BMJ Open

Synthesizing evidence on machine learning methods, applications, and economic analysis to predict heart failure hospitalization risk: a scoping review protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-083188
Article Type:	Protocol
Date Submitted by the Author:	13-Dec-2023
Complete List of Authors:	Seringa, Joana; NOVA University Lisbon NOVA National School of Public Health; NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal Abreu, João; NOVA University Lisbon NOVA National School of Public Health Magalhaes, Teresa; NOVA University Lisbon NOVA National School of Public Health; NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal
Keywords:	Heart failure < CARDIOLOGY, Hospitalization, eHealth, Health Services

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Synthesize evidence on machine learning methods, applications, and economic analysis to predict heart failure hospitalization risk: a scoping review protocol

Joana Seringa[1,2]*, João Abreu[1], Teresa Magalhães[1,2]

[1]NOVA National School of Public Health, Lisbon, Portugal

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

*Corresponding author – jm.seringa@ensp.unl.pt

NOVA National School of Public Health 1600-560 Lisbon, Portugal

Word count: 1745

ABSTRACT

Introduction Machine Learning (ML) has emerged as a powerful tool for autonomously extracting knowledge from extensive datasets without explicit programming. This capacity to discern patterns and generate new information has revolutionized various domains, including the medical field. In cardiology, ML applications have shown promise in predictive risk assessment and phenotypic classification of Heart Failure (HF) patients, a chronic condition with a global prevalence exceeding 64 million individuals. HF imposes significant morbidity, mortality, and economic burden, necessitating innovative approaches for management. This scoping review aims to comprehensively evaluate the potential integration of ML techniques in the management and treatment of HF, to enhance patient outcomes and alleviate the associated economic strain.

Methods and Analysis This scoping review will use the approach described by Arksey and O'Malley [1]. The process consists of five phases: (1) defining the research questions, (2) identifying relevant studies, (3) selecting eligible studies, (4) organizing the data, and (5) compiling, summarizing, and reporting the results. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) was used in this protocol and the PRISMA extension for scoping reviews will be used to present the results. Databases searched are PubMed, SCOPUS, and Web of Science. The database was last searched in November 2023.

Ethics and dissemination This review does not require ethics approval. The dissemination strategy includes peer review publication, conference presentations, and to relevant stakeholders.

Strengths and limitations of this study

A scoping review is a type of literature review that aims to provide an overview of existing research on a particular topic to identify gaps in the literature. While scoping reviews are valuable in many research contexts, they have several limitations. In this specific review, we identified the following limitations (1) the heterogeneity of the studies, which can make it challenging to draw comparisons or conclusions; (2) the fact that we will not include grey literature (e.g., unpublished studies and conference abstracts), which can result in an incomplete picture of the available evidence; (3) the subjectivity in data extraction, which can introduce researcher bias; (4) as a static study, it may not capture the most recent developments in a rapidly evolving field, as it's the case.

Despite these limitations, scoping reviews are valuable tools for mapping the literature 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.[2]

The learning process begins with the acquisition of observations or empirical data, comprising instances, firsthand experience, or instructional input, aiming to discerning patterns within the dataset.[3]

Machine Learning applications have proven to be successful across multiple areas, such as financial, security, industrial, marketing, environmental, and medical areas.[4–7] 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, personalized treatment plans, and enhanced patient outcomes.[7] 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.[6,7]

In cardiology, ML has applications in predictive risk assessment, classification, and identification prognostic phenotypes among patients with Heart Failure (HF).[8,9] Heart Failure is a chronic medical condition where the heart is unable to pump enough blood to meet the body's needs.[10] It's not a sudden stoppage of the heart but rather a progressive condition that can develop over time.[10] 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 hospitalization.[11] This condition affects more than 64 million people worldwide and the global economic burden of HF is estimated at $108 billion per annum.[12]

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 comprehensively assess the potential integration of ML techniques in the management and treatment of heart failure, ultimately aiming to enhance patient outcomes and alleviate the economic strain associated with this condition.[11–13] Previous endeavors, such as those by Mpanya al.[14] and P. M. Croon et al.[15], have sought to explore similar avenues. However, their analyses, limited to data up to 2021, have not encompassed a thorough evaluation of the cost-effectiveness of employing ML models. As the landscape of ML applications continues to evolve, it becomes increasingly crucial to not only extend the temporal scope but also to incorporate a nuanced examination of the cost-benefit dynamics. This study endeavors to bridge this gap, providing a more comprehensive understanding of the role ML can play in mitigating the challenges posed by heart failure, not only from a clinical perspective but also from an economic standpoint.

METHOD AND ANALYSIS

Study design

The scoping review follows the approach outlined by Arksey and O'Malley [1], which involves five key stages: (1) defining the research question, (2) identifying relevant studies, (3) selecting eligible studies, (4) organizing the data, and (5) collecting,

summarizing, and reporting the findings. The protocol was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P), and we will employ the PRISMA-ScR extension [16] specifically designed for scoping reviews when presenting the results.

Stage 1: Defining the research question

The main research question is, "What machine learning models are used to predict hospitalization risk in individuals with heart failure, and what are the economic analyses 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 hospitalization risk in individuals with heart failure;

2. To determine the data sources and variables utilized;

3. To identify the predictive performance of these ML models;

4. To summarize the key findings in the literature regarding the application of machine learning in heart failure hospitalization risk prediction;

5. To identify the economic analysis conducted to determine the cost-effectiveness and economic impact of implementing these models;

6. To provide recommendations for the development and application of ML models to predict HF hospitalizations.

Stage 2: Identifying relevant studies.

The databases will include PubMed, SCOPUS, and Web of Science. The database was last searched in November 2023. Articles will be included if they meet the inclusion criteria. The search terms will be ("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 removing 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 (ML-PICO) [17] format.

The full-text screening phase is the second step. Two independent reviewers will choose the studies that fit the inclusion criteria. During this screening phase, studies that might fit the inclusion criteria will be retrieved in full text. 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. Data will be extracted by two separate reviewers, and inter-rated reliability will be evaluated and discussed in relation to the themes. 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 most current one will be kept. The language is limited to English.

Stage 4: Organizing the data

Critical information will be collected from the relevant studies: title, publication year, country and context, design, methods, sample size, variables, ML algorithms, performance metrics, applications, and the 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, summarizing, and reporting the findings

During this phase, we will gather, condense, and present the data obtained from the scoping review. We'll conduct a descriptive analysis, consolidate related data segments, extract deductive codes aligned with the results, and assess interrater reliability. The information gathered from the studies included in the review will be organized into tables. Analysis of the data collected will provide recommendations to guide the development and implementation of machine learning models for predicting the risk of hospitalization by patients with heart failure.

ETHICS AND DISSEMINATION

This review does not require ethics approval. Our dissemination strategy includes peer review publications, presentations at conferences, and to relevant stakeholders.

Author Contributions

JS conceived the idea for the scoping review and led the design of the protocol and methodology. JS and JB wrote the first draft of the manuscript. TM provided inputs and critically revised the manuscript. All authors provided valuable inputs to the research questions and subject matter.

Acknowledges

Funding

Competing interests statement

None declared.

Patient and public involvement

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

References

1   Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8:19–32.

2   Sen J, Mehtab S, Sen R, *et al. Machine Learning: Algorithms, Models, and Applications*. 2022. https://doi.org/10.5772/intechopen.94615

3   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.

4   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–e187176.

5   De Mauro A, Sestino A, Bacconi A. Machine learning and artificial intelligence use in marketing: a general taxonomy. *Italian Journal of Marketing*. 2022;2022:439–57.

6   Toh C, Brody JP. Applications of Machine Learning in Healthcare. doi: 10.5772/intechopen.92297

7   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. doi: 10.1016/J.COMPBIOMED.2022.105458

8   Seetharam K, Balla S, Bianco C, *et al.* Applications of Machine Learning in Cardiology. *Cardiol Ther*. 2022;11:355–68.

9   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.

10   Inamdar AA, Inamdar AC. Heart Failure: Diagnosis, Management and Utilization. *J Clin Med*. 2016;5. doi: 10.3390/JCM5070062

11   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.

12   Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure: A systematic review 2004-2016. *BMC Cardiovasc Disord*. 2018;18:1–11.

13   Precedence Research. Machine Learning Market Size To Surpass USD 771.38 Bn By 2032. https://www.precedenceresearch.com/machine-learning-market (accessed 7 October 2023)

14   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.

15   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. *European heart journal Digital health*. 2022;3:415–25.

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.

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

SCHOLARONE™
Manuscripts

1
2
3   1   **Machine learning methods, applications, and economic analysis to predict heart**
4
5   2   **failure hospitalization risk: a scoping review protocol**
6
7   3
8
9   4   Joana Seringa[1,2]*, João Abreu[1], Teresa Magalhães[1,2]
10
11   5
12
13   6   [1]NOVA National School of Public Health, Lisbon, Portugal
14
15   7
16
17   8   [2]NOVA National School of Public Health, Public Health Research Centre, Comprehensive
18
19   9   Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal
20
21   10
22
23   11
24
25   12   *Corresponding author – jm.seringa@ensp.unl.pt
26
27   13   NOVA National School of Public Health 1600-560 Lisbon, Portugal
28
29   14
30
31   15
32
33   16
34
35   17
36
37   18
38
39   19
40
41   20
42
43   21
44
45   22
46
47   23
48
49   24
50
51   25
52
53   26
54
55   27
56
57   28
58
59
60

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 synthesize evidence on machine learning methods, applications, and economic analysis to predict heart failure hospitalization risk.

Methods and Analysis This scoping review will use the approach described by Arksey and O'Malley. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) was used in this 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 full-text studies for inclusion and extract the data. All the studies focusing on ML models applied to predict the risk of hospitalization from HF adult patients will be included.

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

Keywords: Scoping Review Protocol; Machine Learning; Heart Failure; Hospitalization

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 (e.g., unpublished studies and conference abstracts), which can result in an incomplete picture of the available evidence;

- As a static study, the study 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 ML methods, applications, and economic analysis to predict heart failure hospitalization 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, firsthand experience, or instructional input, aiming to uncover patterns within the dataset.[2]

Machine Learning 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, personalized 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]

88   In cardiology, ML has applications in predictive risk assessment, classification, and

89   identification of prognostic phenotypes among patients with Heart Failure (HF).[7,8]

90   Heart Failure is a chronic medical condition where the heart is unable to pump enough

91   blood to meet the body's needs.[9] It's not a sudden stoppage of the heart but rather a

92   progressive condition that can develop over time.[9] Additionally, HF is a complex and

93   life-threatening syndrome marked by substantial morbidity and mortality, diminished

94   functional capacity and quality of life, as well as elevated costs stemming from

95   decompensation and subsequent hospitalization.[10] This condition affects more than

96   64 million people worldwide, with a global economic burden estimated at $108 billion

97   per annum.[11]

98   Considering the global burden of HF, its related economic implications, and the

99   successful exponential growth of ML, it is imperative to conduct this scoping review to

100 synthesize evidence on machine learning methods, applications, and economic analysis

101 to predict heart failure hospitalization risk aiming to reduce unplanned hospitalizations,

102 enhance patient outcomes and alleviate the associated financial strain associated with

103 this condition.[10–12] Previous endeavours, such as those by Mpanya al.[13] and P. M.

104 Croon et al.[14], have sought to explore similar avenues. However, their analyses,

105 limited to data up to 2021, have not considered the cost-effectiveness of employing ML

106 models. As the landscape of ML applications continues to evolve, it becomes

107 increasingly crucial to extend the temporal scope but incorporate a nuanced

108 examination of the cost-benefit dynamics. This study aims to bridge this gap, providing

109 a more comprehensive understanding of current ML models applied to predict heart

110 failure hospitalization risk, from a clinical and technical perspective but also from an

111 economic standpoint.

112

113 **METHOD AND ANALYSIS**

114 **Study design**

115 The scoping review follows the approach outlined by Arksey and O'Malley [15], which

116 involves five key stages: (1) defining the research question, (2) identifying relevant

117 studies, (3) selecting eligible studies, (4) organizing the data, and (5) collecting,

118 summarizing, and reporting the findings. The protocol was based on the Preferred

119 Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P), and

120 we will employ the PRISMA-ScR extension [16] specifically designed for scoping reviews

121 when presenting the results.

Stage 1: Defining the research question

123 The main research question is, "What machine learning models are used to predict

124 hospitalization risk in individuals aged 18 years or older with heart failure, and what

125 economic analyses are conducted to assess the cost-effectiveness and economic impact

126 of implementing these predictive models in clinical practice?"

127 The following specific objectives will guide the review:

    1. To identify the ML models used for predicting hospitalization risk in individuals
       with heart failure;

    2. To determine the data sources and variables utilized;

    3. To identify the predictive performance of these ML models;

    4. To summarize the key findings in the literature regarding the application of
       machine learning in heart failure hospitalization 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 hospitalizations.

Stage 2: Identifying relevant studies.

139 The databases will be PubMed, SCOPUS, and Web of Science. Articles will be included if

140 they meet the inclusion criteria. A preliminary search to validate the appropriateness of

141 the search terms was conducted, resulting in the identification of the following terms:

142 ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR

143 "Worsening" [tiab]) AND "Artificial Intelligence" [Mesh] AND ("Heart Failure" [Mesh] OR

144 "Heart Failure, Diastolic" [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure,

145 Chronic" [tiab] OR "Heart Failure, Acute" [tiab]) OR ("Cost-Benefit Analysis" [Mesh] AND

146 ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR

147 "Worsening" [tiab]) AND ("Artificial Intelligence" [Mesh] OR "Machine Learning" [Mesh]

148 OR "Deep Learning" [Mesh]) AND ("Heart Failure" [Mesh] OR "Heart Failure, Diastolic"

149 [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure, Chronic" [tiab] OR "Heart

150 Failure, Acute" [tiab])).

151 The search results will be exported into a Word document for data management,

152 including removing duplicated articles.

153 **Stage 3: Selecting eligible studies**

154 After eliminating duplicates, data will be exported into Rayyan software. One reviewer

155 will do title and abstract screening to select studies related to our Machine Learning,

156 Population, Identification, Crosscheck, and Outcomes (ML-PICO) [17] format.

157 The full-text screening phase is the second step. Two independent reviewers will choose

158 the studies that fit the inclusion criteria. Studies that might fit the inclusion criteria will

159 be retrieved in full text during this screening phase. Studies in full text that do not fit the

160 inclusion criteria will be eliminated, and the final report will include the rationale behind

161 the omission. Two separate reviewers will extract data, and inter-rater reliability will be

162 evaluated and discussed. The complete report on the search results will be included in

163 the final report, along with a PRISMA flow diagram. Disagreements between the

164 reviewers will be settled by conversation or consulting a third reviewer. If a study has

165 several publications, the most current one will be kept. Studies will be excluded if the

166 population is not specifically for adults with heart failure. Studies that do not include

167 economic analysis but focus on ML models to predict HF hospitalizations will be

168 included. The language is limited to English.

169 **Stage 4: Organizing the data**

170 Critical information will be collected from the relevant studies: title, authors, publication

171 year, country and context, study design, sample size, variables, ML algorithms,

172 performance metrics and results, applications, and economic analysis performed and

173 results.

174   Two researchers will extract the data, and disagreements will be discussed to reach a

175   consensus among the team members.

176   **Stage 5: Collecting, summarizing, and reporting the findings**

177   During this phase, we will gather, condense, and present the data obtained from the

178   scoping review. We'll conduct a descriptive analysis, consolidate related data segments,

179   extract deductive codes aligned with the results, and assess interrater reliability. The

180   information gathered from the studies included in the review will be organized into

181   tables. Analysis of the data collected will provide recommendations to guide the

182   development and implementation of machine learning models for predicting the risk of

183   hospitalization by patients with heart failure.

184   **Patient and public involvement**

185   Patients and/or the public were not involved in the design, or conduct, or reporting, or

186   dissemination plans of this research.

187   **ETHICS AND DISSEMINATION**

188   This review does not require ethics approval. Our dissemination strategy includes a

189   peer-reviewed publications, presentations at conferences, and dissemination to

190   relevant stakeholders.

191   **AUTHOR CONTRIBUTIONS**

192   JS conceived the idea for the scoping review and led the design of the protocol and

193   methodology. JS and JA wrote the first draft of the manuscript. TM provided inputs and

194   critically revised the manuscript. All authors approved the final manuscript.

195   **ACKNOWLEDGES**

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

197   **FUNDING**

198 The present publication was funded by Fundação Ciência e Tecnologia, IP national

199 support through CHRC (UIDP/04923/2020).

200 **COMPETING INTERESTS STATEMENT**

201 None declared.

202 **REFERENCES**

203 1   Sen J, Mehtab S, Sen R, *et al. Machine Learning: Algorithms, Models, and*
204   *Applications*. 2022. https://doi.org/10.5772/intechopen.94615
205 2   Abraham Iorkaa A, Barma M, Muazu H. Machine Learning Techniques, methods
206   and Algorithms: Conceptual and Practical Insights. *Int J Eng Res Appl*.
207   2021;11:55–64.
208 3   Rivara FP, Fihn SD, Perlis RH. Advancing Health and Health Care Using
209   Machine Learning: JAMA Network Open Call for Papers. *JAMA Netw Open*.
210   2018;1:e187176–e187176.
211 4   De Mauro A, Sestino A, Bacconi A. Machine learning and artificial intelligence
212   use in marketing: a general taxonomy. *Italian Journal of Marketing*.
213   2022;2022:439–57.
214 5   Toh C, Brody JP. Applications of Machine Learning in Healthcare. doi:
215   10.5772/intechopen.92297
216 6   Shehab M, Abualigah L, Shambour Q, *et al.* Machine learning in medical
217   applications: A review of state-of-the-art methods. *Comput Biol Med*. 2022;145.
218   doi: 10.1016/J.COMPBIOMED.2022.105458
219 7   Seetharam K, Balla S, Bianco C, *et al.* Applications of Machine Learning in
220   Cardiology. *Cardiol Ther*. 2022;11:355–68.
221 8   Cuocolo R, Perillo T, De Rosa E, *et al.* Current applications of big data and
222   machine learning in cardiology. *J Geriatr Cardiol*. 2019;16:601.
223 9   Inamdar AA, Inamdar AC. Heart Failure: Diagnosis, Management and
224   Utilization. *J Clin Med*. 2016;5. doi: 10.3390/JCM5070062
225 10   Savarese G, Becher PM, Lund LH, *et al.* Global burden of heart failure: a
226   comprehensive and updated review of epidemiology. *Cardiovasc Res*.
227   2023;118:3272–87.
228 11   Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure:
229   A systematic review 2004-2016. *BMC Cardiovasc Disord*. 2018;18:1–11.
230 12   Precedence Research. Machine Learning Market Size To Surpass USD 771.38
231   Bn By 2032. https://www.precedenceresearch.com/machine-learning-market
232   (accessed 7 October 2023)
233 13   Mpanya D, Celik T, Klug E, *et al.* Predicting mortality and hospitalization in
234   heart failure using machine learning: A systematic literature review. *IJC Heart &*
235   *Vasculature*. 2021;34:100773.
236 14   Croon PM, Selder JL, Allaart CP, *et al.* Current state of artificial intelligence-
237   based algorithms for hospital admission prediction in patients with heart failure: a
238   scoping review. *European heart journal Digital health*. 2022;3:415–25.
239 15   Arksey H, O'Malley L. Scoping studies: towards a methodological framework.
240   *Int J Soc Res Methodol*. 2005;8:19–32.

241   16   Tricco AC, Lillie E, Zarin W, *et al.* PRISMA extension for scoping reviews
242   (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169:467–73.
243   17   Liu X, Anstey J, Li R, *et al.* Rethinking PICO in the Machine Learning Era: ML-
244   PICO. *Appl Clin Inform*. 2021;12:407.
245

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-083188.R2
Article Type:	Protocol
Date Submitted by the Author:	21-Mar-2024
Complete List of Authors:	Seringa, Joana; NOVA University Lisbon NOVA National School of Public Health; NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal Abreu, João; NOVA University Lisbon NOVA National School of Public Health Magalhaes, Teresa; NOVA University Lisbon NOVA National School of Public Health; NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health informatics, Public health, Cardiovascular medicine
Keywords:	Heart failure < CARDIOLOGY, Hospitalization, eHealth, Health Services, Machine Learning

SCHOLARONE™
Manuscripts

1

2

3 **Machine learning methods, applications, and economic analysis to predict heart**

4 **failure hospitalization risk: a scoping review protocol**

5

6 Joana Seringa[1,2]*, João Abreu[1], Teresa Magalhães[1,2]

7

8 [1]NOVA National School of Public Health, Lisbon, Portugal

9

10 [2]NOVA National School of Public Health, Public Health Research Centre, Comprehensive

11 Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal

12

13 *Corresponding author – jm.seringa@ensp.unl.pt

14 NOVA National School of Public Health 1600-560 Lisbon, Portugal

15

16

17

18

19

20

21

22

23

24

25

26

27

28

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 synthesize evidence on machine learning methods, applications, and economic analysis to predict heart failure hospitalization risk.

Methods and Analysis This scoping review will use the approach described by Arksey and O'Malley. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) was used in this 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 full-text studies for inclusion and extract the data. All the studies focusing on ML models to predict the risk of hospitalization from HF adult patients will be included.

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

Keywords: Scoping Review Protocol; Machine Learning; Heart Failure; Hospitalization

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 (e.g., unpublished studies and conference abstracts) and systematic reviews, which can result in an incomplete picture of the available evidence;

- As a static study, the study 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 ML methods, applications, and economic analysis to predict heart failure hospitalization 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, firsthand experience, or instructional input, aiming to uncover patterns within the dataset.[2]

Machine Learning 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, personalized 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]

88   In cardiology, ML has applications in predictive risk assessment, classification, and

89   identification of prognostic phenotypes among patients with Heart Failure (HF).[7,8]

90   Heart Failure is a chronic medical condition where the heart is unable to pump enough

91   blood to meet the body's needs.[9] It's not a sudden stoppage of the heart but rather a

92   progressive condition that can develop over time.[9] Additionally, HF is a complex and

93   life-threatening syndrome marked by substantial morbidity and mortality, diminished

94   functional capacity and quality of life, as well as elevated costs stemming from

95   decompensation and subsequent hospitalization.[10] This condition affects more than

96   64 million people worldwide, with a global economic burden estimated at $108 billion

97   per annum.[11]

98   Considering the global burden of HF, its related economic implications, and the

99   successful exponential growth of ML, it is imperative to conduct this scoping review to

100 synthesize evidence on machine learning methods, applications, and economic analysis

101 to predict heart failure hospitalization risk aiming to reduce unplanned hospitalizations,

102 enhance patient outcomes and alleviate the associated financial strain associated with

103 this condition.[10–12] Previous endeavours, such as those by Mpanya al.[13] and P. M.

104 Croon et al.[14], have sought to explore similar avenues. However, their analyses,

105 limited to data up to 2021, have not considered the cost-effectiveness of employing ML

106 models. As the landscape of ML applications continues to evolve, it becomes

107 increasingly crucial to extend the temporal scope but incorporate a nuanced

108 examination of the cost-benefit dynamics. This study aims to bridge this gap, providing

109 a more comprehensive understanding of current ML models to predict heart failure

110 hospitalization risk, from a clinical and technical perspective but also from an economic

111 standpoint.

112 **METHOD AND ANALYSIS**

113 **Study design**

114 The scoping review follows the approach outlined by Arksey and O'Malley [15], which

115 involves five key stages: (1) defining the research question, (2) identifying relevant

116 studies, (3) selecting eligible studies, (4) organizing the data, and (5) collecting,

117 summarizing, and reporting the findings. The protocol was based on the Preferred

118 Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P), and

119 we will employ the PRISMA-ScR extension [16] specifically designed for scoping reviews

120 when presenting the results.

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

122 **Stage 1: Defining the research question**

123 The main research question is, "What machine learning models are used to predict

124 hospitalization risk in individuals aged 18 years or older with heart failure, and what

125 economic analyses are conducted to assess the cost-effectiveness and economic impact

126 of implementing these predictive models in clinical practice?"

127 The following specific objectives will guide the review:

128   1. To identify the ML models used for predicting hospitalization risk in individuals

129   with heart failure;

130   2. To determine the data sources and variables utilized;

131   3. To identify the predictive performance of these ML models;

132   4. To summarize the key findings in the literature regarding the application of

133   machine learning in heart failure hospitalization risk prediction;

134   5. To identify the economic analysis conducted to determine the cost-

135   effectiveness of these models;

136   6. To provide recommendations for developing and applying ML models to

137   predict HF hospitalizations.

138 **Stage 2: Identifying relevant studies**

139 The databases will be PubMed, SCOPUS, and Web of Science. Articles will be included if

140 they meet the inclusion criteria. A preliminary search to validate the appropriateness of

141 the search terms was conducted, resulting in the identification of the following terms:

142 ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR

143 "Worsening" [tiab]) AND "Artificial Intelligence" [Mesh] AND ("Heart Failure" [Mesh] OR

144 "Heart Failure, Diastolic" [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure,

145 Chronic" [tiab] OR "Heart Failure, Acute" [tiab]) OR ("Cost-Benefit Analysis" [Mesh] AND

146 ("Hospitalization" [Mesh] OR "Decompensat*" [tiab] OR "Readmission" [tiab] OR

147 "Worsening" [tiab]) AND ("Artificial Intelligence" [Mesh] OR "Machine Learning" [Mesh]

148 OR "Deep Learning" [Mesh]) AND ("Heart Failure" [Mesh] OR "Heart Failure, Diastolic"

149 [Mesh] OR "Heart Failure, Systolic" [Mesh] OR "Heart Failure, Chronic" [tiab] OR "Heart

150 Failure, Acute" [tiab])).

151 The search results will be exported into a Word document for data management,

152 including removing duplicated articles.

153 **Stage 3: Selecting eligible studies**

154 After eliminating duplicates, data will be exported into Rayyan software. One reviewer

155 will do title and abstract screening to select studies related to our Machine Learning,

156 Population, Identification, Crosscheck, and Outcomes (ML-PICO) [17] format.

157 The full-text screening phase is the second step. Two independent reviewers will choose

158 the studies that fit the inclusion criteria. Studies that might fit the inclusion criteria will

159 be retrieved in full text during this screening phase. Studies in full text that do not fit the

160 inclusion criteria will be eliminated, and the final report will include the rationale behind

161 the omission. Two separate reviewers will extract data, and inter-rater reliability will be

162 evaluated and discussed. The complete report on the search results will be included in

163 the final report, along with a PRISMA flow diagram. Disagreements between the

164 reviewers will be settled by conversation or consulting a third reviewer. If a study has

165 several publications, the most current one will be kept. Eligible studies will be

166 considered based on the criteria identified in Table 1.

Table 1. Inclusion and exclusion criteria	
Inclusion criteria	**Exclusion criteria**
Studies focusing on ML models to predict the risk of hospitalization 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 hospitalizations.	Studies whose cause of hospitalization is not HF.

	Grey literature (e.g., unpublished studies and conference abstracts) and systematic reviews.
	Studies in languages other than English.

167

Stage 4: Organizing the data

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

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

Stage 5: Collecting, summarizing, and reporting the findings

176 During this phase, we will gather, condense, and present the data obtained from the
177 scoping review. We'll conduct a descriptive analysis, consolidate related data segments,
178 extract deductive codes aligned with the results, and assess interrater reliability. The
179 information gathered from the studies included in the review will be organized into
180 tables. Analysis of the data collected will provide recommendations to guide the
181 development and implementation of machine learning models for predicting the risk of
182 hospitalization by patients with heart failure.

Patient and public involvement

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

ETHICS AND DISSEMINATION

187 This review does not require ethics approval. Our dissemination strategy includes a

188 peer-reviewed publications, presentations at conferences, and dissemination to

189 relevant stakeholders.

AUTHOR CONTRIBUTIONS

191 JS conceived the idea for the scoping review and led the design of the protocol and

192 methodology. JS and JA wrote the first draft of the manuscript. TM provided inputs and

193 critically revised the manuscript. All authors approved the final manuscript.

ACKNOWLEDGES

COMPETING INTERESTS STATEMENT

200 None declared.

REFERENCES

202 1 Sen J, Mehtab S, Sen R, *et al. Machine Learning: Algorithms, Models, and*
203 *Applications*. 2022. https://doi.org/10.5772/intechopen.94615
204 2 Abraham Iorkaa A, Barma M, Muazu H. Machine Learning Techniques, methods
205 and Algorithms: Conceptual and Practical Insights. *Int J Eng Res Appl.*
206 2021;11:55–64.
207 3 Rivara FP, Fihn SD, Perlis RH. Advancing Health and Health Care Using
208 Machine Learning: JAMA Network Open Call for Papers. *JAMA Netw Open.*
209 2018;1:e187176–e187176.
210 4 De Mauro A, Sestino A, Bacconi A. Machine learning and artificial intelligence
211 use in marketing: a general taxonomy. *Italian Journal of Marketing.*
212 2022;2022:439–57.
213 5 Toh C, Brody JP. Applications of Machine Learning in Healthcare. doi:
214 10.5772/intechopen.92297
215 6 Shehab M, Abualigah L, Shambour Q, *et al.* Machine learning in medical
216 applications: A review of state-of-the-art methods. *Comput Biol Med.* 2022;145.
217 doi: 10.1016/J.COMPBIOMED.2022.105458

218 7   Seetharam K, Balla S, Bianco C, *et al.* Applications of Machine Learning in
219   Cardiology. *Cardiol Ther*. 2022;11:355–68.
220 8   Cuocolo R, Perillo T, De Rosa E, *et al.* Current applications of big data and
221   machine learning in cardiology. *J Geriatr Cardiol*. 2019;16:601.
222 9   Inamdar AA, Inamdar AC. Heart Failure: Diagnosis, Management and
223   Utilization. *J Clin Med*. 2016;5. doi: 10.3390/JCM5070062
224 10 Savarese G, Becher PM, Lund LH, *et al.* Global burden of heart failure: a
225   comprehensive and updated review of epidemiology. *Cardiovasc Res*.
226   2023;118:3272–87.
227 11 Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure:
228   A systematic review 2004-2016. *BMC Cardiovasc Disord*. 2018;18:1–11.
229 12 Precedence Research. Machine Learning Market Size To Surpass USD 771.38
230   Bn By 2032. https://www.precedenceresearch.com/machine-learning-market
231   (accessed 7 October 2023)
232 13 Mpanya D, Celik T, Klug E, *et al.* Predicting mortality and hospitalization in
233   heart failure using machine learning: A systematic literature review. *IJC Heart &*
234   *Vasculature*. 2021;34:100773.
235 14 Croon PM, Selder JL, Allaart CP, *et al.* Current state of artificial intelligence-
236   based algorithms for hospital admission prediction in patients with heart failure: a
237   scoping review. *European heart journal Digital health*. 2022;3:415–25.
238 15 Arksey H, O'Malley L. Scoping studies: towards a methodological framework.
239   *Int J Soc Res Methodol*. 2005;8:19–32.
240 16 Tricco AC, Lillie E, Zarin W, *et al.* PRISMA extension for scoping reviews
241   (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169:467–73.
242 17 Liu X, Anstey J, Li R, *et al.* Rethinking PICO in the Machine Learning Era: ML-
243   PICO. *Appl Clin Inform*. 2021;12:407.
244