BMJ Open

Machine learning methods, applications and economic analysis to predict heart failure hospitalisation risk: a scoping review protocol

Joana Seringa ,[1,2] João Abreu,[1] Teresa Magalhaes [1,2]

To cite: Seringa J, Abreu J, Magalhaes T. Machine learning methods, applications and economic analysis to predict heart failure hospitalisation risk: a scoping review protocol. *BMJ Open* 2024;**14**:e083188. doi:10.1136/bmjopen-2023-083188

► Prepublication history for this paper is available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-083188).

Received 13 December 2023
Accepted 22 March 2024

[1]NOVA National School of Public Health, NOVA University Lisbon, Lisbon, Portugal
[2]NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal

Correspondence to
Joana Seringa;
jm.seringa@ensp.unl.pt

ABSTRACT

Introduction Machine learning (ML) has emerged as a powerful tool for uncovering patterns and generating new information. In cardiology, it has shown promising results in predictive outcomes risk assessment of heart failure (HF) patients, a chronic condition affecting over 64 million individuals globally.

This scoping review aims to synthesise the evidence on ML methods, applications and economic analysis to predict the HF hospitalisation risk.

Methods and analysis This scoping review will use the approach described by Arksey and O'Malley. This protocol will use the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Protocol, and the PRISMA extension for scoping reviews will be used to present the results. PubMed, Scopus and Web of Science are the databases that will be searched. Two reviewers will independently screen the full-text studies for inclusion and extract the data. All the studies focusing on ML models to predict the risk of hospitalisation from HF adult patients will be included.

Ethics and dissemination Ethical approval is not required for this review. The dissemination strategy includes peer-reviewed publications, conference presentations and dissemination to relevant stakeholders.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The heterogeneity of the studies can make it challenging to draw comparisons or conclusions.
⇒ This scoping review will not include grey literature (eg, unpublished studies and conference abstracts) and systematic reviews, which can result in an incomplete picture of the available evidence.
⇒ As a statical study, it may not capture the most recent developments in a rapidly evolving field, as is the case.
⇒ Despite the limitations, this scoping review will be a valuable tool for mapping the current literature on machine learning methods, applications and economic analysis to predict the heart failure hospitalisation risk, drawing recommendations and identifying research gaps.

INTRODUCTION

Machine learning (ML) refers to the capacity of a system to autonomously acquire, integrate and subsequently generate knowledge from extensive datasets, further extending this acquired knowledge by uncovering new information, all without the need for explicit programming.[1]

The learning process begins with acquiring observations or empirical data, comprising instances, first-hand experience or instructional input, aiming to uncover patterns within the dataset.[2]

ML applications have proven successful across multiple areas, such as financial, security, industrial, marketing, environmental and medical.[3–6] In the medical area, the application of ML has been particularly transformative and revolutionary, and it has fundamentally changed how healthcare is delivered, leading to improved diagnostics, personalised treatment plans and enhanced patient outcomes.[6] Some practical examples in the medical area are medical imaging, genetic information, drug combinations, population-wide patient health outcomes and natural language processing of existing medical documentation.[5 6]

In cardiology, ML has applications in predictive risk assessment, classification and identification of prognostic phenotypes among patients with heart failure (HF).[7 8] HF is a chronic medical condition where the heart is unable to pump enough blood to meet the body's needs.[9] It is not a sudden stoppage of the heart but rather a progressive condition that can develop over time.[9] Additionally, HF is a complex and life-threatening syndrome marked by substantial morbidity and mortality, diminished functional capacity and quality of life, as well as elevated costs stemming from decompensation and subsequent hospitalisation.[10] This condition affects more than 64 million people worldwide, with a global economic burden estimated at $108 billion per annum.[11]

Table 1 Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Studies focusing on ML models to predict the risk of hospitalisation from HF patients aged 18 years and older.	Studies whose population is not of adult individuals with HF.
Studies that do not include economic analysis but focus on ML models to predict HF hospitalisations.	Studies whose cause of hospitalisation is not HF.
	Grey literature (eg, unpublished studies and conference abstracts) and systematic reviews.
	Studies in languages other than English.

HF, heart failure; ML, machine learning.

Considering the global burden of HF, its related economic implications, and the successful exponential growth of ML, it is imperative to conduct this scoping review to synthesise evidence on ML methods, applications and economic analysis to predict HF hospitalisation risk aiming to reduce unplanned hospitalisations, enhance patient outcomes and alleviate the associated financial strain associated with this condition.[10–12] Previous endeavours, such as those by Mpanya *et al*[13] and Croon *et al*,[14] have sought to explore similar avenues. However, their analyses, limited to data up to 2021, have not considered the cost-effectiveness of employing ML models. As the landscape of ML applications continues to evolve, it becomes increasingly crucial to extend the temporal scope but incorporate a nuanced examination of the cost-benefit dynamics. This study aims to bridge this gap, providing a more comprehensive understanding of current ML models to predict the HF hospitalisation risk not only from a clinical and technical perspective but also from an economic standpoint.

METHODS AND ANALYSIS

Study design

The scoping review follows the approach outlined by Arksey and O'Malley,[15] which involves five key stages: (1) defining the research question; (2) identifying relevant studies; (3) selecting eligible studies; (4) organising the data and (5) collecting, summarising and reporting the findings. The protocol was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) protocols, and we will employ the PRISMA extension for scoping reviews[16] specifically designed for scoping reviews when presenting the results.

The study is planned to begin by the end of March 2024 and finish by June 2024.

Stage 1: defining the research question

The main research question is, 'What machine learning models are used to predict the hospitalisation risk in individuals aged 18 years or older with heart failure, and what economic analyses are conducted to assess the cost-effectiveness and economic impact of implementing these predictive models in clinical practice?'

The following specific objectives will guide the review:

1. To identify the ML models used for predicting hospitalisation risk in individuals with HF.
2. To determine the data sources and variables used.
3. To identify the predictive performance of these ML models.
4. To summarise the key findings in the literature regarding the application of ML in HF hospitalisation risk prediction.
5. To identify the economic analysis conducted to determine the cost-effectiveness of these models.
6. To provide recommendations for developing and applying ML models to predict HF hospitalisations.

Stage 2: identifying relevant studies

The databases will be PubMed, Scopus and Web of Science. Articles will be included if they meet the inclusion criteria. A preliminary search to validate the appropriateness of the search terms was conducted, resulting in the identification of the following terms: ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR "Worsening" [tiab]) AND "Artificial Intelligence" [Mesh] AND ("Heart Failure" [Mesh] OR "Heart Failure, Diastolic" [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure, Chronic" [tiab] OR "Heart Failure, Acute" [tiab]) OR ("Cost-Benefit Analysis" [Mesh] AND ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR "Worsening" [tiab]) AND ("Artificial Intelligence" [Mesh] OR "Machine Learning" [Mesh] OR "Deep Learning" [Mesh]) AND ("Heart Failure" [Mesh] OR "Heart Failure, Diastolic" [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure, Chronic" [tiab] OR "Heart Failure, Acute" [tiab])).

The search results will be exported into a Word document for data management, including the removal of duplicated articles.

Stage 3: selecting eligible studies

After eliminating duplicates, data will be exported into Rayyan software. One reviewer will do title and abstract screening to select studies related to our Machine Learning, Population, Identification, Crosscheck and Outcomes[17] format.

The full-text screening phase is the second step. Two independent reviewers will choose the studies that fit the inclusion criteria. Studies that might fit the inclusion

criteria will be retrieved in full text during this screening phase. Studies in full text that do not fit the inclusion criteria will be eliminated, and the final report will include the rationale behind the omission. Two separate reviewers will extract the data, and inter-rater reliability will be evaluated and discussed. The complete report on the search results will be included in the final report, along with a PRISMA flow diagram. Disagreements between the reviewers will be settled by conversation or by consulting a third reviewer. If a study has several publications, the latest one will be kept. Eligible studies will be considered based on the criteria identified in table 1.

Stage 4: organizing the data

Critical information will be collected from the relevant studies—title, authors, publication year, country and context, study design, sample size, variables, ML algorithms, performance metrics and results, applications, and economic analysis performed and results.

Two researchers will extract the data, and disagreements will be discussed to reach a consensus among the team members.

Stage 5: collecting, summarising and reporting the findings

During this phase, we will gather, condense and present the data obtained from the scoping review. We will conduct a descriptive analysis, consolidate related data segments, extract deductive codes aligned with the results and assess inter-rater reliability. The information gathered from the studies included in the review will be organised into tables. Analysis of the data collected will provide recommendations to guide the development and implementation of ML models for predicting the risk of hospitalisation by patients with HF.

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Ethics and dissemination

This review does not require ethical approval. Our dissemination strategy includes peer-reviewed publications, conference presentations and dissemination to relevant stakeholders.

Acknowledgements The authors are grateful to Isabel Andrade and Joana Pires.

Contributors JS conceived the idea for the scoping review and led the design of the protocol and methodology. JS and JA wrote the first draft of the manuscript. TM provided inputs and critically revised the manuscript. All authors approved the final manuscript.

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

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Joana Seringa http://orcid.org/0000-0002-1346-570X
Teresa Magalhaes http://orcid.org/0000-0003-3794-1659

REFERENCES

1 Sen J, Mehtab S, Sen R, *et al*. Machine learning: Algorithms models, and applications. *arXiv* 2022.
2 Abraham Iorkaa A, Barma M, Muazu H. Machine learning techniques, methods and Algorithms: conceptual and practical insights. *Int J Eng Res Appl* 2021;11:55–64.
3 Rivara FP, Fihn SD, Perlis RH. Advancing health and health care using machine learning: JAMA network open call for papers. *JAMA Netw Open* 2018;1:e187176.
4 De Mauro A, Sestino A, Bacconi A. Machine learning and artificial intelligence use in marketing: a general Taxonomy. *Ital J Mark* 2022;2022:439–57.
5 Toh C, Brody JP. n.d. Applications of machine learning in Healthcare.
6 Shehab M, Abualigah L, Shambour Q, *et al*. Machine learning in medical applications: A review of state-of-the-art methods. *Comput Biol Med* 2022;145:105458.
7 Seetharam K, Balla S, Bianco C, *et al*. Applications of machine learning in cardiology. *Cardiol Ther* 2022;11:355–68.
8 Cuocolo R, Perillo T, De Rosa E, *et al*. Current applications of big data and machine learning in cardiology. *J Geriatr Cardiol* 2019;16:601–7.
9 Inamdar AA, Inamdar AC. Heart failure: diagnosis, management and utilization. *J Clin Med* 2016;5:62.
10 Savarese G, Becher PM, Lund LH, *et al*. Global burden of heart failure: a comprehensive and updated review of epidemiology. *Cardiovasc Res* 2023;118:3272–87.
11 Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure: A systematic review 2004-2016. *BMC Cardiovasc Disord* 2018;18:74.
12 Precedence Research. Machine Learning Market Size To Surpass USD 771.38 Bn By 2032, Available: https://www.precedenceresearch.com/machine-learning-market [Accessed 7 Oct 2023].
13 Mpanya D, Celik T, Klug E, *et al*. Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. *IJC Heart & Vasculature* 2021;34:100773.
14 Croon PM, Selder JL, Allaart CP, *et al*. Current state of artificial intelligence-based Algorithms for hospital admission prediction in patients with heart failure: a Scoping review *Eur Heart J Digit Health* 2022;3:415–25.
15 Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* 2005;8:19–32.
16 Tricco AC, Lillie E, Zarin W, *et al*. PRISMA extension for Scoping reviews (PRISMA-SCR): checklist and explanation. *Ann Intern Med* 2018;169:467–73.
17 Liu X, Anstey J, Li R, *et al*. Rethinking PICO in the machine learning era: ML-PICO. *Appl Clin Inform* 2021;12:407–16.