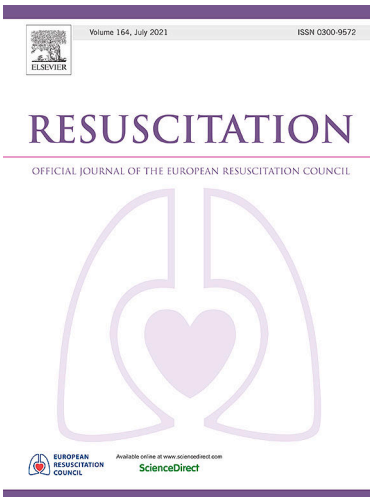
Journal Pre-proofs



Review

Pre-arrest Prediction of Survival Following In-hospital Cardiac Arrest: A Systematic Review of Diagnostic Test Accuracy Studies

Kasper G. Lauridsen, Therese Djärv, Jan Breckwoldt, Janice A. Tjissen, Keith Couper, Robert Greif, on behalf of the Education, Implementation, Team Task Force of the International Liaison Committee on Resuscitation ILCOR,

PII: S0300-9572(22)00632-3
DOI: https://doi.org/10.1016/j.resuscitation.2022.07.041
Reference: RESUS 9563

To appear in: *Resuscitation*

Received Date: 8 July 2022
Revised Date: 28 July 2022
Accepted Date: 29 July 2022

Please cite this article as: K.G. Lauridsen, T. Djärv, J. Breckwoldt, J.A. Tjissen, K. Couper, R. Greif, on behalf of the Education, Implementation, Team Task Force of the International Liaison Committee on Resuscitation ILCOR, Pre-arrest Prediction of Survival Following In-hospital Cardiac Arrest: A Systematic Review of Diagnostic Test Accuracy Studies, *Resuscitation* (2022), doi: https://doi.org/10.1016/j.resuscitation.2022.07.041

Pre-arrest Prediction of Survival Following In-hospital Cardiac Arrest:

A Systematic Review of Diagnostic Test Accuracy Studies

Kasper G. Lauridsen, MD, PhD[1,2,3]; Therese Djärv, MD, PhD[4,5]; Jan Breckwoldt MD, MME[6]; Janice A. Tjissen, MD, MSc[7] Keith Couper RN, PhD[8,9]; Robert Greif, MD, MME[10,11] on behalf of the Education, Implementation and Team Task Force of the International Liaison Committee on Resuscitation (ILCOR)

Kasper G. Lauridsen: Conceptualization, methodology, investigation, formal analysis, project administration, writing - original draft, writing - review & editing.

Therese Djärv: Conceptualization, methodology, investigation, formal analysis, writing - review & editing.

Jan Breckwoldt: Conceptualization, methodology, investigation, formal analysis, writing - review & editing.

Keith Couper: Conceptualization, methodology, investigation, formal analysis, writing - review & editing.

Janice Tjissen: Conceptualization, methodology, investigation, formal analysis, writing - review & editing.

Robert Greif: Conceptualization, methodology, investigation, formal analysis, resources, writing - review & editing.

[1] Research Center for Emergency Medicine, Aarhus University Hospital, Aarhus, Denmark

[2] Emergency Department, Randers Regional Hospital, Randers, Denmark

[3] Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia, Philadelphia, U.S.A

[4] Medical Unit of Emergency Medicine, Karolinska University Hospital, Stockholm, Sweden

[5] Center for Resuscitation Science, Karolinska Institutet, Stockholm, Sweden

[6] Institute of Anesthesiology, Zurich University Hospital, Zurich, Switzerland

[7] Department of Paediatrics, University of Western Ontario, London, Canada

[8] Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry, United Kingdom

[9] Critical care unit, University Hospitals Birmingham NHS Foundation Trust, Birmingham. United Kingdom

[10]Department of Anesthesiology and Pain Therapy, Bern University Hospital, Inselspital, University of Bern, Bern, Switzerland

[11]School of Medicine, Sigmund Freud University Vienna, Vienna, Austria

Word count manuscript: 2,913

Word count abstract: 248

Correspondence: Dr. Kasper Glerup Lauridsen, Emergency Department, Randers Regional Hospital, Skovlyvej 15, 8930 Randers NØ, Denmark

Mail: kglerup@clin.au.dk

Telephone: +45 2246 7815

Abstract

Aim: To evaluate the test accuracy of pre-arrest clinical decision tools for in-hospital cardiac arrest survival outcomes.

Methods: We searched Medline, Embase, and Cochrane Library from inception through January 2022 for randomized and non-randomized studies. We used the Quality Assessment of Diagnostic Accuracy Studies framework to evaluate risk of bias, and Grading of Recommendations Assessment, Development and Evaluation methodology to evaluate certainty of evidence. We report sensitivity, specificity, positive predictive outcome, and negative predictive outcome for prediction of survival outcomes. PROSPERO CRD42021268005.

Results: We searched 2517 studies and included 23 studies using 13 different scores: 12 studies investigating 8 different scores assessing survival outcomes and 11 studies using 5 different scores to predict neurological outcomes. All were historical cohorts/ case control designs including adults only. Test accuracy for each score varied greatly. Across the 12 studies investigating 8 different scores assessing survival to hospital discharge/ 30-day survival, the negative predictive values (NPVs) for the prediction of survival varied from 55.6% to 100%. The GO-FAR score was evaluated in 7 studies with NPVs for survival with cerebral performance category (CPC) 1 ranging from 95.0% to 99.2%. Two scores assessed survival with CPC ≤2 and these were not externally validated. Across all prediction scores, certainty of evidence was rated as very low.

Conclusions: We identified very low certainty evidence across 23 studies for 13 different pre-arrest prediction scores to outcome following IHCA. No score was sufficiently reliable to support its use in clinical practice. We identified no evidence for children.

Introduction

In-hospital cardiac arrest (IHCA) occurs with an incidence of 1-10 per 1,000 hospital admissions.[1,2]

Currently, only 20-30 % of adult IHCA patients survive to hospital discharge.[3–6] Some of these patients survive with unfavourable neurological outcome that may not be valued by the patient.[7–9] Several factors including older age and comorbidities are associated with potential futility of cardiopulmonary resuscitation (CPR).[6,10,11] Therefore, it is necessary for healthcare providers to discuss the appropriateness of attempting cardiopulmonary resuscitation with patients at risk of cardiac arrest.[12]

Do-not-attempt-CPR (DNACPR) decisions provide a process to document a clinical or patient decision that an individual should not receive resuscitation in the event of cardiac arrest. However, previous studies have identified variability in decision-making[13,14] and found DNACPR status to be inappropriately associated with demographic factors such as gender, ethnicity, and language.[15–17]

A key barrier to making DNACPR decisions is that the prediction of outcome following IHCA can be challenging.[13,14] Pre-arrest prediction rules may serve as an important decision aid to facilitate DNACPR discussions and reduce variability in decision making. Accordingly, several pre-arrest prediction rules have been developed over the years.[18,19] However, no systematic reviews have assessed the test accuracy of current pre-arrest prediction rules for IHCA. The International Liaison Committee on Resuscitation (ILCOR) task force on Education, Implementation, and Teams ranked the topic as a high priority and initiated this systematic review in collaboration with the Pediatric Life Support and Advanced Life Support task forces. The aim of this systematic review was to assess whether any pre-arrest prediction rule can predict survival outcomes following IHCA with sufficient precision to support its implementation in clinical practice.

Methods

This systematic review is reported in accordance with the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies.[20] The review was completed as part of the evidence evaluation process of ILCOR's Education, Implementation, and Teams task force and was registered at the International Prospective Registry for Systematic Reviews (PROSPERO CRD42021268005). No ethical approval was required to conduct this study.

In accordance with the review process of ILCOR, we used the PICOST format (Population, Intervention, Comparison, Outcome, Study Design, Timeframe) to frame this research question: For hospitalized adults and children experiencing an in-hospital cardiac arrest (P), does use of any pre-arrest clinical prediction rule (I), compared to no clinical prediction rule (C), predict return of spontaneous circulation, survival to hospital discharge/ 30-days or survival with favorable neurological outcome (O). We included randomized controlled trials and non-randomized studies (non-randomized controlled trials, interrupted time series, controlled before-and-after studies, cohort studies, case series where n ≥ 5) in all languages. We excluded editorials, commentaries, opinion papers, and conference abstracts (S). We searched Medline, Embase, and Cochrane

databases for all years. The search strategy was created and performed by an information specialist on January 8th, 2021, and an updated search was conducted on January 13th, 2022 (T). The search strategy is described in Appendix 1.

Definitions

We included studies on pre-arrest clinical prediction rules aiming to predict the chance of surviving (or not surviving) an IHCA, with or without favorable neurological outcome. We defined IHCA as any cardiac arrest with clinical indication for cardiopulmonary resuscitation (CPR) occurring inside the hospital regardless of the underlying cause of the arrest.[21] Studies on patients with out-of-hospital cardiac arrest being transported to the hospital with ongoing CPR were excluded. We defined pre-arrest clinical prediction rules as a set of clinical variables available before a cardiac arrest to predict the chance of surviving a cardiac arrest (+/- favourable neurological outcome). Studies utilizing termination of resuscitation rules and post-arrest prediction rules were excluded.

We chose to report predicted survival, as opposed to predicted death, as this is most commonly used outcome by ILCOR. Thus, we characterized true positives as a patient surviving that was predicted to survive and valued perfect negative predictive values (i.e. no missed survivors). We included the following outcomes: return of spontaneous circulation (ROSC), survival to hospital discharge/ 30-day survival, and survival with favorable neurological outcome. As studies may use different instruments (e.g. cerebral performance category or modified Ranking scale) with different cut-offs to define favourable neurological outcomes, we did not pre-specify any strict criteria for this outcome. We prospectively defined the following subgroup analyses of interest: paediatric patients vs. adult patients, studies before vs. after 2010, and historical cohorts vs. prospective clinical studies. We chose the cut-off of 2010 as studies have found stagnating survival rates after 2010 and lower survival rates before 2010.[6,22–24]

Data extraction and quality assessment

Following completion of database searches, we reviewed study titles and abstracts, and excluded obviously irrelevant papers. We subsequently reviewed full-text papers against study inclusion criteria. At both the title/ abstract and full-text screening stage, each paper was independently reviewed by two reviewers using Covidence software (Covidence®, Melbourne, Australia). Disagreements were solved by discussion with a third reviewer. In the event that key data were not reported, we contacted the corresponding author by email and sent a reminder two weeks after in case of no response. Data were extracted on a spreadsheet created by the authors to identify study- and patient characteristics and test accuracy outcomes.

Bias assessment

Bias assessment was conducted independently by two reviewers using the revised framework for Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2 tool).[25] This framework comprises 4 domains for study bias: patient selection, index test, reference standard, and flow and timing and 3 domains for applicability: patient selection, index test, reference standard. In case of disagreement regarding risk of bias for any domain, consensus was reached by discussion with a third reviewer. One reviewer (TD) was excluded from bias assessment to mitigate conflicts of interest as she had published studies that were part of the review.

Data analysis and synthesis

We report positive predictive values (PPV), specificity, sensitivity, negative predictive values (NPV), positive likelihood ratios, negative likelihood ratios, and area under receiver operating characteristic curves (AUC) with 95% confidence intervals for patient outcomes when possible. We calculated each diagnostic outcome with 95% confidence intervals (CI) using Stata version 16.0 (StataCorp LP, College Station, TX, USA). We report the AUCs presented in the studies. In case no AUC was presented in the study, we calculated an AUC based

on the sensitivity and specificity for the cut-off used. In accordance with the Cochrane handbook, we decided not to conduct any meta-analysis of studies investigating the Good Outcome Following Attempted Resuscitation (GO-FAR) score due to high risk of bias. Due to serious clinical heterogeneity and high risk of bias, no meta-analysis was conducted for the other scores.

We assessed the overall certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology.[26] We used the bias domains of the QUADAS tool to feed in to the bias part of GRADE, and the applicability part to feed in to directness. We used GRADEpro software (McMaster University, 2014) to synthesize the overall risk of bias across studies.

Results

We identified 2521 studies, of which 23 studies were eligible for inclusion (Fig. 1). Overall, 22 of the included studies were historical cohort studies and 1 study utilized a case-control design. Publication dates ranged from 1989 to 2021, with 17 studies published after 2010. Studies included only adult patients, no studies included paediatric cardiac arrest patients. Studies evaluating scores predicting survival to hospital discharge/ 30-day survival included a median of 205 patients (range: 33-656 patients). Studies evaluating scores predicting survival with favorable neurological outcome included a median of 863 patients (range: 287-62131 patients). We identified no studies of scores predicting return of spontaneous circulation. Study characteristics are presented in Table 1 and Table 2. Characteristics of each score are presented in Supplement 2 and likelihood ratios are presented in Supplement 3.

Risk of bias assessment and certainty of evidence

Bias assessment performed using the QUADAS-2 framework are presented in Table 3. We rated flow and timing as a concern for all studies as factors that contribute to calculation of the clinical prediction (e.g. age, co-morbid state) may have informed the decision to terminate resuscitation efforts, thus creating a self-

fulfilling prophecy.[27] Moreover, there was concern about patient selection (including applicability) in several studies due to missing data, patient exclusions, single-centre designs, and patient cohorts pre-dating 2010, due to lower survival rates compared with more recent cohorts. There was concern about the index test (including applicability) in several studies due to lack of pre-specified cut-offs and physiological parameters that may change frequently, making the score challenging to apply in the clinical setting. Finally, there was risk of bias for the reference standard in studies assessing neurological outcomes as these are at risk of subjectivity and inconsistency in reporting.

The overall certainty of evidence was rated as very low across all scores. Certainty of evidence was downgraded for risk of bias, indirectness, imprecision, and inconsistency for all scores (Appendix 4).

Scores predicting survival to hospital discharge/ 30-day survival

We identified 12 studies using 8 different scores to predict survival to hospital discharge/ 30-day survival finding NPVs of 55.6-100%.[28–39] The pre-arrest morbidity (PAM) score to predict survival to hospital discharge was used in 7 cohort studies. [28,29,32–35,38] The studies evaluated different cut-offs to avoid missing survivors (i.e. no false negatives) and found NPVs of 92.6-100% with 95% confidence intervals ranging from 47.8.-100% (Table 4).

The prognosis after resuscitation (PAR) score to predict survival to hospital discharge was used in 5 smaller cohort studies, 3 older studies published 1994-1999[28,29,33] and 2 more recent studies from 2014 and 2018.[35,38] The studies evaluated different cut-offs to avoid missing survivors and found NPVs of 95.4-100% with 95% confidence intervals ranging from 79.6-100% (Table 1).

Moreover, the following scores were investigated: The modified early warning score (MEWS),[31] the National Early Warning Score (NEWS),[37,39] the Clinical Frailty Scale,[36] the APACHE III score,[28] a neuronal network model,[30] and the modified pre-arrest morbidity (MPI) score (Table 1).[29,38] In addition, several studies were found that did not report data to calculate predictive values with confidence intervals. Limpawattana et al.[38] reported the following predictive values without confidence intervals (predicted death as opposed to

survival): A PPV of 92.2, a specificity of 87.8, a sensitivity of 39.2, and a NPV of 28.1 for a PAM score >6 and PPV of 89.6, a specificity of 80.5, a sensitivity of 45.7, and a NPV of 28.7 for a MPI score >5. In addition, one study reported a PPV of 76, a specificity of 86, a sensitivity of 47, and a NPV of 33 for a MEWS of 7 (prediction of death as opposed to survival).[28,31]

Scores predicting survival with favorable neurological outcome

We identified 11 studies using 5 different scores to predict survival to hospital discharge with favorable neurological outcome showing NPVs of 95.0-99.2%.[18,40–49] Overall, 6 studies investigated the Good Outcome Following Attempted Resuscitation (GO-FAR) score aiming to predict survival to hospital discharge with a Cerebral Performance Category (CPC) of 1 (Table 5).[18,40,43–45,49] Hong et al. showed a better AUC when adding albumin to the GO-FAR score to predict survival with CPC≤2 and reported a sensitivity of 94.1 (95% CI: 87.6-97.8), a specificity of 11.7 (95% CI: 8.5-15.6), a NPV of 87.0 (95% CI: 73.7-95.1), and a PPV of 24.1 (95% CI: 20.0-28.6) for prediction of survival to hospital discharge.[42] One study derived 5 classification and regression trees to predict survival to hospital discharge with a Cerebral Performance Category (CPC) of 1 and two of these models were externally validated in a second study (Table 2).[46,47] Finally, one study used the GO-FAR 2 score and one study used the Prediction of outcome for In-Hospital Cardiac Arrest (PIHCA) score to predict survival to hospital discharge with a CPC≤2.[41,48]

Sub-group analyses

We did not identify sufficient data to undertake our pre-planned sub-group analyses.

Discussion

This is the first systematic review on diagnostic accuracy test studies for pre-arrest prediction of survival for IHCA. We identified 23 studies using 13 different pre-arrest prediction rules. We identified no prospective implementation of any score and the level of evidence was rated as very low certainty. The most extensively validated score was the GO-FAR score that predicted chance of survival to hospital discharge with a CPC of 1

resulting in NPVs of 95-99%, albeit with significant statistical uncertainty around the estimate of NPV in some studies. We found no studies on clinical prediction rules for paediatric patients.

Pre-arrest prediction rules may be used to facilitate DNACPR discussions, empower patients to express their wishes based on objective information, and to make DNACPR decisions. It is widely considered that CPR is sometimes initiated even though a DNACPR decision should have been in place due to futility.[15,50,51] This exposes the patient and their family to the harms of a non-beneficial resuscitation attempt and diverts the hospital resuscitation team from their other clinical duties.

If a pre-arrest prediction tool is reliable, clinical implementation could potentially support DNACPR decision-making and contribute to fewer futile CPR attempts. At the same time, it could reduce variability in decision-making and contribute to equity in decision-making. However, reliance on a pre-arrest prediction tool whose test accuracy is inadequate might lead to more patients not receiving CPR, where it might have been beneficial and in line with patient's values and preferences.

A widely used definition for futility within medical research is a survival chance of <1%.[52–54] Accordingly, a lower boundary of the 95% confidence interval >0.99 for the NPV may be considered acceptable in some instances for a pre-arrest prediction rule.[55] However, use of a pre-arrest prediction resulting in 1% of potential survivors not being resuscitated may not be universally accepted. Notably, none of the included studies had 95% confidence intervals for the NPV or sensitivity >0.99 but the GO-FAR score performed well with a NPV >99% in 3 studies, ranging from 96.2-100% for the point estimate in all studies.[18,40,42–45,49]

An issue for all of the identified studies is that no prospective implementation was used, and several studies were based on patient cohorts from the 1980's, 90's, and 2000's where survival rates were lower compared to contemporary cohorts after 2010.[6,22–24] As clinicians may have inaccurate expectations about survival outcomes and may terminate resuscitative efforts prematurely based on patient characteristics,[56,57] the use of historical cohorts may induce a self-fulfilling prophesy and lead to a critical risk of bias in the studies.

Overall, comorbidity scores such as the PAM score and the PAR score performed differently in patient cohorts from the 1980's, 1990's, and 2000's.[28,29,32,33,35,38] These findings suggest that the scores may not be applicable to contemporary patient cohorts. Moreover, the early warning scores such as the NEWS and MEWS were overall highly inaccurate for prediction of patient survival and may be used to measure patient deterioration but should not be used to predict survival outcomes following cardiac arrest.[31,37,39]

With 7 studies investigating the GO-FAR score, this is the most extensively validated tool. However, several issues should be mentioned in relation to the clinical applicability. First, the score utilizes physiologic measures of hypotension and respiratory insufficiency captured within 4 hours before the cardiac arrest as part of the score. As these components of the score may fluctuate over time, the overall score may also change, making it challenging to use to inform clinical decision-making. Second, the GO-FAR score measures survival to hospital discharge with a CPC of 1. Although this outcome may be considered relevant by patients and clinicians, the generally used definition of favorable neurological outcome include survival with a CPC of 2 may be highly valued by patients and relatives.[21,58] The GO-FAR 2 score and the PIHCA score aim to predict survival with CPC ≤ 2 which resembles the generally used definition of favorable neurological outcome.[41,48] Both scores performed reasonably well in the derivation and internal validation, but the scores have not yet been externally validated and utilize physiologic measures as does the GO-FAR score.

Our review identified several important knowledge gaps. First, there are no prospective implementation studies on any pre-arrest prediction model. As such, it would be premature to consider use of any of these scores in clinical practice even though a perfect prediction may not be needed to initiate a DNACPR discussion. Second, no studies included paediatric patients or were conducted in low resource settings. Third, we found no evidence for return of spontaneous circulation and long-term survival outcomes. Fourth, there is a knowledge gap linking pre-arrest prediction scores to patient/ relative perspectives. We do not know which information the patients would prefer to know regarding the predicted outcomes and which cerebral performance categories that would be valued by different patient groups. Finally, scores that utilize

physiologic measures within short time intervals before a cardiac arrest are difficult to use in clinical practice as they might only indicate the pre-arrest deterioration. A reliable score that may combine elements such as acute and chronic comorbidities, admission blood samples, age, and frailty without using vital signs measured in the hours before the arrest is needed.

Limitations

This systematic review included only historical cohort studies and one case control study. In addition, there were issues with indirectness and clinical application of the clinical prediction rules resulting in very low certainty evidence. There was a large clinical heterogeneity among the included studies and the heterogeneity combined with high risk of bias prevented us from conducting meta-analyses.

Conclusions:

This systematic review identified very low certainty evidence for 13 different scores to predict survival to hospital discharge/ 30 days and favorable neurological outcome. None of these were able to reliably predict no chance of survival or favorable neurological outcome. We identified no evidence for children.

Conflicts of interest: KGL, JB, and RG are members of the ILCOR EIT Task Force (RG as chair). TD is vice chair of the ILCOR first aid task force, JT is member of the ILCOR pediatric life support task force, and KC is member of the ILCOR advanced life support task force. RG is ERC Director of Guidelines and ILCOR.

Acknowledgement:

The following non-task force members are acknowledged for their contributions: Information specialist Jenny Ring.

The following ILCOR EIT Taskforce Members are acknowledged as collaborators on this systematic review: Janet E. Bray, Jonathan P. Duff, Elaine Gilfoyle, Ming-Ju Hsieh, Andrew S. Lockey, Taylor Sawyer, Yiqun Lin, Adam Cheng, Furnan Bhanji, Kathryn Eastwood, Catherine Patocka, Chin Wei Yang, Yasuko Matsuyama, Sebastian Schnaubelt, Jeffrey L. Pellegrino, Kevin Nation, Joyce Yeung. Judith Finn and Peter Morley are acknowledged as members of the ILCOR Scientific Advisory Committee.

Funding: KGL was funded by the Central Denmark Region. Information specialist, Jenny Ring, received payment from the Australian Resuscitation Council to develop the search strategy as an information specialist. The sponsors had no role in designing or executing the study or in the interpretation, writing, or submission of the manuscript.

References

1. Andersen LW, Holmberg MJ, Berg KM, Donnino MW, Granfeldt A. In-Hospital Cardiac Arrest: A Review. *JAMA*. 2019;321(12):1200-1210. doi:10.1001/jama.2019.1696
2. Sandroni C, Nolan J, Cavallaro F, Antonelli M. In-hospital cardiac arrest: incidence, prognosis and possible measures to improve survival. *Intensive Care Med*. 2007;33(2):237-245. doi:10.1007/s00134-006-0326-z [doi]
3. Andersen LW, Holmberg MJ, Lofgren B, Kirkegaard H, Granfeldt A. Adult in-hospital cardiac arrest in Denmark. *Resuscitation*. 2019;140:31-36. doi:10.1016/j.resuscitation.2019.04.046
4. Nolan JP, Soar J, Smith GB, et al. Incidence and outcome of in-hospital cardiac arrest in the United Kingdom National Cardiac Arrest Audit. *Resuscitation*. 2014;85(8):987-992. doi:10.1016/j.resuscitation.2014.04.002
5. Hessulf F, Karlsson T, Lundgren P, et al. Factors of importance to 30-day survival after in-hospital cardiac arrest in Sweden - A population-based register study of more than 18,000 cases. *Int J Cardiol*. 2018;255:237-242. doi:10.1016/j.ijcard.2017.12.068
6. Wiberg S, Holmberg MJ, Donnino MW, et al. Age-dependent trends in survival after adult in-hospital cardiac arrest. *Resuscitation*. 2020;151:189-196. doi:10.1016/j.resuscitation.2020.03.008
7. Witten L, Gardner R, Holmberg MJ, et al. Reasons for death in patients successfully resuscitated from out-of-hospital and in-hospital cardiac arrest. *Resuscitation*. 2019;136:93-99. doi:10.1016/j.resuscitation.2019.01.031
8. Al-Dury N, Rawshani A, Israelsson J, et al. Characteristics and outcome among 14,933 adult cases of in-hospital cardiac arrest: A nationwide study with the emphasis on gender and age. *Am J Emerg*

Med. 2017;35(12):1839-1844. doi:10.1016/j.ajem.2017.06.012

9. Perman SM, Stanton E, Soar J, et al. Location of In-Hospital Cardiac Arrest in the United States-Variability in Event Rate and Outcomes. *J Am Heart Assoc*. 2016;5(10).

10. Fernando SM, Tran A, Cheng W, et al. Pre-arrest and intra-arrest prognostic factors associated with survival after in-hospital cardiac arrest: Systematic review and meta-analysis. *BMJ*. 2019;367:l6373. doi:10.1136/bmj.l6373

11. Piscator E, Hedberg P, Göransson K, Djärv T. Survival after in-hospital cardiac arrest is highly associated with the Age-combined Charlson Co-morbidity Index in a cohort study from a two-site Swedish University hospital. *Resuscitation*. 2016;99:79-83. doi:10.1016/j.resuscitation.2015.11.023

12. Mentzelopoulos SD, Couper K, Voorde P Van de, et al. European Resuscitation Council Guidelines 2021: Ethics of resuscitation and end of life decisions. *Resuscitation*. 2021;161:408-432. doi:10.1016/j.resuscitation.2021.02.017

13. Ebell MH, Bergus GR, Warbasse L, Bloomer R. The inability of physicians to predict the outcome of in-hospital resuscitation. *J Gen Intern Med*. 1996;11(1):16-22. doi:10.1007/BF02603480

14. Miller DL, Gorbien MJ, Simbartl LA, Jahnigen DW. Factors influencing physicians in recommending in-hospital cardiopulmonary resuscitation. *Arch Intern Med*. 1993;153(17):1999-2003.

15. Becker C, Manzelli A, Marti A, et al. Association of medical futility with do-not-resuscitate (DNR) code status in hospitalised patients. *J Med Ethics*. January 2021. doi:10.1136/medethics-2020-106977

16. Perman SM, Beaty BL, Daugherty SL, et al. Do Sex Differences Exist in the Establishment of "Do Not Attempt Resuscitation" Orders and Survival in Patients Successfully Resuscitated From In-Hospital Cardiac Arrest? *J Am Heart Assoc*. 2020;9(4):e014200. doi:10.1161/JAHA.119.014200

17. Piscator E, Göransson K, Forsberg S, Herlitz J, Djärv T. Do-Not-Attempt-Cardiopulmonary-Resuscitation (DNACPR) decisions in patients admitted through the emergency department in a Swedish University Hospital - An observational study of outcome, patient characteristics and changes in DNACPR decisions. *Resusc plus*. 2022;9:100209. doi:10.1016/j.resplu.2022.100209

18. Ebell MH, Jang W, Shen Y, Geocadin RG. Development and validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score to predict neurologically intact survival after in-hospital cardiopulmonary resuscitation. *JAMA Intern Med*. 2013;173(20):1872-1878. doi:10.1001/jamainternmed.2013.10037

19. Beck K, Vincent A, Cam H, et al. Medical futility regarding cardiopulmonary resuscitation in in-hospital cardiac arrests of adult patients: A systematic review and Meta-analysis. *Resuscitation*. 2022;172:181-193. doi:10.1016/j.resuscitation.2021.11.041

20. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA*. 2018;319(4):388-396. doi:10.1001/jama.2017.19163

21. Nolan JP, Berg RA, Andersen LW, et al. Cardiac Arrest and Cardiopulmonary Resuscitation Outcome Reports: Update of the Utstein Resuscitation Registry Template for In-Hospital Cardiac Arrest: A

Consensus Report From a Task Force of the International Liaison Committee on Resuscitation (American . *Resuscitation*. August 2019. doi:10.1016/j.resuscitation.2019.08.021

22. Girotra S, Chan PS. Trends in survival after in-hospital cardiac arrest. *N Engl J Med*. 2013;368(7):680-

23. Holmberg MJ, Wiberg S, Ross CE, et al. Trends in Survival After Pediatric In-Hospital Cardiac Arrest in the United States. *Circulation*. 2019;140(17):1398-1408. doi:10.1161/CIRCULATIONAHA.119.041667

24. Hjalmarsson P, Memar M, Geara S-J, et al. Trends in co-morbidities and survival for in-hospital cardiac arrest -A Swedish cohort study. *Resuscitation*. 2018;124:29-34. doi:10.1016/j.resuscitation.2017.12.025

25. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009

26. Iorio A, Spencer FA, Falavigna M, 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. doi:10.1136/bmj.h870

27. Lauridsen KG, Baldi E, Smyth M, Perkins GD, Greif R. Clinical decision rules for termination of resuscitation during in-hospital cardiac arrest: A systematic review of diagnostic test accuracy studies. *Resuscitation*. 2021;158:23-29. doi:10.1016/j.resuscitation.2020.10.036

28. Ebell MH, Kruse JA, Smith M, Novak J, Drader-Wilcox J. Failure of three decision rules to predict the outcome of in-hospital cardiopulmonary resuscitation. *Med Decis Making*. 1997;17(2):171-177. doi:10.1177/0272989X9701700207

29. Bowker L, Stewart K. Predicting unsuccessful cardiopulmonary resuscitation (CPR): a comparison of three morbidity scores. *Resuscitation*. 1999;40(2):89-95. doi:10.1016/s0300-9572(99)00008-8

30. Ebell MH. Artificial neural networks for predicting failure to survive following in-hospital cardiopulmonary resuscitation. *J Fam Pract*. 1993;36(3):297-303.

31. Stark AP, Maciel RC, Sheppard W, Sacks G, Hines OJ. An early warning score predicts risk of death after inhospital cardiopulmonary arrest in surgical patients. *Am Surg*. 2015;81(10):916-921. doi:10.1177/000313481508101001

32. Cohn EB, Lefevre F, Yarnold PR, Arron MJ, Martin GJ. Predicting survival from in-hospital CPR: Meta-analysis and validation of a prediction model. *J Gen Intern Med*. 1993;8(7):347-353.

33. O'Keeffe S, Ebell MH. Prediction of failure to survive following in-hospital cardiopulmonary resuscitation: comparison of two predictive instruments. *Resuscitation*. 1994;28(1):21-25. doi:10.1016/0300-9572(94)90050-7

34. George ALJ, Folk BP 3rd, Crecelius PL, Campbell WB. Pre-arrest morbidity and other correlates of survival after in-hospital cardiopulmonary arrest. *Am J Med*. 1989;87(1):28-34. doi:10.1016/s0002-9343(89)80479-6

35. Ohlsson MA, Kennedy LM, Juhlin T, Melander O. Evaluation of pre-arrest morbidity score and prognosis after resuscitation score and other clinical variables associated with in-hospital cardiac

arrest in southern Sweden. *Resuscitation*. 2014;85(10):1370-1374. doi:10.1016/j.resuscitation.2014.07.009

36. Ibitoye SE, Rawlinson S, Cavanagh A, Phillips V, Shipway DJHH. Frailty status predicts futility of [illegible] doi:10.1093/ageing/afaa104

37. Haegdorens F, Monsieurs KG, De Meester K, Van Bogaert P. The optimal threshold for prompt clinical review: An external validation study of the national early warning score. *J Clin Nurs*. 2020;29(23-24):4594-4603. doi:10.1111/jocn.15493

38. Limpawattana P, Suraditnan C, Aungsakul W, et al. A comparison of the ability of morbidity scores to predict unsuccessful cardiopulmonary resuscitation in thailand. *J Med Assoc Thail*. 2018;101(9):1231-1236.

39. Roberts D, Djärv T. Preceding national early warnings scores among in-hospital cardiac arrests and their impact on survival. *Am J Emerg Med*. 2017;35(11):1601-1606. doi:10.1016/j.ajem.2017.04.072

40. Rubins JB, Kinzie SD, Rubins DM. Predicting Outcomes of In-Hospital Cardiac Arrest: Retrospective US Validation of the Good Outcome Following Attempted Resuscitation Score. *J Gen Intern Med*. 2019;34(11):2530-2535. doi:10.1007/s11606-019-05314-x

41. George N, Thai TN, Chan PS, Ebell MH. Predicting the probability of survival with mild or moderate neurological dysfunction after in-hospital cardiopulmonary arrest: The GO-FAR 2 score. *Resuscitation*. 2020;146:162-169. doi:10.1016/j.resuscitation.2019.12.001

42. Hong S-I, Kim Y-J, Cho YJ, Huh JW, Hong S-B, Kim WY. Predictive value of pre-arrest albumin level with GO-FAR score in patients with in-hospital cardiac arrest. *Sci Rep*. 2021;11(1):10631. doi:10.1038/s41598-021-90203-9

43. Piscator E, Goransson K, Bruchfeld S, et al. Predicting neurologically intact survival after in-hospital cardiac arrest-external validation of the Good Outcome Following Attempted Resuscitation score. *Resuscitation*. 2018;128:63-69. doi:10.1016/j.resuscitation.2018.04.035

44. Cho YJ, Kim YJ, Kim MY, et al. Validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score in an East Asian population. *Resuscitation*. 2020;150:36-40. doi:10.1016/j.resuscitation.2020.02.035

45. Ohlsson MA, Kennedy LM, Ebell MH, Juhlin T, Melander O. Validation of the good outcome following attempted resuscitation score on in-hospital cardiac arrest in southern Sweden. *Int J Cardiol*. 2016;221:294-297. doi:10.1016/j.ijcard.2016.06.146

46. Guilbault RWR, Ohlsson MA, Afonso AM, Ebell MH. External Validation of Two Classification and Regression Tree Models to Predict the Outcome of Inpatient Cardiopulmonary Resuscitation. *J Intensive Care Med*. 2017;32(5):333-338. doi:10.1177/0885066616686924

47. Ebell MH, Afonso AM, Geocadin RG. Prediction of survival to discharge following cardiopulmonary resuscitation using classification and regression trees. *Crit Care Med*. 2013;41(12):2688-2697. doi:10.1097/CCM.0b013e31829a708c

48. Piscator E, Goransson K, Forsberg S, et al. Prearrest prediction of favourable neurological survival following in-hospital cardiac arrest: The Prediction of outcome for In-Hospital Cardiac Arrest (PIHCA)

score. *Resuscitation*. 2019;143:92-99. doi:10.1016/j.resuscitation.2019.08.010

49. Thai TN, Ebell MH. Prospective validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score for in-hospital cardiac arrest prognosis. *Resuscitation*. 2019;140:2-8.

50. Druwe P, Monsieurs KG, Piers R, et al. Perception of inappropriate cardiopulmonary resuscitation by clinicians working in emergency departments and ambulance services: The REAPPROPRIATE international, multi-centre, cross sectional survey. *Resuscitation*. 2018;132:112-119. doi:10.1016/j.resuscitation.2018.09.006

51. Druwé P, Benoit DD, Monsieurs KG, et al. Cardiopulmonary Resuscitation in Adults Over 80: Outcome and the Perception of Appropriateness by Clinicians. *J Am Geriatr Soc*. 2020;68(1):39-45. doi:10.1111/jgs.16270

52. Schneiderman LJ, Jecker NS, Jonsen AR. Medical futility: its meaning and ethical implications. *Ann Intern Med*. 1990;112(12):949-954. doi:10.7326/0003-4819-112-12-949

53. Morrison LJ, Kierzek G, Diekema DS, et al. Part 3: ethics: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation*. 2010;122(18 Suppl 3):S665-75. doi:10.1161/CIRCULATIONAHA.110.970905

54. Mancini ME, Diekema DS, Hoadley TA, et al. Part 3: Ethical Issues: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation*. 2015;132(18 Suppl 2):S383-96. doi:10.1161/CIR.0000000000000254

55. Nas J, Kleinnibbelink G, Hannink G, et al. Diagnostic performance of the basic and advanced life support termination of resuscitation rules: A systematic review and diagnostic meta-analysis. *Resuscitation*. 2020;148:3-13. doi:10.1016/j.resuscitation.2019.12.016

56. Hansen C, Lauridsen KG, Schmidt AS, Løfgren B. Decision-making in cardiac arrest: physicians’ and nurses’ knowledge and views on terminating resuscitation. *Open Access Emerg Med*. 2018;Volume 11:1-8. doi:10.2147/OAEM.S183248

57. Bradley SM, Liu W, Chan PS, et al. Duration of resuscitation efforts for in-hospital cardiac arrest by predicted outcomes: Insights from Get With The Guidelines - Resuscitation. *Resuscitation*. 2017;113:128-134. doi:10.1016/j.resuscitation.2016.12.017

58. Perkins GD, Jacobs IG, Nadkarni VM, et al. Cardiac Arrest and Cardiopulmonary Resuscitation Outcome Reports: Update of the Utstein Resuscitation Registry Templates for Out-of-Hospital Cardiac Arrest. *Resuscitation*. 2015;96:328-340. doi:10.1016/j.resuscitation.2014.11.002

Figure 1: Flow Diagram

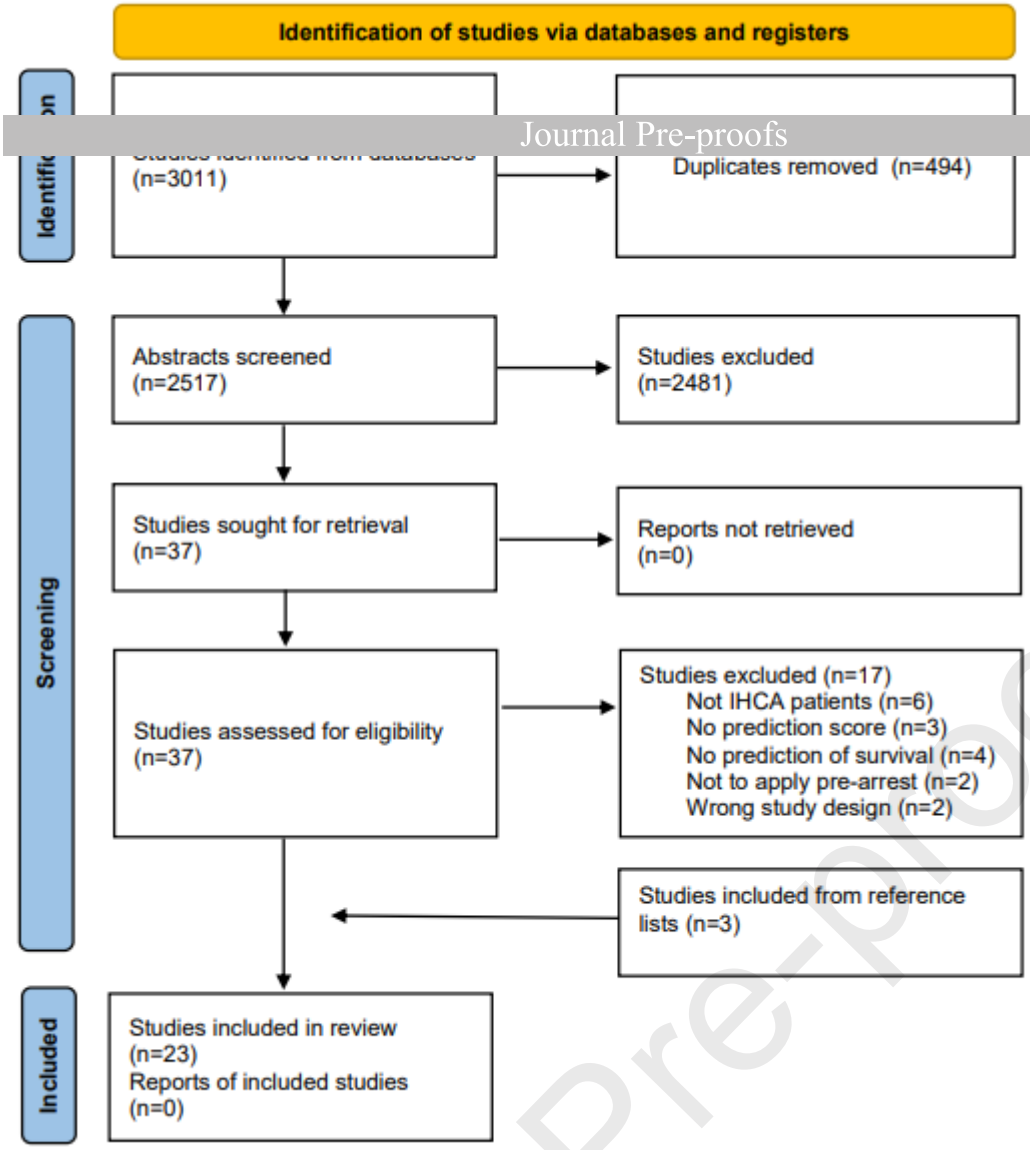


Table legends

Table 1: Characteristics of studies predicting survival to hospital discharge/ 30-day survival. OR: operating room. ED: emergency department. ICU: intensive care unit. IHCA: in-hospital cardiac arrest. MET: medical emergency team. MEWS: modified early warning score. NEWS: national early warning score. PAM: pre-arrest morbidity score. PAR: prognosis after resuscitation. MPI: modified pre-arrest morbidity. CFS: clinical frailty scale.

Study design	Patient inclusion	Pre-arrest prediction model	Survival to hospital discarge	Survival with favourable neurological outcome

Limpawattana 2018:[38] single center historical cohort from Thailand.	192 adults from 2013-2014. Exclusion: Subsequent events during same hospitalization.	PAM ≥6, PAR (n/a), MPI≥5	25.9%	n/a
historical cohort from Sweden.	Excluded suicides.	PAM>7	20.2%	n/a
Ebell 1997:[28] Historical cohort from 3 US hospitals.	656 adults from 1990-1993. Exlusion: arrests in the ED and OR and subsequent events.	PAM >8 & PAR>8	5.3%	n/a
O'Keeffe 1994:[33] Sinlge center historical cohort from Ireland.	274 adults (index events) in wards, ED, and ICU from an unknown 2 year-period. Exclusion: unknown.	PAR >5 and PAM >8	9.1%	n/a
Bowker 1999:[29] Single center historical cohort from the UK.	264 adult index events from 1994-1996. Exclusion: None.	PAM >6, PAR >7, MPI >6	10.6%	n/a
George Jr 1989:[34] Single center historical cohort from the US.	140 adult IHCAs from 1985. Not including ED, OR, cath lab.	PAM >8	24%	n/a
Cohn 1993:[32] Single center historical cohort from the US.	43 survivors and 43 non-survivors from 1986-1991. Exclusion: None mentioned.	PAM >8	50%	n/a
Ibitoye 2020:[36] Single center historical cohort from the UK.	90 adults (>60 years) from 2017-2018 in a single center in UK. Exclusion criteria not described.	CFS: >4	14%	13%
Ebell 1993:[30] Single center historical cohort from the US.	218 adult index IHCAs from 1989-1991. Exclusion: None mentioned.	Neuronal network	15.6%	n/a
Stark 2015:[31] Single center historical cohort from the US.	62 adult, elective surgical patients (includes 10 without CA) from 2013-2014. Exclusion: none mentioned.	MEWS≥7	43.5%	n/a
Roberts 2017:[39] Single center historical cohort from Sweden.	296 adult IHCAs from 2015. Included all patients with calculated NEWS score within 12h.	NEWS≥7	30%	29.5%
Haegdorens 2020:[37] historical cohort from 6 hospitals in Belgium.	33 adults from 2014-2015. Exclusion: Pregnancy.	NEWS ≥5	57.6%	n/a

Table 2: Characteristics of studies predicting survival with favorable neurological outcome. CPC: cerebral performance category. CART: classification and regression tree. GO-FAR: good outcome following attempted resuscitation. PIHCA: prediction of outcome following in-hospital cardiac arrest. OR: operating room. ECMO: extracorporeal membrane oxygenation. GWTG: get-with-the-guidelines. IHCA: in-hospital cardiac arrest. MET: medical emergency team.

Study design	Patient inclusion	Pre-arrest prediction model	Survival to hospital discarge	Survival with favourable

				neurological outcome
Ohlsson 2016:[45] Single center historical cohort study from Sweden.	287 adults (≥18 years) from 2007-2010. Exclusion: suicides.	GO-FAR ≥24		
Thai 2019:[49] Nationwide historical cohort study from the US.	62131 adults from 2010-2016 within GWTG-R. Exclusion: Patients with missing data on CPC or any predictor.	GO-FAR ≥24	23.3%	15.0% (CPC1)
Cho 2020:[44] Single center historical cohort study from the UK.	1011 adults from 2013-2017. Exclusion: those not resuscitated by the MET and traumas.	GO-FAR ≥24	25.4%	16.0% (CPC1)
Piscator 2018:[43] Historical cohort study from 6 hospitals in Stockholm.	717 adults (≥18 years) from 2013-2014. Exclusion: not described.	GO-FAR ≥24	28%	22.3% (CPC1)
Ebell 2013:[18] Nationwide historical cohort using the GWTG registry.	51420 adults (≥18 years) from 2007-2009 within GWTG. Exclusion: not described.	GO-FAR ≥24	18.5%	10.4% (CPC1)
Rubins 2019:[40] Single center historical cohort study.	403 adults from a single center in the US from 2009-2018. Exclusion: ECMO patients.	GO-FAR ≥24	33.0%	17.4% (CPC1)
Hong 2021:[42] Single center historical cohort from Korea. Predicts survival with CPC ≤2.	863 adult IHCAs from 2013-2017. Included patients with data on GO-FAR and p-albumin.	GO-FAR + GO-FAR + albumin	n/a	14.7%
George 2020:[41] Nationwide historical cohort study utilizing the GWTG registry.	52468 adults (≥18 years) from 2012-2017 within GWTG. Exclusion: Arrests in the OR, missing data on CPC.	GO-FAR 2 ≥24	n/a	20.8%
Piscator 2019: Multicenter historical cohort study.	717 adults (≥18 years) from 6 Swedish hospitals in Stockholm, 2013-2014. Exclusion: not described.	PIHCA	27.5%	25.2%
Ebell 2013:[47] Nationwide historical cohort study from the US. Developed 5 CART models to predict CPC 1. The 2 best models shown.	52527 adults from the NRCPR registry in 2007-2009. Exclusion: Missing outcomes.	CART models	n/a	18.1% (CPC1)
Guilbault 2017:[46] Single center historical cohort study from Sweden. Utilized the two CART models from Ebell 2013.	287 adults (≥18 years) from 2007-2010. Exclusion: not described.	2 CART models	n/a	11.9% (CPC1)

Table 3: Quality of bias assessment using the QUADAS-II framework. Green means low risk of bias, yellow means unclear risk, red means high risk of bias.

	Risk of bias	**Applicability concerns**

	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Bow							
Stark 2015[31]	−	?	+	−	−	?	+
O'Keeffe 1994[33]	+	?	+	−	−	+	+
Limpawattana 2018[38]	−	?	+	−	?	+	+
Haegdorens 2020[37]	−	?	+	−	?	+	+
Ohlsson 2014[35]	+	?	+	−	?	+	+
Ebell 1997[28]	−	?	+	−	−	+	+
Rubins 2019[40]	−	+	?	−	?	?	−
Ebell 2013[47]	+	+	?	−	+	?	−
George 2020[41]	−	+	?	−	−	?	+
Thai 2019[49]	−	+	?	−	+	?	−
Cho 2020[44]	−	+	?	−	−	?	−
Ibitoye 2020[36]	−	−	+	−	−	+	+
Piscator 2018[43]	?	?	?	−	+	?	−
Guilbault 2017[46]	?	?	?	−	−	?	−
Ohlsson 2016[45]	+	?	?	−	−	?	−
Piscator 2019[48]	?	?	?	−	+	?	+
Ebell 2013[18]	?	?	?	−	−	?	−
Hong 2021[42]	?	?	?	−	?	?	+
Roberts 2017[39]	?	?	+	−	?	+	+

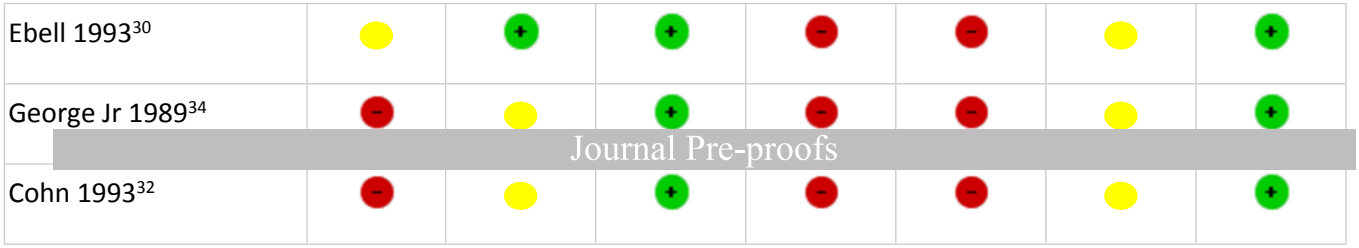


Table 4: Predictive values of scores predicting survival to hospital discharge/ 30-day survival. Data are presented as estimates with 95% confidence intervals. AUC: Area under the receiver operating curve. MEWS: modified early warning score. NEWS: national early warning score. PAM: pre-arrest morbidity score. PAR: prognosis after resuscitation. MPI: modified pre-arrest morbidity. CFS: clinical frailty scale. *The AUC was not reported with a 95% confidence interval in the study.

Study and score	Sensitivity	Specificity	NPV	PPV	AUC
Limpawattana 2018[38] (PAM ≥6, PAR (n/a) or MPI≥5)	n/a	n/a	n/a	n/a	0.62 (0.51 to 0.73)
	n/a	n/a	n/a	n/a	0.60 (0.49 to 0.71)
	n/a	n/a	n/a	n/a	0.63 (0.52 to 0.73)
Ohlsson 2014[35] (PAR >10 or PAM>7)	98.3 (90.8-100)	10.5 (6.8-15.2)	96.0 (79.6-99.9)	21.8 (16.9-27.2)	0.72 (0.65-0.79)
	96.6 (88.1-99.6)	10.9 (7.2-15.7)	92.6 (75.7-99.1)	21.5 (16.7-27.0)	0.60 (0.53–0.67)
Ebell 1997[28] (PAM >8 & PAR>8)	100 (98.5-100)	2.7 (1.4-4.8)	100 (71.5-100)	38.4 (34.7-42.3)	0.52*
	97.6 (94.8-99.1)	30.6 (26.2-35.4)	95.4 (90.3-98.3)	46.1 (41.8-50.5)	0.56*
O'Keeffe 1994[33] (PAR >5 and PAM >8)	100 (86.3-100)	22.8 (17.8-28.4)	100 (93.9-100)	11.1 (7.3-16.0)	0.74*
	100 (86.3-100)	2.0 (0.6-4.5)	100 (47.8-100)	9.1 (6.0-13.2)	0.67*
Bowker 1999[29] (PAM >6, PAR >7, MPI >6)	100 (87.7-100)	19.9 (15.0-25.6)	100 (92.5-100)	12.9 (8.7-18.1)	0.56 (0.54-0.59)
	100 (87.7-100)	28.8 (23.1-35.0)	100 (94.7-100)	14.3 (9.7-20.0)	0.57 (0.55-0.60)
	100 (87.7-100)	22.5 (17.3-28.3)	100 (93.3-100)	13.3 (9.0-18.6)	0.61 (0.59-0.64)
George Jr 1989[34] (PAM >8)	100 (89.7-100)	22.6 (15.1-31.8)	100 (85.8-100)	29.3 (21.2-38.5)	0.61 (0.57-0.65)
Cohn 1993[32] (PAM >8)	100 (92.0-100)	25.0 (12.7-41.2)	100 (69.2-100)	59.5 (47.4-70.4)	0.63 (0.56-0.69)
Ibitoye 2020[36] (CFS:					

>4)	100 (75.3-100)	51.9 (40.3-63.5)	100 (91.2-100)	26.0 (14.6-40.3)	0.76 (0.70-0.82)
(Neuronal network)					
Stark 2015[31] (MEWS≥7)	n/a	n/a	n/a	n/a	0.78*
Roberts 2017[39] (NEWS≥7)	89.3 (80.1-95.3)	31.7 (25.6-38.2)	89.7 (80.8-95.5)	30.7 (24.7-37.3)	0.61 (0.56-0.65)
Haegdorens 2020[37] (NEWS ≥5)	57.9 (33.5-79.7)	71.4 (41.9-91.6)	55.6 (30.8-78.5)	73.3 (44.9-92.2)	0.65 (0.48-0.81)

Table 5: Predictive values of scores predicting survival with favourable neurological outcomes. Data are presented as estimates with 95% confidence intervals. AUC: Area under the receiver operating curve CPC: cerebral performance category. CART: classification and regression tree. GO-FAR: good outcome following attempted resuscitation. PIHCA: prediction of outcome following in-hospital cardiac arrest. OR: operating room. ECMO: extracorporeal membrane oxygenation. GWTG: get-with-the-guidelines. IHCA: in-hospital cardiac arrest. MET: medical emergency team. *The AUC was not reported with a 95% confidence interval in the study.

Study and score	Sensitivity	Specificity	NPV	PPV	AUC
Ohlsson 2016[45] (GO-FAR)	97.8 (88.2-99.9)	10.3 (6.8-14.9)	96.2 (80.4-99.9)	16.9 (12.5-22.0)	0.85 (0.78-0.91)
Thai 2019[49] (GO-FAR)	99.2 (99.0-99.4)	8.2 (7.9-8.4)	98.4 (97.9-98.7)	16.1 (15.8-16.4)	0.75 (0.75-0.76)
Cho 2020[44] (GO-FAR)	99.4 (96.6-100)	11.4 (9.4-13.8)	99.0 (94.4-100)	17.6 (15.2-20.3)	0.81 (0.78-0.84)
Ebell 2013[47] (Two CART	96.0 (94.9-96.9)	24.1 (23.3-24.8)	97.8 (97.2-98.3)	14.6 (13.9-15.2)	0.73*

models)	94.1 (92.9-95.2)	29.5 (28.8-30.3)	97.5 (97.0-98.0)	14.7 (14.1-15.4)	0.71*
Guilbault 2017[46] (Two CART models)	99.5)	34.6)	99.7)	25.9)	0.77*
	95.6 (84.9-99.5)	36.4 (30.3-42.8)	97.8 (92.2-99.7)	21.8 (16.3-28.3)	0.71*
Piscator 2018[43] (GO-FAR)	99.3 (96.1-100.)	9.7 (6.9-13.1)	97.4 (86.2-99.4)	28.9 (24.9-33.1)	0.8 (0.76-0.84)
Ebell 2013[18] (GO-FAR)	99.3 (99.0-99.5)	10.4 (10.1-10.7)	99.2 (98.9-99.5)	11.4 (11.1-11.7)	0.78*
Rubins 2019[40] (GO-FAR)	95.7 (88.0-99.1)	17..1 (13.2-21.6)	95.0 (86.1-99.0)	19.5 (15.5-24.1)	0.68 (0.62-0.73)
Hong 2021[42] (GO-FAR& GO-FAR + Albumin)	n/a	n/a	n/a	n/a	0.79 (0.74–0.85)
	n/a	n/a	n/a	n/a	0.80 (0.75–0.85)
George 2020[41] (GO-FAR2)	98.9 (98.6-99.1)	6.7 (6.4-6.9)	95.7 (94.9-96.4)	21.8 (21.4-22.2)	0.69*
Piscator 2019 (PIHCA)	99.4 (96.8-100)	8.4 (6.0-11.3)	97.4 (86.5-99.9)	29.4 (25.7-33.2)	0.81 (0.81-0.81)

Conflicts of interest: KGL, JB, and RG are members of the ILCOR EIT Task Force (RG as chair). TD is vice chair of the ILCOR first aid task force, JT is member of the ILCOR pediatric life support task force, and KC is member of the ILCOR advanced life support task force. RG is ERC Director of Guidelines and ILCOR.

Supplement legends

Supplement 2: Characteristics of each of the included scores. MEWS: modified early warning score. NEWS: national early warning score. PAM: pre-arrest morbidity score. PAR: prognosis after resuscitation. MPI: modified pre-arrest morbidity. CFS: clinical frailty scale. CART: classification and regression tree. GO-FAR: good outcome following attempted resuscitation. PIHCA: prediction of outcome following in-hospital cardiac arrest.

Supplement 3: True positives, false positives, true negatives, false negatives, positive likelihood ratio, and negative likelihood ratio for studies predicting survival to hospital discharge (Supplement 3A) and studies predicting survival with favourable neurological outcome (Supplement 3B). Studies without data on positives and negatives were marked as n/a for likelihood ratios.

Supplement 4: Evidence tables for each score describing the downgrading for risk of bias, indirectness, inconsistency, imprecision, and other considerations.

Search strategy EMBASE

No.	Query
#1	resuscitation:ti,ab,kw OR cpr:ti,ab,kw
#2	'do not resuscitate':ti,ab,kw OR dnacpr:ti,ab,kw OR dnar:ti,ab,kw
#3	'cardiac arrest':ti,ab,kw OR 'heart arrest':ti,ab,kw OR 'cardiopulmonary arrest':ti,ab,kw OR 'cardio-pulmonary arrest':ti,ab,kw OR 'pre-arrest$':ti,ab,kw OR 'prearrest':ti,ab,kw OR 'arrest in hospital':ti,ab,kw OR 'in-hospital arrest$':ti,ab,kw OR ihca:ti,ab,kw OR 'ih-ca':ti,ab,kw
#4	#1 OR #2 OR #3
#5	#4 NOT ([conference abstract]/lim OR [conference review]/lim OR [editorial]/lim OR [erratum]/lim OR [letter]/lim OR [note]/lim OR [book]/lim OR 'case report'/de)
#6	#5 NOT ([animals]/lim NOT [humans]/lim)
#7	#6 AND [1990-2021]/py
#8	(((clinical OR predict* OR prognos*) NEAR/2 (rule$ OR score$ OR tool$ OR decision OR aid$ OR instrument$ OR model$)):ti,ab,kw) OR 'decision aid':ti,ab,kw
#9	go-far':ti,ab,kw OR 'good outcome following attempted resuscitation':ti,ab,kw
#10	#8 OR #9
#11	#7 AND #10
#12	#11 AND ('systematic review'/de OR 'systematic review':ti)

#13	#11 NOT #12

Study and score	True positives	True negatives	False positives	False negatives	Negative likelihood ratio	Positive likelihood ratio
Ibitoye 2020[31] (CFS: >4)	13	40	37	0	0.07 (0.00-1.05)	2.01 (1.56-2.58)
Limpawattana 2018[33] (PAM ≥6, PAR (n/a) or MPI≥5)					n/a	n/a
					n/a	n/a
					n/a	n/a
Haegdorens 2020[32] (NEWS ≥5)	11	10	4	8	0.59 (0.32-1.10)	2.03 (0.81-5.05)
Ohlsson 2014[30] (PAR >10 or PAM>7)	57	24	205	1	0.16 (0.02-1.19)	1.10 (1.04-1.16)
	56	25	204	2	0.32(0.08-1.30)	1.08 (1.01-1.16)
Ebell 1997[23] (PAM >8 & PAR>8)	35	11	610	0	0.75 (0.05-12.50)	1.00 (0.97-1.05)
	29	125	496	6	0.85 (0.40-1.79)	1.04 (0.89-1.21)
O'Keeffe 1994[28] (PAR >5 and PAM >8)	25	59	200	0	0.08 (0.01-1.32)	1.27 (1.17-1.39)
	25	5	249	0	0.89 (0.05-15.68)	1.00 (0.95-1.06)
Bowker 1999[24] (PAM >6, PAR >7, MPI >6)	47	28	189	0	0.08 (0.00-1.28)	1.15 (1.09-1.21)
	68	28	168	0	0.05 (0.00-0.81)	1.17 (1.10-1.24)
	53	28	183	0	0.08 (0.00-1.20)	1.29 (1.20-1.38)
Stark 2015[26] (MEWS≥7)					n/a	n/a
Roberts 2017[34] (NEWS≥7)	67	70	151	8	0.34 (0.17-0.67)	1.31 (1.16-1.47)
Ebell 1993[25] (Neuronal network)	31	96	88	3	0.17 (0.06-0.50)	1.91 (1.59-2.29)
George Jr 1989[29] (PAM >8)	34	24	82	0	0.06 (0.00-1.00)	1.29 (1.17-1.43)
Cohn 1993[27] (PAM >8)	44	10	30	0	0.04 (0.00-0.72)	1.33 (1.11-1.59)

outcome

Study and score	True positives	True negatives	False positives	False negatives	Negative likelihood ratio	Positive likelihood ratio
Ohlsson 2016[40] (GO-FAR)	44	25	217	1	0.22 (0.03-1.55)	1.09 (1.03-1.16)
Thai 2019[44] (GO-FAR)	9275	4313	48471	72	0.09 (0.07-0.12)	1.08 (1.08-1.08)
Cho 2020[39] (GO-FAR)	161	97	752	1	0.05 (0.01-0.38)	1.12 (1.09-1.15)
Ebell 2013[42] (Two CART models)	1649	3063	9654	69	0.17 (0.13-0.21)	1.26 (1.25-1.28)
	1617	3926	8791	101	0.19 (0.16-0.23)	1.36 (1.34-1.38)
Guilbault 2017[41] (Two CART models)	43	69	173	2	0.16 (0.04-0.61)	1.34 (1.21-1.48)
	43	88	154	2	0.12 (0.03-0.48)	1.50 (1.34-1.68)
Piscator 2018[38] (GO-FAR)	140	37	345	1	0.07 (0.01-0.53)	1.10 (1.06-1.14)
Ebell 2013[18] (GO-FAR)	5293	4762	41149	37	0.07 (0.05-0.09)	1.11 (1.10-1.11)
Rubins 2019[35] (GO-FAR)	67	57	276	3	0.25 (0.08-0.78)	1.15 (1.08-1.24)
Hong 2021[37] (GO-FAR& GO-FAR + Albumin)					n/a	n/a
					n/a	n/a
George 2020[36] (GO-FAR2)	10808	2771	38765	124	0.17 (0.14-0.20)	1.06 (1.06-1.06)
Piscator 2019 (PIHCA)	173	38	416	1	0.07 (0.01-0.50)	1.09 (1.05-1.12)