

Risk Prediction Models for Patients With Acute Heart Failure: A Protocol for Systematic Review, Critical Appraisal, and Meta-Analysis

Xuecheng Zhang

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Kehua Zhou

ThedaCare Regional Medical Center

Jingjing Zhang

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Ying Chen

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Hengheng Dai

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Siqi Wan

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Zhiyue Guan

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Mingzhi Hu

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Jing Kang

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Yan Liu (✉ sasliu@yeah.net)

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Hongcai Shang

Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital

Protocol

Keywords: acute heart failure, risk prediction model, meta-analysis, systematic review

Posted Date: October 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-948336/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Nearly a third of patients with acute heart failure (AHF) die or are readmitted within three months after discharge, accounting for the majority of costs associated with heart failure-related care. A considerable number of risk prediction models, which predict outcomes for mortality and readmission rates, have been developed and validated for patients with AHF. These models could help clinicians stratify patients by risk level and improve decision making, and provide specialist care and resources directed to high-risk patients. However, clinicians sometimes reluctant to utilize these models, possibly due to their poor reliability, the variety of models, and/or the complexity of statistical methodologies. Here, we describe a protocol to systematically review extant risk prediction models. We will describe characteristics, compare performance, and critically appraise the reporting transparency and methodological quality of risk prediction models for AHF patients.

Method

Embase, Pubmed, Web of Science, and the Cochrane Library will be searched from their inception onwards. A back word will be searched on derivation studies to find relevant external validation studies. Multivariable prognostic models used for AHF and mortality and/or readmission rate will be eligible for review. Two reviewers will conduct title and abstract screening, full-text review, and data extraction independently. Included models will be summarized qualitatively and quantitatively. We will also provide an overview of critical appraisal of the methodological quality and reporting transparency of included studies using the Prediction model Risk of Bias Assessment Tool (PROBAST tool) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD statement).

Discussion

The result of the systematic review could help clinicians better understand and use the prediction models for AHF patients, as well as make standardized decisions about more precise, risk-adjusted management.

Systematic review registration

: PROSPERO registration number CRD42021256416.

1 Background

Heart failure is a rapidly growing health issue associated with high mortality and readmission rates[1, 2]. Nearly a third of patients with acute heart failure (AHF) die or are readmitted three months after discharge, accounting for the majority of costs associated with heart failure-related care[3–5]. Although

some improvements have been made in AHF fields, these achievements were realized via adherence to existing chronic heart failure care and improvement in the quality of patient care[6], but not because of new therapeutic developments [7–9]. In this context, risk stratifications of AHF patients would be useful for more effective, risk-adjusted management[1, 10]. Accurate risk predictions could lead to timely decisions at emergency department (ED) patient triage (discharge with close outpatient follow up versus hospitalization) and thus appropriate level of care (outpatient care, ward, telemetry, or intensive care) and its corresponding treatments [11]. Risk prediction models help clinicians stratify patients by their severity of disease or general health condition [12]. Therefore, specialist care and resources could be directed to high risk patients, while unnecessary testing and procedures could be avoided for low risk patients, leading to overall healthcare cost savings[10].

Risk prediction models usually encompass a combination of predictors such as vital signs, biomarkers, and demographics; they provide estimates of future outcomes of AHF [13]. These statistical models, which yield a so-called risk score for mortality, hospital readmission and other outcomes in AHF. A considerable number of multivariable prognostic models have been developed for patients with AHF, and the outcomes most frequently predicted are mortality and readmission rate[12]. Several scores, such as MEESSEI-AHF[5] and ADHF/NTproBNP[14], have been confirmed to achieve reliable risk stratification in patients with AHF, while others showed less satisfactory performance[15, 16]. On the other side, clinicians and other associated health care professionals are usually unable to accurately judge the probability of death or severe complications or predict hospital readmission in AHF. An ideal combination of the best prognostic models and clinical judgment can supplement each other and optimize clinical care. However, improvements remains in huge demand by hospital administrators and policy makers as clinicians are reluctant to utilize these models, possibly due to their poor reliability, the variety of models, and/or the complexity of statistical methodologies.

Although variable systematic reviews are available in the literature on prognostic models for outcomes in heart failure [12, 17], few have been specifically dedicated for the AHF population or the AHF subgroups. Therefore, a new systematic review of models evaluated both in the ED and hospitalization settings will be a necessary supplement. is in urgent need. Here we describe our protocol to systematically review the prognostic models for patients with AHF. Our specific objectives are: (1) to identify prediction models for mortality and readmission in patients with AHF and qualitatively describe characteristics of these models, (2) to compare the performance of identified models quantitatively across different populations and settings with meta-analysis if appropriate, and (3) to critically appraise the reporting transparency and methodological quality of these studies on prediction models.

2 Methods

The protocol is reported with reference to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement[18, 19] (the checklist is shown in the Supplementary file in Appendix 1) and is conducted using the methodology recommended for systematic review and meta-analysis of prediction models[20]. We will conduct the framing of the review question (Table 1), study

design, data extraction, data analysis and critical appraisal under the relevant guidelines[21–23]. We have registered the study on PROSPERO(CRD42021256416).

2.1 Eligibility criteria

Studies will be screened on the basis of the following criteria: population (P), index (I), comparator (C), outcomes (O), timing (T), and setting (S). The PICOTS is an amendment of the PICOS system specific to systematic reviews of prediction models with extra consideration for timing and clinical setting[20, 24].

2.1.1 Patients

Studies on prediction models for patients with AHF (or acute decompensated heart failure) be considered for inclusion, but those focus on exclusive patients with specific morbidities or groups of children will be excluded.

2.1.2 Index/Potential prognostic models

Studies reporting the derivation and validation of a multivariable prediction model will be eligible for this review. The derived model should include at least 50 patients who experienced an event during the period of observation, because studies with fewer cases may not be sufficient for convincing administrative or clinical use[17].

2.1.3 Outcome, timing and setting

The primary outcomes predicted by the models under review are mortality and readmission rates of AHF patients. Readmission is defined as all-cause or heart failure-related readmission, which includes emergency department revisits or re-hospitalizations. The studies included in this review should report at least one of the primary outcomes. Studies for prediction models in inpatient and emergency department settings at any period will be included.

2.1.4 Types of studies and limits

Cohort, nested case-control, case-cohort, or registry studies using any type of data source (e.g., administrative databases and electronic medical records) will be included in this review. Only those original research articles reported in peer-reviewed publications will be eligible for this review. Secondary research, reviews, conference proceedings, dissertations, editorials, expert opinions, or consensus paper abstracts will be excluded.

2.1.5 Information sources and search strategy

Embase, Pubmed, Web of Science, and the Cochrane Library will be searched for results from their inception onwards. Detailed search strategies are presented in Supplementary file in Appendix 2, and the search terms cover expressions for acute heart failure, prediction model, and mortality/readmission[25]. We will focus on studies published in English. We will search a backward citation on all model derivation studies and identify relevant external validation studies via these studies.

2.2 Data management and study selection

Reviewers will undergo formal training before the formal selection of studies, considering different backgrounds and knowledge of the reviewers. Eligibility criteria will be explained and discussed in detail to ensure the reviewers have the same understanding. Search results will be combined using Endnote X9, and duplicates will be removed. Abstracts and titles of each study will be screened by two reviewers (selected among XZ, SW, ZG, MH), excluding articles based on eligibility criteria.

2.3 Data collection and analysis

2.3.1 Data extraction

Two independent reviewers (selected among XZ, SW, ZG, MH) will perform the data extraction, using a standardized data extraction form for all included studies. The form was developed based on the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (the CHARMS checklist)[22]. For each included study, the following information, but not limited to, will be extracted: study information (title, authors, countries and regions, published year); type of prediction model (regression-based model or risk score, derivation and/or validation); source of data (cohort, registry, case-control or randomized trial participants data); participants (consecutive/inconsecutive participants, location, number of centers, ED/hospitalization/combination/unclear, acute heart failure/acute decompensate heart failure, age limitation, mean age, percent of male patients, eligibility criteria, study dates); outcomes to be predicted (types of outcome, the time point and over what time period the outcome is predicted); candidate predictors (number and type of predictors, timing of predictor measurement); missing data; sample size; model derivation; model performance; model evaluation and results. Full details of the form are presented in the Supplementary file in Appendix 3.

Disagreements will be resolved via discussion or consultation with a third reviewer when a consensus cannot be reached. We will obtain missing data from the authors if possible. A study will be excluded if the key missing data could not be supplemented.

2.3.2 Critical appraisal

We will provide an overview critical appraisal of the methodological quality and reporting transparency of included studies. To critically appraise the methodological quality of included studies, which will include the risk of bias and applicability, we will use the PROBAST tool[23, 24]. This tool consists of 20 items structured in four domains: participants, predictors, outcome, and analysis. After 4 steps of an overall assessment, each domain will be ranked as “high”, “low” or “unclear” for both risk of bias and applicability. The final results of the PROBAST assessments will be presented in a summary table (Supplementary file in Appendix 4).

The TRIPOD statement[21, 26], which provides a checklist of 22 items considered indispensable for reporting transparency of a prediction model study, will be used for evaluating transparency of the reporting of the included studies. Each element of the TRIPOD statement could be answered with “yes” or

“no”, depending on whether the element was reported in the studies. Each “yes” answer will receive 1 point, and each “no” answer will receive 0 points. Generally, when several aspects are described within a TRIPOD item, all aspects has to be reported to obtain a point for that item. For elements that do not apply to a specific situation, they can be marked as “not applicable”[27]. We will report the overall TRIPOD adherence score of each study, which is developed to assess the uniformity in measuring adherence to the TRIPOD statement. The computational method of TRIPOD adherence score is calculated by dividing the sum of the adhered TRIPOD items by the total number of applicable TRIPOD items[28].

To ensure consistence in evaluating the parameters, reviewers will participate in a pre-study training before the critical appraisal. Any disagreement will be handled as described previously.

2.3.3 Qualitative data synthesis of the prediction models

Characteristics of the included studies will be described systematically by a narrative synthesis, and quantitative data from the included studies will be presented. Key findings, such as outcomes to be predicted, predictors, performance measures, and the predictive accuracy of the model, will be tabulated to facilitate comparisons. We will report uncertain measures in the way they were published or approximated using published methods[20].

2.3.4 Preparation for quantitative synthesis

Quantitative synthesis of predictive performance of the included models will be based on the results of performance measures and their precision values [20]. However, the types of performance measures vary among different studies and are often inconsistent or even not reported for further analyses. We will focus on discrimination and calibration, which are considered as the two most common performance measures of predictive models. Discrimination refers to a prediction model’s ability to distinguish between patients developing and not developing the target outcome and is often quantified by the concordance (C) statistic[29]. The C-statistic will be estimated from the reported measure when missing, if necessary[20]. Calibration refers to a model’s accuracy of the predicted risk probabilities and indicates the degree of which expected outcomes and observed outcomes agree with each other[29]. Previous studies offer different statistic measures for calibration, such as calibration plot, calibration slope, Hosmer-Lemeshow test, and others. If the data on the total number of expected (E) and observed (O) events are available for extraction among the included models, the O:E ratio will be used for further analysis and roughly provide an indication of the overall model calibration[30]. Extracted C-statistics and total O:E ratios will be rescaled before further meta-analysis to improve the validity of their underlying assumptions, according to statistic models provided in the literature [31, 32].

2.3.5 Quantitative syntheses

The nature of quantitative analyses will depend on the number of prediction models of the included studies. Derivation and validation of models will be considered separately.

Data will be synthesized by a meta-analysis if the number (at least 5) of studies included is large enough in a subset with an analogous clinical question appraised repeatedly [33]. A meta-analysis of validation

studies for a common prediction model will also be performed if feasible.

2.3.6 Meta-analyses and investigation of heterogeneity

Performance measures of discrimination and calibration will be analyzed by a random-effects model of meta-analysis to estimate the average performance for all included models, if appropriate[20]. The meta-analysis will be performed using Softwares STATA 16.0 (Stata Corp, College Station, TX, USA) with several packages. To better handle the uncertainty in the estimated heterogeneity among studies, we will adopt the restricted maximum likelihood (REML) estimation and use the Hartung-Knapp-Sidik-Jonkman (HKSJ) method when calculating 95% confidence intervals for the average performance[34, 35].

Heterogeneity in the pooled results is usually dependent on differences in the design and populations across the validation studies, such as changes in case mix variation or baseline risk[36, 37]. The case-mix variation of each study will be quantified by estimating the standard deviation of the linear predictor[20]. Cochran's Q and the I^2 statistic will be calculated for statistical homogeneity assessment of heterogeneity[38]. Potential sources of heterogeneity will be explored using meta-regression analyses if there are enough studies included in the meta-analyses (≥ 10 studies)[39, 40].

2.3.7 Subgroup analysis

If the number of included studies is large enough for specific subgroups, we will perform the following subgroup analyses: (1) type of prediction model; (2) source of data; (3) outcomes to be predicted (hospitalization and readmission); (4) method of predictive model building; (5) location; (6) timing (at which time point and over what period the outcome is predicted); (7) settings; and (8) others. Further subgroup analysis will be dependent on the final data extraction.

2.3.8 Sensitivity analysis

Sensitivity analyses will be conducted by excluding studies with a high risk of bias (at least 4/7 domain determined using the PROBAST tool) and studies with low reporting transparency (the overall TRIPOD adherence score $< 50\%$), to explore their influence on effect size.

2.4 Reporting and dissemination

We will report our review per guidance by the PRISMA statement[41]. The GRADE approach (grading of recommendations, assessment, derivation, and evaluation)[42] will be used to appraise confidence in estimates. Informed consent and ethical approval are not required because all data will be acquired from published studies. The results of the study will be published in peer-reviewed journals and presented at conferences if possible. Any important protocol modifications will be presented and made available on the PROSPERO registration.

3 Discussion

This systematic review protocol will identify existing prognostic models for mortality and readmission in patients with AHF. These prognostic models will be comprehensively summarized and critically

appraised, and a meta-analysis will be conducted on their performance compared across pre-defined subgroups if appropriate.

The methodological quality (which includes applicability and risk of bias) and reporting transparency of the available reviews seem to be suboptimal[30, 43]. Good methodological quality and transparent reporting are essential for clinicians to accurately judge the performance and applicability of prognostic models in real clinical settings of AHF for individualized predictions. Without these criteria, the use of prediction models will be limited due to lack of availability, and sometimes can even be misleading[44, 45]. Importantly, the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) [21] and the Prediction model Risk Of Bias Assessment Tool (PROBAST)[23] were published in 2015 and 2019 to improve reporting transparency and methodological quality of prediction. A 2020 systematic review regarding risk prediction models in patients with AHF have focused on the models utilized in the ED setting; however, models exclusively applied in hospitalized patients were excluded from the review[25].

These aforementioned limitations compromise the relevancy of the model and the comprehensiveness of the review. Our review will include the models utilized in both ED and inpatient settings and compare the performance of two groups with meta-analysis if feasible. Thus, our review will complement the existing reviews in this area. Additionally, we will review the risk of bias and relevance using PROBAST, and the adherence to the reporting guideline using TRIPOD, which have not been previously reported in reviews. The study may provide a structural reference of data extraction and critical appraisal for future reviews. Finally, our review will utilize a formal training to standardize article selection, data extraction, and assessment.

We acknowledge several limitations in our review protocol. First, the included studies may report the model performance using different kinds of measures, and the procedure of merging data may lead to additional bias. Secondly, there may be limited studies meeting our eligibility criteria, which is not sufficient for further meta-analyses. Thirdly, we will exclude models focusing purely on other outcomes apart from mortality and hospitalization, such as adverse events[46]. Future studies may extend the scope to all models for other kinds of outcomes and review all AHF prediction models.

Ultimately, our systematic review will identify prognostic models, which may help clinicians to better allocate and utilize medical resources and contribute to more precise, risk-adjusted management, allowing personalized prevention and therapeutic options while decreasing mortality and hospital readmission in AHF.

Table 1

Framing of this systematic review using key items identified by the CHARMS checklist[22]

Items	Comments
1. Prognostic versus diagnostic prediction mode	Aim is to review models predicting future events (prognostic prediction model)
2. Intended scope of the review	Models to inform physicians' therapeutic decision making
3. Type of prediction modelling studies	Prediction models for acute heart failure patients (derivation and validation), the performance of acute heart failure models (validation), and added value of specific predictor or updating of a specific model (derivation and validation)
4. Target population to whom the prediction model applies	Patients with diagnosed acute heart failure
5. Outcome to be predicted	Mortality and readmission rate
6. Time span of prediction	Within an unlimited time interval (e.g. 7 days, 30days, 180days, 1 year or any other specific time interval)
7. Intended moment of using the model	At any time during patients in ED or hospitalization

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that the research was conducted without any commercial or financial relationships that could be considered as potential conflicts of interest.

Funding

The study is supported by National Key R&D Program of China(2017YFC1700400, 2017YFC1700403) and National Natural Science Foundation of China(82174233). The funder played no role in the research.

Author Contributions

HS, YL and XZ conceived the study. KZ, HD, JZ and YC contributed to the design study. YL advised on statistical analysis. XZ drafted the manuscript. All authors provided suggestions during the drafting of manuscript.

Acknowledgments

The authors would like to thank all members in the group for their support and help.

References

1. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail.* 2016;18(8):891–975.
2. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA, Maddox TM, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circulation Heart failure.* 2013;6(3):606–19.
3. Bueno H, Ross JS, Wang Y, Chen J, Vidan MT, Normand SL, Curtis JP, Drye EE, Lichtman JH, Keenan PS, et al. Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure, 1993-2006. *Jama* 2010;303(21):2141-2147.
4. Krumholz HM, Normand SL, Wang Y. Trends in hospitalizations and outcomes for acute cardiovascular disease and stroke, 1999-2011. *Circulation.* 2014;130(12):966–75.
5. Miró Ò, Rossello X, Gil V, Martín-Sánchez FJ, Llorens P, Herrero-Puente P, Jacob J, Bueno H, Pocock SJ. Predicting 30-Day Mortality for Patients With Acute Heart Failure in the Emergency Department: A Cohort Study. *Ann Intern Med.* 2017;167(10):698–705.
6. Pang PS, Collins SP, Miró Ò, Bueno H, Diercks DB, Di Somma S, Gray A, Harjola VP, Hollander JE, Lambrinou E et al Editor's Choice-The role of the emergency department in the management of acute heart failure: An international perspective on education and research. *European heart journal Acute cardiovascular care* 2017;6(5):421-429.
7. Packer M, O'Connor C, McMurray JJV, Wittes J, Abraham WT, Anker SD, Dickstein K, Filippatos G, Holcomb R, Krum H, et al. Effect of Ularitide on Cardiovascular Mortality in Acute Heart Failure. *N Engl J Med.* 2017;376(20):1956–64.

8. Metra M, Teerlink JR, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Pang PS, Ponikowski P, Voors AA, et al. Effects of Serelaxin in Patients with Acute Heart Failure. *N Engl J Med*. 2019;381(8):716–26.
9. Gong B, Li Z, Yat Wong PC. Levosimendan Treatment for Heart Failure: A Systematic Review and Meta-Analysis. *J Cardiothorac Vasc Anesth*. 2015;29(6):1415–25.
10. Shiraishi Y, Kohsaka S, Abe T, Nagai T, Goda A, Nishihata Y, Nagatomo Y, Saji M, Toyosaki Y, Takei M, et al. Derivation and Validation of Clinical Prediction Models for Rapid Risk Stratification for Time-Sensitive Management for Acute Heart Failure. *Journal of clinical medicine* 2020;9(11).
11. Lee DS, Ezekowitz JA. Risk stratification in acute heart failure. *Can J Cardiol*. 2014;30(3):312–9.
12. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. *JACC Heart failure*. 2014;2(5):429–36.
13. Wussler D, Michou E, Belkin M, Kozhuharov N, Diebold M, Gualandro DM, Breidhardt T, Mueller C. Mortality prediction in acute heart failure: scores or biomarkers? *Swiss Med Wkly*. 2020;150:w20320.
14. Scrutinio D, Ammirati E, Guida P, Passantino A, Raimondo R, Guida V, Sarzi Braga S, Pedretti RF, Lagioia R, Frigerio M, et al. Clinical utility of N-terminal pro-B-type natriuretic peptide for risk stratification of patients with acute decompensated heart failure. Derivation and validation of the ADHF/NT-proBNP risk score. *Int J Cardiol*. 2013;168(3):2120–6.
15. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *Jama*. 2005;293(5):572–80.
16. Collins SP, Jenkins CA, Harrell FE Jr, Liu D, Miller KF, Lindsell CJ, Naftilan AJ, McPherson JA, Maron DJ, Sawyer DB, et al. Identification of Emergency Department Patients With Acute Heart Failure at Low Risk for 30-Day Adverse Events: The STRATIFY Decision Tool. *JACC Heart failure*. 2015;3(10):737–47.
17. Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, Woodward M, Patel A, McMurray J, MacMahon S. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart failure*. 2014;2(5):440–6.
18. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, JSr. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. 2015;4:1.
19. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA, Group P-P. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;350:g7647.
20. Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, Riley RD, Moons KG. A guide to systematic review and meta-analysis of prediction model performance. *BMJ*. 2017;356:i6460.
21. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.

22. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med.* 2014;11(10):e1001744.
23. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S, Groupdagger P. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med.* 2019;170(1):51–8.
24. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med.* 2019;170(1):W1–33.
25. Miró Ò, Rossello X, Platz E, Masip J, Gualandro DM, Peacock WF, Price S, Cullen L, DiSomma S, de Oliveira MT Jr, et al. Risk stratification scores for patients with acute heart failure in the Emergency Department: A systematic review. *European heart journal Acute cardiovascular care.* 2020;9(5):375–98.
26. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1–73.
27. Zamanipoor Najafabadi AH, Ramspek CL, Dekker FW, Heus P, Hooft L, Moons KGM, Peul WC, Collins GS, Steyerberg EW, van Diepen M. TRIPOD statement: a preliminary pre-post analysis of reporting and methods of prediction models. *BMJ open.* 2020;10(9):e041537.
28. Heus P, Damen J, Pajouheshnia R, Scholten R, Reitsma JB, Collins GS, Altman DG, Moons KGM, Hooft L. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. *BMJ open.* 2019;9(4):e025611.
29. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, McGinn T, Guyatt G. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *Jama.* 2017;318(14):1377–84.
30. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, Altman DG, Moons KG. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med.* 2012;9(5):1–12.
31. Gengsheng Q, Hotalovac L. Comparison of non-parametric confidence intervals for the area under the ROC curve of a continuous-scale diagnostic test. *Stat Methods Med Res.* 2008;17(2):207–21.
32. van Klaveren D, Steyerberg EW, Perel P, Vergouwe Y. Assessing discriminative ability of risk models in clustered data. *BMC medical research methodology.* 2014;14:5.
33. Zhong J, Si L, Zhang G, Huo J, Xing Y, Hu Y, Zhang H, Yao WJSr. Prognostic models for knee osteoarthritis: a protocol for systematic review, critical appraisal, and meta-analysis. 2021;10(1):149.
34. Cornell JE, Mulrow CD, Localio R, Stack CB, Meibohm AR, Guallar E, Goodman SN. Random-effects meta-analysis of inconsistent effects: a time for change. *Ann Intern Med.* 2014;160(4):267–70.
35. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird

- method. BMC medical research methodology 2014;14.
36. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol.* 2015;68(3):279–89.
 37. Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: Use of benchmark values to disentangle a case-mix effect from incorrect coefficients. *Am J Epidemiol.* 2010;172(8):971–80.
 38. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods.* 2006;11(2):193–206.
 39. Ban JW, Emparanza JI, Urreta I, Burls A. Design Characteristics Influence Performance of Clinical Prediction Rules in Validation: A Meta-Epidemiological Study. *PloS one.* 2016;11(1):e0145779.
 40. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ.* 2011;342:d549.
 41. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.
 42. Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, McGinn T, Hayden J, Williams K, Shea B, et al. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ.* 2015;350:h870.
 43. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, Voysey M, Wharton R, Yu LM, Moons KG, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC medical research methodology.* 2014;14:40.
 44. Andaur Navarro CL, Damen J, Takada T, Nijman SWJ, Dhiman P, Ma J, Collins GS, Bajpai R, Riley RD, Moons KG, et al. Protocol for a systematic review on the methodological and reporting quality of prediction model studies using machine learning techniques. *BMJ open.* 2020;10(11):e038832.
 45. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine.* 1996;15(4):361–87.
 46. Chang KF, Lin G, Huang PC, Juan YH, Wang CH, Tsai SY, Lin YC, Wu MT, Liao PA, Yang LY, et al. Left ventricular function and myocardial triglyceride content on 3T cardiac MR predict major cardiovascular adverse events and readmission in patients hospitalized with acute heart failure. *Journal of clinical medicine* 2020;9(1).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfile14.docx](#)