors (22)	Conceptual Model of Evidence-Based Practice Implementation in Public Services Sectors (22), the CAIMeR Theory (52)

an evidence base that needs to be translated into real world settings (33, 39, 44, 46). Other models incorporate a research development phase (38, 46, 47) in which best practices are still to be defined.

We find variation among models as to what is viewed as the main process of implementation. In some models such process takes the form of a stepwise approach to ensure successful implementation of an intervention (30–32, 42, 49, 51). Nilsen (25) calls these action models. They are built upon critical steps or phases that need to be followed or focused upon in order to reach successful implementation. These main phases or steps can either be aimed at the implementation process itself (31, 32) or at the process of using research to initiate change (42, 51). In such models, key drivers or components tend to be highlighted that are necessary for change (33, 49) and/or they have a thorough focus on those strategies that will lead to sustainable change, which is referred to as general implementation strategies (31), transfer strategies (30), capacity-building strategies (32) et cetera.

Another approach for describing a process is to have models differentiate between the main phases of how implementation efforts takes form, in order to make sense of the implementation process itself (22, 38, 43, 45–47). These models describe similar phases. They distinguish between either a development (38, 47), preparation (43) or exploration phase (22, 45), a pre-adoption phase [such as piloting (38), installation (45), or the intent/decision to adopt (22, 43, 47)], an actual implementation- (22, 38, 45, 47) or translation phase (43, 46) and a sustainment (22) or institutionalization (46, 47) phase. We reduce these models to three core phases: a development phase, a translation phase and a sustainment phase—as depicted in **Table 6**. This is a simplification that is relevant for intervention designers and practitioners, as these phases make most sense to them as distinct phases that require other types of action from them.

Development Phase

The development phase is the initial phase in which preparatory activities are conducted in order to successfully introduce the intervention. In the different models, various elements are considered to be relevant in this initial phase, which leads to a variety of actions that can be taken to prepare for and develop an intervention. Overall, the development phase comprises:

1) Synthesizing or collecting research evidence on which an intervention can be based;
2) Exploring the host setting;
3) Considering the overall fit of an intervention within a particular setting;
4) Ensuring readiness and intend to adopt the intervention.

Most models require that intervention designers synthesize existing evidence (38, 39, 42, 44), or that they conduct their own (discovery) research (44, 46). This will lead to either a theory (38), approach or practice (46), or research findings that can be translated into an evidence based practice (EBP) standard (42) or guidelines (44). Other models have a different focus and depart from the idea of planning (26) for an intervention or a more general exploration phase (22, 45). This is less focused on research translation and more intended to gain awareness of an issue (22), and to explore practices and implementation strategies that might respond to this issue (22). Exploration could also refer to assessing the feasibility of implementation

TABLE 4 | Overview of process models with their main phases, steps, or requirements.

	Framework	Phases/steps/requirements
Models who distinguish between phases of the implementation process	Medical Research Council guidance, Craig et al. (38)	Development Feasibility and piloting *Evaluation* Implementation
	Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors, Aarons et al. (22)	Exploration Adoption decision/Preparation Active implementation Sustainment
	Consolidated Framework for Implementation Research (CFIR), Damschroder et al. (26)	Planning Engaging Executing *Reflecting and evaluating*
	NCCDPHP Knowledge to Action Framework for Public Health, Wilson et al. (46)	Research phase Translation phase Institutionalization phase
	Research Utilization Model (modified from Rogers), Davis et al. (47)	Stage 0. Research Development *Stage 1. Dissemination* Stage 2. Intent to adopt Stage 3.a Implementation Stage 3.b Adaptation Stage 4. Institutionalization Stage 5. Diffusion and replication
	Active Implementation Frameworks, Blanchard et al. (45)	Exploration Installation Initial implementation Full implementation
	Stetler Model of Research Utilization, Stetler (43)	Phase 1: Preparation Phase 2: Validation Phase 3: Comparative Evaluation Phase 4: Decision making Phase 5: Translation/application Phase 6: Evaluation
	Generic Implementation Framework (GIF), Moullin et al. (19)	Pre-implementation Process of implementation Post-implementation
Action models with a step-wise approach	ACE Star Model of Knowledge Transformation, Stevens (44)	Discovery Research Evidence Summary Translation to Guidelines Practice Integration *Process, Outcome Evaluation*
	A model for large scale knowledge translation, Pronovost et al. (39)	1. Summarize the evidence *2*. Identify local barriers to implementation *3. Measure performance* *4*. Ensure all patients receive the interventions
	Advancing understanding of mechanism of change in implementation science, Lewis et al. (31)	*Step 1: Specifying implementation strategies* *Step 2: Generating strategy-mechanism linkages* *Step 3: Identifying proximal and distal outcomes* *Step 4: Articulating effect modifiers*
	Organizational model for transformational change in health care systems, Lukas et al. (49)	Impetus to Transform *Leadership* Improvement Initiatives Alignment Integration
	The ottawa model of health care research, Logan et al. (30)	1. Assess: Practice environment, potential adopters, evidence-based innovation 2. Monitor: Transfer strategies, adoption 3. Evaluate: Outcomes
	Quality Implementation Framework, Meyers et al. (32)	1. Initial considerations regarding the host setting *2*. Creating a structure for implementation *3*. Ongoing structure once implementation begins *4*. Improving future applications

(Continued)

TABLE 4 | Continued

Framework	Phases/steps/requirements
IOWA Model, Doody and Doody (42)	1. Selection of a topic *2*. Forming a team *3*. Evidence retrieval *4*. Grading the evidence *5*. Developing an EBP Standard *6*. Implement the EBP *7. Evaluation*

**The underlined phases/steps/requirements are those that are directly incorporated into the three main phases we put forward as common thread in these models.*
**The phases/steps/requirements in italics are linked to the three main components as described in Framework components.*

intentions or examining the readiness of the setting in which an intervention should take place (45). This is in line with Meyers, Durlak and Wandersman (32) who mention the importance of "*initial considerations regarding the host setting*," which refers to exploring whether there is a fit between an intervention and the host setting.

The fit between an intervention and the host setting (32, 45) can be linked to the need to asses contextual factors in this initial development phase (10, 53). The Ottawa Model of Health Care Research refers to "*assess*" as a first step, which means that the implementation environment, potential adopters, and the evidence-based innovation itself have to be examined (30). This is relevant when trying to assess the feasibility and compatibility of the intervention within a specific context. Pronovost et al. (39) mention a barrier assessment, which is a similar approach as the Conceptual Model of Evidence-Based Practice Implementation in which much emphasis is placed on mapping various hindering and promoting context variables in order to increase implementation success (22). These models recognize the importance of scanning contextual variables to identify barriers and facilitators that will affect implementation efforts.

Lastly, some models incorporate the decision or intend to adopt as a key element of the initial phase (22, 30, 43, 47). Lukas et al. (49) refer to this as the "*impetus to transform,*" which indicates that the decision to adopt a certain intervention is affected by various elements (22). This relates back to overall practitioner readiness (45), and the fit between the intervention and the setting in which it will be implemented (32, 43). According to the Quality Implementation Framework, a key step in the initial phase is also to create a structure for implementation (32). This can mean having a plan for implementation (32), but also to form a team that is dedicated to ensure implementation of an intervention (32, 42). The CFIR also recognizes the importance of engaging different actors that are involved in the intervention and views it as one of the core activities in the first phases of developing an intervention (26).

Translation Phase

Many frameworks refer to an implementation phase (22, 38, 42, 45, 47). It can also be called executing (26), adoption (30), improvement initiatives (49), or practice integration (44). Following the definition of Blanchard et al. (45), the core of this phase is to integrate the intervention into everyday practice, relying on the preparatory work started in the initial phase. We decided to follow the approach of the NCCDPHP Knowledge to Action Framework for Public Health (46) and the Stetler Model of Research Utilization (43) in which this phase is called the translation phase. They view the implementation process as translating research into practice. The core of these phases is however similar: it refers to the entire process of putting research into practice (46), thus implementing change into real world settings. In short, the actions that are key within the translation phase are:

1) Introducing the intervention by applying the strategies as defined in the development phase;
2) Monitoring how different components interact with each other to ensure continuous improvement.

All models with a translation phase will agree that key activities within this phase are applying those strategies (30, 45) or types of support (42, 46) that have been defined in the development phase, in order to introduce the intervention. For example, training or coaching is organized (45, 46), leadership- or communication structures are put in place (42, 49), technical assistance is provided or financial resources are made available (46). The Ottawa Model of Health Care Research (30) sees this as a monitoring phase, which means that strategies for introducing and implementing the intervention are to be observed and adjusted if necessary. Within the Research Utilization Model (47), the term "adaptation" is introduced, which means that "*over time, an innovation, the social system into which it is introduced, or both, may change or be modified to facilitate use of the innovation.*" This suggests that interaction is expected between the intervention, the strategies used and the context or setting in which the intervention takes place.

Sustainment Phase

Seven process models that we included in our analysis mention some form of sustainment phase. Aarons et al. (22) directly incorporate a sustainment phase and define it as "*the continued use of an innovation in practice*." This corresponds with what is named the "institutionalization phase" in the NCCDPHP Knowledge to Action Framework for Public Health (46) and the Research Utilization Model (47). Institutionalization of an intervention means that the intended change within an intervention becomes an established activity or norm within the setting it is implemented (46). It becomes integrated into

TABLE 5 | Overview of determinant frameworks with their main components, constructs, or elements.

Determinant framework	Components/constructs/elements
Consolidated Framework for Implementation Research (CFIR), Damschroder et al. (26)	Intervention Characteristics *Individuals involved* Inner setting Outer setting *Process*
Integrated Promoting Action Research in Health Services Framework (i-PARiHS), Stetler et al. (33)	Evidence/Evidence and EBP characteristics (revised version) Context/Contextual readiness for targeted EBP implementation (revised version) Facilitation *Successful implementation (revised version)*
CAIMeR theory, Blom and Morén (52)	Contexts *Actors* Interventions *Mechanisms* *Results*
Barrier assessment, Cochrane et al. (20)	Cognitive-behavioral barriers Attitudinal or rational-emotional barriers Professional barriers Barriers embedded in the guidelines or evidence Patient barriers Support or resources System and process barriers
Ecological Framework—Interactive Systems Framework for Dissemination and Implementation, Durlak and DuPre (48)	Community level factors Provider characteristics Characteristics of the innovation Factors relevant to the prevention delivery system Organizational capacity Factors related to the prevention support system
Conceptual model for considering the determinants of diffusion, dissemination, and implementation of health service delivery and organization, Greenhalgh et al. (50)	The innovation System antecedents for innovation System readiness for innovation *Adopter* Assimilation *Implementation process* *Linkage* Outer context Communication and influence Diffusion and dissemination
Understanding user context framework for knowledge translation, Jacobson et al. (34)	*The user group* The issue The research The researcher-user relationship Dissemination strategies
The interdisciplinary conceptual framework of clinicians' compliance with evidence-based guidelines, Gurses et al. (35)	System characteristics Provider characteristics Guideline characteristics Implementation characteristics
Four levels of change for improving quality, Ferlie and Shortell (40)	Individual change Group/team change Organizational change Larger system/environment change
A practical, robust implementation and sustainability model (PRISM), Feldstein and Glasgow (36)	Program (Interventions) External environment Implementation and sustainability infrastructure *Recipients*
Translating research into practice, Bradley et al. (41)	Top-down support Leadership Credibility of evidence-based practice Organizational culture Coordination of different stakeholders Intervention infrastructure Dissemination Diffusion

(Continued)

TABLE 5 | Continued

Determinant framework	Components/constructs/elements
Determinants and consequences of implementation effectiveness, Klein and Sorra (37)	Climate for implementation Skills Incentives and disincentives Absence of obstacles Innovation values fit Commitment Strategic accuracy of innovation adoption *Implementation effectiveness* *Innovation effectiveness*
Conceptual framework, Lau et al. (3)	External context Organization Professional Intervention
Generic Implementation Framework (GIF), Moullin et al. (19)	Innovation Context domains Strategies Factors *Evaluations*
The ottawa model of health care research, Logan et al. (30)	Practice environment *Potential adopters* Evidence-based innovation Transfer strategies Adoption *Outcomes*
Theoretical domains framework (v2.0), Atkins et al. (51)	Knowledge, skills, social/professional role and identity, beliefs about capabilities, optimism, beliefs about consequences, reinforcement, intentions, goals, memory, attention and decision processes, environmental context and resources, social influences, emotion, behavioral regulation
Conceptual model of evidence-based practice implementation in public service sectors, Aarons et al. (22)	Outer context Inner context *Interconnections*

**The underlined components/constructs/elements are those that are directly incorporated into the three main components we put forward as common thread in these frameworks.*
**The components/constructs/elements in italics are linked to the three main components in Framework components, as either outcomes or evaluation (linked to intended change), actors (linked to context) or process (linked to strategies).*

the routines and practices of this setting (47), and it should be ensured that the intervention is applied to all of whom it is aimed (39). Central in the sustainment phase is:

1) Applying the strategies as defined in the development phase to help sustain the intervention;
2) Reflecting upon the actions taken and ensuring continuous improvement.

Indeed, the aim of the sustainment phase of an intervention is that the intended change is maintained and becomes part of the daily routines and practices. This goes beyond a mere adoption of an intervention. The Organizational model for transformational change in health care systems (49) incorporates a similar idea, which is referred to as integration. Blanchard et al. (45) also ∗speak of integration of new learnings into practice, which they call full implementation. All of these imply that an intended change is adopted and in time harmonizes with, or replaces previously existing practices and activities.

A sustainment phase is also the phase in which continuous improvements ensure a fit between the intervention and the setting in which it is implemented. The Quality Implementation Framework (32) sees the improvement of future applications as the core of this final phase. This is learning from experience. Through reflection and feedback from the setting in which the intervention is introduced, strengths, and weaknesses of the intervention can be detected and acted upon (32). For Blanchard et al. (45) this implies achieving fidelity and improving outcomes. This phase can directly be linked to evaluation, which four process models include as a separate phase.

This notion of continuous improvement can be linked to reflection and evaluation as a part of the process. Several process models include evaluation or measuring performance and outcomes as a phase of the implementation process, for example in the Medical Research Council guidance (38), the CFIR (26), the Ottawa Model of Health Care Research (30), the IOWA model (42), the ACE Star Model of Knowledge Transformation (44), the Stetler Model of Research Utilization (43) and Advancing Understanding of Mechanism of Change in Implementation Science (31), and the model for large scale knowledge translation (39). These frameworks or models generally include minor guidance about how to assess success or failure. There are however also frameworks that are designed specifically to guide the evaluation process, examples of which are the RE-AIM framework (54), the PRECEDE-PROCEED model (55) and the Implementation Outcomes Framework (IOF) (56).

TABLE 6 | Overview of process models in relation to a development phase, translation phase and sustainment phase.

	Framework	Development phase	Translation phase	Sustainment phase
Models who distinguish between phases of the implementation process	Medical research council guidance, Craig et al. (38)	Development	Feasibility and piloting Implementation	–
	Conceptual model of evidence-based practice implementation in public service sectors, Aarons et al. (22)	Exploration Adoption Decision/Preparation	Active implementation	Sustainment
	Consolidated Framework for Implementation Research (CFIR), Damschroder et al. (26)	Planning Engaging	Executing	–
	NCCDPHP knowledge to action framework for public health, Wilson et al. (46)	Research phase	Translation phase	Institutionalization phase
	Research utilization model (modified from Rogers), Davis et al. (47)	Research Development Intent to adopt	Implementation Adaptation	Institutionalization Diffusion and replication
	Active implementation frameworks, Blanchard et al. (45)	Exploration	Installation Initial implementation	Full implementation
	Stetler model of research utilization, Stetler (43)	Preparation validation comparative evaluation decision making	Translation/application	–
	Generic Implementation Framework (GIF), Moullin et al. (19)	Pre-implementation	Process of implementation	Post-implementation
	ACE star model of knowledge transformation, Stevens (44)	Discovery Research Evidence Summary Translation into guidelines	Practice integration	–
	A model for large scale knowledge translation, Pronovost et al. (39)	Summarize the evidence Identify local barriers to implementation	–	Ensure all patients receive the interventions
Action models with a step-wise approach	Advancing understanding of mechanism of change in implementation science, Lewis et al. (31)	–	–	–
	Organizational model for transformational change in health care systems, Lukas et al. (49)	Impetus to Transform	Improvement Initiatives Alignment	Integration
	The ottawa model of health care research, Logan et al. (30)	Assess (practice environment, potential adopters, evidence-based innovation)	Monitor (transfer strategies, adoption)	-
	Quality implementation framework, Meyers et al. (32)	Initial considerations regarding the host setting Creating a structure for implementation	Ongoing structure once implementation begins	Improving future applications
	IOWA model, Doody and Doody (42)	Selection of a topic Forming a team Evidence retrieval Grading the evidence Developing an EBP standard	Implement the EBP	–

Framework Components

Throughout the three phases of the implementation process, we distinguish components that have to be taken into account within each phase. Therefore, we looked into what Nilsen (25) calls determinant frameworks. These are designed with the intent to understand and explain what influences implementation outcomes, and thus provide information on which components to focus for implementation success. Some frameworks tend to mainly focus on enlisting relevant context variables [e.g., Theoretical Domains Framework 2.0 (51)], while others also specify the relationships and interactions between types of determinants (25). These frameworks provide valuable input when describing different types of context variables that might hinder or facilitate intervention efforts.

Table 7 gives an overview of how various determinant frameworks refer to the three components that we have extracted:

intended change, context and/or strategies. They will provide further guidance on how to understand and work with these elements and how they can affect implementation outcomes.

Intended Change

The intended change deals with any conscious change into current practices of primary care providers or any actions that actors undertake (57), which are expected to solve a care or quality gap (58). This can take the form of a task-oriented change in practice (33), require behavioral change (59) either at individual or group/team level (40) and/or have a broader organizational impact whereby a more complex transformational change is initiated (33). The intended change derives from the objectives of the intervention, with the assumption that the initiated change will contribute to realizing these objectives (58).

Twelve of the determinant frameworks mention a component similar to the intended change as part of the implementation process. This is referred to as (characteristics of) an intervention (3, 26, 52), innovation (19, 30, 48, 60), change (40), program (36) or issue (34), or involves an evidence based practice (20, 30, 33, 41) or guidelines (20, 35). Determinant frameworks that do not mention the intervention as a separate component either focus on context variables (40), domains (51) or barriers (20), or incorporate intervention aspects in general implementation characteristics (35, 37).

The CFIR (26), the Interactive Systems Framework for Dissemination and Implementation (48) and i-PARiHS specifically zoom in on the characteristics of such an intended change (in these models referred to as intervention, innovation, or evidence-based practice). This indicates that an intervention or intended change is complex, multi-faceted, and different components will be interacting with each other (26). Characteristics that are mentioned are among others compatibility (33, 35, 48), adaptability (36, 48), complexity (26, 33, 35, 36), and/or relative advantage (33, 35). Such inherent characteristics of the intervention will have an impact on its overall implementation success.

As the intended change is expected to contribute to realizing the objectives of the intervention, it is important to define what outcomes are expected from the intended change. Four determinant frameworks incorporate *results* (52), *output* (52), *outcomes* (30), (implementation or innovation) *effectiveness* (37), or *successful implementation* (33) as separate components. This helps focusing on the objectives that are set when defining an intervention and the benefits that arise when implementation is successful (37). The time frame in which results can be observed, can differ majorly. Certain results are obtained early on, while others only exist in the long-term even after the intervention is finished (52). When defining the intended change, it is thus key to not only define the behavioral or organizational change that is expected, but also the expected results and how this can be evaluated.

Context

Context variables can be defined as *"the set of circumstances or unique factors that surround a particular implementation effort* (26)." They are dynamic factors that interact, influence, modify, and facilitate or constraint intervention and implementation efforts (53). Context variables are most prominent in what Nilsen (25) defines as determinant frameworks, in which the main objective is to gain insight in those barriers and facilitators that impact implementation outcomes (25). Some are built with the interaction of context variables (40), context domains (51), or barriers (20) as a main focus. Most frameworks indeed incorporate some form of context variables as an essential part of the implementation process. i-PARiHS (33), the Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors (22), the CFIR (26), the CAIMeR theory (52), and the GIF (19) directly incorporate context, contextual readiness, inner and outer context, context domains, setting, or factors as a component of the framework. A distinction is sometimes made between inner- and outer context or setting (22, 26), which mentions inner context variables as being specific to a person, team our organization (on micro and meso level), while outer context variables are broader in nature such as socio-economic or policy variables (on macro level).

When referring to context, some frameworks only incorporate context variables on the macro level. They zoom in on the so called outer context (50), external context (3), or external environment (36). Elements on an organizational or individual-adopter level are then incorporated under a different name. For example, organizational aspects can also be referred to as system characteristics (35), system antecedents or system readiness for innovation (50), practice environment (30), system and process barriers (20), implementation and sustainability infrastructure (36), organizational culture (41) or climate for implementation (37), intervention infrastructure (41), or factors relevant to the prevention delivery system (48).

When it comes to the micro context, individual adopter characteristics are mentioned by fewer frameworks. They are referred to as professional (3), or provider characteristics (35, 48), or more specifically as cognitive-behavioral barriers, attitudinal, or rational-emotional barriers or professional barriers (20), which indicates that individual adopter characteristics can cover a wide range of micro level aspects. This is also noticeable in the Theoretical Domains Framework (51), in which a wide variety of "domains" is mentioned, many of which are individual adopter characteristics such as professional role, beliefs about capabilities, etc.

On the micro level, context variables highly relate to the actors to which the intended change concerns. Greenhalgh et al. (50) state that "*people are not passive recipients of innovations*." The dynamic interplay of how individuals relate to the organization in which they work (26) and their general assumptions about people, society and their profession (52) influences their perception and the way in which they make sense of an intended change. Six determinant frameworks include *actors* (52), *individuals involved* (26), *potential adopters* (30, 50), *recipients* (36), or *the user group* (34) as a core component. Incorporate actors as one of the components strengthens the view that actors have an impact on the way an intervention is realized. In five determinant frameworks, the influence actors have on implementation success is recognized by including individual attitudes, cognitions, or professional characteristics

TABLE 7 | Overview of determinant frameworks that incorporate intended change, context, and strategies as components.

Determinant framework	Intended change	Context	Strategies
Consolidated Framework for Implementation Research (CFIR), Damschroder et al. (26)	Intervention characteristics	Inner setting Outer setting	-
Integrated Promoting Action Research in Health Services Framework (i-PARiHS), Stetler et al. (33)	Evidence/Evidence and EBP characteristics	Context/Contextual readiness for targeted EBP implementation	Facilitation
CAIMeR theory, Blom amd Morén (52)	Interventions	Contexts	–
Barrier assessment, Cochrane et al. (20)	*Barriers embedded in the guidelines or evidence*	Cognitive-behavioral barriers Attitudinal or rational-emotional barriers Professional barriers Patient barriers System and process barriers	Support or resources
Ecological framework—interactive systems framework for dissemination and implementation, Durlak and DuPre (48)	Characteristics of the innovation	Community level factors Provider characteristics Factors relevant to the prevention delivery system: organizational capacity	Factors related to the prevention support system
Conceptual model for considering the determinants of diffusion, dissemination, and implementation of health service delivery and organization, Greenhalgh et al. (50)	The innovation	System antecedents for innovation System readiness for innovation Outer context	Communication and influence Diffusion and dissemination
Understanding user context framework for knowledge translation, Jacobson et al. (34)	The issue	–	Dissemination strategies
The interdisciplinary conceptual framework of clinicians' compliance with evidence-based guidelines, Gurses et al. (35)	Guideline characteristics	System characteristics Provider characteristics	Implementation characteristics
Four levels of change for improving quality, Ferlie and Shortell (40)	Individual change Group/team change Organizational change Larger system/environment change	–	–
A practical, robust implementation and sustainability model (PRISM), Feldstein and Glasgow (36)	Program (interventions)	External environment Implementation and sustainability infrastructure	-
Translating research into practice, Bradley et al. (41)	Credibility of evidence-based practice	Top-down support Leadership Organizational culture Intervention infrastructure	Coordination of different stakeholders Dissemination Diffusion
Determinants and consequences of implementation effectiveness, Klein and Sorra (37)	–	Climate for Implementation Innovation values fit	Skills/Incentives and disincentives/Absence of obstacles
Conceptual framework, Lau et al. (3)	Intervention	External context Organization Professional	-
Generic Implementation Framework (GIF), Moullin et al. (19)	Innovation	Context domains Factors	Strategies
The ottawa model of health care research, Logan et al. (30)	Evidence-based innovation	Practice environment	Transfer strategies
Theoretical domains framework (v2.0), Atkins et al. (51)	–	*Provides a list of domains that can be incorporated as context variables.*	–
Conceptual model of evidence-based practice implementation in public service sectors, Aarons et al. (22)	–	Outer context Inner context	–

as a context variable (3, 20, 35, 48, 51). The component actors can thus be incorporated as a separate component of an implementation model, but it can also be included as a micro level context variable.

Overall, there is a wide belief that the context in which a primary care intervention takes place highly determines implementation success (10). This makes scanning and taking into account the context key for each phase of the implementation process. When determining implementation strategies, context variables must be taken into account in order for strategies to be tailored and fit local circumstances (10). This is in line with realist evaluation, whereby the general aim is to find out *"what works, for whom, and under what conditions?"* (6). In this approach, context variables are the conditions in which an intervention takes place.

Strategies

Implementation strategies can be defined as the approach(es) and means that are used to ensure or enhance the adoption of the target behaviors and other requirements of the primary care intervention by the targeted actors (10, 61). Whereas, the intended change refers to *what* is to be implemented, the strategies refer to *how* they are to be implemented and is linked to the process or mechanism that intervention designers want to trigger in order to accomplish implementation.

Implementation strategies are directly referred to in few process models, such as Advancing Understanding of Mechanism of Change in Implementation Science (31), whereby a first step to implementation is to specify the implementation strategies; the Ottawa Model of Health Care Research (30) in which transferring strategies is a part of monitoring the uptake of the intervention and in the GIF (19), in which the strategies are viewed as the approaches to respond to barriers and facilitators. Throughout other determinant frameworks, a component similar to implementation strategies is included in eight of the models we included in our analysis, either in the form of facilitation (33), support (e.g., training, assistance) (20, 48), implementation characteristics (35) and dissemination and/or diffusion of strategies (34, 41, 50). Frameworks also tend to incorporate those elements that are considered to be most decisive as strategies, such as communicational aspects (50), coordination of different stakeholders (41), or the use of incentives and disincentives (37).

Implementation strategies are discussed more in-depth in the Expert Recommendations for Implementing Change (ERIC) study, in which a compilation of 73 implementation strategies was made (62, 63). This can serve as a guide for when the most fitting implementation strategies have to be selected for the implementation of a certain intervention. To make more sense of the wide diversity of implementation strategies, they often are categorized. For example, Powell et al. (64) distinguishes between strategies that are related to either planning, educating, financing, restructuring, managing quality, and/or attending to policy context. Another categorization can be found in Charif et al. (65), who differentiate strategies that are related to either the health infrastructure, policy and regulation, financing, human resource, or patients (65).

Implementation strategies can be very different depending on the type of change that is initiated, and should ideally be tailored to fit the inner and outer context (10, 66), making use of the facilitators or barriers that are observed in order to ensure a fit between the intervention and its context (3). When defining implementation strategies to implement one's intervention, Proctor et al. (67) have set up guiding principles to name, define, and operationalize implementation strategies by firstly specifying the following elements: (1) actor, (2) action, (3) action target, (4) temporality, (5) dose, (6) implementation outcome affected, and (7) justification. These can support intervention designers in defining implementation strategies.

In short, implementation strategies are expected to lead to an intended change in a given context. This means that there is an underlying process that will bring about this change. Three determinant frameworks include this *(implementation) process* (26, 60) or *mechanism* (52) as one of the core components. These frameworks have a more explanatory approach and put more emphasis on understanding the process of change. For complex interventions, this consists of many interdependent sub-processes that may or may not follow a clear path to success (26). The process involves decision making activities, the use of resources, communication, and collaboration (50). Blom and Morén (52) view this as an either social, socio-psychological, or psychological mechanism that is at the base of change. Greenhalgh et al. (50) and Lewis et al. (31) also refer to *linkages* or *effect modifiers* and Aarons et al. (22) speak about *interconnections,* referring to the fit between an innovation and a system or organization that comes into play when introducing a change. These frameworks thus incorporate the process or mechanism of change as a core element that needs to be understood in order to fully know how to target certain interventions in specific settings. When choosing implementation strategies, it is thus recommended to make explicit the assumptions of how a certain strategy will lead to the intended change in a given context.

DISCUSSION

We have identified the core building blocks of an overarching implementation framework for complex interventions in primary care services. Throughout our narrative review, three core phases are detracted that describe the process of implementation in relation to three core components. This process can roughly be divided in a development phase, a translation phase, and a sustainment phase. For each phase, three main components are essential to define, tailor, and manage to successfully implement an intervention in a specific setting. These are the intended change, the context, and the implementation strategies. Other related components that are closely linked to these three components may still be relevant,

such as actors, the process or mechanism, and the outcomes and evaluation of the intervention.

An overarching implementation framework is needed to transcend the solely theoretical models and to aim for a model that is both explanatory as well as actionable. Context variables should be given a prominent place in this, as tailoring interventions to local circumstances is considered key for reaching implementation success (9, 10). By focusing on the core components intended change, context, and strategies we propose meaningful concepts to intervention designers and practitioners for reflecting upon the interactions of these components. The next step is synthesizing these core building blocks into a framework that consists of a clear and actionable pathway for intervention designers, and which enables them to prioritize and reflect upon those actions that need to be taken for the implementation of complex interventions.

Our research is part of a larger project that intends to make progress in three main research areas: to improve goal oriented care, self-management, and inter-professional collaboration. In each of the three areas, one or more interventions will be used for developing and evaluating the implementation of interventions in these three areas. The model that we will further develop will allow to develop and implement interventions with broad consideration of the setting or context in which they will be introduced, and how this interacts with the intended change and the implementation strategies that are used.

A limitation of our review is that we did not gather and include our sources in a systematic way. We used a more intuitive approach whereby sources were gathered mainly through expertise from our research team, by database searches with a set of different key words and by further use of a snowball approach that lead to the most prominent frameworks and models that exist. Furthermore, as we have only included English literature, there seems to be a slight overrepresentation of literature deriving from native English authors and/or institutions. Moreover, we have no view on gray literature or literature written in foreign languages, which might further limit our scope.

Although there is no assurance that we have covered all relevant literature, the methodology of a narrative review allowed us to explore the broad range of implementation literature and interpret various approaches in the light of interventions that aim toward pro-active, person-centered primary care. This way, we could harmonize literature into insightful constructs and phases which are to be made concrete when further applying them in the defining and execution of interventions.

CONCLUSION

An overarching implementation model is needed to bridge the gap between scientific evidence and actual practice in primary care. Through a narrative review, we have identified the core building blocks that form the common thread of existing implementation frameworks or models and we synthesized it in three core phases (a development phase, a translation phase and a sustainment phase) and three core components (the intended change, the context and the implementation strategies). These core building blocks can be used to develop an overarching implementation model that is both explanatory, as well as actionable. The main phases and components are the basis on which further guidance for intervention designers will be elaborated. A strength of the model that we will develop based upon this research is that it will be further developed and refined in collaboration with three research teams that will actively use the model to develop and introduce one or more interventions in primary care. This allows for direct feedback on its applicability and therefore ensures its actionability.

AUTHOR CONTRIBUTIONS

IH wrote the main manuscript text. AD, EV, PR, and SA contributed to the different steps of the making of this manuscript. All authors reviewed the manuscript.

FUNDING

IH and EV received funding from the Koning Boudewijnstichting.

ACKNOWLEDGMENTS

This article was written on behalf of the Primary Care Academy.

REFERENCES

1. Crabtree BF, Nutting PA, Miller WL, McDaniel RR, Stange KC, Jaen CR, et al. Primary care practice transformation is hard work: insights from a 15-year developmental program of research. *Med Care.* (2011) 49:S28. doi: 10.1097/MLR.0b013e3181cad65c
2. Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, et al. Can we systematically review studies that evaluate complex interventions? *PLoS Med.* (2009) 6:e1000086. doi: 10.1371/journal.pmed.1000086
3. Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care—causes of the evidence to practice gap: systematic reviews of reviews. *Implement Sci.* (2015) 11:40. doi: 10.1186/s13012-016-0396-4
4. Epstein RM, Fiscella K, Lesser CS, Stange KC. Why the nation needs a policy push on patient-centered health care. *Health Affairs.* (2010) 29:1489–95. doi: 10.1377/hlthaff.2009.0888
5. Kitson A, Marshall A, Bassett K, Zeitz K. What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. *J Adv Nurs.* (2013) 69:4–15. doi: 10.1111/j.1365-2648.2012.06064.x
6. Salter KL, Kothari A. Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review. *Implement Sci.* (2014) 9:115. doi: 10.1186/s13012-014-0115-y
7. Grol R, Wensing M, Eccles M, Davis D. *Improving Patient Care: The Implementation of Change in Health Care*. New Jersey: John Wiley & Sons (2013). doi: 10.1002/9781118525975

8. Olswang LB, Prelock PA. Bridging the gap between research and practice: implementation science. *J Speech Lang Hear Res.* (2015) 58:S1818–26. doi: 10.1044/2015_JSLHR-L-14-0305
9. May CR, Johnson M, Finch T. Implementation, context and complexity. *Implement Sci.* (2016) 11:141. doi: 10.1186/s13012-016-0506-3
10. Pfadenhauer LM, Gerhardus A, Mozygemba K, Lysdahl KB, Booth A, Hofmann B, et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. *Implement Sci.* (2017) 12:21. doi: 10.1186/s13012-017-0552-5
11. Eccles MP, Mittman BS. *Welcome to Implementation Science*. New York, NY: Springer (2006). doi: 10.1186/1748-5908-1-1
12. Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. *BMC Med.* (2019) 17:88. doi: 10.1186/s12916-019-1322-9
13. Hull L, Goulding L, Khadjesari Z, Davis R, Healey A, Bakolis I, et al. Designing high-quality implementation research: development, application, feasibility and preliminary evaluation of the implementation science research development (ImpRes) tool and guide. *Implement Sci.* (2019) 14:80. doi: 10.1186/s13012-019-0897-z
14. National Clinical Effectiveness Committee. *Implementation Guide and Toolkit for National Clinical Guidelines: Department of Health (DoH)*, (2018). Available online at: http://hdl.handle.net/10147/623839
15. Registered Nurses' Association of Ontario (RNAO). *Toolkit: Implementation of Best Practice Guidelines (2nd ed)*, Toronto, ON (2012).
16. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. Standards for Reporting Implementation Studies (StaRI) statement. *BMJ.* (2017) 356:i6795. doi: 10.1136/bmj.i6795
17. Smith JD, Li DH, Rafferty MR. The implementation research logic model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implement Sci.* (2020) 15:1–12. doi: 10.1186/s13012-020-01041-8
18. Birken SA, Powell BJ, Shea CM, Haines ER, Kirk MA, Leeman J, et al. Criteria for selecting implementation science theories and frameworks: results from an international survey. *Implement Sci.* (2017) 12:124. doi: 10.1186/s13012-017-0656-y
19. Moullin JC, Sabater-Hernández D, Fernandez-Llimos F, Benrimoj SI. A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. *Health Res Policy Syst.* (2015) 13:16. doi: 10.1186/s12961-015-0005-z
20. Cochrane LJ, Olson CA, Murray S, Dupuis M, Tooman T, Hayes S. Gaps between knowing and doing: understanding and assessing the barriers to optimal health care. *J Continuing Educ Health Professions.* (2007) 27:94–102. doi: 10.1002/chp.106
21. Davies P, Walker AE, Grimshaw JM. A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. *Implement Sci.* (2010) 5:14. doi: 10.1186/1748-5908-5-14
22. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Adm Policy Ment Health.* (2011) 38:4–23. doi: 10.1007/s10488-010-0327-7
23. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. *J Chiropractic Med.* (2006) 5:101–17. doi: 10.1016/S0899-3467(07)60142-6
24. Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews? *Eur J Clin Invest.* (2018) 48:e12931. doi: 10.1111/eci.12931
25. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci.* (2015) 10:53. doi: 10.1186/s13012-015-0242-0
26. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* (2009) 4:50. doi: 10.1186/1748-5908-4-50
27. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* (2005) 8:19–32. doi: 10.1080/1364557032000119616
28. Rogers EM. *Diffusion of Innovations: Simon and Schuster*, New York, NY (2010).
29. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology.* (2009) 43:535–54. doi: 10.1177/0038038509103208
30. Logan J, Graham ID. Toward a comprehensive interdisciplinary model of health care research use. *Sci Commun.* (1998) 20:227–46. doi: 10.1177/1075547098020002004
31. Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, et al. From classification to causality: advancing understanding of mechanisms of change in implementation science. *Front Public Health.* (2018) 6:136. doi: 10.3389/fpubh.2018.00136
32. Meyers DC, Durlak JA, Wandersman A. The quality implementation framework: a synthesis of critical steps in the implementation process. *Am J Community Psychol.* (2012) 50:462–80. doi: 10.1007/s10464-012-9522-x
33. Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A guide for applying a revised version of the PARIHS framework for implementation. *Implement Sci.* (2011) 6:99. doi: 10.1186/1748-5908-6-99
34. Jacobson N, Butterill D, Goering P. Development of a framework for knowledge translation: understanding user context. *J Health Serv Res Policy.* (2003) 8:94–9. doi: 10.1258/135581903321466067
35. Gurses AP, Marsteller JA, Ozok AA, Xiao Y, Owens S, Pronovost PJ. Using an interdisciplinary approach to identify factors that affect clinicians' compliance with evidence-based guidelines. *Critic Care Med.* (2010) 38:S282–91. doi: 10.1097/CCM.0b013e3181e69e02
36. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. *Joint Commission J Qual Patient Safety.* (2008) 34:228–43. doi: 10.1016/S1553-7250(08)34030-6
37. Klein KJ, Sorra JS. The challenge of innovation implementation. *Acad Manage Rev.* (1996) 21:1055–80. doi: 10.5465/amr.1996.9704071863
38. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ.* (2008) 337:a1655. doi: 10.1136/bmj.a1655
39. Pronovost PJ, Berenholtz SM, Needham DM. Translating evidence into practice: a model for large scale knowledge translation. *BMJ.* (2008) 337:a1714. doi: 10.1136/bmj.a1714
40. Ferlie EB, Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *Milbank Quart.* (2001) 79:281–315. doi: 10.1111/1468-0009.00206
41. Bradley EH, Webster TR, Baker D, Schlesinger M, Inouye SK, Barth MC, et al. Translating research into practice: speeding the adoption of innovative health care programs. *Issue Brief.* (2004) 724:1–12. doi: 10.1111/j.1532-5415.2004.52510.x
42. Doody CM, Doody O. Introducing evidence into nursing practice: using the IOWA model. *Br J Nurs.* (2011) 20:661–4. doi: 10.12968/bjon.2011.20.11.661
43. Stetler CB. Updating the Stetler model of research utilization to facilitate evidence-based practice. *Nurs Outlook.* (2001) 49:272–9. doi: 10.1067/mno.2001.120517
44. Stevens K. The impact of evidence-based practice in nursing and the next big ideas. *Online J Issues Nurs.* (2013) 18:4. doi: 10.3912/OJIN.Vol18No02Man04
45. Blanchard C, Livet M, Ward C, Sorge L, Sorensen TD, McClurg MR. The active implementation frameworks: a roadmap for advancing implementation of comprehensive medication management in primary care. *Res Soc Admin Pharmacy.* (2017) 13:922–9. doi: 10.1016/j.sapharm.2017.05.006
46. Wilson KM, Brady TJ, Lesesne C, Translation NWGo. Peer reviewed: an organizing framework for translation in public health: the knowledge to action framework. *Prev Chronic Dis.* (2011) 8:A46. Available online at: http://www.cdc.gov/pcd/issues/2011/mar/10_0012.htm
47. Davis SM, Peterson JC, Helfrich CD, Cunningham-Sabo L. Introduction and conceptual model for utilization of prevention research. *Am J Prevent Med.* (2007) 33:S1–5. doi: 10.1016/j.amepre.2007.04.004
48. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* (2008) 41:327–50. doi: 10.1007/s10464-008-9165-0
49. Lukas CV, Holmes SK, Cohen AB, Restuccia J, Cramer IE, Shwartz M, et al. Transformational change in health care systems: an organizational model. *Health Care Manage Rev.* (2007) 32:309–20. doi: 10.1097/01.HMR.0000296785.29718.5d
50. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic

review and recommendations. *Milbank Quart.* (2004) 82:581–629. doi: 10.1111/j.0887-378X.2004.00325.x
51. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implement Sci.* (2017) 12:77. doi: 10.1186/s13012-017-0605-9
52. Blom B, Morén S. Explaining social work practice — the CAIMeR theory. *J Soc Work.* (2009) 10:98–119. doi: 10.1177/1468017309350661
53. Pfadenhauer LM, Mozygemba K, Gerhardus A, Hofmann B, Booth A, Lysdahl KB, et al. Context and implementation: a concept analysis towards conceptual maturity. *Z Evid Fortbild Qual Gesundhwes.* (2015) 109:103–14. doi: 10.1016/j.zefq.2015.01.004
54. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health.* (1999) 89:1322–7. doi: 10.2105/AJPH.89.9.1322
55. Crosby R, Noar SM. What is a planning model? An introduction to PRECEDE-PROCEED. *J Public Health Dentistry*. (2011). 71:S7–15. doi: 10.1111/j.1752-7325.2011.00235.x
56. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health.* (2011) 38:65–76. doi: 10.1007/s10488-010-0319-7
57. Chen HT. Theory-driven evaluation: conceptual framework, application and advancement. In: Strobl R, Lobermeier O, Heitmeyer W, editors. *Evaluation von Programmen und Projekten für eine demokratische Kultur*. Wiesbaden: Springer (2012). p. 17–40. doi: 10.1007/978-3-531-19009-9_2
58. Crable EL, Biancarelli D, Walkey AJ, Allen CG, Proctor EK, Drainoni ML. Standardizing an approach to the evaluation of implementation science proposals. *Implement Sci.* (2018) 13:71. doi: 10.1186/s13012-018-0770-5
59. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med.* (2013) 46:81–95. doi: 10.1007/s12160-013-9486-6
60. Greenhalgh T, Robert G, Bate P, Macfarlane F, Kyriakidou O. *Diffusion of Innovations in Health Service Organisations: A Systematic Literature Review*. New York, NY: John Wiley & Sons (2008).
61. Neta G, Brownson RC, Chambers DA. Opportunities for epidemiologists in implementation science: a primer. *Am J Epidemiol.* (2017) 187:899–910. doi: 10.1093/aje/kwx323
62. Waltz TJ, Powell BJ, Matthieu MM, Damschroder LJ, Chinman MJ, Smith JL, et al. Use of concept mapping to characterize relationships among implementation strategies and assess their feasibility and importance: results from the Expert Recommendations for Implementing Change (ERIC) study. *Implement Sci.* (2015) 10:109. doi: 10.1186/s13012-015-0295-0
63. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci.* (2015) 10:21. doi: 10.1186/s13012-015-0209-1
64. Powell BJ, McMillen JC, Proctor EK, Carpenter CR, Griffey RT, Bunger AC, et al. A compilation of strategies for implementing clinical innovations in health and mental health. *Med Care Res Rev.* (2012) 69:123–57. doi: 10.1177/1077558711430690
65. Charif AB, Zomahoun HTV, LeBlanc A, Langlois L, Wolfenden L, Yoong SL, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implement Sci.* (2017) 12:139. doi: 10.1186/s13012-017-0672-y
66. Bunger AC, Powell BJ, Robertson HA, MacDowell H, Birken SA, Shea C. Tracking implementation strategies: a description of a practical approach and early findings. *Health Res Policy Syst.* (2017) 15:15. doi: 10.1186/s12961-017-0175-y
67. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. *Implement Sci.* (2013) 8:139. doi: 10.1186/1748-5908-8-139

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Public Health

ORIGINAL RESEARCH
published: 17 November 2021
doi: 10.3389/fpubh.2021.774907

The Implementation of a Primary Care-Based Integrated Mobile Health Intervention for Stroke Management in Rural China: Mixed-Methods Process Evaluation

Enying Gong[1,2,3], *Lixin Sun*[2], *Qian Long*[2], *Hanzhang Xu*[4], *Wanbing Gu*[2], *Janet Prvu Bettger*[4,5], *Jingru Tan*[2,3], *Jixiang Ma*[6], *Tazeen Hasan Jafar*[4,7], *Brian Oldenburg*[2,8] *and Lijing L. Yan*[2,4,9,10]*

[1] School of Population Medicine and Public Health, China Academy of Medical Science & Peking Union Medical College, Beijing, China, [2] Global Health Research Center, Duke Kunshan University, Suzhou, China, [3] School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia, [4] Duke Global Health Institute, Duke University, Durham, NC, United States, [5] Department of Orthopedic Surgery, Duke University, Durham, NC, United States, [6] Shandong Center for Disease Control and Prevention, Shandong, China, [7] Program in Health Services & Systems Research, Duke-NUS Medical School, Singapore, Singapore, [8] Non-communicable Disease Unit, Baker Institute, Melbourne, VIC, Australia, [9] School of Health Sciences, Wuhan University, Wuhan, China, [10] Peking University School of Global Health and Development, Peking University, Beijing, China

OPEN ACCESS

Edited by:
Michele Mario Ciulla, University of Milan, Italy

Reviewed by:
Eron Grant Manusov, The University of Texas Rio Grande Valley, United States
Michael Pollack, University of Newcastle, Australia

***Correspondence:**
Lijing L. Yan
lijing.yan@dukekunshan.edu.cn

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Public Health

Received: *13 September 2021*
Accepted: *11 October 2021*
Published: *17 November 2021*

Citation:
Gong E, Sun L, Long Q, Xu H, Gu W, Bettger JP, Tan J, Ma J, Jafar TH, Oldenburg B and Yan LL (2021) The Implementation of a Primary Care-Based Integrated Mobile Health Intervention for Stroke Management in Rural China: Mixed-Methods Process Evaluation. Front. Public Health 9:774907. doi: 10.3389/fpubh.2021.774907

Background: There is a lack of evidence concerning the effective implementation of strategies for stroke prevention and management, particularly in resource-limited settings. A primary-care-based integrated mobile health intervention (SINEMA intervention) has been implemented and evaluated *via* a 1-year-long cluster-randomized controlled trial. This study reports the findings from the trial implementation and process evaluation that investigate the implementation of the intervention and inform factors that may influence the wider implementation of the intervention in the future.

Methods: We developed an evaluation framework by employing both the RE-AIM framework and the MRC process evaluation framework to describe the implementation indicators, related enablers and barriers, and illustrate some potential impact pathways that may influence the effectiveness of the intervention in the trial. Quantitative data were collected from surveys and extracted from digital health monitoring systems. In addition, we conducted quarterly in-depth interviews with stakeholders in order to understand barriers and enablers of program implementation and effectiveness. Quantitative data analysis and thematic qualitative data analysis were applied, and the findings were synthesized based on the evaluation framework.

Results: The SINEMA intervention was successfully implemented in 25 rural villages, reached 637 patients with stroke in rural Northern China during the 12 months of the trial. Almost 90% of the participants received all follow-up visits per protocol, and about half of the participants received daily voice messages. The majority of the intervention components were adopted by village doctors with some adaptation made. The interaction between human-delivered and technology-enabled components reinforced the program implementation and effectiveness. However, characteristics of the participants, doctor-patient relationships, and the healthcare system context attributed to the variation of program implementation and effectiveness.

Conclusion: A comprehensive evaluation of program implementation demonstrates that the SINEMA program was well implemented in rural China. Findings from this research provide additional information for program adaptation, which shed light on the future program scale-up. The study also demonstrates the feasibility of combining RE-AIM and MRC process evaluation frameworks in process and implementation evaluation in trials.

Clinical Trial Registration: www.ClinicalTrials.gov, identifier: NCT03185858.

Keywords: stroke, implementation evaluation, mobile health, rural China, RE-AIM (reach, effectiveness, adoption, implementation and maintenance)

INTRODUCTION

Stroke is the second leading cause of mortality and disability worldwide. An estimation from the Global Burden of Disease study found that 77% of stroke survivors were from low- and middle-income countries (LMICs), where individuals with lower socioeconomic status suffer more, and effective strategy in secondary prevention is far lacking (1, 2). In China, the burden of stroke is substantial, with 11.1 million stroke survivors national wide (3, 4) and disproportionally higher in Northern China and rural regions (5). The recurrence rate was as high as 11.2% for all stroke survivors and was 5.7% within 1 year and 22.5% within 5 years among the low-income rural population (5, 6). The limited capacity of primary healthcare system and the overburden of secondary and tertiary hospitals attributed to the fragmented care for stroke prevention in rural China. Thus, effective strategies to improve stroke management are in great need.

Evidence on effective secondary prevention exists, but challenges lie in translating these strategies into routine practice, especially in resource-constrained settings. A primary care-based integrated mobile health intervention (SINEMA intervention) has been designed and implemented in rural China (7). This intervention applied tailored intervention strategies tested in previous programs, such as task shifting, task sharing, and digital health technologies (8–11), and targeted primary healthcare providers and community-dwelling patients who suffered stroke to address the barriers in stroke management. The effectiveness of the SINEMA intervention has been demonstrated, with improvements in blood pressure control, medication adherence and quality of life, and a reduction in disability, stroke recurrence, and deaths were also observed at 12 months among stroke survivors, allocated to the intervention arm compared with the participants who received usual care (12).

Further investigation of the implementation of the SINEMA intervention is very important in order to uncover the implementation outcomes and understand the extent to which effectiveness was affected by other factors (13, 14). Such findings will also help inform the future optimization of implementation of the SINEMA intervention in other settings. Many different frameworks have been proposed for guiding the implementation evaluation (14–18). For example, the RE-AIM evaluation framework proposes five key dimensions—with these being, reach, effectiveness, adoption, implementation, and maintenance—to inform the future implementation, generalizability, and scalability of effective programs (17). The MRC process evaluation framework is another commonly used framework for complex intervention that emphasizes the implementation, mechanisms of impact, contextual factors, and the relationships between these dimensions (14). While many different community-based interventions have adopted a single framework to illustrate certain aspects of program implementation (15), few studies have undertaken a comprehensive evaluation by utilizing quantitative and qualitative data to describe the implementation outcomes and explain how the program was implemented.

This current study examines the implementation of the SINEMA intervention to provide further information for researchers, practitioners, and policymakers. We developed an implementation and process an evaluation framework that combined both the RE-AIM and MRC process evaluation frameworks and utilized both quantitative and qualitative data. This paper reports the findings on implementation outcomes, relates enablers and barriers, and illustrates some potential impact pathways that may influence the effectiveness of the SINEMA intervention and its wider implementation.

METHODS

SINEMA Trial, Study Setting, and Intervention Components

The SINEMA trial was a Hybrid II effectiveness-implementation trial (19). The effectiveness of the SINEMA program was investigated by a cluster randomized controlled trial conducted among 50 villages of rural Hebei Province, Northern China (7). A total of 1,299 rural stroke survivors (an average of 25.5 participants per village) were recruited in the trial. Twenty-five villages, including 637 patients, were randomly allocated in the intervention arm and implemented the SINEMA program over 12 months (12).

The study was conducted in a resource-limited rural county with doubled stroke burden and less than half of the annual disposable income *per capita* than the national average (20–22). In rural China, the general practice and preventive care

are mainly delivered by primary healthcare providers, including village doctors who have received minimum basic medical and pharmaceutical training and can prescribe medications (23). The acute-stage stroke treatments are mainly delivered at county hospitals, while rehabilitative care and follow-up visits are largely unavailable. The outpatient services are paid out of pocket, although the zero-Markup drug policy allows low price of medications in the village clinics and the NCD insurance package reduced the catastrophic health expenditure by reimbursing some outpatient services if patients received care from county hospitals (24).

Built on such context, the SINEMA intervention was designed, and a detailed description of the intervention and trial design has already been published elsewhere (7). In brief, the SINEMA intervention included both provider-facing and patient-facing components (**Figure 1**, left panel). As both receptors and providers of the intervention, village doctors received training, performance-based financial support, and virtual-group peer support. They delivered monthly face-to-face follow-up visits to participants with support from an android-based mobile application (*SINEMA App*). The participants received monthly follow-up visits and daily voice messages dispatched automatically at no cost if they had a phone available. A digital health system, consisting of the SINEMA App and voice messages dispatching system, was developed to support the program delivery (25, 26). Besides, five physicians from township hospitals and one county manager also facilitated the program implementation by providing support and performing quality control.

Evaluation Framework and key Measurement

The implementation and process evaluation was performed based on an evaluation framework that derived from both the RE-AIM framework (17) and the MRC process evaluation framework (14), as described in **Figure 1**. The RE-AIM framework was used to inform the measurement of implementation outcomes, covering program reach, adoption, implementation, and maintenance. Reach was assessed by the absolute proportion and the representativeness of individuals involved in the trial among those identified with stroke history during the village-wide screening. Adoption was measured from the provider perspective by considering the acceptance and uptake of the intervention among village doctors. Implementation was evaluated both quantitatively and qualitatively, including the intensity and the quality of services delivered and the adaptation made by providers. Maintenance was defined as perceived willingness of providers and participants to maintain or scale up the SINEMA program post-trial. We also identified facilitators and barriers that may influence each RE-AIM dimensions. The MRC process evaluation framework was used to investigate the interactive relationship between context,

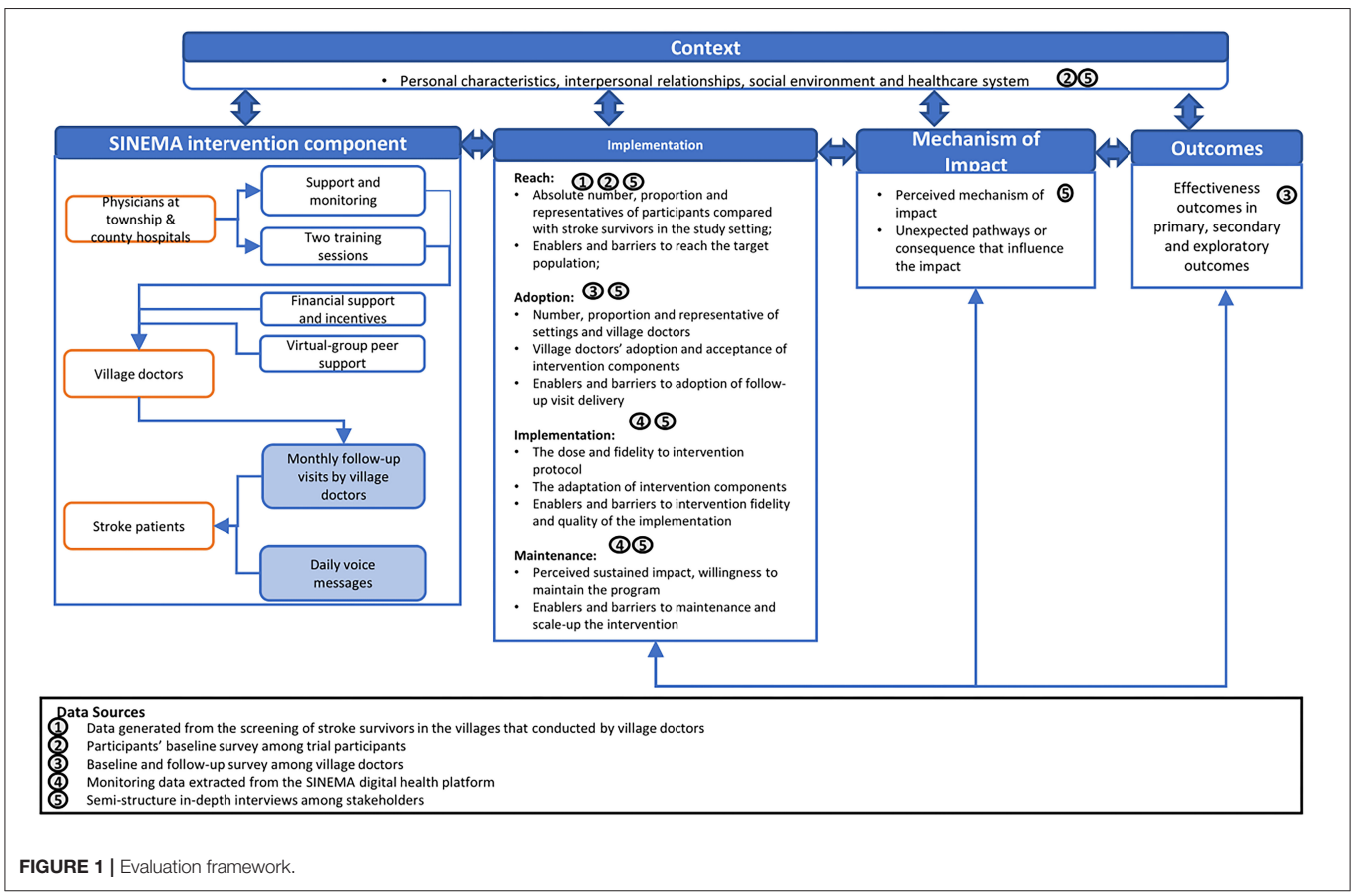


FIGURE 1 | Evaluation framework.

TABLE 1 | Data sources and data collection approach.

Stage of the program implementation	Timepoint of data collection	Type of data collection	Information collected	Data collection approach	Type of data
Pre-implementation	Screening (before 0 months)	Screening data among potential participants	Basic information (socio-demographic, stroke history) of stroke survivors (60 villages, $n = 2333$)	Village doctors filled in the form based on existing health records or interviews with potential participants.	Quan
	Baseline (0 months)	Survey among recruited participants	Basic information and major outcome indicators of recruited participants' (50 villages, $n = 1,299$)	A questionnaire administered by trained assessors via face-to-face interviews.	Quan
	Baseline (0 months)	Self-administered surveys among village doctors	Basic information, attitude and practice of stroke prevention among village doctors ($n = 50$)	Village doctors self-administered the online survey.	Quan
Implementation	Throughout program implementation (0–12 months)	Monitoring data from digital health system	The number of follow-up visits delivered, and voice messages received.	Extracted from digital health platforms.	Quan
	Quarterly (3, 6, 9 months)	Semi-structured in-depth interviews among stakeholders	Implementation situation and relevant factors among participants, village doctors, township physicians.	Project-related researchers conducted semi-structure in-depth interviews.	Qual
Post-implementation	Quarterly (12 months)	Semi-structured in-depth interviews among stakeholders	Implementation situation and relevant factors among participants, village doctors, township physicians.	Researchers who have not involved in the program design and implementation conducted semi-structured in-depth interviews.	Qual
	Follow-up survey (12 months)	Self-administered surveys among village doctors	Attitude and practice of stroke management among 50 village doctors, acceptance and adoption of intervention components among 25 village doctors in intervention arm.	Village doctors self-administered the online survey.	Quan

intervention components, and implementation indicators. The MRC framework also enabled us to assess some of the unexpected pathways or consequences that have not been previously considered.

Data Collection

To obtain a comprehensive overview of program implementation, we collected data in multiple phases from various stakeholders. Multiple data collection approaches were adopted, including self-administered surveys among village doctors, face-to-face interview-based surveys among participants, monitoring data from the digital health system, and in-depth interviews among stakeholders. **Table 1** summarizes the data source and data collection approaches.

Self-Administered Surveys Among Village Doctors

Fifty village doctors participated in a self-administered survey dispatched *via* an online survey platform (Qualtrics, Provo, UT) at both baseline and 12 months. The survey asked about their sociodemographic information and their knowledge, attitude, and practice in the secondary prevention of stroke. Questions about the acceptance and preference in intervention components were also asked at the 12 months among village doctors allocated in the intervention arm.

Screening Data About Potential Participants

Before recruitment of participants, village doctors conducted a screening of residents in the villages to identify people who had been diagnosed with stroke. Village doctors reviewed health records of existing residents or conducted door-to-door screening and provided a list of potential participants with detailed information on age, gender, stroke history, basic communication abilities, and disabilities. This information guided the invitation of the potential participants and provided the information to understand the reach of the program.

Surveys Among Recruited Participants

Among participants, trained assessors performed a face-to-face baseline survey. These data covered a broad range of information, including their sociodemographic characteristics and a series of indicators on health history, health behaviors, and health conditions.

Monitoring Data From Digital Health System

The digital health system, which consisted of the SINEMA mobile application and voice-message dispatching system, also tracked and monitored the delivery of follow-up visits and voice messages. The data on follow-up were extracted from the SINEMA server, which tracked the number of follow-up visits

delivered by each village doctor and for each patient. A third-party message-dispatching system captured the number of voice messages distributed and recorded the number of voice messages answered by the participants on a given day.

Semi-structured In-depth Interviews Among Stakeholders

Semi-structured in-depth interviews were conducted for four waves at the 3rd, 6th, 9th, and 12th months from the initial of the intervention with slightly different purposes and adjusted interview guides. **Supplementary Table 1** reported the number of interviews conducted in each wave and by type of stakeholders. The first two waves focused on how the SINEMA intervention was adopted and implemented and the potential enablers and barriers that may influence the implementation. The third and fourth waves of in-depth interviews focused on the impact of intervention components and the key enablers and barriers that may influence the implementation, effectiveness, and maintenance. The first three waves of interviews were conducted by the research team, while the last wave was performed by an independent investigator who was not involved in any stage of the intervention design and implementation to ensure the objectivity of the evaluation. Stakeholders involved in the intervention implementation (including patients who suffered stroke, village doctors, physicians at township hospitals, and a county coordinator) were invited to participate in the interviews in all four waves. The village doctors who allocated in the control arm were also invited at the fourth wave.

We used purposive sampling to ensure the diversity of the participants and the coverage across villages and townships. At each wave of data collection, the research team identified the villages to ensure that at least one village from each township was selected. All intervention villages were covered across the four waves of data collection. Within each village, the research team interviewed the village doctor and selected one or two participants based on their availability, willingness, ability to communicate, and demographic characteristics to ensure representatives of the participants. Interviews were conducted either in the village clinics or at the homes of the participants. The physicians at township hospitals and the county coordinator who involved in the study were also interviewed at their working places. All interviews lasted between 20 and 40 min and were audio-recorded with verbatim transcripts for data analysis.

Data Analysis

The data analysis of quantitative and qualitative data was performed independently, and then the findings were embedded within the designed evaluation framework (27). This approach gathered the quantitative and qualitative data by not only demonstrating the key dimensions of program implementation but also exploring its variations, facilitators, and barriers. For quantitative data, descriptive analysis was performed by using STATA 15.0 software. *T*-test and chi-square test were used for comparison between groups of participation status.

For qualitative data, the thematic analysis approach was applied with the following steps: First, we familiarized ourselves with the data by reading all transcripts. Due to the large numbers of interviews administered during the four waves ($n = 98$), we went through all the transcripts and classified the quality of the transcripts based on the quality of the interview, the amount of information contained in the conversation, the types of stakeholders, and the wave of data collection. Forty-three (43.8%) transcripts were classified into the high-quality group and received full analysis with line-by-line coding from two researchers, while others received rapid coding from one researcher (**Supplementary Table 1** reports the number of transcripts involved in the full analysis). Second, we developed the coding structure. Researchers (EG, LS) coded at least one transcript from each type of stakeholders from all four waves. Data were coded on a line-by-line basis, and data were initially organized according to the topic of questions from the interview guide. We then inductively derived codes with a more elaborate hierarchical coding scheme by considering different intervention components and the dimensions of the evaluation framework. Researchers discussed the coding structure, and issues were resolved by consensus. Third, transcripts classified as high-quality transcripts were coded by two researchers independently, with at least half transcripts were double coded to ensure the objectivity and transparency of the process. Any discrepant interpretations were discussed between the researchers and across a broader research team. The researchers also scanned the remaining transcripts that were classified as low-quality transcripts to avoid missing information. Fourth, themes were developed to map each dimension of the framework by reading the coded data and the original transcripts to ensure that the themes were authentic and rooted in the data. All the quotes involved in the manuscript were translated from Chinese to English.

RESULTS

Utilizing the evaluation framework described in **Figure 1**, we summarized the context, implementation outcomes, and impact pathways of the SINEMA intervention. Characteristics of stakeholders involved in the in-depth interviews are summarized in **Tables 2**, **3**.

The Context for Implementing SINEMA

Characteristics of Villages and Healthcare System in the Region

The SINEMA intervention was implemented in Nanhe County, rural Hebei, China. The 50 villages in the rural regions scattered around urban areas of the county, where the county hospitals located with an average distance of 14.5 km (**Supplementary Table 2**). During the interview, majority of the participants identified village clinics as their first contact point of the healthcare system to address their day-to-day health needs; some participants also mentioned that they sought care from other healthcare facilities, including private clinics and pharmacies within the village or nearby villages, township hospitals, county hospitals, and hospitals in nearby cities. Several factors may determine different choices of healthcare facilities, such as healthcare needs of the participants, trust and

TABLE 2 | The characteristics of the participants who were involved in the screening, recruited in the trial, and participated in the in-depth interviews.

Characteristics of participants	Stroke survivors screened [a] (*n* = 2,081)	Stroke survivors participated in the trial (*n* = 1,299)	Stroke survivors involved in the interviews (*n* = 51)
Sex, % female	909 (43.7%)	553 (42.6%)	23 (45.1%)
Mean age at baseline (SD)	67.1 (9.2)	65.7 (8.2)	65.6 (7.7)
Stroke type [b]			
Ischemic	1,731 (83.2%)	1,119 (86.1%)	42 (82.4%)
Hemorrhage	331 (15.9%)	176 (13.6%)	9 (17.6%)
Not specified	19 (0.9%)	4 (0.3%)	0 (0%)
Self-report medicine taking during screening or baseline survey			
Antiplatelet	1,357 (65.2%)	852 (65.6%)	19 (37.3%)
Satin	699 (33.6%)	340 (26.2%)	11 (21.6%)
Anti-hypertensive medicines	1,675 (80.5%)	1,030 (79.3%)	42 (82.4%)
Had experienced stroke recurrence	603 (29.0%)	378 (29.1%)	10 (19.6%)
Visited to village clinics in the past month	1,600 (76.9%)	795 (61.2%)	40 (78.4%)
Difficult to get out of bed [c]	160 (7.7%)	27 (2.1%)	0 (0%)
Having basic communication ability	1,919 (92.2%)	1,259 (96.9%)	50 (98.0%)
Duration since the first stroke event			
<3 years	549 (26.4%)	357 (27.5%)	10 (19.6%)
3–5 years	443 (21.3%)	329 (25.3%)	18 (35.3%)
6–9 years	479 (23.0%)	257 (19.8%)	7 (13.7%)
≥ 10 years	610 (29.3%)	356 (27.4%)	16 (31.4%)

[a] *Only stroke survivors who were from the 50 eligible villages were accounted. Stroke survivors from 10 villages that did not meet the cluster eligible criteria were excluded from the analysis.*

[b] *For the participants who had multiple stroke experiences, the type of stroke accounted for the latest stroke event.*

[c] *The participants who had limited walkability but able to visit to the village clinics with support of family caregivers were included in the trial, otherwise were excluded.*

relationship with doctors, insurance coverage, and the quality of available services.

> *"I visited village clinics quite often. If I am available, I will come here to measure my blood pressure. Not every day, but one time per 3 to 5 days." –Participant, 3-month interview*
> *"I got my medicines from the No. 2 county hospital because I could get reimbursement from the hospital. I need to pay out of pocket if I get medicine from the village clinic."*
> *- Participant, 6-month interview*

Interpersonal Relationships and Support for Care

According to the baseline survey, 410 (64.3%) participants had a family caregiver, mainly their spouses, daughters or daughters-in-law, as most of their sons or sons-in-law were working outside of villages. Family caregivers played the most critical role in daily life and treatment adherence of the participants. Although the participants mentioned that they know other village residents pretty well, people who had insufficient support from the family members could not get extra help from other neighbors or friends as stigma related to stroke existed. Some participants mentioned that they were unwilling to depend on other people or discuss their health conditions with other neighbors in the villages.

> *"I don't know other people who also had this disease (stroke), but there should be some. I don't like to bother others. I could do most of the things by myself. I don't like to talk too much with others as I don't want to become a topic of their gossips."– Participant, 6-month interview*

Personal Characteristics

At the personal level, the study participants were a vulnerable population group with low socioeconomic status. Among the participants who received the SINEMA intervention, 264 (41.1%) had received no formal schooling at all, 276 (43.3%) had more than two other chronic diseases, and 179 (28.1%) were experiencing moderate to severe disabilities (**Table 2**). About 7.3% participants experienced depression at the baseline, and many of them experienced various levels of cognitive impairment or other issues related to stroke or aging, which brought further obstacles to understanding and accepting the intervention, building their confidence, and improving self-efficacy and self-management behaviors.

> *"My health condition is getting poorly, and I don't think I could be fully recovered as I am getting old anyhow."—Participant, 3-month interview*
> *"Some stroke patients had impaired brain function and poor memory. They thought there was no big difference whether they take medicine or not."—Village doctor, 6-month interview*

Implementation Outcomes

This section presents the results on implementation outcomes regarding program reach, adoption, implementation, and

TABLE 3 | Characteristics of village doctors who were involved in trial, implemented the SINEMA intervention, and participated in the in-depth interviews and their perceptions of stroke care and SINEMA intervention components.

Characteristics and perceptions of village doctors	Involved in the trial ($n = 50$)	Allocated in intervention arm ($n = 25$)	Participated in-depth interviews ($n = 27$)
Age, mean (SD), years	46.0 (6.4)	46.1 (7.3)	46.0 (7.5)
Sex, % female	8 (16.0%)	3 (12.0%)	3 (11.1%)
Education, *n* (%)			
High school or equivalent	29 (58.0%)	15 (60.0%)	16 (59.3%)
Junior college	18 (36.0%)	8 (32.0%)	9 (33.3%)
College	3 (6.0%)	2 (8.0%)	2 (7.4%)
Years as village doctor, mean (SD)	24.3 (7.4)	24.2 (8.6)	24.2 (8.8)
Self-evaluation on existing workload at baseline			
Very high	12 (24.0%)	6 (24.0%)	6 (22.2%)
High	26 (52.0%)	12 (48.0%)	13 (48.1%)
Acceptable	12 (24.0%)	7 (28.0%)	8 (29.6%)
Agreement on the top three most essential intervention components			
Face-to-face training	NA	20 (80.0%)	NA
SINEMA App (training module)	NA	13 (52.0%)	NA
SINEMA App (follow-up module)	NA	14 (56.0%)	NA
SINEMA App (reminder module)	NA	5 (20.0%)	NA
SINEMA App (performance statistics module)	NA	3 (12.0%)	NA
Financial compensations and incentives	NA	4 (16.0%)	NA
Reminders and feedbacks from township physicians	NA	7 (28.0%)	NA
Agreement on the following statements related to the perceptions and attitude on stroke care at 12 months of the intervention implementation			
I am aware of the health conditions of stroke patients in my village.	33 (67.4%)	18 (75.0%)	19 (70.4%)
I am confidence in prescribing the most appropriate medicines for stroke patients.	41 (83.7%)	21 (87.5%)	21 (77.8%)
I am confident in providing support and guidance to stroke patients.	42 (85.7%)	24 (100.0%)	25 (92.6%)
My patients trusted me.	39 (79.6%)	22 (91.7%)	23 (85.2%)
All my stroke patients could adhere to my suggestions and prescriptions.	24 (49.0%)	15 (62.5%)	15 (55.6%)
Agreement on the following statements related to the impact of the intervention at the 12 months of the intervention implementation			
The frequency of getting blood pressure monitoring had improved among my patients.	NA	22 (91.7%)	NA
The program led to a clear improvement in blood pressure control among my patients.	NA	23 (95.8%)	NA
The program led to a clear improvement in medication adherence among my patients.	NA	23 (95.8%)	NA
The program made more patients in my villages proactively do physical activities.	NA	24 (100.0%)	NA
My patients rely on me more after the project.	NA	21 (87.5%)	NA
The project improved my authority in the village.	NA	21 (87.5%)	NA

NA. not applicable, as the questions were only for village doctors in the intervention arm.

maintenance. The enablers and barriers that influence each domain of the implementation outcomes are summarized in **Table 4** and described briefly below.

Reach

A median of 1.7% [an interquartile range (IQR): 1.3, 2.4%] residents in recruited villages was screened with stroke history within 50 eligible villages. At the village level, the proportion of people recruited in the trial accounted for a median of 70.1% (IQR: 58.3, 87.%) among all stroke survivors screened in the villages. The recruited participants were similar to those screened with self-reported stroke history in the region, except stroke survivors who reported bedridden were generally not recruited. The participants could represent a general group of rural community-dwelling stroke survivors with a median of 5.3 (IQR: 2.3, 9.8) years of stroke history since the first event, stable health conditions, and basic communication abilities (**Table 2**). The findings from the in-depth interviews revealed that knowledgeable village doctors who reviewed existing health records and performed door-to-door screening enabled the reach of the program to the targeted population in a timely fashion. People employed outside of the villages or having families living outside may be left out of the program (**Table 4**).

Adoption of the SINEMA Program Among Village Doctors

Fifty eligible villages out of 109 villages within five townships were formally invited and recruited in the study. These villages represented typical middle-to-large-sized rural villages of Northern China with a median of 2,422.5 residents per village

TABLE 4 | Enablers and barriers on implementation indicators.

Indicators	Enablers or Barriers	Some selected quotes
Reach	***Enablers:*** • Knowledgeable community healthcare workers (village doctors) • Existing health records and door-to-door screening	• *"I had screened 35 potential participants, but there were a few patients who I failed to contact (during the screening and recruitment stage). A few of them were not at home when I tried to contact them, after several try, I gave them up. There is another patient who was hospitalized; thus, it ended up with 23 participants (after the eligibility screening)."—Vilage doctor 6, 9- month*
	Barriers: • Participants' migration for working and family purposes	
Adoption of the SINEMA program among providers	***Enablers:*** • Knowledge, capabilities and confidence gained from training and previous experience • Perceived benefits on residents and themselves • Perceived credibility from the top-down approach	• *"The training sessions were helpful. I learned some new knowledge, which is helpful in my work. Especially the knowledge on the effect and side-effect of these essential medicines. Following what the chief physician told in training, I started to adjust the medicines for patients during the follow-up visits."—Village doctor 2, 3-month* • *"I was in charge of the blood pressure management service (as part of the Basic Public Health Services). Even if there is no such program focus on stroke patients, I need to manage more than a hundred of patients with hypertension. Delivering follow-up visits (quarterly) is my job. Thus, for some of our participants, this intervention is an add-on service to the Basic Public Health Services."—Village doctor 10, 12-months* • *"The financial compensation did not attract me at all. I participated because the township physicians invited me and said this program is good for our residents and I could also learn a lot from the program."—Village doctor 8, 9-month*
	Barriers: • Poor technology literacy • Lacking needs and motivations in specific tasks	• *"At the beginning, I wasn't too familiar with the app and the procedure, but I grasp the skills and could deliver follow-up visits smoothly after several rounds (of monthly visits)."—Village doctor 1, 6-month* • *"It (APP) reminded the date of follow-up visits for each patient, but I started follow-up visits before the system reminds me. It could be useful, but I don't need it."— Village doctor 5, 12-month*
Implementation and fidelity of follow-up visits	***Enablers:*** • Patients' needs and willingness in improving health and Providers' responsibilities and efforts • Trusted doctor-patient relationships • SINEMA App standardized procedures • Support and quality control from township physicians	• *"People won't reject help on improving their health. So there is no much difficulties for me to implement the follow-up visits."—Village doctor5, 12- month* • *"I manage all stroke patients in this village. They seldomly visit other providers. I know their health condition quite well and patients relies on me, and they get used to this relationship… This relationship was not established in a day; it has been several years."—Village doctor 3, 12-month* • *"Before the program, I delivered services to patients, but there is no standard. I am now delivering follow-up visits by following the app. It is simple and comprehensive. The focus is not on the diagnosis, but on follow-up visits to communicate about their conditions. Having this procedure is helpful."—Village doctor 1, 12-month* • *"For some village doctors, I need to remind them multiple times (to complete the follow-up visits). I checked their follow-up records to see whether there is any patient with extremely high blood pressure that needs more attention. I also talk to village doctors, if I found some unreasonable records." —Township physician, 3-months*
	Barriers: • Pre-existing heavy workload and competing programs • Participants' low awareness, adherence or cooperation • Technical difficulties (unstable internet access)	• *"The workload was increased, but I conduct follow-up visits when I was not too busy."—Village doctor 3, 12-month* • *"For some patients, they don't give enough attention to it. You call them several times, but they do not come. For some people who are elderly or have low awareness, they don't care it too much. You have to visit them and let them know the importance."—Village doctor 5, 3-month* • *"Sometimes, the internet was not stable, and the data cannot be uploaded. Thus, it is better to take the information down and then upload later."—Village doctor 5, 6-month*
Implementation and fidelity of voice-message components	***Enablers:*** • Perceived benefits • Free at no cost • Simplicity in content and dispatch way • Nudges and suggestions from village doctors	• *"I received it (voice message) every day. I seldom had a phone call, but I receive your message every day. I learn things from listening to it, and there is no cost."—Participant 14, 12-month* • *"I can receive it every day. It was useful for me. I put my phone near my bed. It reminds me of taking medicines and do exercise. It (voice messages) told me many knowledges and I followed it."—Participant 3, 6-month* • ***"****During each follow-up visits, I will remind them to continue taking medicines and pick up the voice messages if they can."—Village doctor 2, 12-month*
	Barriers: • Phone ownership and use pattern • Lacking individualized contents to meet diverse needs • Hearing problems	• *"I have received voice messages, but not every day. Sometimes, I went out but didn't bring my phone with me; then I could not receive the messages."—Participant 1, 3-month* • *"I don't use the phone quite often. I don't know how to use it. My phone shared with my family members. When it put here, I could listen to it; otherwise, I cannot. I have received some…. There was once when I clicked the button, but I cannot hear it."- Participant 5, 3-month* • *"Some people they seldom pick up phone calls. Especially the elderly, they have hearing issues, and they cannot pick up the call."—Township physician 1, 6-month*

(Continued)

TABLE 4 | Continued

Indicators	Enablers or Barriers	Some selected quotes
Maintenance	***Enablers:*** • Perceived cost-benefits on patients and providers • Supportive environment Integration with other services	• *"I don't foresee too much challenge (in continuing the service). It is a basic service at no cost for patients. They will accept it… It (the SINEMA program) also didn't bring too much burden on me. At least, it helped the patient-doctor relationship. I also improved my skills by communicating more frequently with many patients."—Village doctor 1, 12-month* • *"It also depends on the village doctors. Some village doctors who are more responsive to the patients in the village will continue doing it even without the program and financial support." —County manager, 12-month* • *"This program should be integrated with other activities in both urban and rural settings. Such as the program should be integrated with the basic public health services so that the program could get more policy support." —Township physician, 12-month*
	Barriers: • Workloads and competing programs • Lacking mechanism to share the implementation cost	• *"If we included it (the SINEMA program) as routine services, the workload could be heavy. There are patients with stroke, coronary heart diseases, and mental disorders. If all counted, the workload will be heavy to maintain the current frequency of follow-up visits."—Village doctor 4, 12-month* • *"It is quite important to identify how the cost could be sustainably covered. Maybe the program could be integrated with public health programs or other existing programs. It is hard to allocate funding if there is no support from the county or above authorities."—Township physician, 12-month*

(IQR: 1,772, 3,600) (**Supplementary Table 2**). All village doctors ($n = 25$, mean age: 46 ± years old, 12% females, **Table 3**) from the intervention arm adopted the SINEMA intervention. A few village doctors expressed that they had seldomly used mobile apps before the program and experienced a learning curve. Some village doctors did not use all modules of the SINEMA App due to lacking needs and motivations.

> *"I seldomly used the performance module and checked these statistics. I only check whether I missed any follow-up visits. I completed my tasks while evaluating my work is other people's tasks." —Village doctor, 6-month interview*

Twenty village doctors (83.3%) considered the training sessions as the most valuable and essential component. They believed that knowledge and skills gained through training sessions and previous experience in delivering similar services mitigated the learning curves of program adoption and enabled them to deliver the follow-up visit to the participants. The follow-up visits module and the training modules of the SINEMA App received 14 and 13 votes, ranked as the second and third most essential components by village doctors. Financial incentives were not ranked highly as essential intervention components, as some village doctors mentioned that the financial incentives did not impact much on their decisions of program adoption or the amount was not high enough to be a driver; rather, the perceived benefits to the residents that village doctors learned from the communication with county and township physicians, as well as the opportunities of receiving training and guidance from experts, were major enablers that determined their adoption.

> *"The payment didn't influence me much. It was not the case that if I got more money, I could work better.… Even if you stop paying me, it will not influence me much. Similar to Basic Public Health Services, the project brought benefits to our residents. After participating in this project, I could better manage my patients; I met them face-to-face monthly. If someone comes to check my work, I do not need to lie to him or her, as I have done this work as required. There is a benefit."—Village doctor, 9-month interview*

Implementation and Fidelity of Follow-Up Visits

Twenty-five village doctors in the intervention arm performed an average of 291.5 (SD: 29.5) follow-up visits over 12 months. Among 637 participants, 564 (88.5%) received no <12 follow-up visits as full dose per protocol. The participants who received the full dose were more likely to be those without family caregivers and had multiple chronic disease conditions (**Supplementary Table 3**). Although the quantity of the follow-up visits was high, the quality of follow-up visits varied. Follow-up visits were delivered mainly at the village clinics or homes of patients. Some village doctors scheduled all follow-up visits on certain days of a month; others performed the follow-up visits as an add-on service once participants came to clinics. Based on descriptions of village doctors of the key steps of follow-up visits, we found that many village doctors adjusted the procedure by skipping some steps.

> *"I opened the app, measured the blood pressure, and asked questions based on the app. If needed, I asked all questions (on the app), or I selected key questions to ask.… For example, if the patient has a stable situation, or I know him or her quite well, I may skip the question about hospitalization and only emphasize medicine use. For the side effect of medicines, it does not need to be asked each time."—Village doctor, 6-month interview*

Factors that influenced the quantity and the quality of follow-up visits are summarized in **Table 4**. Internal enablers included perceived responsibilities from providers, good preexisting patient-doctor relationship, and strong willingness from the participants. The SINEMA App and the quality control from the township physicians promoted the intervention fidelity. Village doctors mentioned that the designed SINEMA App played a supportive role in standardizing follow-up visits and assisting the information management. However, the required internet access

may also bring some barriers when there is no stable internet access. Village doctors and township physicians also stated that preexisting heavy workload and lacking compliance of patients might limit the quantity and the quality of follow-up visits, but the top-down support and quality control may encourage high fidelity.

"He (the township physician) is playing supervision and encouraging role. For example, if he found that there are certain patients that left without follow-up visits, he will remind me to finish at my earliest." -Village doctor, 12-month interview

Implementation and Acceptance of Voice Messages

About half of the participants answered the voice messages during the program implementation on a given day among those who agreed to receive voice messages. The answering rate was maintained over the 12-month implementation period (**Supplementary Figure 1**). The implementation of voice messages was influenced by phone use patterns and characteristics of the participants. Both quantitative and qualitative data indicated that the participants who had their phones without sharing with family members picked up more voice messages (**Supplementary Table 3**). Most participants mentioned that they considered voice messages as good reminders and a reliable source of getting information and favored the simple content, repeated leading sentences, and local dialect. However, some participants who experienced cognitive declines could only recall the leading sentences. In contrast, a few participants stated that they dropped out of the voice-message component halfway through as the contents were too simple for them.

"It (the voice messages) says about taking medicines and doing exercise. I cannot remember other details.... Nevertheless, it is useful as it reminds me in the morning, and it shows care about me." –Participant, 12-month interview

Maintenance

During the interview, most of the village doctors and the participants expressed their willingness to continue the program. Some village doctors suggested expanding the participants to individuals with other chronic conditions, but others also expressed concerns about the workloads if the program expanded to a larger population group. Village doctors perceived that impacts of intervention could be maintained as participants have established good habits in visiting village clinics and taking medicines. Some village doctors also expressed a spillover effect of the program to other existing services and programs.

"During these months, they have developed a habit. Participants have kept a good relationship with me, and I expect they will continue visiting me as often.... I also used this program approach for the Basic Public Health Services. I planned for the follow-up visits regularly so that I don't need to be rushed or to lie to people who check my work."- Village doctor, 9-month interview

The county and township managers also expressed that the maintenance of the program may be influenced by the scope of targeted participants, workloads of village doctors, and the financial mechanisms to cover the cost of the program delivery. Integrating the program with the existing programs and information systems was also mentioned as critical factors for future maintenance and scaling-up.

"This program should be integrated with other activities in both urban and rural settings. Such as the program should be integrated with the basic public health services so that the program could get more policy support." -Township physician, 12-month interview

Effectiveness and Perceived Mechanism of Change

The trial results on the effectiveness of the SINEMA intervention have been detailed elsewhere (12). In brief, the intervention achieved a significant reduction in systolic blood pressure (between-arm difference: −2.8 mmHg, 95% CI: −4.8, −0.9; $p = 0.005$), and improved medication adherence, physical activities and quality of life as secondary outcomes and reduced stroke recurrence, hospitalization, and deaths as exploratory outcomes. The impact pathways of the intervention were further revealed through in-depth interviews.

The Influence on Confidence and Practice of Village Doctors

Village doctors acknowledged that they prescribed medicines mainly by following the previous prescriptions before the intervention. Training and support improved their awareness of clinical guidelines and encouraged them to provide more guidance and suggestions during visits of patients. Some village doctors also mentioned that they considered adherence and long-term benefits of participants while prescribing medicines. The survey at 12 months post-baseline also indicated increased perceived confidence in prescribing evidence-based medicines and supporting patients (**Table 3**).

"Through this project and the training session, I changed my mind by considering not only the medicine price but also the effect. For patients who had poor adherence, I suggested them to change to prolonged antihypertensive medicines. Although the cost is a little bit higher, the effect in controlling blood pressure was largely improved."—Village doctor, 9-month interview

Interactive Impact of Voice Messages and Follow-Up Visits on Self-Management and Doctor-Patient Relationships of Patients

An interactive function between voice messages and follow-up visits has been identified from interviews. We found that the face-to-face communication between providers and participants about the content of voice messages facilitated the general adoption and acceptance of the voice message component; meanwhile, the daily voice messages reinforced adherence of patients to the suggestions of the doctors delivered during the follow-up visit component. Village doctors stated that voice messages supplemented their role to provide extra assistant for the participants to improve their treatment adherence. Some participants mentioned that they discussed the contents of voice

messages with village doctors, which reinforced them to pick up voice messages and translate health education information to self-management activities, as they received consistent information from the village doctors and voice messages.

> *"The voice messages improved their awareness. We (village doctors) cannot observe and remind patients every day, but we can only remind patients when we have a face-to-face appointment. The voice messages add to that to remind patients so that they consider how they should take medicines and do exercise as the doctors told them; then, they establish a good habit."—Village doctor, 9-month interview*

The frequent face-to-face visit and daily voice messages also help village doctors build or maintain a good relationship with the participants. About 88% of the village doctors acknowledged that their patients relied more on them due to the intervention, and the project improved their authorities in the village (**Table 3**). The patients also mentioned that they felt more care from the village doctors through voice messages and follow-up visits.

> "*Through the follow-up visits, patients trust me more. (I conducted) follow-up visits once a month, without asking them to pay, and, if they don't come, I will call them to remind them. There were several patients who sought care from other providers if they were sick (before the intervention). But I followed up with them and they said they were willing to my suggestions because they think my suggestions were helpful. Now, our relationship is not bad. I say "hi" to them if I meet them on the street."—Village doctor, 12-month interview*
> *"The doctor measured my blood pressure frequently. He also reminded me to pick up the call or visit him if I forgot. I could feel that he cares about me."—Participant, 9-month interview*

The Influence of Personal and Interpersonal Characteristics on Program Effectiveness

The influence of follow-up visits and voice messages varied by person. Personal factors, such as personal education background, cognitive functions, and self-efficacy, may influence their acceptance of information delivered, the impact on behavior changes as intermediate outcomes, and the long-term effects on health outcomes.

> *"Almost all of them could understand the content, but some people, such as those who have received education or with strong willingness to improve their health, could absorb more information from voice messages; for others, they may just listen to them but not take anything from them."—Village doctor, 3-month interview*
> *"This disease cannot be fully treated. I am becoming older, and I don't think I could get better. Even if I don't have a disease, my health condition will get worse anyhow."- Participant, 12-month interview*

Village doctors also mentioned that family caregivers are another channel that they exchange information with if family caregivers could become the extra support to encourage participants to adhere to the treatment and self-management activities.

> *"Some stroke patients had impaired brain function. It is hard to communicate with them… For these patients, I told their family members about their medicine prescriptions to remind them. But there is also a case that the patients' wife spends the whole day playing mahjong and seldomly take care of the patient. For this type of patients, their family are not helpful at all." -Village doctors, 12-month interview*

The Impact of Other Healthcare Services or Programs

Some participants mentioned that they also sought healthcare services from other healthcare providers. Inconsistent information gained from different sources brought obstacles for participants to adhere to the suggestions provided by village doctors during the follow-up visits, thus may limit the role of the village doctors.

> *"I take medicines from several places (providers). I could pay less (out of pocket) from other places if I know the person."—Participant, 12-month interview*
> *"About half of participants visited county hospitals and got their medicines prescribed there because they could get reimbursement if they have special non-communicable disease insurance. For them, I don't interrupt their medicine use—I only encourage them to continue taking medicine."—Village doctor, 9-month interview*

Besides, some other existing services or programs had some overlap with the SINEMA intervention, as the participants with hypertension and diabetes could have been targeted by these programs already. Although not all the participants involved in the trial could access other services, it may attenuate the observed between-arm differences of the intervention.

> *"There are other health education and follow-up visits services as part of the Essential Public Health Services. We provided village-wide health education sessions and advocated basic public health services, including blood pressure management. Our residents (including stroke patients) may improve their health literacy and be aware of the benefit of taking medicines and keep a good diet (during the study period due to these programs)." —Village doctors from the control arm, 12-month interview*

DISCUSSION

Based on the RE-AIM and MRC Process Evaluation frameworks, this study gives new evidence concerning implementation outcomes and the factors relevant to the implementation and effectiveness of the SINEMA program. The SINEMA intervention successfully reached a representative group of community-dwelling stroke survivors and brought significant benefits to health and wellbeing of participants. Although some of the village doctors made adaptations to the program delivery approach, all of them adopted the intervention and delivered with high fidelity. Some potential impact pathways were identified, such as empowerment among village doctors in clinical decision-making and the interactive impact on stroke survivors *via* both human-delivered services and technology-enabled components. The contextual factors, including personal and interpersonal characteristics and healthcare system and environment, also

interacted with intervention components and provided some further explanations about program effectiveness.

For complex interventions, implementation and process evaluation is very important to provide additional information on how the different intervention components were implemented and interacted with the context. We innovatively derived a framework from both the RE-AIM and the MRC process evaluation frameworks to provide a more comprehensive view of program implementation and effectiveness. Following the recent suggestion about the use of RE-AIM framework (17), we used mixed methods to report both the findings on each dimension of RE-AIM and illustrated the facilitators and barriers that may influence these dimensions. The MRC framework provided another lens through which we are able to understand the context and potential impact pathways for the effects of the SINEMA program. The study findings provide learnings for policymakers, health practitioners, and researchers regarding the future adaptation, optimization, and implementation of the SINEMA program.

Our findings demonstrate that future adaptation of the SINEMA intervention needs to consider the coherence and relationship of its intervention components. Many of the SINEMA intervention components, such as capacity building, task shifting, home-based follow-up visits, and technology-enabled tools, have been investigated previously (9, 28–32). However, the impact of these components was not always positive and additive in the real-world setting (15). Indeed, our study demonstrates that the intervention components were not independent of one another but rather interacted and were synergistic with each other. The components of training, financial compensation, and top-down support from township physicians became the facilitators for adopting and implementing the follow-up visit component. The follow-up visits and voice message components interactively influenced behaviors of providers and patients, and improved doctor-patient relationships. This finding re-emphasized the core concept of the Chronic Care Model, that is, the importance of building a provider-patient alliance with productive information exchange (33, 34). This suggests that future implementation of community-based services should address the barriers of program adoption at levels of both providers and patients by combining effective strategies into a streamlined program to flat the adoption curve and to overcome implementation challenges.

Our study also emphasized the unique role of information and communication technologies in supporting program delivery and program evaluation. The use of information and communication tools generated data to support intensive monitoring of program implementation. Future studies could further investigate the best approach of using digital solution-generated data for providing real-time feedback about program implementation. The digital health system also supported intervention delivery. In line with previous studies (35), we also identified barriers to digital health adoption among individuals who shared devices or had low-technology literacy. Interestingly, our study also showed a reinforcement loop between the technology-enabled component and the human-delivered component that were not considered in many previous studies (36–38). Such findings indicated some spillover effect of digital health solutions beyond service delivery if the technology-enabled component could be embedded into the healthcare system. The results added further evidence on mHealth-enabled interventions on chronic disease management, whose importance was amplified due to the COVID-19 pandemic (39).

Findings from our study also illustrated the complex impact of stakeholders and the context on program implementation and effectiveness. Variability of program implementation across villages was attributed to attitude, capacities, and practice of village doctors. In line with previous studies that identified barriers to normalizing new interventions when a competing program exists (15, 40, 41), we also noticed the challenge to sustain SINEMA intervention when primary healthcare providers rated their workloads as high with limited capabilities and incentives in stretching to other new tasks. However, our study illustrated the benefits of the SINEMA intervention to existing services in reducing the learning curve and the spillover effect of the SINEMA intervention on improving the quality of existing preventive services. This finding highlighted the importance of considering the synergies and integration across programs when introducing and implementing a new program.

The impact of the intervention was different among subgroups. For example, the quantitative subgroup analysis showed that the impact of the SINEMA intervention was consistently positive except for males and those who were <65 years (12), while the qualitative data added further information by showing that different self-efficacy, doctor-patient relationships, and preference in service utilization may explain the variation in health outcomes. Males and younger participants are more likely to take some casual work outside of villages and had more chance to interact with other healthcare providers beyond the intervention scope, such as private health clinics or pharmacies. As the qualitative findings suggested, services provided from these facilities may disrupt the uptake or the intervention impact pathway when conflicting information was delivered to the participants. These findings provide more explanations about program effectiveness and offer suggestions for future program optimization to deliver more personalized information and involve other available private services in the program.

Our study has several strengths. We developed an evaluation framework that combined the RE-AIM and MRC frameworks and used data collected at four time points to comprehensively understand the implementation of the SINEMA intervention. Data extracted from the digital health system provided additional real-time monitoring on program uptake and implementation beyond the traditional observational approach. In addition, we combined the findings from quantitative and qualitative data to demonstrate how the program was implemented and to identify the facilitators and barriers that may influence the program implementation and effectiveness.

There were also some limitations. First, the definitions and measurement of RE-AIM indicators were defined based on the best estimation of available data. For example, we measured the program reach by analyzing the representativeness of our participants among screened stroke survivors in

the study settings rather than more broad scope due to the limited access of data. The measurement of program maintenance is also limited to willingness of the participants to maintain the program, which may overestimate the long-term program maintenance in actual practice. Second, the trial was conducted in 50 villages of one county in rural Northern China. Villages and participants recruited in the trial shared many similarities, limiting the observed diversity and the external validity of the study findings. Future studies could explore the adaptation and implementation of the program when the program is disseminated to diverse settings and populations. Third, as a common limitation of complex intervention, we could not fully explain the impact mechanism or distinguish separate effects from different components. However, we illustrated some potential impact pathways that were not explained by quantitative intermediate outcomes and provided some unexpected pathways or consequence that influenced the impact. Such findings are important for considering the scale-up and adaptation of the program to other contexts and settings.

CONCLUSION

The SINEMA intervention reached a representative stroke patient group in rural China, adopted by village doctors and implemented with relatively high fidelity. The program benefited both providers and patients, but the impact was diverse by characteristics of individual, interpersonal relationships, and other services in the setting. There is a need to explore the adaption of the SINEMA model in other settings and for other chronic diseases.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Duke University Ethical Review Board, Duke Kunshan University Ethical Review Board, Beijing Tiantan Hospital Ethical Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EG drafted the manuscript. EG, WG, QL, and LY contributed to the conceptualization, design of the study, and data collection. JB and LY contributed to acquiring study funding. EG, LS, and JT performed data analysis. EG, LS, HX, QL, and LY were involved in the data interpretation. QL, HX, JB, JM, TJ, BO, and LY contributed to the revision of the manuscript. All authors read and approved the final manuscript.

FUNDING

The study is funded by the United Kingdom Medical Research Council, Economic and Social Research Council, Department for International Development, and Welcome Trust (Grant No. MR/N015967/1). EG is supported by the Melbourne Graduate Scholarship. EG and BO were also supported by the National Health Medical Research Council (Grant 1170937). The publication of this article is supported by Shandong Center for Disease Control and Prevention, School of Population Medicine and Public Health, Peking Union Medical College.

ACKNOWLEDGMENTS

We would like to thank Dejin Dong and his team (Xingtai City Center for Disease Prevention and Control), Zhenli Xu and his team (Nanhe County Center for Disease Prevention and Control), Congxuan Wu and his team (Ren County Center for Disease Prevention and Control), and all staff members from township health care centers who support the program implementation and data collection. We would also like to thank Shengjie Zhang, Duan Zhao, Ying Yang, and Wenbo Song in support of the data collection. We would also like to thank all the study participants (including patients, village doctors, and township physicians in the region).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.774907/full#supplementary-material

REFERENCES

1. Avan A, Digaleh H, Di Napoli M, Stranges S, Behrouz R, Shojaeianbabaei G, et al. Socioeconomic status and stroke incidence, prevalence, mortality, and worldwide burden: an ecological analysis from the global burden of disease study 2017. *BMC Med.* (2019) 17:191. doi: 10.1186/s12916-019-1397-3
2. Krishnamurthi RV, Ikeda T, Feigin VL. Global, regional and country-specific burden of ischaemic stroke, intracerebral haemorrhage and subarachnoid haemorrhage: a systematic analysis of the global burden of disease study 2017. *Neuroepidemiology.* (2020) 54:171–9. doi: 10.1159/000506396
3. Li Q, Wu H, Yue W, Dai Q, Liang H, Bian H, et al. Prevalence of stroke and vascular risk factors in china: a nationwide community-based study. *Scientific Rep.* (2017) 7:6402. doi: 10.1038/s41598-017-06691-1
4. Wang W, Jiang B, Sun H, Ru X, Sun D, Wang L, et al. Prevalence, incidence, and mortality of stroke in China. *Circulation.* (2017) 135:759-71. doi: 10.1161/CIRCULATIONAHA.116.025250
5. Liu L, Wang D, Wong KSL, Wang Y. Stroke and stroke care in China: huge burden, significant workload, and a national priority. *J Cerebr Circul.* (2011) 42:3651-4. doi: 10.1161/STROKEAHA.111.635755
6. Han J, Mao W, Ni J, Wu Y, Liu J, Bai L, et al. Rate and determinants of recurrence at 1 year and 5 years after stroke in a low-income population in Rural China. *Front Neurol.* (2020) 11(2). doi: 10.3389/fneur.2020.00002
7. Gong E, Gu W, Sun C, Turner EL, Zhou Y, Li Z, et al. System-integrated technology-enabled model of care to improve the health of stroke patients in Rural China: protocol for SINEMA—a cluster-randomized controlled trial. *Am Heart J.* (2019) 207:27–39. doi: 10.1016/j.ahj.2018.08.015

8. Pandian JD, Gall SL, Kate MP, Silva GS, Akinyemi RO, Ovbiagele BI, et al. Prevention of stroke: a global perspective. *Lancet.* (2018) 392:1269–78. doi: 10.1016/S0140-6736(18)31269-8
9. Kalkonde Yogeshwar V, Alladi S, Kaul S, Hachinski V. Stroke prevention strategies in the developing world. *Stroke.* (2018) 49:3092–7. doi: 10.1161/STROKEAHA.118.017384
10. Parappilly BP, Field TS, Mortenson WB, Sakakibara BM, Eng JJ. Effectiveness of interventions involving nurses in secondary stroke prevention: a systematic review and meta-analysis. *Euro J Cardiovascul Nurs.* (2018) 17:728–36. doi: 10.1177/1474515118779732
11. Bridgwood B, Lager KE, Mistri AK, Khunti K, Wilson AD, Modi P. Interventions for improving modifiable risk factor control in the secondary prevention of stroke. *Cochr Database Systematic Rev.* (2018) 5:pub3. doi: 10.1002/14651858.CD009103.pub3
12. Yan LL, Gong E, Gu W, Turner EL, Gallis JA, Zhou Y, et al. Effectiveness of a primary care-based integrated mobile health intervention for stroke management in Rural China (SINEMA): A cluster-randomised controlled trial. *PLOS Med.* (2021) 18:31003582. doi: 10.1371/journal.pmed.1003582
13. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new medical research council guidance. *Bmj.* (2008) 337:A1655. doi: 10.1136/bmj.a1655
14. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: medical research council guidance. *Bmj.* (2015) 350:1258. doi: 10.1136/bmj.h1258
15. Limbani F, Goudge J, Joshi R, Maar MA, Miranda JJ, Oldenburg B, et al. Process evaluation in the field: global learnings from seven implementation research hypertension projects in low-and middle-income countries. *BMC Public Health.* (2019) 19:953. doi: 10.1186/s12889-019-7261-8
16. Liu H, Mohammed A, Shanthosh J, Laba T-L, Hackett ML, Peiris D, et al. Process evaluations of primary care interventions addressing chronic disease: a systematic review. *BMJ Open.* (2019) 9:E025127. doi: 10.1136/bmjopen-2018-025127
17. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM Planning and evaluation framework: adapting to new science and practice with a 20-year review. *Front Public Health.* (2019) 7:64. doi: 10.3389/fpubh.2019.00064
18. May C. A Rational model for assessing and evaluating complex interventions in health care. *BMC Health Serv Res.* (2006) 6:86. doi: 10.1186/1472-6963-6-86
19. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care.* (2012) 50:217–26. doi: 10.1097/MLR.0b013e3182408812
20. Xu G, Ma M, Liu X, Hankey GJ. Is there a stroke belt in China and why? *Stroke.* (2013) 44:1775–83. doi: 10.1161/STROKEAHA.113.001238
21. Wu S, Wu B, Liu M, Chen Z, Wang W, Anderson CS, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. *Lancet Neurol.* (2019) 18:394–405. doi: 10.1016/S1474-4422(18)30500-3
22. Hebei Economic Yearbook 2016 [Internet]. *China Statistic Press.* (2016) [Cited 2019.01.01]. Available online at: http://www.Hetj.gov.cn/res/Nj2016/Indexch.htm.
23. Li X, Lu J, Hu S, Cheng K, De Maeseneer J, Meng Q, et al. The primary health-care system in China. *The Lancet.* (2017) 390:2584–94. doi: 10.1016/S0140-6736(17)33109-4
24. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 Years of Health-care reform in China: progress and gaps in universal health coverage. *The Lancet.* (2019) 394:1192–204. doi: 10.1016/S0140-6736(19)32136-1
25. Gong E, Gu W, Luo E, Tan L, Donovan J, Sun C, et al. Development and local contextualization of mobile health messages for enhancing disease management among community-dwelling stroke patients in rural china: multimethod study. *JMIR Mhealth Uhealth.* (2019) 7:E15758. doi: 10.2196/15758
26. Wu N, Gong E, Wang B, Gu W, Ding N, Zhang Z, et al. A Smart and multifaceted mobile health system for delivering evidence-based secondary prevention of stroke in rural china: design, development, and feasibility study. *JMIR MHealth UHealth.* (2019) 7:E13503. doi: 10.2196/13503
27. Creswell JW. *A Concise Introduction to Mixed Methods Research.* New York, NY: SAGE Publications (2014).
28. Anand T, Joseph LM, Geetha A, Prabhakaran D, Jeemon P. Task sharing with non-physician health-care workers for management of blood pressure in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet Global Health.* (2019) 7:E761–E71. doi: 10.1016/S2214-109X(19)30077-4
29. Schwalm J-D, McCready T, Lopez-Jaramillo P, Yusoff K, Attaran A, Lamelas P, et al. A community-based comprehensive intervention to reduce cardiovascular risk in hypertension (hope 4): a cluster-randomised controlled trial. *Lancet.* (2019) 394:1231–42. doi: 10.1016/S0140-6736(19)31949-X
30. Jafar TH, Gandhi M, de Silva HA, Jehan I, Naheed A, Finkelstein EA, et al. A community-based intervention for managing hypertension in rural South Asia. *NE J Med.* (2020) 382:717-26. doi: 10.1056/NEJMoa1911965
31. Maria JL, Anand T, Dona B, Prinu J, Prabhakaran D, Jeemon P. Task-sharing interventions for improving control of diabetes in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Glob Health.* (2020) 30:449. doi: 10.1016/S2214-109X(20)30449-6
32. Pandian JD, William AG, Kate MP, Norrving B, Mensah GA, Davis S, et al. Strategies to improve stroke care services in low- and middle-income countries: a systematic review. *Neuroepidemiology.* (2017) 49:45–61. doi: 10.1159/000479518
33. Coleman K, Austin BT, Brach C, Wagner EH. Evidence on the chronic care model in the new Millennium. *Health Affair.* (2009) 28:75-85. doi: 10.1377/hlthaff.28.1.75
34. Gee PM, Greenwood DA, Paterniti DA, Ward D, Miller LMS. The ehealth enhanced chronic care model: a theory derivation approach. *J Med Int Res.* (2015) 17:e86. doi: 10.2196/jmir.4067
35. Palacholla RS, Fischer N, Coleman A, Agboola S, Kirley K, Felsted J, et al. provider- and patient-related barriers to and facilitators of digital health technology adoption for hypertension management: scoping review. *JMIR Cardio.* (2019) 3:E11951. doi: 10.2196/11951
36. He J, Irazola V, Mills KT, Poggio R, Beratarrechea A, Dolan J, et al. Effect of a community health worker–led multicomponent intervention on blood pressure control in low-income patients in argentina: a randomized clinical trial. *JAMA.* (2017) 318:1016–25. doi: 10.1001/jama.2017.11358
37. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mhealth interventions: systematic review of systematic reviews. *JMIR MHealth UHealth.* (2018) 6:e23. doi: 10.2196/mhealth.8873
38. Chen S, Gong E, Kazi DS, Gates AB, Bai R, Fu H, et al. Using mobile health intervention to improve secondary prevention of coronary heart diseases in china: mixed-methods feasibility Study. *JMIR MHealth UHealth.* (2018) 6:e9. doi: 10.2196/mhealth.7849
39. Keesara S, Jonas A, Schulman K. Covid-19 and Health care's digital revolution. *N Engl J Med.* (2020) 382:e82. doi: 10.1056/NEJMp2005835
40. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? a qualitative study. *Implement Sci.* (2011) 6:6. doi: 10.1186/1748-5908-6-6
41. May CR, Cummings A, Girling M, Bracher M, Mair FS, May CM, et al. Using normalization process theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. *Implement Sci.* (2018) 13:80. doi: 10.1186/s13012-018-0758-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

PERSPECTIVE
published: 03 February 2022
doi: 10.3389/fmed.2021.763533

A Novel Framework for Understanding the Pattern Identification of Traditional Asian Medicine From the Machine Learning Perspective

Hyojin Bae[1], *Sanghun Lee*[2,3], *Choong-yeol Lee*[1] *and Chang-Eop Kim*[1]*

[1] *Department of Physiology, Gachon University College of Korean Medicine, Seongnam, South Korea,* [2] *Korean Medicine Data Division, Korea Institute of Oriental Medicine, Daejeon, South Korea,* [3] *Department of Korean Convergence Medical Science, University of Science and Technology, Daejeon, South Korea*

Pattern identification (PI), a unique diagnostic system of traditional Asian medicine, is the process of inferring the pathological nature or location of lesions based on observed symptoms. Despite its critical role in theory and practice, the information processing principles underlying PI systems are generally unclear. We present a novel framework for comprehending the PI system from a machine learning perspective. After a brief introduction to the dimensionality of the data, we propose that the PI system can be modeled as a dimensionality reduction process and discuss analytical issues that can be addressed using our framework. Our framework promotes a new approach in understanding the underlying mechanisms of the PI process with strong mathematical tools, thereby enriching the explanatory theories of traditional Asian medicine.

Keywords: pattern identification, machine learning, dimensionality reduction, diagnostic system, traditional Asian medicine, traditional Chinese medicine, syndrome differentiation

OPEN ACCESS

Edited by:
Michele Mario Ciulla, University of Milan, Italy

Reviewed by:
Young-Bae Park, Nubebe Mibyeong Research Institu, South Korea
Wonmo Jung, Hamkkekiugi Inc., South Korea

***Correspondence:**
Chang-Eop Kim
eopchang@gachon.ac.kr

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *24 August 2021*
Accepted: *23 December 2021*
Published: *03 February 2022*

Citation:
Bae H, Lee S, Lee C-y and Kim C-E (2022) A Novel Framework for Understanding the Pattern Identification of Traditional Asian Medicine From the Machine Learning Perspective. Front. Med. 8:763533. doi: 10.3389/fmed.2021.763533

INTRODUCTION

Pattern identification (PI), a distinctive diagnostic system found in traditional Asian medicine (TAM), is a clinical reasoning process that uses the signs and symptoms of patients to identify diagnostic patterns (1). These patterns convey information about the nature of the disease or the location of lesions and serve as a guide for treatment selection (2) (e.g., drain for a "excess" pattern and tonify for a "deficient" pattern). Notably, patterns in TAM are pragmatic concepts that are widely accepted as a useful treatment target rather than actual pathogens or objectively measurable states (3). It can be said that PI is a strategy chosen to make diagnostic decisions based on naked sense observations and to determine corresponding treatments. Despite their centrality in theory and practice, the information processing principles of PI have remained relatively superficial. Additionally, abstract descriptions make it difficult to objectively describe the PI process, resulting in a low level of consistency between practitioners (4–6).

In recent years, approaches based on machine learning (ML) have demonstrated remarkable performance in a variety of tasks, including image classification, speech processing, and natural language processing, all of which are difficult to solve using knowledge-based approaches (7). Interestingly, this success has spawned approaches in systems neuroscience that use ML to study how the brain works (8–11). The strategy is to use ML algorithms as a computational model of

the brain and to benchmark this model in order to gain a better understanding of how the brain represents, learns, and flexibly processes high-dimensional information.

Inspired by the idea that ML models can help capture critical aspects of the brain's computation, we present a novel framework for explaining how information is processed in the PI system and why it is effective. Within our framework, we model the PI system as a dimensionality-reduction algorithm and propose several research questions. By leveraging ML's framework, we can adopt powerful mathematical tools, broaden the scope of inquiry, and enrich explanatory theory in TAM.

MANUSCRIPT FORMATTING

A Brief Introduction to Dimensionality Reduction

In this paper, we view the PI system through the lens of dimensionality reduction process, which reduces high-dimensional data to a low-dimensional representation. To that end, we'll discuss high-dimensional data and dimensionality reduction briefly. Rather than providing strict mathematical definitions, we will explain these concepts with examples to aid intuitive understanding.

The dimensionality of data is defined as the number of features (attributes) that describe the observations in data (**Figure 1A**) (assuming that the number of rows (observations) exceeds the number of columns (features/attributes) and the data matrix is full-rank, which is easily satisfied in noisy, real-world datasets). A larger number of features leads to a more detailed representation of the observation (i.e., high representational power) (12). Additionally, when compared to low-dimensional space, high-dimensional space makes data classification easier (13). For instance, while classification in a low-dimensional space requires non-linear and complex decision boundaries, data can be made linearly separable by adding additional dimensions (axes) (**Figure 1B**).

However, high-dimensional space does not come without drawbacks. Due to the fact that more than four-dimensional space is beyond human cognition, high-dimensional data are unintuitive, making it difficult to interpret or derive insights. More importantly, as the input dimension increases, the classifier's performance on unseen data typically degrades rather than improves. A common explanation for this is the "curse of dimensionality" (14). As the dimension increases, the volume of space in which data are represented increases exponentially, to the point where available data become sparse (**Figure 1C**) (15). In this case, the model is likely to miss generalizable patterns in the data. One solution is to increase the size of the training data until the density is sufficient, while another is to reduce the dimensionality of the data, which is usually the more practical option (16).

Apart from these disadvantages of high-dimensional data, the typical motivation for dimensionality reduction is that the genuine dimension (i.e., degree of freedom) of the space may be significantly less than the number of features due to feature dependencies (17). That is, even if the dataset contains hundreds or even millions of features, the majority of variation may be explained by a handful of latent variables. There are numerous dimensionality-reduction algorithms, and which one to use depends on the nature of the data and the research objective. For instance, principal component analysis (PCA), one of the most widely used linear dimensionality-reduction techniques, seeks to identify orthogonal axes [i.e., principal components (PCs)] that best account for the variance of the data via a linear combination of existing axes (18). By projecting the data into a subspace of leading PCs, we can obtain a compact representation of the data, albeit with some information loss (**Figure 1D**). There are also non-linear techniques such as Isomap (19), t-stochastic neighbor embedding (20), uniform manifold approximation and projection (21) that capture non-linear relations between variables. Overall, the motivations for dimensionality reduction in dealing with high-dimensional data are as follows: first, high-dimensional data are unintuitive; second, they are prone to the curse of dimensionality; and third, a dataset's dimensionality may be artificially high.

Modeling the PI as a Dimensionality-Reduction Process From the Symptom Space

One of the most distinctive characteristics of TAM in clinical practice is the use of patterns to identify and treat the patient. TAM physicians evaluate patient's clinical symptoms and signs and classify them according to specific pattern groups (4). The identified patterns provide basis for prescribing treatments including herbal formula (22). Each patient can be thought of as a point in a multidimensional symptom space, with each dimension corresponding to a distinct symptom. If the total number of symptoms is p, the patient is represented as a p-dimensional vector whose elements are the coordinate values on each symptom axis. Similarly, the herbal space can be defined in the same way, with each dimension representing an individual herb. If the total number of herbs is q, a herbal prescription (a mixture of herbs) is represented as a q-dimensional vector whose elements are the coordinate values for each herbal axis. Following that, treatment selection can be formulated as a mapping from the symptom space to the herbal space (To keep the discussion concise, treatment is limited to herbal prescriptions). From the doctor's perspective, there are several motivations to reduce the dimension of the input data to perform this task successfully. Assuming the symptom and herbal space have dimensions of p and q, respectively, the number of theoretically possible mappings is q^p. Even if each p and q are on a tens-scale, they are already beyond the cognitive capacity of any single human memory. In this case, shrinking the input space's dimension can exponentially reduce the number of available alternatives (**Figure 2A**).

Additionally, there is frequently a high degree of correlation and redundancy among individual symptoms, limiting possible patterns of variation (e.g., fever may have a positive correlation with thirst and a negative correlation with a pale face). In other words, a small number of independent patterns can effectively describe the system's behavior, resulting in symptom data that

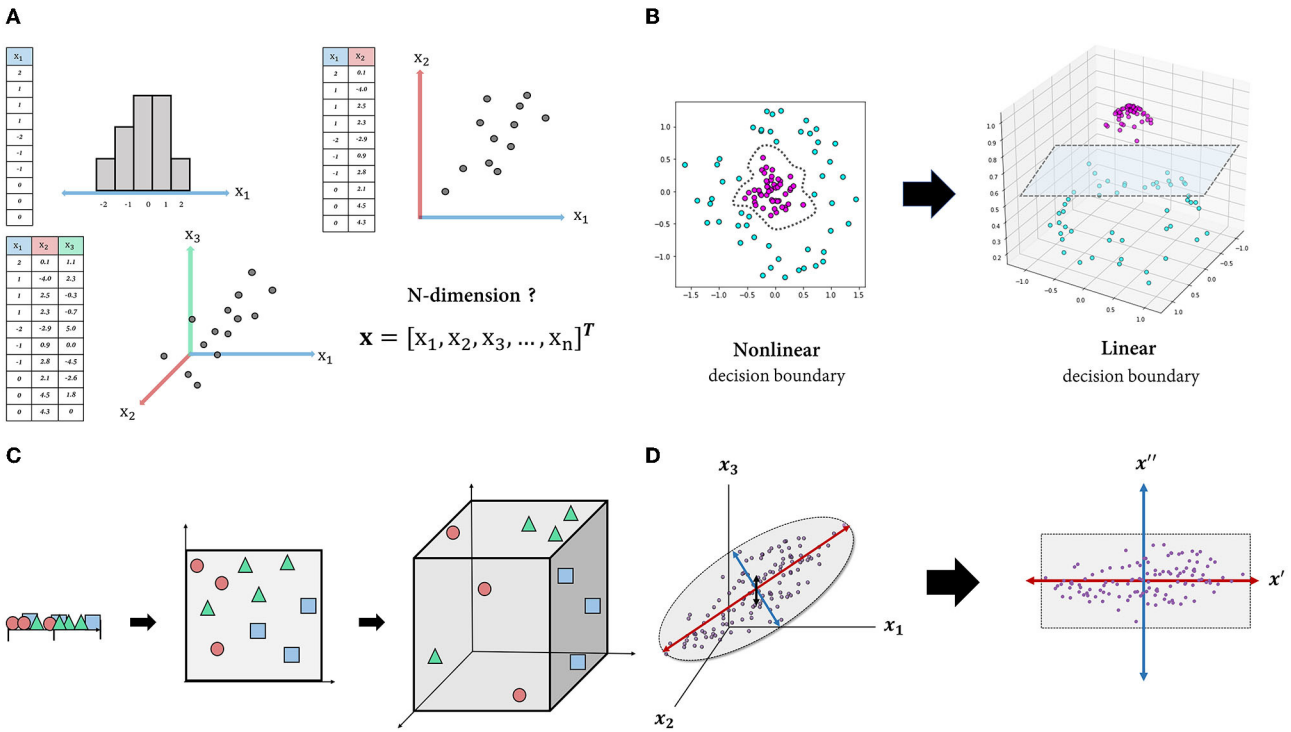


FIGURE 1 | Schematic figures explaining the features of high-dimensional data. **(A)**: Intuitive understanding of multidimensional data. Each observation (row) in each table is described by one or more features (columns) and visualized as a point in one-, two- or three-dimensional space. This representation is easily extended to four-dimensional space or higher, but it cannot be visualized. **(B)**: By transforming the dataset into high-dimensional space, the data can be separated using a linear decision boundary. **(C)**: Curse of dimensionality. As the dimension of the space increases, the volume of the space expands exponentially, and the density of the space becomes increasingly sparse. **(D)**: Projection into an intrinsic-dimensional space. Data laid out in three-dimensional space can be approximated by a two-dimensional plane composed of newly discovered axes that account for majority of the data variability.

may span only a constrained, low-dimensional subset of the entire space. In this case, the overall structure of the symptom data, which may not be visible at the individual symptom level, may be more important for treatment selection. Disentangling the data based on latent patterns may aid in revealing the data's intrinsic structure (**Figure 2B**).

Given this perspective, the PI system can be modeled as the process of representing high-dimensional symptom data in a low-dimensional space defined by a few latent patterns. Assuming that the patient records contain p symptom variables and are represented by r pattern variables, PI process can be described as follows: Given a set of n patient vectors $X = \{x_i\}_{i=1}^{n}$ [i.e., the ith training sample is a vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]^T$, where x_{ij} is jth feature of ith sample], the aim is to transform each vector $x_i \in R^p$ into a new vector $z_i \in R^r = [z_{i1}, z_{i2}, \ldots, z_{ir}]^T$ where $r \ll p$. The mapping function $f: X \subset R^p \rightarrow Z \subset R^r$ can be estimated differently depending on the specific forms of the objective function.

Interestingly, human-inferred latent patterns may not always be the optimal solution in terms of information loss minimization, which is the primary goal of dimensionality-reduction algorithms such as PCA. This is because, given human expert's inductive reasoning, reduced representations must not only deliver compact information but also be cognitively efficient. Indeed, the patterns in TAM, such as heat, cold, deficiency, and excess, are primarily intuitive and metaphorical concepts that are embodied in daily life (**Figure 2C**). Inferring patterns from experiences or observations of the physical world and explaining the physiological and pathological phenomena of the human body in terms of these conceptual patterns are key characteristics of TAM theory (23). While this approach may appear crude and ideological in comparison to pathogen-based diagnosis, it provides an intuitive foundation for inductive reasoning (24, 25).

Research Questions in PI Systems That Can Be Addressed Using Mathematical Metrics Developed in ML

In this section, we raise several research questions that can be addressed by utilizing our novel framework that models the PI system in terms of the ML perspective. In particular, we focus on the topics for specifying the dimensionality reduction properties of the PI system.

Is PI a Linear or Non-linear Process?

Dimensionality-reduction algorithms can be classified mathematically as linear or non-linear, which is critical for implementation. Linear techniques such as PCA,

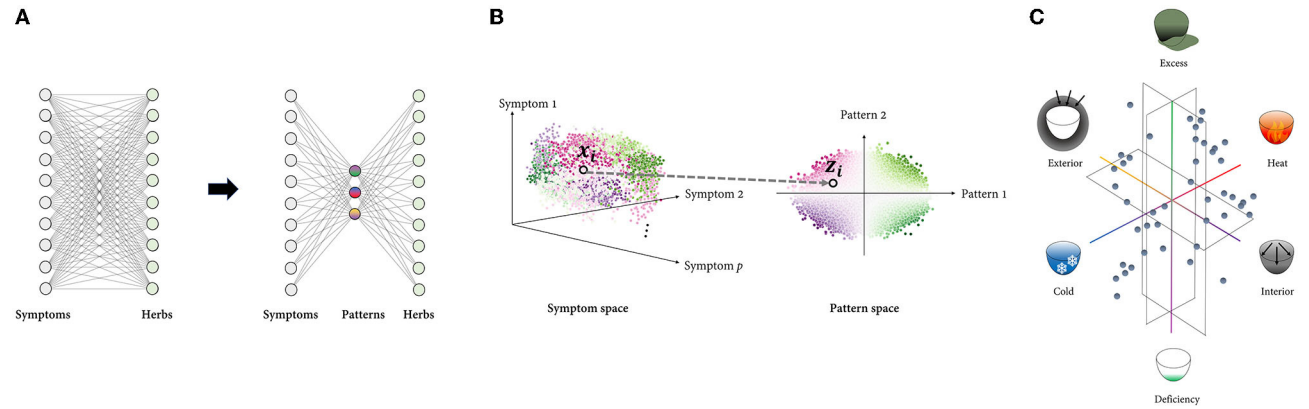


FIGURE 2 | Framework modeling the PI with dimensionality reduction. **(A)**: Rather than mapping high-dimensional spaces directly, the number of cases can be reduced exponentially by first projecting the input space to low-dimensional space composed of multiple latent variables and then mapping it to the output space. **(B)**: Representing the data using a few underlying patterns reveals the intrinsic structure of the data, which is difficult to capture in a high-dimensional space where distinct factors of variations are highly entangled. Each point represents sample data, and the points denoted by a black circle represent the *ith* sample, which is represented in both the symptom space ($x_i \in R^p$) and the pattern space ($z_i \in R^r$). The points are color-coded according to the identified patterns. **(C)**: TAM's low-dimensional pattern space is constructed from metaphorical concepts that are embodied in everyday life. The pattern space in eight-principle PI, the most comprehensive type of PI, is visualized as an example. Six of eight principle patterns are composed of the exterior, interior, cold, heat, deficiency, and excess, while the other two are Yin and Yang, which are higher concepts that encompass the other six patterns. Six principle patterns are grouped in pairs of mutually opposing properties: exterior-interior, cold-heat, and deficiency and excess. These three pairs represent the extent to which external pathogens penetrate the body, the nature of the disease, and the relative superiority of the body's resistance to pathogenic factors and the pathogenic qi, respectively. These concepts' familiar and abstract characteristics enable robust inference of the pathological pattern from a myriad of symptom phenotypes.

multidimensional scaling, and factor analysis are widely used in a variety of fields. They employ straightforward linear algebraic techniques that are easy to implement and provide clear geometric interpretations (26). In the real world, however, data may form a highly non-linear manifold. Low-dimensional embeddings obtained via methods assuming a linear submanifold may be unsatisfactory in this case (27).

Whether to use a linear or non-linear technique should be determined by the nature of the data being analyzed, as well as the nature of the problem being solved. The PI process should compress the symptom space while retaining the information required for treatment selection, but its linear or non-linear nature has not been investigated. For instance, the probability of being identified as a particular pattern can increase supra-linearly when a particular symptom pair appears concurrently, whereas the probability may be negligible in the absence of a single symptom.

Numerous techniques exist for quantifying non-linearity in operations (28–30). Quantifying non-linearity may allow for the assessment of the adequacy of currently developed tools supporting clinical PI. For example, a questionnaire based on linear regression may be ineffective for a disease in which significant non-linear associations exist between symptoms and patterns.

How and to What Extent the PI Abstracts Information

The core characteristic of human intelligence is to learn from small samples to deal with previously unknown situations, which are often linked with the critical challenges raised in ML (31). For the brain to learn efficiently within its limited resources, it is necessary to draw general conclusions from individual experiences rather than memorize them all (32). Abstraction and hierarchical information processing are critical capabilities that contribute to the human brain's remarkable capacity for generalization (7, 33).

Given that PI is the process of representing patients' clinical symptoms as metaphorical patterns, it is fundamentally an abstraction process. Abstract representations may aid physicians in robustly inferring pathological patterns from a wide variety of symptom combinations, thereby simplifying patient classification. It is critical to investigate how and at what level abstractions are made and how they contribute to patient classification and/or treatment selection in order to gain a better understanding of information processing in PI.

Obtaining abstract (high-level) representations while ignoring irrelevant details is also critical in artificial intelligence (AI). Deep neural networks, in particular, such as convolutional neural networks and autoencoders are thought to learn abstract representations, and abstraction in representations can be quantified in various ways, for example, the degree of dichotomy or the capacity for generalization (34–38). Similarly, abstractions in PI can be explicitly quantified using the PI model's representation. Whether or not TAM concepts with varying levels of abstraction are hierarchically encoded in the system, or whether the level of abstraction varies between different types of PI that employ distinct conceptual patterns, such as Qi and blood, viscera and bowels (zangfu), or the five phases, could be specific research topics. This would enable us to assess the appropriate level of abstraction as well as its advantages and disadvantages.

What Is the Objective Function of the PI System?

The objective function specifies how a model's performance/cost is calculated, and a model is trained to maximize or minimize it. In other words, the objective function represents the model's learning goal, which is a critical component that must be specified in ML practice along with the learning rule and the architecture (39). Similarly, we can consider the PI system's objective function. Investigating the objective function that led the development of TAM's clinical decision-making model into its current form will give insight on the information processing strategy of PI system.

We can start with a common objective function of ML to determine that of the PI system. In supervised learning, the most widely used objective function is as follows:

$$\hat{f} = \underset{f \epsilon \mathcal{H}}{arg\ min} \left[\frac{1}{n} \sum_{i=1}^{n} L(y_i, f(x_i)) + \lambda J(f) \right]$$

$\mathcal{H}$ denotes the function space of f, and the function $\hat{f}$ is found by the minimization of the cost (inside the square bracket), which is composed of the loss function $L\left(y_i, f\left(x_i\right)\right)$ and the regularization function $J(f)$ with its associated regularization weight λ. y_i and $f\left(x_i\right)$ denotes the ground-truth treatment and model prediction for the *ith*-sample x_i, respectively. To minimize the loss, the model should fit the training data as closely as possible. However, the complexity of the model is constrained by the penalty imposed by the regularization term. The strategy of having two conflicting components in the objective function enables the designer to consider a reasonable bias-variance trade-off (i.e., enhancing the model's reliability in the face of unseen data at the expense of greater bias) (40). In other words, the objective function formulation expresses explicitly which characteristics the system values and penalizes.

It will be important to investigate which type of performance or penalty should be assessed by the loss and regularization functions in order to induce the current PI system. For example, when describing the long-term evolution of a PI system, the regularization function may be used to constrain the agent's cognitive and/or computational load rather than to prevent overfitting (In the long run, variance shrinks because $\lim_{n \to \infty} \sigma^2 = 0$, where σ^2 denotes the model variance).

DISCUSSION

Earlier research on developing an AI-based diagnostic system for TAM was primarily focused on developing an expert system that makes use of expertise and ontology (41–45), whereas in recent years, a bottom-up approach that generates knowledge from the data has become more prevalent. The majority of recent PI studies utilizing ML have attempted to develop predictive models capable of reproducing a physician's diagnosis (46–48). While these studies explored the clinical applicability of ML algorithms based on their predictive performance, there were also studies examining the PI theory itself. One study validated TAM pattern types statistically by demonstrating that patient clusters in the data set correspond well to theoretical pattern types (49, 50), and another used a decision tree algorithm to extract a collection of symptoms indicative of a pattern in a particular disease (51) [For a more comprehensive and systematic review of the application of quantitative models in traditional medicine, see (52, 53)]. Our study is unique in that it presents a broader framework for explaining and analyzing PI system's information processing strategy from a ML perspective. Additionally, while we explained the PI process as dimensionality reduction, it is not exclusive to other ML algorithms such as clustering.

When dimensionality reduction is used to extract latent features, the process is comparable to that of theorization or modeling. Both involve deriving fundamental principles or patterns from massive and disordered data at the expense of detailed information. A model that fully describes all data samples is merely an enumeration of facts and is incapable of conveying generalized knowledge. Instead, we require a simple explanation to make sense of the data despite the presence of residuals that the model cannot account for. This aspect of dimensionality reduction is consistent with TAM's distinctive way of thinking, which seeks to interpret changes of the patient's symptoms and discomfort using abstract concepts that describe the dynamic nature of the micro-environment of the human body (54). By grasping the generalizable principles underlying individual observations, we can explain, predict, and manipulate the observed system's behavior beyond the scope of our experience.

According to cognitive psychology, humans frequently employ heuristic strategies that arrive at satisfactory solutions with a modest amount of computation to make decisions within their cognitive capacity and time constraints (55–57). Numerous models have been proposed to explain the strategies employed by the human brain, and dimensionality-reduction model in this paper is in line with such models. However, the constraint on computing resources in the dimensionality reduction of ML is not severe, resulting in differences between the human and machine computation. Additionally, it is expected that extensive feature selection will occur prior to dimensionality reduction in the actual PI process, based on cues such as the patient's chief complaint. This procedure would be based on the physician's prior knowledge, which correspond to the Bayesian prior. It is also noteworthy that reduced representations in PI systems must be interpretable because they are the product of conscious reasoning, unlike many ML algorithms, including PCA.

We combine the ingredients of systems neuroscience and ML to propose a conceptual framework for investigating the PI system, based on TAM domain knowledge. The introduction of a new perspective leads to the emergence of novel research questions and methodologies, opening a novel field of investigation. By implementing mathematical tools developed in ML, we will be able to verify a variety of hypotheses to which qualitative approaches have been applied primarily and contribute to the development of shareable explicit knowledge. This may help overcome one of TAM theory's primary flaws, namely that it is subjective and difficult to articulate. While this framework leaves room for elaboration, we believe it will serve as the foundation for developing interpretable AI for the medical domain.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

C-EK, C-yL, and SL: conceptualization. HB and C-EK: investigation. HB: writing—original draft. C-EK and SL: writing—review and editing and funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2020R1F1A107317912) and the Collection of Clinical Big Data and Construction of Service Platform for Developing Korean Medicine Doctor with Artificial Intelligence research project (KSN2021110).

REFERENCES

1. Birch S, Alraek T. Traditional East Asian medicine: how to understand and approach diagnostic findings and patterns in a modern scientific framework? *Chin J Integ Med.* (2014) 20:336–40. doi: 10.1007/s11655-014-1809-3
2. Textbook Compilation Committee of Pathology in Korean Medicine. *Pathology in Korean Medicine*. Seoul: Hanuimunhwasa (2010).
3. Birch S, Alraek T, Bovey M, Lee MS, Lee JA, Zaslawsk C, et al. Overview on pattern identification – history, nature and strategies for treating patients: a narrative review. *Eur J Integ Med*. (2020) 35:101101. doi: 10.1016/j.eujim.2020.101101
4. Zhang GG, Lee W, Bausell B, Lao L, Handwerger B, Berman B. Variability in the traditional Chinese medicine (TCM) diagnoses and herbal prescriptions provided by three TCM practitioners for 40 patients with rheumatoid arthritis. *J Altern Compl Med*. (2005) 11:415–21. doi: 10.1089/acm.2005.11.415
5. Berle CA, Cobbin D, Smith N, Zaslawski C. A novel approach to evaluate traditional chinese medicine treatment outcomes using pattern identification. *J Altern Compl Med*. (2010) 16:357–67. doi: 10.1089/acm.2009.0367
6. Zhou X, Li Y, Peng Y, Hu J, Zhang R, He L, et al. Clinical phenotype network: the underlying mechanism for personalized diagnosis and treatment of traditional Chinese medicine. *Front Med.* (2014) 8:337–46. doi: 10.1007/s11684-014-0349-8
7. Saxe A, Nelli S, Summerfield C. If deep learning is the answer, what is the question?. *Nat Rev Neurosci*. (2021) 22:55–67. doi: 10.1038/s41583-020-00395-8
8. Glaser JI, Benjamin AS, Farhoodi R, Kording KP. The roles of supervised machine learning in systems neuroscience. *Prog Neurobiol.* (2019) 175:126–37. doi: 10.1016/j.pneurobio.2019.01.008
9. Williamson RC, Doiron B, Smith MA, Yu BM. Bridging large-scale neuronal recordings and large-scale network models using dimensionality reduction. *Curr Opin Neurobiol.* (2019) 55:40–7. doi: 10.1016/j.conb.2018.12.009
10. Marblestone AH, Wayne G, Kording KP. Toward an integration of deep learning and neuroscience. *Front Comput Neurosci.* (2016) 10:94. doi: 10.3389/fncom.2016.00094
11. Stevenson IH, Kording KP. How advances in neural recording affect data analysis. *Nat Neurosci.* (2011) 14:139. doi: 10.1038/nn.2731
12. Bengio Y, Delalleau O. *On the Expressive Power of Deep Architectures, International Conference on Algorithmic Learning Theory*. Heidelberg: Springer (2011). p. 18–36.
13. Tarca AL, Carey VJ, Chen X-w, Romero R, Drăghici S. Machine learning and its applications to biology. *PLoS Comput Biol.* (2007) 3:e116. doi: 10.1371/journal.pcbi.0030116
14. Friedman JH. On bias, variance, 0/1—loss, the curse-of-dimensionality. *Data Min Knowl Dis*. (1997) 1:55–77. doi: 10.1023/A:1009778005914
15. Altman N, Krzywinski M. The curse (s) of dimensionality. *Nat Methods.* (2018) 15:399–400. doi: 10.1038/s41592-018-0019-x
16. Verleysen M, François D. The curse of dimensionality in data mining and time series prediction. In: Cabestany J, Prieto A, Sandoval F, editors. *Computational Intelligence and Bioinspired Systems*. Berlin: Springer Berlin Heidelberg (2005). p. 758–70.
17. Cayton L. Algorithms for manifold learning. *Univ California San Diego Tech Rep.* (2005) 12:1.
18. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans Royal Soc A Mathe Phys Eng Sci.* (2016) 374:20150202. doi: 10.1098/rsta.2015.0202
19. Tenenbaum JB, Silva Vd, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science.* (2000) 290:2319. doi: 10.1126/science.290.5500.2319
20. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* (2008) 9.
21. McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv*. (2018) 03426: doi: 10.21105/joss.00861
22. Wang J, Cai C, Kong C, Cao Z, Chen Y. A computer method for validating traditional Chinese medicine herbal prescriptions. *Am J Chin Med.* (2005) 33:281–97. doi: 10.1142/S0192415X05002825
23. Lee CY. Discussion on the issues of the modernization of the fundamental theories and terms in korean medicine. *J Physiol Pathol Korean Med.* (2013) 27:540–52.
24. Overholser JC. Elements of the socratic method: II. Inductive reasoning. *Psychoth Theory Res Prac Train.* (1993) 30:75. doi: 10.1037/0033-3204.30.1.75
25. Heit E. Properties of inductive reasoning. *Psychonomic Bull Rev.* (2000) 7:569–92. doi: 10.3758/BF03212996
26. Cunningham JP, Ghahramani Z. Linear dimensionality reduction: survey, insights, and generalizations. *J Mach Learn Res.* (2015) 16:2859–900.
27. Lee JA, Verleysen M. *Nonlinear Dimensionality Reduction*. New York, NY: Springer Science & Business Media (2007).
28. Emancipator K, Kroll MH. A quantitative measure of nonlinearity. *Clin Chem.* (1993) 39:766–72. doi: 10.1093/clinchem/39.5.766
29. Weilin H, Yin H. *Linear and Nonlinear Dimensionality Reduction for Face Recognition, 2009 16th IEEE International Conference on Image Processing (ICIP)* (Piscataway, NY). (2009). p. 3337–40.
30. Nagamine T, Seltzer ML, Mesgarani N. *On the Role of Nonlinear Transformations in Deep Neural Network Acoustic Models, Interspeech*. Rhineland-Palatinate: dblp computer science bibliography (2016). p. 803–7.
31. Cortese A, De Martino B, Kawato M. The neural and cognitive architecture for learning from a small sample. *Curr Opin Neurobiol.* (2019) 55:133–41. doi: 10.1016/j.conb.2019.02.011
32. Chollet F. The measure of intelligence. *arXiv*. (2019) 01547.
33. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND. How to grow a mind: statistics, structure, and abstraction. *Science.* (2011) 331:1279. doi: 10.1126/science.1192788
34. Saitta L, Zucker J-D. *Abstraction in Artificial Intelligence and Complex Systems*. New York, NY: Springer (2013).
35. Zucker J-D. A grounded theory of abstraction in artificial intelligence. *Philos Trans Royal Soc London Series B Biol Sci.* (2003) 358:1293–309. doi: 10.1098/rstb.2003.1308
36. Wutz A, Loonis R, Roy JE, Donoghue JA, Miller EK. Different levels of category abstraction by different dynamics in different prefrontal areas. *Neuron.* (2018) 97:716–26.e8. doi: 10.1016/j.neuron.2018.01.009
37. Saez A, Rigotti M, Ostojic S, Fusi S, Salzman CD. Abstract context representations in primate amygdala and prefrontal cortex. *Neuron.* (2015) 87:869–81. doi: 10.1016/j.neuron.2015.07.024

38. Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, Daniel Salzman C. The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell*. (2020) 183:954–67. doi: 10.1016/j.cell.2020.09.031
39. Richards BA, Lillicrap TP, Beaudoin P, Bengio Y, Bogacz R, Christensen A, et al. A deep learning framework for neuroscience. *Nat Neurosci.* (2019) 22:1761–70. doi: 10.1038/s41593-019-0520-2
40. Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning requires rethinking generalization. *arXiv*. (2016) 03530.
41. Lukman S, He Y, Hui S-C. Computational methods for traditional Chinese medicine: a survey. *Comp Meth Prog Biomed.* (2007) 88:283–94. doi: 10.1016/j.cmpb.2007.09.008
42. Huang M-J, Chen M-Y. Integrated design of the intelligent web-based Chinese medical diagnostic system (CMDS)–Systematic development for digestive health. *Exp Syst Appl.* (2007) 32:658–73. doi: 10.1016/j.eswa.2006.01.037
43. Xiao S, Peng C, Wang Z, Wang F, Zhiwei N. Using the algebraic sum method in medical expert systems. *IEEE Eng Med Biol Mag.* (1996) 15:80–2. doi: 10.1109/51.499763
44. Gu P, Chen H, Yu T. Ontology-oriented diagnostic system for traditional Chinese medicine based on relation refinement. *Comp Mathemat Meth Med.* (2013). doi: 10.1155/2013/317803
45. Wang X-w, Qu H-b, Liu X-s, CHENG Y-y. Bayesian network approach to knowledge discovery in traditional Chinese medicine. *J Zhej Univ Eng Sci*. (2005) 39:948.
46. Xu Q, Tang W, Teng F, Peng W, Zhang Y, Li W, et al. Intelligent syndrome differentiation of traditional chinese medicine by ann: a case study of chronic obstructive pulmonary disease. *IEEE Access.* (2019) 7:76167–75. doi: 10.1109/ACCESS.2019.2921318
47. Liu G-P, Yan J-J, Wang Y-Q, Zheng W, Zhong T, Lu X, et al. Deep learning based syndrome diagnosis of chronic gastritis. *Comp Mathemat Meth Med.* (2014) 2014:8. doi: 10.1155/2014/938350
48. Zhang Q, Zhang W-T, Wei J-J, Wang X-B, Liu P. Combined use of factor analysis and cluster analysis in classification of traditional Chinese medical syndromes in patients with posthepatitic cirrhosis. *J Chin Integ Med.* (2005) 3:14–18. doi: 10.3736/jcim20050105
49. Zhang NL, Yuan S, Chen T, Wang Y. Latent tree models and diagnosis in traditional Chinese medicine. *Art Intell Med.* (2008) 42:229–45. doi: 10.1016/j.artmed.2007.10.004
50. Zhang NL, Fu C, Liu TF, Chen B-x, Poon KM, Chen PX, et al. A data-driven method for syndrome type identification and classification in traditional Chinese medicine. *J Integ Med*. (2017) 15:110–23. doi: 10.1016/S2095-4964(17)60328-5
51. Wang Y, Ma L, Liao X, Liu P. Decision tree method to extract syndrome differentiation rules of posthepatitic cirrhosis in traditional Chinese medicine. In: *2008 IEEE International Symposium on IT in Medicine and Education* (IEEE). (2008). p. 744–8. doi: 10.1109/ITME.2008.4743965
52. Arji G, Safdari R, Rezaeizadeh H, Abbassian A, Mokhtaran M, Hossein Ayati M. A systematic literature review and classification of knowledge discovery in traditional medicine. *Comp Meth Prog Biomed*. (2019) 168:39–57. doi: 10.1016/j.cmpb.2018.10.017
53. Chu X, Sun B, Huang Q, Peng S, Zhou Y, Zhang Y. Quantitative knowledge presentation models of traditional Chinese medicine (TCM): a review. *Art Intell Med.* (2020) 103:101810. doi: 10.1016/j.artmed.2020.101810
54. Lee C-Y. Investigating 'model-dependent realism' from the viewpoint of the traditional medical theories research. *J Physiol Pathol Korean Med.* (2015) 29:353–60. doi: 10.15188/kjopp.2015.10.29.5.353
55. Shah AK, Oppenheimer DM. Heuristics made easy: an effort-reduction framework. *Psychol Bull.* (2008) 134:207. doi: 10.1037/0033-2909.134.2.207
56. Lieder F, Griffiths TL, Huys QJ, Goodman ND. The anchoring bias reflects rational use of cognitive resources. *Psychon Bull Rev*. (2018) 25:322–49. doi: 10.3758/s13423-017-1286-8
57. Morsanyi K, Handley SJ. How smart do you need to be to get it wrong? The role of cognitive capacity in the development of heuristic-based judgment. *J Exp Child Psychol.* (2008) 99:18–36. doi: 10.1016/j.jecp.2007.08.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership