PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prediction models for hospital readmissions in patients with heart disease: a systematic review and meta-analysis
AUTHORS	Van Grootven, Bastiaan; Jepma, Patricia; Rijpkema, Corinne; Verweij, Lotte; Leeflang, Mariska; Daams, Joost; Deschodt, Mieke; Milisen, Koen; Flamaing, Johan; Buurman, Bianca

VERSION 1 – REVIEW

REVIEWER	Robinson, Robert Southern Illinois University School of Medicine
REVIEW RETURNED	18-Jan-2021

GENERAL COMMENTS	Abstract - Well written, restates content of article well. includes PROSPERO ID for easy identification. Introduction - Sufficiently detailed to justify the research question and explains gaps in current knowledge. Methods - Methods are conventional and not controversial. Sufficiently detailed to allow independent repetition of the study. Results - Results are complex, but nicely summarized with tables and text. The characteristics, performance, and limitations of the identified risk prediction tools are clearly presented - allowing the reader to compare and contrast these tools. Discussion - The discussion covers the limitations, biases, and shortfalls of the included trials and suggests directions for future research that may address these concerns. The limitations of the research methods were well covered by the authors. Conclusions were based on the reported data and are not overly broad or optimistic. Overall, this manuscript is well written and addresses an important area of patient care - predicting what patients with acute cardiac issues are at higher risk of hospital readmission. This study identifies significant shortcomings in the current state of knowledge in terms of utility, bias, and generalization.

REVIEWER	Wessler, Benjamin Tufts Medical Center
REVIEW RETURNED	08-Apr-2021

GENERAL COMMENTS	The authors describe a systematic review and meta-analysis of clinical prediction models for patients with acute cardiac

	syndromes. 43 newly developed clinical predictive models were included. The most notable finding is that the majority of models have a high risk of bias. Discrimination was modest for most models and predictors of discriminatory performance include study population, data source, and number of predictors. Calibration was not reported reliably. Overall this paper represents an important assessment of risk of bias of available prediction models for acute cardiac conditions. I have concerns about methods—in particular the lack of attention to how the competing risk of death is handled. The authors come to the same conclusion that has been reached in the past—mainly that discrimination is modest and there is little consistency in predictors that are used. As a result the impact of this study is modest. Major 1. The Results section of the abstract needs to be re-worked. The last sentence appears incomplete. What are the authors trying to share with the comment "eighteen predictors were pooled" 2. The methods for identifying validations is under-described. It is likely that more model validations would be found if citation searches of the included models were performed. 3. The objective of the paper is to describe performance of clinical prediction models for unplanned hospital admissions. Authors should be more specific about their focus on acute heart disease prediction models. 4. What did authors do with Cox models or risk scores which might not have included C stastitic? These models are often usable but may have been excluded from the review. This should be discussed. 5. How was competing risk of death handled? This is incredibly important for this type of analysis. As presented, these results are not terribly helpful for a practicing clinician or methodologist. 6. What percent of models reported some measure of calibration? 7. Conflation of c-statistic from derivation and c-statistic from validation presented at the time of model development. These are not interchangeable. 8. "The c-statistics of specific prediction models that were evaluated in multiple studies were pooled for the endpoint 30 days follow-up" I don't understand this. 9. When performance metrics are assessed across different conditions (e.g. AMI, TAVR, HF, surgical) please describe if there is a statistical difference in the apparent discriminatory performance. 10. How is adequate discrimination defined? 11. Was there a statistical difference in the effect size of variables stratified by clinical condition (ie. The effect of predictors was mostly smaller in the HF samples) Minor 1. Page 11, line 19 and 26: *till* should be *to*

	2. 60 vs sixty. Please be consistent

REVIEWER	Roberts, Derek University of Calgary and the Foothills Medical Centre, Division of Vascular and Endovascular Surgery, Department of Surgery
REVIEW RETURNED	10-Jun-2021

GENERAL COMMENTS	I thank the Editors, BMJ Open, and especially the authors for the privilege of reviewing the manuscript by van Grootven et al. The authors sought to evaluate prediction models for hospital readmissions in patients with heart disease by conducting a large systematic review and meta-analysis. I am impressed by the degree of methodologic rigour with which this study was conducted. It was a pleasure to read. As someone who values methodology highly, the authors have done a wonderful job addressing this complex issue. Further, the detail provided on each study, prediction model, and predictors is top notch. I have several suggestions that are largely unrelated to the strong methodology used to complete this study. They more relate to helping others to understand the results. I am both a clinician and a researcher, but many who want to use the results of this systematic review will be purely clinicians, policy-makers, or administrators. Specifically: 1. The introduction would benefit from a single paragraph describing how prediction models are commonly created. I would keep this brief (e.g., 4 sentences), but I think it would help to set the stage. 2. Why did the authors search for study protocols? 3. Please describe the "Critical Appraisal and Data Extraction for Systematic Reviews" checklist. I regularly do systematic reviews and do not recognize this tool. 4. What is the PROBAST tool and what "quality" elements does it contain? 5. Under summary measures, please describe what the c-statistic means. Please also define how the calibration slope, calibration in large, and the Hosmer-Lemeshow test are interpreted. 6. Please describe in the methods, what internal and external validation are, derivation and validation cohorts are, etc. This will help readers interpret the study's findings. 7. In the discussion, the authors talk a lot about how no model is very predictive. I agree with their interpretation of the data. However, science must move forward. The authors have compiled the best evidence on this topic. Please outline which models they think should be externally validated or updated. Provide guidance to researchers interested in doing these studies. Also, please include a summary table or paragraph that explains exactly which predictors are key to include in any future models. Help guide the literature.

	8. In accordance with the above and to circle back to my first comment, please insert a paragraph telling researchers, in addition to what predictors must be contained in future prediction models, exactly how you would suggest they study their models (derive and validate them). Be specific and guide future researchers. It was a pleasure to review this manuscript. I congratulate the authors on completing a great study and contribution to the literature.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Robert Robinson, Southern Illinois University School of Medicine Comments to the Author:
Abstract - Well written, restates content of article well. includes PROSPERO ID for easy identification.

Introduction - Sufficiently detailed to justify the research question and explains gaps in current knowledge.

Methods - Methods are conventional and not controversial. Sufficiently detailed to allow independent repetition of the study.

Results - Results are complex, but nicely summarized with tables and text. The characteristics, performance, and limitations of the identified risk prediction tools are clearly presented - allowing the reader to compare and contrast these tools.

Discussion - The discussion covers the limitations, biases, and shortfalls of the included trials and suggests directions for future research that may address these concerns. The limitations of the research methods were well covered by the authors.

Conclusions were based on the reported data and are not overly broad or optimistic.

Overall, this manuscript is well written and addresses an important area of patient care - predicting what patients with acute cardiac issues are at higher risk of hospital readmission. This study identifies significant shortcomings in the current state of knowledge in terms of utility, bias, and generalization.

Our answer: We thank the reviewer for reviewing our manuscript and the positive comments about our paper.

Reviewer: 2
Dr. Benjamin Wessler, Tufts Medical Center Comments to the Author:
The authors describe a systematic review and meta-analysis of clinical prediction models for patients with acute cardiac syndromes. 43 newly developed clinical predictive models were included. The most notable finding is that the majority of models have a high risk of bias. Discrimination was modest for most models and predictors of discriminatory performance include study population, data source, and number of predictors. Calibration was not reported reliably. Overall this paper represents an important assessment of risk of bias of available prediction models for acute cardiac conditions. I have concerns about methods—in particular the lack of attention to how the competing risk of death is handled. The authors come to the same conclusion that has been reached in the past—mainly that discrimination is modest and there is little consistency in predictors that are used.

Our answer: We thank the reviewer for reviewing our manuscript. We agree that competing risk of death is an important topic in prediction models for readmission. We used the Probast tool[1] to assess the risk of bias. In the analysis domain, the complexity of data analysis was assessed including how authors have accounted for competing risk. Many studies did not use appropriate statistics for the development and validation of the prediction models which resulted in 97.8% of the studies that had high risk of bias on this domain. For example, we observed that many studies did not mentioned if and how competing risk was analyzed. We have clarified this in the risk of bias paragraph of the results. Please see page 21: "For example a description on how complexities in data were handled (e.g. competing risk of death) was often missing and relevant performance measures were incomplete (e.g. calibration)."

Reviewer: 3

Dr. Derek Roberts, University of Calgary and the Foothills Medical Centre Comments to the Author:
I thank the Editors, BMJ Open, and especially the authors for the privilege of reviewing the manuscript by van Grootven et al. The authors sought to evaluate prediction models for hospital readmissions in patients with heart disease by conducting a large systematic review and meta-analysis.

I am impressed by the degree of methodologic rigour with which this study was conducted. It was a pleasure to read. As someone who values methodology highly, the authors have done a wonderful job addressing this complex issue. Further, the detail provided on each study, prediction model, and predictors is top notch.

I have several suggestions that are largely unrelated to the strong methodology used to complete this study. They more relate to helping others to understand the results. I am both a clinician and a researcher, but many who want to use the results of this systematic review will be purely clinicians, policy-makers, or administrators. Specifically:

1. The introduction would benefit from a single paragraph describing how prediction models are commonly created. I would keep this brief (e.g., 4 sentences), but I think it would help to set the stage.

Our answer: We thank the reviewer for reviewing our manuscript and this suggestion. We have extended the first paragraph of the introduction with more information about the creation of prediction models. Please see page 5: ". Prediction models guide healthcare providers in daily practice to assess patients' probability of readmission within a certain time frame and include candidate variables identified by clinical perspectives, literature or data-driven approaches, e.g. using machine learning techniques.[2] Data are often collected from observational cohorts of intervention studies and subsequently analyzed to examine what set of predictors best predict the risk of readmission."

2. Why did the authors search for study protocols?

Our answer: We have searched for study protocols to identify all potential eligible studies. We additionally searched for study results of the identified protocols and included these studies when they met the inclusion criteria.

3. Please describe the "Critical Appraisal and Data Extraction for Systematic Reviews" checklist. I regularly do systematic reviews and do not recognize this tool.

Our answer: Thank you for this suggestion. We have added an explanation in the text. Please see page 8: "The checklist includes items on 11 relevant domains, including source of data, participants, outcomes, candidate predictors, sample size, missing data, model development, model performance, model evaluation, results, and interpretation."

4. What is the PROBAST tool and what "quality" elements does it contain?
Our answer: Thank you for this suggestion. We have added this to the text. Please see x: "The Prediction model Risk Of Bias ASsessment Tool (PROBAST) tool[1] was used to assess the risk of bias (RoB) for four 'quality' domains, i.e. the participants, predictors, outcome and analysis for each model." We have opted to only describe the major domains, because the tool contains a lot of items. We believe that summarising all the individual items would make the text hard to read.

5. Under summary measures, please describe what the c-statistic means. Please also define how the calibration slope, calibration in large, and the Hosmer-Lemeshow test are interpreted.

Our answer: Thank you for this suggestion. We have added a text box in the manuscript with definitions. Please see page 10.

6. Please describe in the methods, what internal and external validation are, derivation and validation cohorts are, etc. This will help readers interpret the study's findings.

Our answer: Thank you for this suggestion. We have added a text box in the manuscript with definitions. Please see page 10.

7. In the discussion, the authors talk a lot about how no model is very predictive. I agree with their interpretation of the data. However, science must move forward. The authors have compiled the best evidence on this topic. Please outline which models they think should be externally validated or updated. Provide guidance to researchers interested in doing these studies. Also, please include a summary table or paragraph that explains exactly which predictors are key to include in any future models. Help guide the literature.

Our answer: Based on the suggestion of the reviewer, we have reflected on some promising prediction models in the discussion. Our recommendations for key predictors are described in comment 8.

Page 34/35: "Therefore, attention might be shifted from developing new risk prediction models to updating and externally validating existing prediction models in different populations with heart disease. For example, the Adventist Health Off-the-shelf model[3] showed high discrimination rates in both the development (0.86) and validation cohort (0.85). External validation is recommended to examine the generalizability of this model in other settings. In addition, the AMI READMITS score[4], full-stay AMI readmission model[4], pre-PCI model[5], motor and cognitive Functional Independence Measure (FIM)[6], READMIT[7], 30-day readmission model of Huynh et al.[8], and the model of Engoren et al.[9] were examined in one study and showed reasonable c-statistics in the development (0.68 – 0.82) and validation cohorts (0.64 – 0.78). For these studies, model updating recalibration and external validation is recommended to improve the predictive performance and generalizability of these prediction models."

8. In accordance with the above and to circle back to my first comment, please insert a paragraph telling researchers, in addition to what predictors must be contained in future prediction models, exactly how you would suggest they study their models (derive and validate them). Be specific and guide future researchers.

Our answer: We agree with the reviewer that more guidance for future researchers might help to move the science forward. Please see page 34 for our recommendations regarding key predictors in future prediction models for readmission:

"Based on our insights, we believe that models could be improved by incorporating some key predictors, i.e. age, gender, comorbidity scores (or at least heart failure, COPD, cardiovascular disease, diabetes mellitus), admission status, readmission history, and the geriatric profile (e.g. functional status, cognitive status). Because there are a still a large number of potential predictors, a large sample size is needed to estimate the coefficients with sufficient precision, and to prevent against overfitting the models. Some selection of predictors may still be warranted, and penalized techniques (e.g. lasso regression) should be preferred over traditional selection based on p-values."

References

1. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med 2019;170(1):51-58.

2. Shipe ME, Deppen SA, Farjah F, Grogan EL. Developing prediction models for clinical use using logistic regression: an overview. J Thorac Dis 2019;11(Suppl 4):S574-S584.

3. Reed J, Bokovoy J, Doram K. Unplanned readmissions after hospital discharge among heart failure patients at risk for 30-day readmission using an administrative dataset and “off the shelf” readmission models. Internet J Cardiovasc Res 2014;9(1):2020-07-15.

4. Nguyen OK, Makam AN, Clark C, Zhang S, Das SR, Halm EA. Predicting 30-Day Hospital Readmissions in Acute Myocardial Infarction: The AMI "READMITS" (Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus, Nonmale Sex, Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure) Score. J Am Heart Assoc 2018;7(8):e008882. doi: 10.1161/JAHA.118.008882.

5. Wasfy JH, Rosenfield K, Zelevinsky K, et al. A prediction model to identify patients at high risk for 30-day readmission after percutaneous coronary intervention. Circ Cardiovasc Qual Outcomes 2013;6(4):429-435.

6. Kitamura M, Izawa KP, Taniue H, et al. Relationship between Activities of Daily Living and Readmission within 90 Days in Hospitalized Elderly Patients with Heart Failure. Biomed Res Int 2017;2017:7420738.

7. Ferraris VA, Ferraris SP, Harmon RC, Evans BD. Risk factors for early hospital readmission after cardiac operations. J Thorac Cardiovasc Surg 2001;122(2):278-286.

8. Huynh Q, Negishi K, De Pasquale CG, et al. Validation of Predictive Score of 30-Day Hospital Readmission or Death in Patients With Heart Failure. Am J Cardiol 2018;121(3):322-329.

9. Engoren M, Habib RH, Dooner JJ, Schwann TA. Use of genetic programming, logistic regression, and artificial neural nets to predict readmission after coronary artery bypass surgery. J Clin Monit Comput 2013;27(4):455-464.

VERSION 2 – REVIEW

REVIEWER	Roberts, Derek University of Calgary and the Foothills Medical Centre, Division of Vascular and Endovascular Surgery, Department of Surgery
REVIEW RETURNED	18-Jul-2021

GENERAL COMMENTS	I thank the authors for their detailed response to my original suggestions. I think the added clarity will help move the field forward. I congratulate them on an excellent contribution to the literature and the completion of a large amount of work.