Protocol

Comparing Decentralized Learning Methods for Health Data Models to Nondecentralized Alternatives: Protocol for a Systematic Review

José Miguel Diniz[1,2], MSc, MD; Henrique Vasconcelos[1], MSc, MD; Júlio Souza[1,3], PhD; Rita Rb-Silva[3], MSc, MD, PhD; Carolina Ameijeiras-Rodriguez[3], DPharm, PhD; Alberto Freitas[1,3], PhD

[1]CINTESIS—Centre for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

[2]PhD Program in Health Data Science, Faculty of Medicine, University of Porto, Porto, Portugal

[3]MEDCIDS—Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

Corresponding Author:
José Miguel Diniz, MSc, MD
PhD Program in Health Data Science
Faculty of Medicine
University of Porto
Rua Dr Plácido da Costa
4200-450
Porto
Portugal
Phone: 351 225 513 622
Email: jmdiniz.med@gmail.com

Abstract

Background: Considering the soaring health-related costs directed toward a growing, aging, and comorbid population, the health sector needs effective data-driven interventions while managing rising care costs. While health interventions using data mining have become more robust and adopted, they often demand high-quality big data. However, growing privacy concerns have hindered large-scale data sharing. In parallel, recently introduced legal instruments require complex implementations, especially when it comes to biomedical data. New privacy-preserving technologies, such as decentralized learning, make it possible to create health models without mobilizing data sets by using distributed computation principles. Several multinational partnerships, including a recent agreement between the United States and the European Union, are adopting these techniques for next-generation data science. While these approaches are promising, there is no clear and robust evidence synthesis of health care applications.

Objective: The main aim is to compare the performance among health data models (eg, automated diagnosis and mortality prediction) developed using decentralized learning approaches (eg, federated and blockchain) to those using centralized or local methods. Secondary aims are comparing the privacy compromise and resource use among model architectures.

Methods: We will conduct a systematic review using the first-ever registered research protocol for this topic following a robust search methodology, including several biomedical and computational databases. This work will compare health data models differing in development architecture, grouping them according to their clinical applications. For reporting purposes, a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram will be presented. CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies)–based forms will be used for data extraction and to assess the risk of bias, alongside PROBAST (Prediction Model Risk of Bias Assessment Tool). All effect measures in the original studies will be reported.

Results: The queries and data extractions are expected to start on February 28, 2023, and end by July 31, 2023. The research protocol was registered with PROSPERO, under the number 393126, on February 3, 2023. With this protocol, we detail how we will conduct the systematic review. With that study, we aim to summarize the progress and findings from state-of-the-art decentralized learning models in health care in comparison to their local and centralized counterparts. Results are expected to clarify the consensuses and heterogeneities reported and help guide the research and development of new robust and sustainable applications to address the health data privacy problem, with applicability in real-world settings.

Conclusions: We expect to clearly present the status quo of these privacy-preserving technologies in health care. With this robust synthesis of the currently available scientific evidence, the review will inform health technology assessment and evidence-based decisions, from health professionals, data scientists, and policy makers alike. Importantly, it should also guide the development and application of new tools in service of patients' privacy and future research.

Trial Registration: PROSPERO 393126; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=393126

International Registered Report Identifier (IRRID): PRR1-10.2196/45823

Introduction

Background

The current health paradigm challenges are unprecedented in their nature and scope. Stemming from a growing [1], aging [1], and comorbid [2] global population, disability is becoming an increasingly large share of the burden of disease and of the already unsustainable health care costs [3-6]. Faced with the need to invest in more effective and preventive strategies, improved research and development are essential to address these unwavering issues [2].

To do so, robust evidence-based health knowledge is needed, as a way of enabling improved planning and provision of care. Presently, many data-driven approaches are commonplace in biomedical sciences and clinical research. These include examples from epidemiological surveillance [7] and cancer prognosis [8] to drug discovery [9,10] and mortality prediction [11].

Thus, access to data has been the foundation for better health models, both in their precision and validity to represent different medical conditions [12,13] and patients [14,15]. In parallel, alongside recent digital transitions and new tools and infrastructures, data analysis has become more powerful and faster than ever [16]. This gave rise to data science [17], resulting from the complex merger of traditional statistics disciplines combined with other subjects.

In particular, data mining techniques have elevated the computational functionality, especially for cognitive and analytic processes that are hard to develop algorithmically [18]. Thus, machine learning models, such as decision trees, linear regression, and support vector machines are now abundant in health research [19,20]. Deep neural networks brought to light even more sophisticated applications of natural language processing and computer vision, both effective and powerful [21,22].

Supported by robust software and hardware, these new model development approaches rely on the availability of large and high-quality training databases. Hence, the concept of big data emerged, referring to the need for comprehensive data sets to sustain the inferential process in both their internal and external validities. While it is often characterized by a few key dimensions (volume, variety, velocity, value, veracity, and variability), many other features (eg, venue and volatility) can

be of relevance [23,24]. However, concerns regarding privacy protection—a fundamental human right [25]—are rising amid increasing numbers of misconducts and violations [26,27].

Limitations of Current Strategies

In response to both data demands and privacy challenges, 2 groups of arguments can be made in favor of transitioning traditional approaches toward a new data science paradigm.

First, to generate and use high-quality big data to support precision and generalizability assumptions, findable, accessible, interoperable, and reusable (FAIR) principles [28] ought to be adopted. In theory, following these criteria is important to develop new scientific studies and validate, or otherwise reproduce, at least part of already available works [29,30]. In practice, FAIR principles are hard to comply with, for several reasons.

Starting with "findable," few health data catalogs are available, and some are no longer being updated [31-33]. Considering access, while new instruments [34-37] have presented legislation and guiding frameworks on how data should be used and shared, their impact is controversial and their implementation is often complex, especially for medical research purposes [38-44]. Even follow-up developments, like the European Health Data Space [45], are viewed with skepticism by some member states regarding its practical feasibility and results [46,47].

Furthermore, data use is often limited by its interoperable characteristics. Technical heterogeneity due to different electronic health record systems, data standards, and data exchange protocols makes it difficult to share and integrate health data across multiple parties. Given all the challenges stated, reusability is practically impossible, outside some rare contexts [48,49].

Additionally, another set of arguments can be presented referring to the need for a systematic approach that does not rely on individuals' actions nor benefits from their limitations. First, we recognize that individuals are not capable of always acting with their own (or collective) interests in mind [50]. Specifically in their health data sharing attitudes, even though people may be aware of the value of data and potential privacy issues, their actions are often contrasting to their stated beliefs [51,52].

Then, we place the onus and burden of sustainable privacy protection on the system itself and are not reliant on the constant and best behavior of every individual agent [53]. Some

successful global data sharing efforts [48,54,55] have already proven that after the demanding set up, it is possible to not only improve access to data but also protect privacy by design [56].

New Technological Solutions

In response to these demands, recent breakthroughs in some computational domains, including secure processing units [57,58], differential privacy [59,60], and homomorphic encryption [61,62], have offered technical alternatives for the use of data in a privacy-preserving fashion. Being one of the more interesting approaches, decentralized learning architectures [63,64] enable data scattered across different silos (eg, health care providers) to be used to develop or validate pre-existing models. By combining the information derived from data present in each silo, it is possible to create more precise and generalizable models.

In general, local models are first developed using the party's own data. Second, only the model parameters (ie, information) are shared, usually with a central coordinator, responsible for aggregating the different local instances to create a new decentralized model. Throughout this process, data remain unmoved and are not accessed or manipulated by third parties.

Accordingly, such models can be produced without mobilizing or otherwise sharing the data set itself by parceling out the inferential process, on top of distributed computation principles [65]. By cyclically repeating this process, we can improve the model performance and include newly available data, fostering the development of continuously updated real-world evidence-based knowledge.

The potential for developing health care applications and generating value is clear [66]. Sharing models developed using different data sets can make the distributed solutions more comprehensive and adequate, due to robust internal and external validations. These approaches are useful to study medical conditions [67-69], especially those with limited prevalence or few observations, and prevent inadequate care due to mis- or underrepresentation of certain groups of patients [70].

Subsequently, significant agreements, like the one achieved between the United States and the European Union [71], promise a new platform for these technologies to be implemented while respecting their differing legislative frameworks. Such consensus can have a seismic impact on the way data science is conducted.

In the meantime, there still are many ongoing challenges regarding decentralized learning [18,72]. Some relevant issues are a lack of objective and measurable standard definitions of privacy and security [73], and server-client trust and honesty assumptions to computationally intensive and energetically demanding tasks [61,74]. Other important problems are the heterogeneity in distributed data and environments as well as fairness and respect for individual and local preferences [73].

However, above all else, the most pressing undertaking remains assessing the validity, relevance, and applicability of already published and available tools to inform and justify subsequent health technological appraisals and their implementation in real-world settings.

To do so, we must address the following questions: are decentralized health data models' performance superior (or noninferior) to current (centralized and local) approaches? and what are the reported privacy gains and the main resource demands?

Aims and Objectives

The systematic review based on this protocol aims to, first, compare the performance among health data models developed or validated using decentralized learning approaches (eg, federated and blockchain) to those developed using nondecentralized methods (eg, centralized and local). These can include applications such as automated diagnosis, segmentation of lesions or features, as well as mortality prediction.

The performance metrics used for model comparison will be the following: area under the receiver operating characteristic (AUROC) curve, F_1 score, Jensen-Shannon distance, sensitivity (or recall), specificity, accuracy, precision (or positive predictive value), negative predictive value, Dice score, as well as any metrics regarding the convergence step. Comparisons will be made only among models using the same type of data (eg, tabular and images) and clinical application (eg, diagnostic and survival). Our secondary aims are to compare the privacy compromise (eg, privacy budget) and resource use (eg, computation power and wall-time) among these health data models using different architectures.

With this review study, our goal is to summarize the progress and findings from state-of-the-art decentralized learning models in health care, in comparison to their currently used counterparts. These results are expected to clarify the consensuses and heterogeneities reported and help guide the research and development of new robust and sustainable applications to address the health data privacy problem, with applicability in real-world (clinical) settings.

Methods

Eligibility Criteria

Regarding the inclusion criteria, relevant studies are original research papers (including published, unpublished, and preprints) targeting one or multiple specific human medical conditions. They should be comparing 2 types of health model learning approaches—one decentralized (eg, federated and blockchain) and other nondecentralized (eg, centralized and local). For this review, decentralized learning architectures are defined as a machine learning approach to use data available from multiple parties, without sharing them with a single entity, to extract information [18].

As there may be some confusion regarding the terms "decentralized" and "distributed," we consider the first as the most appropriate and rigorous designation of our study field, including federated and blockchain architectures. The latter is broader in scope and refers to a computational subject that includes, but also precedes, the current advances and innovations [65]. Nevertheless, studies will be included regardless of the adopted terminology if the definition or methodologies used

are in accordance with the above definition of decentralized learning architectures. Both types of models must report at least one of the following model performance metrics: AUROC curve, F_1 score, Jensen-Shannon distance, sensitivity (or recall), specificity, accuracy, precision (or positive predictive value), negative predictive value, Dice score, as well as any metrics regarding the convergence step.

Regarding the exclusion criteria, papers published before 2012 will not be considered for this analysis, due to the following reasons. First, the earliest decentralized learning framework proposals in health care are only introduced [75,76] or applied [77] after this year. Second, seminal studies detailing and developing the current definition of these concepts were published in 2016 [78,79], with related works being as available as early 2014 [63,80,81]. Moreover, in no systematic review for health care implementations are there primary studies published before 2016 [82-86].

Each synthesis will consider studies with common types of clinical applications (intervention, diagnosis, etiology, prevention, prognosis or prediction, quality of life or meaning, and therapy), types of models, and types of data used. Whenever possible, they will be grouped by health problems.

Information Sources

Recognizing the interdisciplinarity nature of the research being made on health data models, several databases will be queried—some more specific to biomedical scientific research (ie, PubMed, SpringerLink, and Lippincott Williams & Wilkins), some more specific to computer science and informatics engineering (ie, the Association for Computing and Machinery Digital Library or Guide to Computing Literature and IEEE Xplore), while others were more general (ie, Wiley Online Library, Scopus, Web of Science, and Lens).

Moreover, 2 registries for systematic reviews (Cochrane Database of Systematic Reviews and PROSPERO) were surveyed for submissions related to these topics, in order to look for additional primary papers. Furthermore, queries were

also conducted in databases, which included research papers not peer-reviewed or otherwise unpublished (eg, medRxiv and arXiv). For every listed source, searches are expected to be conducted during March 2023.

Five experts in relevant scientific fields (from data science, health informatics, and decentralized learning approaches) will be contacted to ask for suggestions of additional bibliography not included in the selection process, without the knowledge of the selected or rejected papers. Such recommendations will be considered worthy of consideration and included if eligibility criteria are met. Due to the reasons stated in the eligibility criteria, it was deemed appropriate to restrict the search to papers published in 2012 or later.

Search Strategy

Overview

Given the recency of this research domain and the expected limited number papers, it was imperative to devise a broad search strategy. This was materialized in choices such as including several databases, as well as using synonyms and wildcards in the query.

However, due to the popularity of some of the query terms—for example, distributed, model, training, and health—some procedures were adopted to filter noise. For instance, words like "distributed" and "model" should have a limited number of words in between, for the finding to be relevant. Moreover, search engines have heterogeneous features, which make it difficult to conduct the desired exploration.

Hence, a composite search strategy was adopted. The first part, optimizing for comprehensiveness, was focused on writing the query and electing relevant filters for each database used (see "Part 1—Database Query" section, Table 1, and Figure 1). Subsequently, a filtration process was applied, using regular expressions (RegEx) code, to make up for the lacking features of the databases used—such as word proximity limits, operators, and metadata fields searched (see "Part 2—Results Filtration" section).

Table 1. General query terms by group.

Group	Terms
A—Model architecture	decentrali*, distributed, federated, central*, multi-party computation, blockchain
B—Model synonym	learn*, model*, network*, AI[a], artificial intelligence, ML[b], machine learning, train*, tensor*, perceptron, algorithm*
C—Health related	health*, medic*, clinic*, patient*, physician*, doctor*
D—Performance metrics	AUROC[c], ROC[d], receiver operating characteristic curve, F_1, Jensen-Shannon, sensitivity, recall, specificity, accuracy, precision, predictive value, Dice, conversion

[a]AI: artificial intelligence.

[b]ML: machine learning.

[c]AUROC: area under the receiver operating characteristic.

[d]ROC: receiver operating characteristic.

Figure 1. Example of search query as used on Lens database. AI: artificial intelligence; AUROC: area under the receiver operating characteristic; ML: machine learning; ROC: receiver operating characteristic.

Part 1—Database Query

First, a simpler version of the query, suitable for all search engines, was used to retrieve a less specific group of abstracts.

For groups A, B, and C, the fields Title, Abstract, Keyword, and Field of Study, when available, will be searched. Terms from group A and B must be near each other, with a maximum of 2 words in between them. For terms in group D, the full-text document will be searched. The query will not be case-sensitive. The * symbol represents the wild card.

As per eligibility criteria, only primary papers from 2012 and beyond will be relevant. Thus, the search query will look for papers with at least 1 term, within the considered fields, from every group.

For each source, a specific query will be produced. Documentation will be made available providing the exact search string, a URL (if possible), and other details, such as filters applied.

Part 2—Results Filtration

As some indispensable terms are very prevalent in publications, such as "model" and "distribution," increasing the relative useful yield of the query and the number of studies retrieved, we conducted a processing task to filter irrelevant studies.

To do so, using RegEx code in R (R Foundation for Statistical Computing), we simulated a "within" operator. It was developed to only capture studies in which the term referring to the model architecture (group A) and the one referring to the model synonym (group B) have no more than 2 other terms separating them. Finally, while this process does not perfectly compensate for the limitation in the databases search features and variation, it is expected to offset the most significant differences and not significantly compromise the pursuit of relevant primary papers.

Selection Process

The selection of the primary studies will comprise 2 moments: the screening phase—when the papers are appraised using only their title and abstract, and the inclusion decision phase—when the papers are appraised using their full-text versions. To manage the appraisal of the retrieved primary studies, the Rayyan [87] software suit (Rayyan) will be used. Before the screening phase, exact matches and additional duplicates will be removed, quantifying the number of papers ruled out.

All papers retrieved through the application of the search methods detailed above will be screened using their title and abstract by 3 researchers acting independently and blinded to each other's decisions. Excluded papers should be labeled using the first unmet criteria of the inclusion criteria.

When there is not a complete agreement on the inclusion (or exclusion) decision, the evaluating researchers will discuss and attempt to achieve a consensus, with potential consultation with the other authors. If that is not possible, the majority decision will be chosen.

After the screening phase, the same sequence of study appraisal and disagreement resolution will be conducted for the full-text versions of the papers. The flow of papers included and excluded will be represented in a diagram, where quantity, source (search method and database), and reason for decision will be explicit.

The flow diagrams for the papers included and excluded will be represented according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. In the end, a final list of all the primary studies selected to be included in the review will be presented with a complete reference and, when possible, a DOI link.

Data Collection Process

For each study, data extraction will be conducted by 3 researchers, who will work independently, in a blinded fashion, by reading the full-text versions (or other versions if full text is not available or does not exist). They will use custom-made web-based forms, including the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) checklist items for reporting quality and

risk of bias assessment [88]. These forms will be piloted before the data collection process.

After completion, retrieved data will be compared to check for errors or inconsistencies and discuss any doubts. The researchers will decide by consensus on the last version of the database of the data collected.

Data Items

A pilot study was conducted to refine the list of relevant data variables to collect, using 100 papers of a developing query. The complete list of variables for collection (whenever available) is the following.

First, general attributes are to be collected, namely, title, author or authors, abstract, publication date, country or countries of the research institutions, type of publication, the journal or publisher, as well as the PICO (Population, Intervention, Comparison, and Outcomes) question.

Specific data points regarding health topics, the type of clinical application (eg, diagnostic and survival), and the applicable medical domain will also be registered.

Concerning the data used, information about the data set size, namely, the number of observations or cases, the number of variables, and the volume in megabytes or gigabytes will be detailed. As far as the nature of the data goes, it will be marked as either synthetic or real, and, in the latter, whether the data were collected for the study (primary source) or not (secondary source). The original data holders will be described by their number, type, and localization, as well as the storage architecture used (eg, centralized, decentralized, and local). Other data points will be the places (geographical and institutional) where data were analyzed, the ethics and legal permissions reported, as well as the data type used (eg, text and images), alongside their format (eg, tabular), and conversion processes.

The type of all models reported are, for example, deep neural network and decision trees, and their performance metrics are specifically AUROC curve, F_1 score, Jensen-Shannon distance, sensitivity or recall, specificity, accuracy, precision or positive predictive value, negative predictive value, Dice score, as well as any metrics regarding the convergence step. The training methods will also be registered, including their update routine (rounds, epochs, quorum, update frequency, update content, and protocols for data communication) and statistical methods used. The validation process used will also be described.

Regarding the architectures used, their type, data flow, and client type (cross-silo vs cross-device vs both) will be extracted, as well as the personalization or customization step and the aggregation methods used. Importantly, we will detail all the model architectures and types compared, and their hypothesis tests.

For our secondary aim, we will collect data on the privacy cost measures (privacy budget, k-anonymity, fingerprinting, entropy, or others) and resource consumption (from computer resources, hardware and software specs, and energy to time and number of rounds), as well as additional security and privacy protection measures.

Lastly, we will record the reason for using decentralized approaches, whether the data and code used are available, and the reported challenges and limitations. No assumptions will be made regarding missing or unclear information—those findings will be reported as such.

Study Risk of Bias Assessment

For each selected paper, the CHARMS [88] and PROBAST (Prediction Model Risk of Bias Assessment Tool) [89] checklists will be used to assess the risk of bias of included research works. The full results of such an appraisal will be presented on a table and considered for the discussion of the results. If deemed relevant, depending on the final list of results, other more specific tools may be used.

Effect Measures

All the effect measures in the original studies will be presented. Whenever unavailable, and if possible, the difference between the AUROC curve estimates for the models of different architectures will be calculated. When multiple rounds of model development, validation, or application exist, the difference in time and rounds to a set performance target will be calculated.

If multiple values for each model are available, the median and mean values will also be used to calculate the differences. If AUROC curve values are not available, other commonly found metrics may also be used.

Synthesis Methods

A qualitative analysis of the evidence will be conducted alongside a descriptive synthesis of the results. Each synthesis will consider the types of clinical applications, models, and data used. Whenever possible they will be grouped by health problems.

Missing data will be reported as "Missing." Whenever a synthesis method is not applicable, it will be labeled as "Not Applicable."

Due to the heterogeneity of the applications and model characteristics, it is not expected to be able to synthesize results in a quantitative fashion. Accordingly, neither a meta-analysis nor a sensitivity analysis will be performed.

Reporting Bias Assessment

For every eligible study, the authors will look up preresearch registers of protocols and check for differences in the published work. Additionally, they will assess whether any pertinent statistics or analyses are unreported. All corresponding authors will be contacted to assess if they have any nonreporting experience with any version of their published work or regarding other unpublished research, as suggested by Ammenwerth and de Keizer [90].

Questions may include the following: Which information systems did you evaluate in the last 3 years? Where did you publish the results? and If you did not publish them, what were the reasons for that decision? (here, some common reasons could be selected or added using free text fields).

Moreover, specific efforts will be made to identify omissions of some measured outcomes, as well as selective reporting of

only "significant" findings from among several analyses undertaken.

Certainty Assessment

Certainty assessment procedures will be performed if appropriate instruments are available at the time of review completion.

Other Information

Efforts will be made to make available the query links (or prompts) for each database used, as well as the RegEx filtering code, the templates for data collection forms, and the data extracted from the included studies, and any other resources, which might be used for the subsequent review.

Results

The queries and data extractions are expected to start on February 28, 2023, and end by July 31, 2023. The research protocol was registered with PROSPERO, under the number 393126, on February 3, 2023.

It is expected that the systematic review will summarize the progress and findings from state-of-the-art decentralized learning models in health care, in comparison to their local and centralized counterparts. These results will help in clarifying the consensus and heterogeneities reported among different models and studies, as well as guide the research and development of new robust and sustainable applications to address the health data privacy problem, with applicability in real-world (clinical) settings.

Discussion

Principal Findings

As the systematic review is yet to be conducted, no specific results can be reported at this time. However, this will be the first systematic review on the comparison of decentralized health data models to more common local or centralized approaches that is both comprehensive and focused on objective performance metrics. Due to our exhaustive search strategy and the plurality of data sources identified, this work will present the clearest situational assessment yet of the application of these technologies in health care.

We hope that, by highlighting effective models and their developments, as well as identifying those which underperformed, we can shed light on more fruitful and interesting directions for our peers. In addition, by considering the development costs and privacy gains, as reported in the primary papers, it is expected that this review may be useful for health technology assessment and evidence-based decisions, from health professionals, data scientists, and policy makers alike.

By providing the original search queries and documentation on the methods used, we open the possibility for researchers to use this protocol and upcoming supplemental materials to not only audit and validate our work but also conduct updated versions of this review as new evidence is published.

Strengths and Limitations

While decentralized health data models are a nascent and growing field of research, all appropriate steps will be taken to ensure a comprehensive and exhaustive search of available literature. To do so, we will consider a variety of sources (both specific to bio- and computational sciences and generalist databases), as well as preprint works. Alongside these efforts, a substantive yet clear search and selection procedures have been detailed, with 3 authors selecting primary papers for inclusion in the review. Together, these will also increase the sensitivity and specificity of the search results and the studies included for analysis.

Moreover, this protocol is already registered with PROSPERO and is written in accordance with the PRISMA 2020 guidelines, which will confer more validity and accountability for the results provided. In conjunction with the review, these actions will allow for reproducible and auditable work. For instance, it may be useful, with the necessary adaptations, to repeat this study, when even more evidence is available.

Some limitations of this work include the expected heterogeneity among primary studies, complicating synthesis and comparability among models, and the rapid evolution of the field. Given the unusual population (the P in PICO) of this systematic review, the appraisal of the primary studies may be made difficult by a lack of appropriate tools to assess the risk of bias, especially in reporting, and the certainty of the findings.

Comparison With Prior Work

While there are some systematic reviews on the topic already available, they present several important shortcomings in the size and scope [82-84], the specificity of health care applications [85], and the capability and comprehensiveness of query prompts [86]. Moreover, to the best of our knowledge, none of them were accompanied by a protocol publication or registry before the corresponding review.

Therefore, a more robust and valid synthesis of the currently available scientific evidence must be conducted. We consider the proposed systematic review is capable of achieving that result.

Conclusions

Increasing pressure to develop better, more effective, and cost sensitive care, juxtaposed with the privacy-preserving principles and methodologies, has created a considerable demand for alternative health data model development, validation, and application procedures.

While decentralized approaches, such as those built with federated and blockchain architectures, promise considerable gains for extracting information out of health data, there is much uncertainty regarding how they compare to current centralized and local models, as well as their associated privacy gains and their resource consumptions.

This protocol is the first, at the time of this writing, to outline a systematic review on health data decentralized models, aiming not only to capture the rich variety and complexity of available

research but also to generate a rigorous and comprehensive assessment of the synthesis of their results and conclusions.

It is expected that such work will have implications for this budding research field and policy making, especially for those working with health data privacy matters. This review will highlight the advances and shortcomings of these approaches to better inform the development and application of new tools in service of patients' privacy while hoping to guide future research.

Data Availability

Documentation will be provided regarding the specific queries used, tailored to each source, including their adapted formulation and filters, to ease reproducibility. Whenever possible, a direct URL link will be included. In addition, the code used to compile and filter the results of the various extraction files will be freely available in a GitHub repository. The original lists of all papers (as extracted) as well as other reports with all those subjected to the screening process will be accessible in the same repository. Moreover, the original data collected from each of the studies included in the review will be shared in a Research Data Repository.

Conflicts of Interest

None declared.

References

1. World population prospects 2022: summary of results. United Nations Department of Economic and Social Affairs, Population Division. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [accessed 2023-05-30]
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease study 2019. Lancet 2020;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Europe 2022: state of health in the EU cycle. Organisation for Economic Co-operation and Development (OECD). URL: https://read.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2022_507433b0-en [accessed 2023-01-11]
4. Ortiz-Ospina E. Long-term perspective on government healthcare spending. Our World in Data. URL: https://ourworldindata.org/when-did-the-provision-of-healthcare-first-become-a-public-policy-priority [accessed 2023-01-11]
5. National health expenditures 2021 highlights. Centers for Medicare & Medicaid Services. URL: https://www.cms.gov/files/document/highlights.pdf [accessed 2023-05-30]
6. Brouwer W, van Baal P, van Exel J, Versteegh M. When is it too expensive? Cost-effectiveness thresholds and health care decision-making. Eur J Health Econ 2019;20(2):175-180 [FREE Full text] [doi: 10.1007/s10198-018-1000-4] [Medline: 30187251]
7. Abdulkareem M, Petersen S. The promise of AI in detection, diagnosis, and epidemiology for combating COVID-19: beyond the hype. Front Artif Intell 2021;4:652669 [FREE Full text] [doi: 10.3389/frai.2021.652669] [Medline: 34056579]
8. Jochems A, Deist TM, El Naqa I, Kessler M, Mayo C, Reeves J, et al. Developing and validating a survival prediction model for NSCLC patients through distributed learning across 3 countries. Int J Radiat Oncol Biol Phys 2017;99(2):344-352 [FREE Full text] [doi: 10.1016/j.ijrobp.2017.04.021] [Medline: 28871984]
9. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine learning methods in drug discovery. Molecules 2020;25(22):52-77 [FREE Full text] [doi: 10.3390/molecules25225277] [Medline: 33198233]
10. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Mol Divers 2021;25(3):1315-1360 [FREE Full text] [doi: 10.1007/s11030-021-10217-3] [Medline: 33844136]
11. Benedetto U, Dimagli A, Sinha S, Cocomello L, Gibbison B, Caputo M, et al. Machine learning improves mortality risk prediction after cardiac surgery: systematic review and meta-analysis. J Thorac Cardiovasc Surg 2022;163(6):2075-2087 [doi: 10.1016/j.jtcvs.2020.07.105] [Medline: 32900480]
12. Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. Lancet Infect Dis 2022;22(9):1293-1302 [FREE Full text] [doi: 10.1016/S1473-3099(22)00320-6] [Medline: 35753318]

13. GBD 2019 Colorectal Cancer Collaborators. Global, regional, and national burden of colorectal cancer and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. Lancet Gastroenterol Hepatol 2022;7(7):627-647 [FREE Full text] [doi: 10.1016/S2468-1253(22)00044-9] [Medline: 35397795]

14. SCORE2-OP Working Group and ESC Cardiovascular Risk Collaboration. SCORE2-OP risk prediction algorithms: estimating incident cardiovascular event risk in older persons in four geographical risk regions. Eur Heart J 2021;42(25):2455-2467 [FREE Full text] [doi: 10.1093/eurheartj/ehab312] [Medline: 34120185]

15. Kline JA, Mitchell AM, Kabrhel C, Richman PB, Courtney DM. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. J Thromb Haemost 2004;2(8):1247-1255 [FREE Full text] [doi: 10.1111/j.1538-7836.2004.00790.x] [Medline: 15304025]

16. Roser M, Ritchie H, Mathieu E. Technological change. Our World in Data. 2013. URL: https://ourworldindata.org/technological-change [accessed 2023-01-13]

17. Cao L. Data science: a comprehensive overview. ACM Comput Surv 2017 Jun 29;50(3):1-42 [FREE Full text] [doi: 10.1145/3076253]

18. Ludwig H, Baracaldo N, editors. Federated Learning: A Comprehensive Overview of Methods and Applications. Cham: Springer; 2022.

19. Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. Comput Struct Biotechnol J 2021;19:2833-2850 [FREE Full text] [doi: 10.1016/j.csbj.2021.05.010] [Medline: 34025952]

20. Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: current status and future challenges—a systematic review. Artif Intell Med 2021;120:102164 [FREE Full text] [doi: 10.1016/j.artmed.2021.102164] [Medline: 34629152]

21. Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. J Biomed Inform 2017;73:14-29 [FREE Full text] [doi: 10.1016/j.jbi.2017.07.012] [Medline: 28729030]

22. D'Antoni F, Russo F, Ambrosio L, Vollero L, Vadalà G, Merone M, et al. Artificial intelligence and computer vision in low back pain: a systematic review. Int J Environ Res Public Health 2021;18(20):10909 [FREE Full text] [doi: 10.3390/ijerph182010909] [Medline: 34682647]

23. Farooqi MM, Shah MA, Wahid A. Big data in healthcare: a survey. In: Khan F, Jan MA, Alam M, editors. Applications of Intelligent Technologies in Healthcare. Cham: Springer; 2019:143-152

24. Hussein AA. Fifty-six big data V's characteristics and proposed strategies to overcome security and privacy challenges (BD2). J Inf Secur 2020;11(4):304-328 [FREE Full text] [doi: 10.4236/jis.2020.114019]

25. Universal declaration of human rights. United Nations. URL: https://www.un.org/en/about-us/universal-declaration-of-human-rights [accessed 2023-01-13]

26. List of GDPR fines. GDPR Enforcement Tracker. URL: https://www.enforcementtracker.com [accessed 2023-01-13]

27. Healthcare data breach statistics. HIPAA Journal. URL: https://www.hipaajournal.com/healthcare-data-breach-statistics/ [accessed 2023-02-10]

28. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016;3(1):160018 [FREE Full text] [doi: 10.1038/sdata.2016.18] [Medline: 26978244]

29. Nosek BA, Errington TM. What is replication? PLoS Biol 2020 Mar;18(3):e3000691 [FREE Full text] [doi: 10.1371/journal.pbio.3000691] [Medline: 32218571]

30. Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, et al. Investigating the replicability of preclinical cancer biology. eLife 2021;10:e71601 [FREE Full text] [doi: 10.7554/eLife.71601] [Medline: 34874005]

31. Oliveira JL, Trifan A, Silva LAB. EMIF catalogue: a collaborative platform for sharing and reusing biomedical data. Int J Med Inform 2019;126:35-45 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.02.006] [Medline: 31029262]

32. Resources database. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. URL: https://www.encepp.eu/encepp/resourcesDatabase.jsp [accessed 2023-02-08]

33. Queralt-Rosinach N, Kaliyaperumal R, Bernabé CH, Long Q, Joosten SA, van der Wijk HJ, BEAT-COVID Group, COVID-19 LUMC Group. Applying the FAIR principles to data in a hospital: challenges and opportunities in a pandemic. J Biomed Semantics 2022;13(1):12 [FREE Full text] [doi: 10.1186/s13326-022-00263-7] [Medline: 35468846]

34. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (text with EEA relevance). Vol 119. EUR-Lex. 2016. URL: http://data.europa.eu/eli/reg/2016/679/oj/eng [accessed 2023-01-11]

35. Health Insurance Portability and Accountability Act of 1996 (HIPAA). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/phlp/publications/topic/hipaa.html [accessed 2023-01-11]

36. Office of the Attorney General. California Consumer Privacy Act (CCPA). State of California—Department of Justice. URL: https://oag.ca.gov/privacy/ccpa [accessed 2023-01-13]

37. Sharing and reuse of health-related data for research purposes: WHO policy and implementation guidance. World Health Organization. URL: https://www.who.int/publications-detail-redirect/9789240044968 [accessed 2023-01-13]

38. Peloquin D, DiMaio M, Bierer B, Barnes M. Disruptive and avoidable: GDPR challenges to secondary research uses of data. Eur J Hum Genet 2020;28(6):697-705 [FREE Full text] [doi: 10.1038/s41431-020-0596-x] [Medline: 32123329]

39. Staunton C, Slokenberga S, Mascalzoni D. The GDPR and the research exemption: considerations on the necessary safeguards for research biobanks. Eur J Hum Genet 2019;27(8):1159-1167 [FREE Full text] [doi: 10.1038/s41431-019-0386-5] [Medline: 30996335]

40. Christofidou M, Lea N, Coorevits P. A literature review on the GDPR, COVID-19 and the ethical considerations of data protection during a time of crisis. Yearb Med Inform 2021;30(1):226-232 [FREE Full text] [doi: 10.1055/s-0041-1726512] [Medline: 34479394]

41. Molnár-Gábor F, Sellner J, Pagil S, Slokenberga S, Tzortzatou-Nanopoulou O, Nyström K. Harmonization after the GDPR? Divergences in the rules for genetic and health data sharing in four member states and ways to overcome them by EU measures: Insights from Germany, Greece, Latvia and Sweden. Semin Cancer Biol 2022 Sep;84:271-283 [FREE Full text] [doi: 10.1016/j.semcancer.2021.12.001] [Medline: 34896635]

42. Lawlor RT. The impact of GDPR on data sharing for european cancer research. Lancet Oncol 2023;24(1):6-8 [doi: 10.1016/S1470-2045(22)00653-2] [Medline: 36400103]

43. Clarke N, Vale G, Reeves EP, Kirwan M, Smith D, Farrell M, et al. GDPR: an impediment to research? Ir J Med Sci 2019;188(4):1129-1135 [FREE Full text] [doi: 10.1007/s11845-019-01980-2] [Medline: 30734900]

44. Mandl KD, Perakslis ED. HIPAA and the leak of "deidentified" EHR data. N Engl J Med 2021;384(23):2171-2173 [doi: 10.1056/NEJMp2102616] [Medline: 34110112]

45. European health data space. European Commission. URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en [accessed 2023-01-11]

46. Pištorová B, Plevák O. Stakeholders doubtful EU health data space will launch on schedule. Euractiv. URL: https://www.euractiv.com/section/health-consumers/news/stakeholders-doubtful-eu-health-data-space-will-launch-on-schedule/ [accessed 2023-01-30]

47. van Kessel R, Wong BLH, Forman R, Gabrani J, Mossialos E. The European health data space fails to bridge digital divides. BMJ 2022;378:e071913 [doi: 10.1136/bmj-2022-071913] [Medline: 35803600]

48. European Health Data Evidence Network (EHDEN). URL: https://www.ehden.eu/ [accessed 2023-01-13]

49. Doiron D, Burton P, Marcon Y, Gaye A, Wolffenbuttel BHR, Perola M, et al. Data harmonization and federated analysis of population-based studies: the BioSHaRE project. Emerg Themes Epidemiol 2013;10(1):12 [FREE Full text] [doi: 10.1186/1742-7622-10-12] [Medline: 24257327]

50. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974;185(4157):1124-1131 [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]

51. Li XB, Liu X, Motiwalla L. Valuing personal data with privacy consideration. Decis Sci 2021;52(2):393-426 [FREE Full text] [doi: 10.1111/deci.12442] [Medline: 34732907]

52. Waldman AE. Cognitive biases, dark patterns, and the 'privacy paradox'. Curr Opin Psychol 2020;31:105-109 [FREE Full text] [doi: 10.1016/j.copsyc.2019.08.025] [Medline: 31590106]

53. Kindervag J. No more chewy centers: introducing the zero trust model of information security. Forrester Research, Inc. 2010. URL: https://media.paloaltonetworks.com/documents/Forrester-No-More-Chewy-Centers.pdf [accessed 2023-05-30]

54. Data standardization. Observational Health Data Sciences and Informatics (OHDSI). URL: https://www.ohdsi.org/data-standardization/ [accessed 2023-01-13]

55. Mitigating the COVID-19 outbreak through global data sharing. World Health Organization. URL: https://www.who.int/teams/health-care-readiness/covid-19/data-platform/mitigating-the-covid-19-outbreak-through-global-data-sharing [accessed 2023-01-13]

56. Cavoukian A. The 7 foundational principles. Privacy by Design. URL: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf [accessed 2023-05-30]

57. Joint Action Towards the European Health Data Space. URL: https://tehdas.eu/ [accessed 2023-01-13]

58. The Scottish National Safe Haven—a secure research environment for health data research. Health Data Research UK. URL: https://www.hdruk.ac.uk/news/the-scottish-national-safe-haven-a-secure-research-environment-for-health-data-research/ [accessed 2023-01-13]

59. Wasserman L, Zhou S. A statistical framework for differential privacy. ArXiv Preprint posted online on October 2, 2009 [FREE Full text] [doi: 10.48550/arXiv.0811.2501]

60. Choudhury O, Gkoulalas-Divanis A, Salonidis T, Sylla I, Park Y, Hsu G, et al. Differential privacy-enabled federated learning for sensitive health data. ArXiv Preprint posted online on February 27, 2020 [FREE Full text] [doi: 10.48550/arXiv.1910.02578]

61. Kucherov NN, Deryabin MA, Babenko MG. Homomorphic encryption methods review. 2020 Presented at: 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus); January 27-30, 2020; St. Petersburg and Moscow, Russia p. 370-373 URL: https://ieeexplore.ieee.org/document/9039110 [doi: 10.1109/eiconrus49466.2020.9039110]

XSL•FO

RenderX

62. Gentry C. Fully homomorphic encryption using ideal lattices. In: Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing. 2009 Presented at: STOC '09; May 31-June 2, 2009; Bethesda, MD p. 169-178 URL: https://dl.acm.org/doi/proceedings/10.1145/1536414 [doi: 10.1145/1536414.1536440]

63. Konečný J, McMahan B, Ramage D. Federated optimization: distributed optimization beyond the datacenter. ArXiv Preprint posted online on November 11, 2015 [FREE Full text] [doi: 10.48550/arXiv.1511.03575]

64. Kuo TT, Ohno-Machado L. ModelChain: decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks. ArXiv Preprint posted online on 6 Feb 2018 [FREE Full text] [doi: 10.48550/arXiv.1802.01746]

65. van Steen M, Tanenbaum AS. Distributed Systems. 3rd edition. Scotts Valley, CA: CreateSpace Independent Publishing Platform; 2017.

66. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. NPJ Digit Med 2020;3(1):1-7 [FREE Full text] [doi: 10.1038/s41746-020-00323-1] [Medline: 33015372]

67. Pati S, Baid U, Edwards B, Sheller M, Wang S, Reina GA, et al. Federated learning enables big data for rare cancer boundary detection. Nat Commun 2022;13(1):7346 [FREE Full text] [doi: 10.1038/s41467-022-33407-5] [Medline: 36470898]

68. Choudhury O, Park Y, Salonidis T, Gkoulalas-Divanis A, Sylla I, Das AK. Predicting adverse drug reactions on distributed health data using federated learning. AMIA Annu Symp Proc 2020;2019:313-322 [FREE Full text] [Medline: 32308824]

69. Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, et al. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach. JMIR Med Inform 2021;9(1):e24207 [FREE Full text] [doi: 10.2196/24207] [Medline: 33400679]

70. Ku E, Amaral S, McCulloch CE, Adey DB, Li L, Johansen KL. Comparison of 2021 CKD-EPI equations for estimating racial differences in preemptive waitlisting for kidney transplantation. Clin J Am Soc Nephrol 2022;17(10):1515-1521 [doi: 10.2215/CJN.04850422] [Medline: 36122938]

71. U.S. and EU to launch first-of-its-kind AI agreement. Reuters. URL: https://www.reuters.com/technology/white-house-european-commission-launch-first-of-its-kind-ai-agreement-2023-01-27/ [accessed 2023-01-30]

72. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, et al. Advances and open problems in federated learning. ArXiv Preprint posted online on March 9, 2021 [FREE Full text] [doi: 10.48550/arXiv.1912.04977]

73. Jin Y, Zhu H, Xu J, Chen Y. Federated Learning: Fundamentals and Advances. Berlin-Heidelberg: Springer; 2022.

74. Cai TT, Wang Y, Zhang L. The cost of privacy: optimal rates of convergence for parameter estimation with differential privacy. Ann Statist 2021;49(5):2825-2850 [doi: 10.1214/21-aos2058]

75. El Emam K, Samet S, Arbuckle L, Tamblyn R, Earle C, Kantarcioglu M. A secure distributed logistic regression protocol for the detection of rare adverse drug events. J Am Med Inform Assoc 2013;20(3):453-461 [FREE Full text] [doi: 10.1136/amiajnl-2011-000735] [Medline: 22871397]

76. Doiron D, Marcon Y, Fortier I, Burton P, Ferretti V. Software application profile: opal and mica: open-source software solutions for epidemiological data management, harmonization and dissemination. Int J Epidemiol 2017;46(5):1372-1378 [FREE Full text] [doi: 10.1093/ije/dyx180] [Medline: 29025122]

77. Wolfson M, Wallace SE, Masca N, Rowe G, Sheehan NA, Ferretti V, et al. DataSHIELD: resolving a conflict in contemporary bioscience—performing a pooled analysis of individual-level data without sharing the data. Int J Epidemiol 2010;39(5):1372-1382 [FREE Full text] [doi: 10.1093/ije/dyq111] [Medline: 20630989]

78. Konečný J, McMahan H, Yu F, Richtárik P, Suresh A, Bacon D. Federated learning: strategies for improving communication efficiency. ArXiv Preprint posted online on October 30, 2017 [FREE Full text] [doi: 10.48550/arXiv.1610.05492]

79. McMahan HB, Moore E, Ramage D, Hampson S, Agüera y Arcas B. Communication-efficient learning of deep networks from decentralized data. ArXiv Preprint posted online on January 26, 2017 [FREE Full text] [doi: 10.48550/arXiv.1602.05629]

80. McMahan B, Streeter M. Delay-tolerant algorithms for asynchronous distributed online learning. 2014 Presented at: Advances in Neural Information Processing Systems 27 (NIPS 2014); 2014; Montreal, QB URL: https://proceedings.neurips.cc/paper/2014/hash/5cce8dede893813f879b873962fb669f-Abstract.html

81. McMahan HB. A survey of algorithms and analysis for adaptive online learning. ArXiv Preprint posted online on November 9, 2015 [FREE Full text] [doi: 10.48550/arXiv.1403.3465]

82. Zerka F, Barakat S, Walsh S, Bogowicz M, Leijenaar RTH, Jochems A, et al. Systematic review of privacy-preserving distributed machine learning from federated databases in health care. JCO Clin Cancer Inform 2020;4:184-200 [FREE Full text] [doi: 10.1200/CCI.19.00047] [Medline: 32134684]

83. Agbo CC, Mahmoud QH, Eklund JM. Blockchain technology in healthcare: a systematic review. Healthcare (Basel) 2019;7(2):56 [FREE Full text] [doi: 10.3390/healthcare7020056] [Medline: 30987333]

84. Crowson MG, Moukheiber D, Arévalo AR, Lam BD, Mantena S, Rana A, et al. A systematic review of federated learning applications for biomedical data. PLOS Digit Health 2022;1(5):e0000033 [FREE Full text] [doi: 10.1371/journal.pdig.0000033] [Medline: 36812504]

85. Qammar A, Karim A, Ning H, Ding J. Securing federated learning with blockchain: a systematic literature review. Artif Intell Rev 2023;56(5):3951-3985 [FREE Full text] [doi: 10.1007/s10462-022-10271-9] [Medline: 36160367]

86. Antunes RS, André da Costa C, Küderle A, Yari IA, Eskofier B. Federated learning for healthcare: systematic review and architecture proposal. ACM Trans Intell Syst Technol 2022;13(4):1-23 [doi: 10.1145/3501813]

87.   Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev
       2016;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
88.   Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data
       extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11(10):e1001744
       [FREE Full text] [doi: 10.1371/journal.pmed.1001744] [Medline: 25314315]
89.   Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias
       and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019 Jan 01;170(1):W1-W33
       [FREE Full text] [doi: 10.7326/M18-1377] [Medline: 30596876]
90.   Ammenwerth E, de Keizer N. A viewpoint on evidence-based health informatics, based on a pilot survey on evaluation
       studies in health care informatics. J Am Med Inform Assoc 2007;14(3):368-371 [FREE Full text] [doi: 10.1197/jamia.M2276]
       [Medline: 17329724]

Abbreviations

AUROC: area under the receiver operating characteristic
CHARMS: Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies
FAIR: findable, accessible, interoperable, and reusable
PICO: Population, Intervention, Comparison, and Outcomes
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROBAST: Prediction Model Risk of Bias Assessment Tool
RegEx: regular expression