BMJ Open

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045849
Article Type:	Original research
Date Submitted by the Author:	13-Oct-2020

Complete List of Authors:	Alhmoud, Baneen; University College London Bonnici, Tim; University College London Patel, Riyaz; UCL, Farr Institute Melley, Daniel; Barts Health NHS Trust Williams, Bryan; University College London, Institute of Cardiovascular Science; Banerjee, Amitava; University College London, Farr Institute of Health Informatics Research

Keywords:	Adult intensive & critical care < ANAESTHETICS, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™ Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review

Baneen Alhmoud[1,2] *research fellow/PhD candidate*
Timothy Bonnici[1,2] *consultant in intensive care medicine*
Riyaz Patel[1,2,3] *professor of cardiology and honorary consultant cardiologist*
Daniel Melley[3] *consultant in intensive care medicine and honorary senior lecturer*
Bryan Williams[1,2] *professor of medicine and consultant physician*
Amitava Banerjee[1,2,3] *associate professor in clinical data science and honorary consultant cardiologist*

Authors' Address:
[1]University College London, London
[2]University College London Hospitals NHS Trust, London
[3]Barts Health NHS Trust, London

Corresponding author: Dr Amitava Banerjee
Institute of Health Informatics, 222 Euston Road, London. NW1 2DA
ami.banerjee@ucl.ac.uk

Keywords: prediction, early warning score, prognosis, disease, clinical setting, systematic review
Abstract word count: 254
Total word count: 3208

Abstract

Objective:
To assess the predictive performance of early warning scores (EWS) in different disease subgroups and clinical settings.
Design:
Systematic review.
Data sources:
Medline, CINAHL, EMBASE and Cochrane database of systematic reviews from 1997 to 2019.
Inclusion criteria:
Randomised trials and observational studies of internal or external validation of EWS, used to predict deterioration (mortality, ICU transfer and cardiac arrest), in any disease subgroups or clinical setting were included.
Results:
Our search identified 770 studies, of which 108 were included. Study designs and methods used to measure predictive accuracy were inconsistent. Risk of bias was significant (high: n = 26 and unclear: n = 58 and low risk: n =19). Research was predominantly observational with only two randomised trials. Predictive accuracy was highest in medical and surgical settings and respiratory diseases of AUC mean (95%CI): 0.74 (0.74–0.75), 0.77 (0.75–0.80), and 0.77 (0.75–0.80), respectively. There were few studies evaluating EWS in specific diseases, e.g. in cardiology (n = 1), and respiratory (n = 7). Mortality and ICU transfer are the most studied outcomes, and cardiac arrest was least examined (n = 8). EWS integration in electronic health records (EHRs) was found in only nine studies.
Conclusion:
Predictive performance of EWS varies by disease and setting. The methodology and the quality of validation studies of EWS is insufficient to recommend their use in all diseases and all clinical settings. There is an urgent need for consistency in methods, and study design, following consensus guidelines for predictive risk scores. Further research should consider specific diseases and settings, utilising EHR data, prior to large-scale implementation.

Systematic review registration:
PROPSERO CRD42019143141

Strengths and limitations

- The first systematic review to investigate the performance of EWS in different patient disease subgroups and clinical settings.
- The study highlights gaps in EWS research in different disease subgroups and clinical settings.
- This study is limited to specific diseases and settings and does not consider the use of EWS in the general population.
- Analysis of EWS' predictive accuracy is based on AUC results only; the most commonly used measure. Results by other validation measures have not been analysed due to their limitations and differences.

Introduction

Across diseases, patient deterioration can range from critical care review and sepsis, to cardiorespiratory arrest and death, resulting in strain on healthcare resources(1,2). Delays or failures in timely detection of deterioration adversely affect prognosis, morbidity, mortality, and healthcare burdens(3). For example, the 20, 000 in-hospital cardiac arrests per year in England are associated with costs of £50 million for resuscitation and post-arrest care(4).

Specific characteristics have long been known to be associated with deteriorating patient health(2, 5–8), including physiological parameters, such as heart rate and blood pressure (5, 9–11). Early warning scores (EWS), widely used in high-income countries, were borne out of the need for early detection and use simple algorithms based on physiological parameters to help clinicians to recognise any worsening in patient status. Standardised tools, such as the modified early warning score (MEWS) (12) were developed for use across different hospital settings, but specialised tools were also designed for particular subgroups, e.g. Rapid Emergency Medicine Score (REMS) (13) and Quick Sequential Organ Failure Assessment (qSOFA) (14) for patients with infections. In recognising different settings, EWS may have compromised simplicity and timeliness of assessment (15). A number of EWS rely on parameters that do not exist in the first hours of assessment, such as blood investigations and imaging (1,16,17).

From fragmented implementation and inadequate early assessment via specialised tools, EWS have shifted back to the standardised prediction models, particularly, the national early warning score (NEWS)(18), followed by NEWS2 (19). NEWS2 was endorsed by NHS England (20), but concerns have included excessive calls to clinicians, and administrative workload. Moreover, symptoms can vary greatly across diseases and settings; partly due to differing pathophysiology depending on the body system affected (21). Therefore, effective EWS may have to be developed for specific disease populations(22).

Systematic reviews have evaluated EWS in pre-hospital settings, ICU and general wards (3,23,24), and patients with sepsis (12), with narrow inclusion criteria and poor methodological quality of included studies. A recent systematic review evaluated development and validation of EWS in general patients, but did not include studies in specific disease subgroups or settings(25).

Objective

In a systematic review, we aim to describe the performance of EWS in different diseases and different clinical settings.

Methods

Search strategy

The protocol adhered to guidelines of PRISMA-P(26). Published articles were identified by searching MEDLINE, CINHAL and EMBASE, between 1997 (initial development of EWS) and 2019. The Cochrane database was searched for systematic reviews (CDSR) and trials (CENTRAL). For grey literature, Google Scholar was searched. During the screening procedure, studies were added from references in review articles and studies. Search strategies were developed by two authors (BA and AB) and reviewed by a third author (TB). Terms used for searching databases include vocabulary terms for early warning or track and trigger scores and acronyms, identified subgroups and settings (e.g., MeSH) and free-text search terms (Figure 1; see Appendix 1).

Inclusion and exclusion criteria

Patient subgroups are identified according to the disease categories and the clinical settings (Appendix 2).

Studies were included when: (1) validation of EWS in adult patients was in a specific setting or disease; (2) it examined the performance of the score, or the impact on mortality, transfer to higher care and cardiac arrest; (3) studies were prospective and retrospective cohort, cross-sectional, case-control studies and trials.

Studies were excluded when: (1) patients were less than 16 years of age; (2) EWS performance was examined in derivation, not validation; (3) non-standard EWS developed for a specific subgroup; (4) EWS validation was performed in a general patient dataset or setting, e.g. validation in a general hospital without consideration of hospital subgroups.

Data extraction

Articles were screened by title and abstract by one author (BA), then full-text screening was by two reviewers (BA and AB). Data was extracted independently by two reviewers (BA and AB) using a standardised and piloted data form, and any disagreements were resolved by a third reviewer (TB). Items for extraction for studies examining predictive accuracy were based on the CHARMS (27) checklist, except for tool derivation that was excluded. For studies addressing clinical outcomes, data extracted were adapted from Agency for Healthcare Research and Quality criteria (28). Data extraction was by two reviewers (AB and TB). When uncertainty occurred, it was resolved by discussion with the study team. Quantitative analysis was conducted where possible, as well as narrative synthesis.

Quality assessment

Risk of biases in validation studies was assessed using PROBAST (29), which classifies studies as low, unclear, or high risk of bias in four aspects: participant selection, predictors, outcomes and analysis within the overall risk of bias and the study applicability domains. For studies examining the clinical outcomes of EWS, ROBINS-I (30) was used.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research

Results

Studies characteristics

From a total of 16,181 articles identified via databases, 1,355 articles' titles and abstracts were screened, and 770 articles were assessed in full for eligibility. A total of 108 articles were included in the final stage; 103 articles assessed the predictive accuracy of EWS, and five articles pertained to the impact of EWS in various diseases and settings. These studies were predominately observational (retrospective= 65, prospective= 36 and RCT=2). Emergency department (ED) (*n* = 47) was the most common clinical setting, followed by medical (*n* = 13), intensive care unit (ICU) (*n* = 13), then surgical (*n* = 9) settings. Sepsis (*n* = 33) was the commonest disease subgroup, and other subgroups ranged from respiratory (n=8) to renal (n=1) (Figure 1 and 2).

Mortality was the main studied outcome, and cardiac arrest was found in a small number of studies (*n* = 8). The effect of EWS on the long-term clinical outcomes was assessed in clinical settings (*n* = 5): including ICU (*n* = 1), surgical (*n* = 1) and medical settings (*n* = 3).

Quality assessment

There was a significant risk of bias found in majority of studies (high risk of bias=26 and unclear risk of bias = 58), while low risk of bias in only 19 studies. In terms of applicability, the narrow inclusion of examined conditions in a certain disease group commonly related to the risk of bias, while in general assessments, biases were commonly related to low sample size or unspecified timing of EWS assessment. There was a wide variation between studies sample sizes (median: 551 and range: 43 - 920029). Studies varied in defining study populations by number of patients, hospital admissions or not specifying the particular study sample. Almost half of the studies (*n* = 49, 48%) have validated their score on a sample of fewer than 500 patients with either multiple or a single observation set (table 1 and 2). Articles investigating the clinical outcomes in different settings were either of low risk (*n* = 2) or moderate risk of bias (*n* =3). External validation was more common (*n* = 83) than internal validation (*n* = 18) and two studies included internal and external validation (see Appendix 3).

EWS validation in patients' subgroups

- Subgroups and EWS

In the studies validating EWS, there was heterogeneity in subgroup definitions, tools, and methods of measuring predictive accuracy was observed. There was overlap commonly between studies of patients with infections receiving care in emergency settings (31–33) and patients with sepsis admitted to intensive care settings (34,35).

EWS models that were integrated with electronic health records (EHR) were examined in recent studies (*n* = 9). Research on datasets utilising EWS-embedded EHRs had larger sample sizes, ranging from 504 (36) to 13,014 patients (37) (Table 1 and 2), with moderate to high predictive ability (area under the curve, AUC: 0.65–0.85). Several studies included comparison between different EWS on the same cohort was in (32,35,38) (see Appendix 2).

- Methodology

There was significant heterogeneity in methods across studies. The majority of studies were observational. Evaluation of predictive accuracy of different EWS in the same study was common (39–42). To measure accuracy of EWS, AUC was most commonly used (*n* = 94), especially when comparing different EWS in the same study (41,43). Presentation of results was variable; for example, confidence intervals were missing in many studies. Other measures, such as analysing sensitivity and specificity, prognostic index and odds ratios, were found in only eight studies (Table 1 and 2). Consequently, it was only feasible to analyse predictive accuracy in studies where AUC was the chosen measure.

Timing from EWS assessment to endpoints was variable. Many studies included (*n* = 43) measured AUC within 24 to 48 hours, while 11 studies had endpoints more than 48-hrs after EWS. However, the majority (*n* = 65, 63%) did not specify the time horizon or in-hospital outcome.

- Predictive performance of EWS

Outcomes were most commonly mortality, transfer to ICU, developing sepsis (in patients with infections), and cardiac arrest. Few studies examined other outcomes, such as respiratory arrest (*n* = 1) and organ failure (*n* = 4). Mortality, ICU admission and cardiac arrest were best predicted in medical (AUC mean: 0.74, 0.75 and 0.74)(44–46) and surgical settings (0.80, 0.79 and 0.75)(47,48), and respiratory diseases (0.75, 0.80 and 0.75) respectively. EWS prediction of sepsis had reasonable predictive performance in all subgroups (AUC: 0.71–0.79), and infectious diseases in particular (AUC: 0.79). Certain outcomes related to specific disease groups were not studied, e.g. cardiac arrest was not studied in cardiac patients (41); respiratory arrest was not tested in respiratory patients (44,49,50).

In disease groups, the best predictive performance was found in the studies examining cardiac (44), stroke (44,51) and renal (44) diseases (AUC: 0.93, 0.88 and 0.87 respectively). In emergency settings, predictive accuracy was variable (AUC: 0.56–0.91) (52–56). In the haematology and oncology diseases, EWS predictive ability was suboptimal in mortality, cardiac arrest and ICU transfer (AUC: 0.52-0.69; Figures 3 and 4) (57–59). EWS prediction of ICU transfer showed acceptable results in the emergency department settings (55,60), infectious diseases (61,62), and where both groups overlap (39,63), but not in gastroenterology and haematology studies (AUC: 0.64 and 0.60) (58,64). Cardiac arrest was the least examined outcome among the three cut points (*n* = 8) and unstudied in cardiac diseases. (Figures 3 and 4; Appendix 3)

From the diseases and settings explored in this systematic review, the long-term outcomes following EWS implementation were narrowly explored in five studies in the ICU, medical and surgical settings. Results were mixed: mortality rate was reduced in three of the studies, in ICU (8) and medical settings (65); and no improvement was observed in a medical setting , yet the study period was undoubtedly inadequate: four months in each study arm within the same year (66). The ICU transfer and cardiac arrest rates improved in a study in a surgical (67) and a medical area (65), while deteriorated in another medical setting study (66). In the surgical sites, ICU admission rate improved in one study (67).

Discussion

In this comprehensive review of EWS across all diseases and settings, we had three main findings. First, EWS studies in different diseases and clinical settings were heterogeneous in methodology, predictive performance measures, and number of studies in each subgroup, with evidence of suboptimal performance of EWS. Second, validation of EWS is limited in specialised settings, including cardiac disease. Third, despite widespread EHR and EWS integration, few studies have explored EHR-based EWS.

Inconsistency in evaluation and the lack of good-quality validation makes the evidence of validity questionable, which ultimately affects how EWS can and should be used in clinical practice, e.g. predicting risk of future deterioration versus actual deterioration(25). The role of multiple observations and change over time is poorly evaluated. For example, a single observation is generally associated with high AUC compared to multiple observations (44,68). Moreover, AUC, the most commonly used measure of predictive performance, has limitations and other metrics, including positive predictive value, should also be assessed (69, 70). Recording observations at an agreed threshold point before the event in a standardised method is necessary to evaluate EWS performance effectively.

EWS were primarily designed for the general patient populations in wards and emergency departments, and remain under-evaluated in specialised diseases and settings. Critical events are commonly associated with cardiovascular diseases, but EWS are poorly validated in this subgroup. Specific disease areas may show unique alarm signs when critical events are anticipated, which may not be captured by standardised EWS, such as NEWS2, where prediction of deterioration is based on predefined thresholds in all patients (20). Thus, some of the parameters in the EWS might not be applicable, and the score could be unrepresentative of the critical state of these patients (22). A recent study of NEWS2 in patients with coronavirus infection found poor performance in severity prediction (71), despite pre-existing conditions being common and predictive in patients with severe outcomes. EWS may need to take account of disease-specific risk factors and comorbidities, not to mention the changing organisation within hospitals and re-allocation of staff and patients in the current COVID-19 context. .

Widespread uptake of EHR and digitisation of patient observations are expected to contribute to efficient use of EWS, by reducing human errors in documentation and calculation, and delays in escalation of care, as well as better evaluation studies of EWS. However, relatively few studies have considered EHR-based EWS, and those studies have lacked clarity as to whether predictive performance of EWS is related to EHR use, diseases or settings. Investigating the implementation and adoption is necessary to understand the application of EWS. Predictive algorithms derived by machine learning have been successfully used in developing and validating different derived EWS (38,72, 73), but will require robust evaluation. Studying the implementation process of EWS within EHR, providing opportunity for qualitative and quantitative insights into escalation of care and facilitators and barriers to use of EWS in routine practice.

This is a comprehensive systematic review across diseases and clinical settings. We looked across methodologies of evaluation as well as performance of EWS. However, we did not include articles in languages other than English, or studies of EWS in children or EWS derivation. We were concerned with the use of general EWS in particular patient subgroups and did not assess EWS developed specifically for particular subgroups or settings.

In terms of research, validation of EWS in disease subgroups should consider similarities and differences across diseases, sample size, and include measures of model discrimination and calibration. Further research should adhere to established guidelines on clinical outcomes and predictive clinical scoring for decision-making, such as the PROGRESS framework (74). In terms of clinical practice, evidence for use of EWS in specialised settings is currently deficient. Both health professionals and healthcare management teams need to be aware of the limitations of EWS and ensure appropriate specialist nursing and care to fully understand patients in particular subgroups and setting rather than relying on generic EWS.

Conclusion

Early warning scores in specific patient subgroups and settings require further prospective validation of their performance in detecting worsening patient outcomes.

Ethics and dissemination: No ethical approval was required.

Contributions
Study design: AB and BA
Search: BA and AB
Review: BA, AB, TB
Data analysis: BA
Data interpretation: AB, BA, TB
Initial draft of the manuscript: BA and AB
Critical review of final version: All authors.

Competing interests: AB has received research grants from Astra Zeneca. All other authors report no competing interests.

Funding

BA is supported by a grant from the Saudi Arabian Cultural Bureau. AB is supported by research funding from NIHR, British Medical Association, Astra-Zeneca, UK Research and Innovation, and the Innovative Medicines Initiative-2 (BigData@Heart Consortium, under grant agreement No. 116074, supported by the European Union's Horizon 2020 research and innovation programme and EFPIA; chaired by DE Grobbee and SD Anker, partnering with 20 academic and industry partners and ESC). BW is an National Institute for Health Research (NIHR) Senior Investigator and is funded by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

References

1. Cetınkaya HB, Koksal O, Sigirli D, Leylek EH, Karasu O. The predictive value of the modified early warning score with rapid lactate level (ViEWS-L) for mortality in patients of age 65 or older visiting the emergency department. Intern Emerg Med. 2017 Dec;12(8):1253–7.
2. Cei M, Bartolomei C, Mumoli N. In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: A prospective study. Int J Clin Pract. 2009.
3. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PWB. The impact of the use of the Early Warning Score (EWS) on patient outcomes: A systematic review. Resuscitation. 2014.
4. Hogan H, Hutchings A, Wulff J, Carver C, Holdsworth E, Welch J, et al. Interventions to reduce mortality from in-hospital cardiac arrest: a mixed-methods study. Heal Serv Deliv Res. 2019;7(2):1–110.
5. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: A retrospective case record review study. BMJ Qual Saf. 2012.
6. De Meester K, Das T, Hellemans K, Verbrugghe W, Jorens PG, Verpooten GA, et al. Impact of a standardized nurse observation protocol including MEWS after Intensive Care Unit discharge. Resuscitation. 2013.
7. Paterson R, MacLeod DC, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: Clinical audit. Clin Med J R Coll Physicians London. 2006.
8. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted

to a tertiary referral intensive care unit after CPR. Resuscitation. 2011.
9. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K. A comparison of Antecedents to Cardiac Arrests, Deaths and EMergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom - The ACADEMIA study. Resuscitation. 2004.
10. Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, et al. Duration of life-threatening antecedents prior to intensive care admission. Intensive Care Med. 2002.
11. Wilkinson K, Martin IC, Gough MJ. National confidential enquiry into patient outcome and death. An age old problem. A review of the care received by elderly patients undergoing surgery. NCEPOD, London. 2011.
12. Hamilton F, Arnold D, Baird A, Albur M, Whiting P. Early Warning Scores do not accurately predict mortality in sepsis: A meta-analysis and systematic review of the literature. J Infect. 2018.
13. Wuytack F, Meskell P, Conway A, McDaid F, Santesso N, Hickey FG, et al. The effectiveness of physiologically based early warning or track and trigger systems after triage in adult patients presenting to emergency departments: A systematic review. BMC Emerg Med. 2017.
14. Plevin R, Callcut R. Update in sepsis guidelines: what is really new? Trauma Surg Acute Care Open. 2017 Sep 7 [cited 2019 Jul 23];2(1):e000088.
15. Morgan RJM, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. Clin Intensive Care. 1997;8(2):100.
16. Mohammed MA, Rudge G, Watson D, Wood G, Smith GB, Prytherch DR, et al. Index blood tests and national early warning scores within 24 hours of emergency admission can predict the risk of in-hospital mortality: a model development and validation study. PLoS One. 2013 May 29;8(5):e64340–e64340.
17. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J EMJ. 2009 Apr;26(4):254–8.
18. Royal College of Physicians of London. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. R Coll Physician. 2012.
19. Royal College of Physicians of London. NHS England approves use of National Early Warning Score (NEWS) 2 to improve detection of acutely ill patients. R Coll Physician. 2017.
20. Inada-Kim M, Nsutebu E. NEWS 2: an opportunity to standardise the management of deterioration and sepsis. BMJ. 2018 Mar 20 [cited 2019 Jul 23];360:k1260.
21. Direkze S, Jain S. Time to intervene? lessons from the NCEPOD cardiopulmonary resuscitation report 2012. Br J Hosp Med. 2012 Oct 16 [cited 2019 Jul 23];73(10):585–7.
22. Badreldin AMA, Doerr F, Bender EM, Bayer O, Brehm BR, Wahlers T, et al. Rapid clinical evaluation: An early warning cardiac surgical scoring system for hand-held digital devices*. Eur J Cardio-thoracic Surg. 2013.
23. Smith MEB, Chiovaro JC, O'Neil M, Kansagara D, Quiñones AR, Freeman M, et al. Early warning system scores for clinical deterioration in hospitalized patients: A systematic review. Annals of the American Thoracic Society. 2014.
24. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: A systematic review.

Resuscitation. 2016.
25. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. bmj. 2020;369.
26. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009.
27. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med. 2014.
28. Chang SM. The Agency for Healthcare Research and Quality (AHRQ) effective health care (EHC) program methods guide for comparative effectiveness reviews: keeping up-to-date in a rapidly evolving field. J Clin Epidemiol. 2011 Nov;64(11):1166–7.
29. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019 Jan 1 [cited 2019 Jul 18];170(1):51.
30. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016 Oct 12 [cited 2019 Jul 17];i4919.
31. Brink A, Alsma J, Verdonschot RJCG, Rood PPM, Zietse R, Lingsma HF, et al. Predicting mortality in patients with suspected sepsis at the Emergency Department; A retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One. 2019 Jan 25;14(1):e0211133–e0211133.
32. Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. Crit Care Med. 2017 Nov;45(11):1805–12.
33. Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA, et al. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department. Intern Emerg Med. 2013 Apr;8(3):255–60.
34. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and q SOFA scores for predicting mortality and length of stay in ICU. J Crit Care. 2017 Oct;41:191–3.
35. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. Hear Lung J Crit Care. 2019 May;48(3):240–4.
36. Vaughn JL, Kline D, Denlinger NM, Andritsos LA, Exline MC, Walker AR. Predictive performance of early warning scores in acute leukemia patients receiving induction chemotherapy. Leuk Lymphoma. 2018 Jun;59(6):1498–500.
37. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015 Aug 5;7(299):299ra122-299ra122.
38. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2018 Oct;131:N.PAG-N.PAG.
39. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-

related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017 Apr 1;195(7):906–11.

40. de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scand J Trauma Resusc Emerg Med. 2017 Sep 11;25(1):91.
41. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013 Apr;84(4):465–70.
42. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted "track and trigger" systems. Resuscitation. 2008.
43. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. Resuscitation. 2008.
44. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012 Mar;83(3):297–302.
45. Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. BJA Br J Anaesth. 2007 May 22;98(6):769–74.
46. Abbott TEF, Torrance HDT, Cron N, Vaid N, Emmanuel J. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med. 2016 Nov;35:78–82.
47. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007 Feb;35(2):402–9.
48. Bartkowiak B, Snyder AM, Benjamin A, Schneider A, Twu NM, Churpek MM, et al. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–63.
49. Qin Q, Xia Y, Cao Y. Clinical study of a new Modified Early Warning System scoring system for rapidly evaluating shock in adults. J Crit Care. 2017 Feb;37:50–5.
50. Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open. 2016 Sep 28;6(9):e011021–e011021.
51. Liljehult J, Christensen T. Early warning score predicts acute mortality in stroke patients. Acta Neurol Scand. 2016 Apr;133(4):261–7.
52. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. Medicine (Baltimore). 2019 Feb;98(6):e14197–e14197.
53. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a prospective observational study. Scand J Trauma Resusc Emerg Med. 2016

Jun 2;24:80.
54. Goulden R, Hoyle M-C, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J EMJ. 2018 Jun;35(6):345–9.
55. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med Off J Eur Soc Emerg Med. 2016 Dec;23(6):406–12.
56. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J. 2014 Jun;31(6):476–81.
57. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted "track and trigger" systems. Resuscitation. 2008.
58. Mulligan A. Validation of a physiological track and trigger score to identify developing critical illness in haematology patients. Intensive Crit Care Nurs. 2010 Aug;26(4):196–206.
59. Cooksley T, Kitlowski E, Haji-Michael P. Effectiveness of Modified Early Warning Score in predicting outcomes in oncology patients. QJM Mon J Assoc Physicians. 2012 Nov;105(11):1083–8.
60. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. BMJ Open. 2019 Jan 17;9(1):e024636–e024636.
61. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care. 2011;15(2):R95–R95.
62. Albur M, Hamilton F, MacGowan AP. Early warning score: a dynamic marker of severity and prognosis in patients with Gram-negative bacteraemia and sepsis. Ann Clin Microbiol Antimicrob. 2016 Apr 12;15:23.
63. Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med. 2018 Apr;13(3):405–12.
64. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. PLoS One. 2016 Aug 17;11(8):e0161401–e0161401.
65. Subbe CP, Davies RG, Williams E, Rutherford P, Gemmell L. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. Anaesthesia. 2003 Aug;58(8):797–802.
66. Sutherasan Y, Theerawit P, Suporn A, Nongnuch A, Phanachet P, Kositchaiwat C. The impact of introducing the early warning scoring system and protocol on clinical outcomes in tertiary referral university hospital. Ther Clin Risk Manag. 2018 Oct;14:2089–95.
67. Heller AR, Mees ST, Lauterwald B, Reeps C, Koch T, Weitz J. Detection of Deteriorating Patients on Surgical Wards Outside the ICU by an Automated MEWS-Based Early Warning System With Paging Functionality. Ann Surg. 2018 May 16.

68. Jarvis SW, Kovacs C, Briggs J, Meredith P, Schmidt PE, Featherstone PI, et al. Are observation selection methods important when comparing early warning score performance? Resuscitation. 2015 May 1 [cited 2020 Jan 7];90:1–6.
69. Romero-Brufau S, Huddleston JM, Naessens JM, Johnson MG, Hickman J, Morlan BW, et al. Widely used track and trigger scores: Are they ready for automation in practice? Resuscitation. 2014 Apr 1 [cited 2020 Jan 9];85(4):549–52.
70. Romero-Brufau S, Huddleston JM, Escobar GJ, Liebow M. Why the C-statistic is not informative to evaluate early warning scores and what metrics to use. Crit Care. 2015;19(1):285.
71. Carr E, Bendayan R, Bean D, O'Gallagher K, Pickles A, Stahl D, et al. Supplementing the National Early Warning Score (NEWS2) for anticipating early deterioration among patients with COVID-19 infection. medRxiv. 2020 Jan 1;2020.04.24.20078006.
72. Churpek MM, Yuen TC, Park SY, Gibbons R, Edelson DP. Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards*. Crit Care Med. 2014 Apr;42(4):841–8.
73. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. N Engl J Med. 2017 Jun 29;376(26):2507–9.
74. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. BMJ Br Med J. 2013 Feb 5;346:e5595.

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review.

Figures and Tables

Figure 1. Search strategy diagram using PRISMA (Preferred reporting items for systematic reviews and meta-analyses).

Figure 2. Number of studies on EWS performance in different subgroups and settings.

Figure 3. EWS average performance average (measured by AUC) in different disease subgroups

Figure 4: EWS average performance (measured by AUC) in different disease clinical settings.

Table 1. Characteristics of included studies on EWS' predictive performance in patients' subgroups and settings (from largest to smallest sample in each subgroup)

Table 2. Characteristics of included studies on EWS' predictive performance in clinical settings (from largest to smallest sample in each setting).

Figure 1. Search strategy diagram using PRISMA (Preferred reporting items for systematic reviews and meta-analyses).

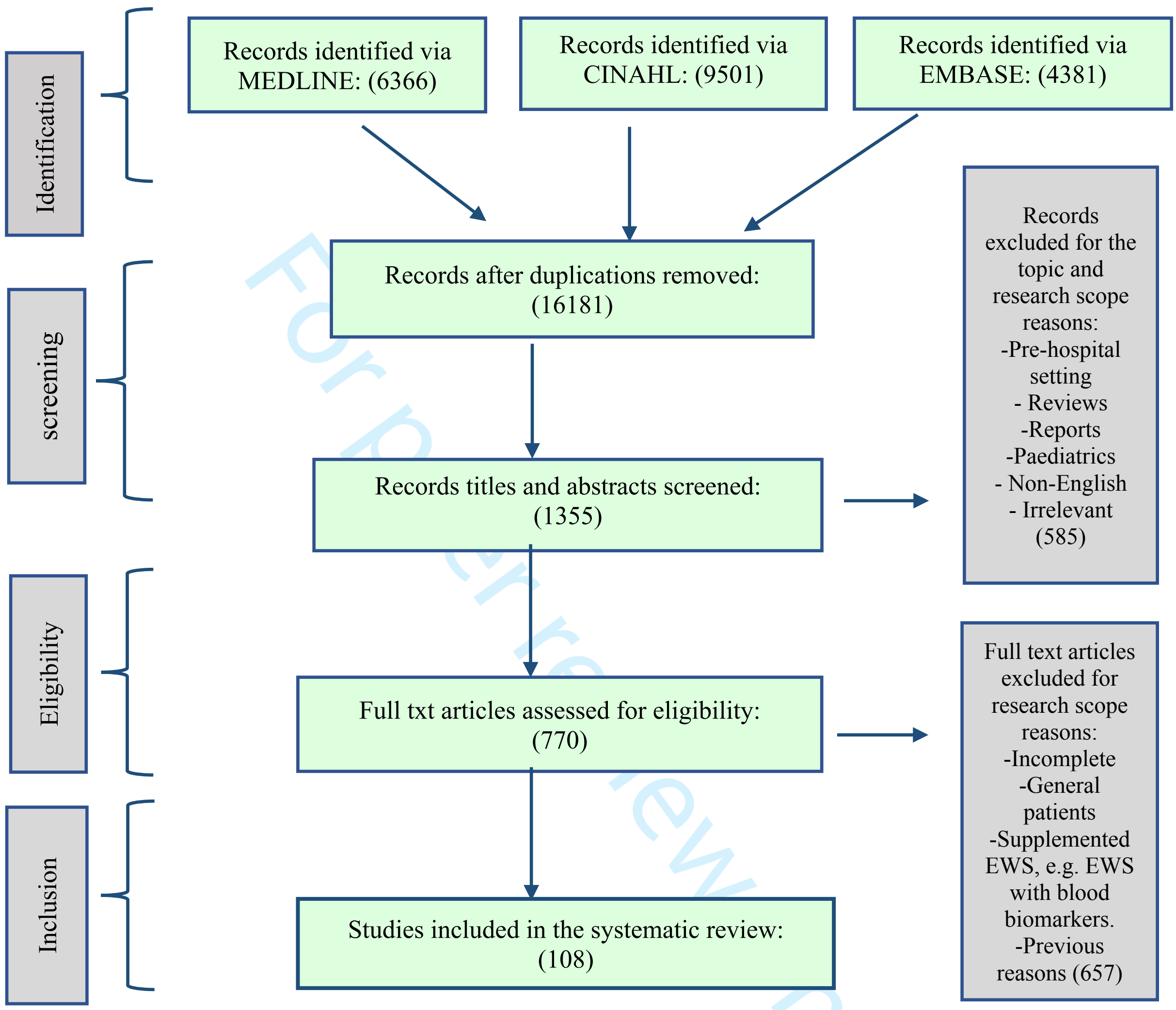
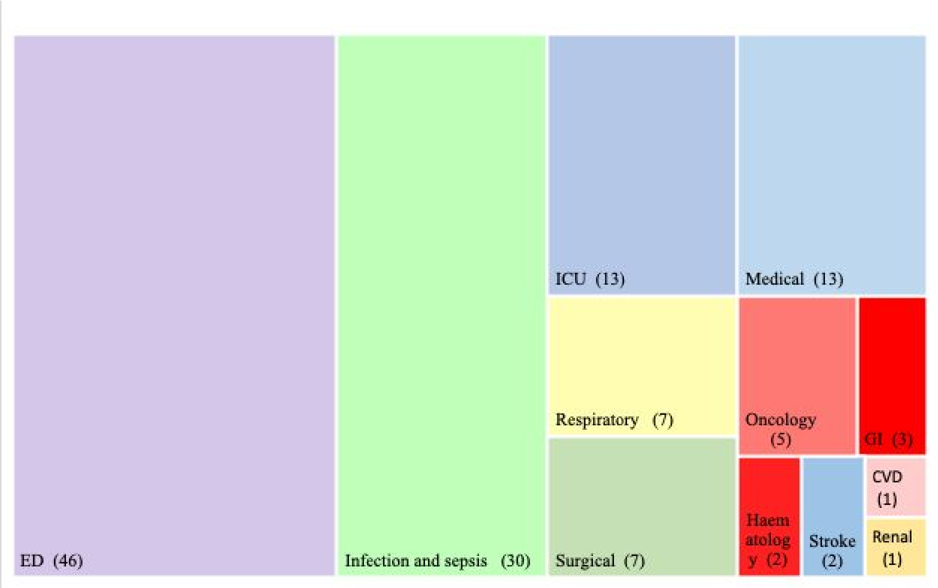
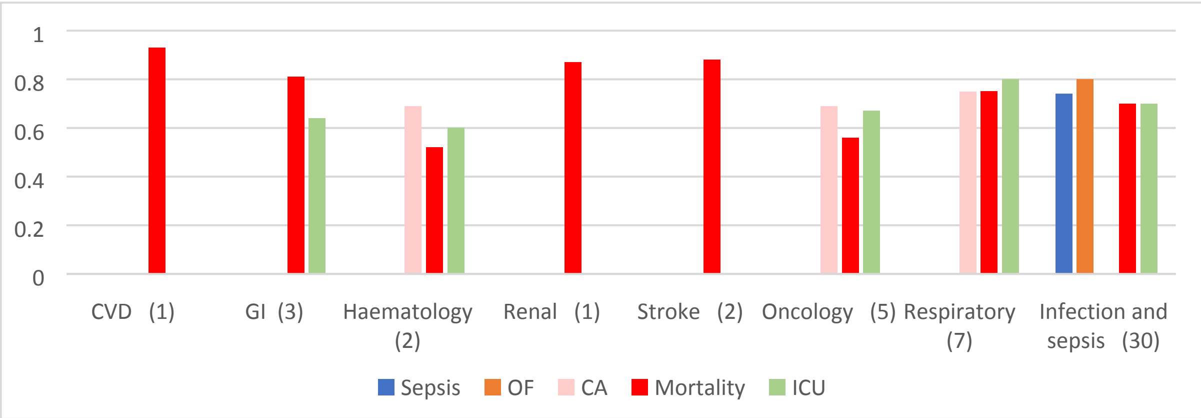


Figure 2. Number of studies on EWS performance in different subgroups and settings.



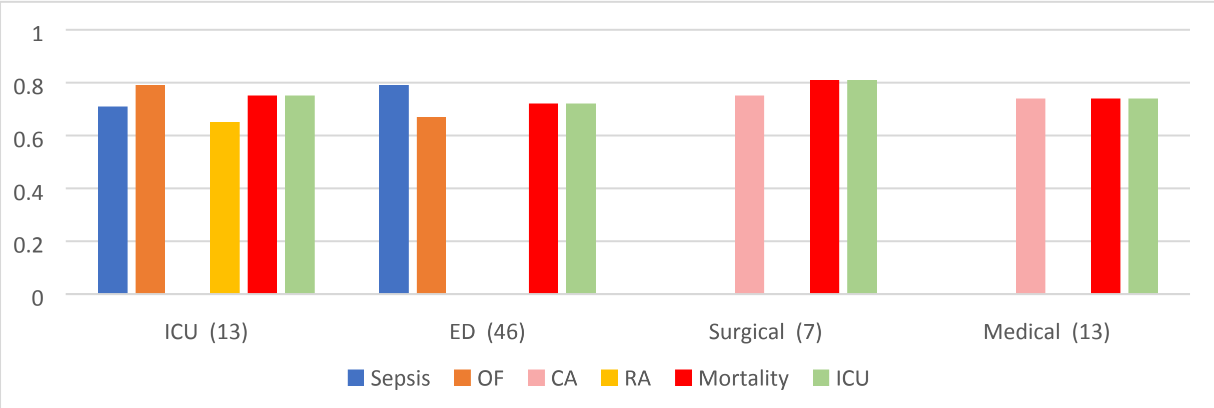
Abbreviations: ED: emergency departments, ICU: intensive care units, GI: gastroenterology, CVD: cardiovascular diseases, (n): indicates the number of studies in each group.

Figure 3. EWS average performance average (measured by AUC) in different disease subgroups



Note: AUC describes EWS ability to predict an outcome accurately, the higher than 0.5, the better the predictability, and 0.5 indicates no ability to predict an

Figure 4: EWS average performance (measured by AUC) in different disease clinical settings.



Note: AUC describes EWS ability to predict an outcome accurately, the higher than 0.5, the better the predictability, and 0.5 indicates no ability to predict an

Table 1. Characteristics of included studies on EWS' predictive performance in patients' subgroups and settings (from largest to smallest sample in each subgroup)

Author, year	Country	Subgroups								Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		Cardiac	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infect/sepsis	ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	CA	RA	Sepsis
Kellett, 2012	Canada	●	○	○	●	●	●	○	○	●	○	●	●	●	○	○	○	10007	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kim, 2017	Korea	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	2172	✓	●	○	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Bozkurt, 2015	Turkey	○	●	○	○	○	○	○	○	○	○	○	○	○	●	○	○	202	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Seak, 2017	Taiwan	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	66	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Hu, 2016	USA	○	○	●	○	○	●	○	○	○	○	○	○	●	○	○	○	565	✓	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Liljehult, 2016	Denmark	○	○	●	○	○	○	○	○	○	○	○	○	●	○	○	○	274	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Mulligan,2010	UK	○	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	71	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cooksley, 2012	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	840	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Vaughn, 2018	USA	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	504	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Young, 2014	USA	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	61	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Von, 2007	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	43	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Pedersen, 2018	Denmark	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	11266	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Forster, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	8812	✓	○	○	○	●	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Pimentel, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	1394	✓	○	○	○	●	●	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Sbiti-rohr, 2016	Switzerland	○	○	○	○	○	○	●	○	○	●	○	○	○	○	●	○	925	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Brabrand, 2017	Denmark	○	○	○	○	○	○	●	○	○	○	○	●	○	●	○	○	570	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2016	Korea	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	553	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Barlow, 2007	UK	○	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	419	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bilben, 2016	Norway	○	○	○	○	○	○	●	○	○	●	○	○	○	●	○	○	246	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Delahanty, 2019	USA	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	920026	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Redfern, 2018	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	241996	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Churpek, Sokol 2017	USA	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	53849	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Faisal, 2019	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	36161		○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓

Author, year	Country	Subgroups: Cardiac	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infect/sepsis	Settings: ICU	ED	Surgical	Medical	Study design: Retrospective	Prospective	RCT	Case Control	Number of patients	EHR	EWS: VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS	Predictive measure	Outcomes studied: Mortality	ICU	C A	R A	Sepsis
Churpek,2017	USA	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	18523	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Henry, 2015	USA	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	13014	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Brink,2019	Netherlands	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	8204	X	○	○	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
De Groot, 2017	Netherlands	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	2280	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	X	X	X
Corfield, 2014	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	2003	X	○	○	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Goulden, 2018	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	1818	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	X	X	X
Khwannimit, 2019	Thailand	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	1589	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Ghanem, 2011	Israel	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	1072	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Saeed, 2019	UK, France, Italy, Sweden & Spain	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	1058	X	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	X	X	X
Innocenti, 2018	Italy	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	742	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Camm, 2018	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	533	X	○	○	○	●	○	○	○	○	○	○	Sens &Spec	✓	✓	X	X	X
Tirotta, 2017	Italy	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	526	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Fong, 2019	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	364	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Prabhakar, 2019	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	343	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Martino, 2018	Italy	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	310	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	X	X	X
Vorwerk, 2009	UK	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	308	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Qin, 2017	China	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	292	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Schmedding, 2019	Gabon	○	○	○	○	○	○	○	●	●	○	○	○	○	●	○	○	277	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Albur, 2016	UK	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	245	X	○	○	●	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Cildir, 2013	Turkey	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	230	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Chiew, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	214	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	X	X	X	X
Samsudin, 2018	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	214	X	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	X	X	X
Chang, 2018	China	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	152	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X
Geier, 2013	Germany	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	151	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	✓
Asiimwe, 2015	Uganda	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	150	X	○	●	○	○	○	○	○	○	○	○	Prognostic index	✓	X	X	X	X
Hung, 2017	Taiwan	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	114	X	○	●	○	○	○	○	○	○	○	○	AUC	✓	X	X	X	X

Garcea, 2006	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	110	✗	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yoo, 2015	Korea	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	100	✗	○	●	○	○	○	○	○	○	○	OR	✓	✓	✗	✗	✗
Siddiqui,2017	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	58	✗	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗

Table 2.

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	CA	RA	Sepsis
Calvert 2016	Israel	●	○	○	○	●	○	○	○	29083	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Awad, 2017	UK	●	○	○	○	●	○	○	○	11722	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Reini, 2012	Sweden	●	○	○	○	○	●	○	○	518	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Chen, 2019	Taiwan	●	○	○	○	●	○	○	○	370	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✓	✗
Baker, 2015	Tanzania	●	○	○	○	○	●	○	○	269	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Gök, 2019	Turkey	●	○	○	○	●	○	○	○	250	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Moseson, 2014	USA	●	○	○	○	○	●	○	○	227	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2013	South Korea	●	○	○	○	●	○	○	○	151	✗	●	●	○	○	○	○	○	●	○	○	AUC	✓	✗	✗	✗	✗
Kown, 2018	Korea	○	●	○	○	●	○	○	○	1986334	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Usman, 2019	USA	○	●	○	○	●	○	○	○	115734	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Jang, 2019	Korea	○	●	○	○	●	○	○	○	56368	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✓	✗	✗
Wei, 2019	China	○	●	○	○	●	○	○	○	39977	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lee, 2019	Korea	○	●	○	○	●	○	○	○	27173	✗	○	●	○	●	○	○	○	○	●	○	AUC	✓	✗	✗	✗	✗
Singer, 2017	USA	○	●	○	○	●	○	○	○	22530	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Eick, 2015	Germany	○	●	○	○	○	●	○	○	5730	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bulut, 2014	Turkey	○	●	○	○	○	●	○	○	2000	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kivipuro, 2018	Finland	○	●	○	○	○	●	○	○	1354	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Eckart, 2019	USA	○	●	○	○	●	○	○	○	1303	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Ho, 2013	Malaysia	○	●	○	○	●	○	○	○	1024	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Skitch, 2018	Canada	○	●	○	○	●	○	○	○	845	✗	○	○	○	●	○	○	○	○	○	●	AUC	✗	✗	✗	✗	✓
Liu, 2014	Malaysia	○	●	○	○	○	●	○	○	702	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✓	✗	✗
Dundar, 2016	Turkey	○	●	○	○	○	●	○	○	671	✗	●	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yuan., 2018	China	○	●	○	○	○	●	○	○	621	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Naidoo, 2014	South Africa	○	●	○	○	●	○	○	○	590	✗	○	○	○	○	○	○	○	○	●	○	Sens & Spec	✓	✗	✗	✗	✗
Liu F.Y, 2015	China	○	●	○	○	○	●	○	○	551	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
So, 2015	China	○	●	○	○	○	●	○	○	544	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Dundar, 2019	Turkey	○	●	○	○	●	○	○	○	455	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lam, 2006	China	○	●	○	○	○	●	○	○	425	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Xie, 2018	China	○	●	○	○	○	●	○	○	383	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cattermole, 2009	China	○	●	○	○	○	●	○	○	330	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Heitz, 2010	USA	○	●	○	○	●	○	○	○	280	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Sirivilaithon, 2019	Thailand	○	●	○	○	○	○	○	●	250	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✗
Cattermole, 2014	China	○	●	○	○	○	●	○	○	230	✗	○	●	○	●	○	○	●	○	○	○	AUC	✓	✗	✗	✗	✗
Najafi, 2018	Iran	○	●	○	○	○	●	○	○	185	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bartkowiak, 2019	USA	○	○	●	○	●	○	○	○	32537	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Kovacs, 2016	UK	○	○	●	●	●	○	○	○	20626	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Plate, 2018	Netherlands	○	○	●	○	○	●	○	○	1782	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Sarani, 2012	Netherlands	○	○	●	○	○	●	○	○	572	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✓	✗	✗	✗

Characteristics of included studies on EWS' predictive performance in clinical settings (from largest to smallest sample in each setting).

9

Hollis, 2016	USA	○	○	●	○	●	○	○	○	522	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Gardner-Thorpe 2006	UK	○	○	•	○	○	•	○	○	334	✗	○	•	○	○	○	○	○	○	○	[illegible]	Sens & Spec	✓	✓	✗	✗	✗
Garcea, 2010	UK	○	○	•	○	•	○	○	○	280	✗	○	○	•	○	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Cuthbertson, 2007	UK	○	○	•	○	•	○	○	○	136	✗	○	•	•	○	○	○	○	○	○	[illegible]	AUC	✗	✓	✗	✗	✗
Prytherch, 2010	UK	○	○	○	•	•	○	○	○	35585	✗	•	○	○	○	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Smith, 2013	UK	○	○	○	•	•	○	○	○	35585	✗	○	○	○	•	○	○	○	○	○	[illegible]	AUC	✓	✓	✓	✗	✗
Rasmussen, 2018	Denmark	○	○	○	•	•	○	○	○	17312	✗	○	○	○	•	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Ghosh, 2018	USA	○	○	○	•	•	○	○	○	2097	✓	○	•	○	•	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Duckitt, 2007	UK	○	○	○	•	○	•	○	○	1102	✗	○	○	•	○	○	○	•	○	○	[illegible]	AUC	✓	✓	✗	✗	✗
Colombo, 2017	Italy	○	○	○	•	•	○	○	○	471	✗	○	•	○	○	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Abbot, 2016	UK	○	○	○	•	○	•	○	○	322	✗	○	○	○	•	○	○	○	○	○	[illegible]	AUC	✓	✗	✗	✗	✗
Wheeler, 2013	Malawi	○	○	○	•	○	•	○	○	302	✗	○	•	○	○	○	○	○	•	○	[illegible]	AUC	✓	✗	✗	✗	✗
Graziadio, 2019	UK	○	○	○	•	○	•	○	○	292	✗	○	○	○	•	○	○	○	○	○	[illegible]	AUC	✓	✓	✗	✗	✗

Abbreviations: VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score; EHR: electronic health records; AUC: area under the curve; Sens and Spec: sensitivity and specificity; OR: odds ratios; ICU: transfer to intensive care unit; CA: cardiac arrest; RA: respiratory arrest.

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review.

(Supplementary data: Appendix)

1 Patients' subgroups

1- Cardiology patients
2- Neurology patients
3- Orthopaedic patients
4- Renal patients
5- Haematology patients
6- Respiratory patients
7- Gastroenterology patients
8- Oncology patients
9- Emergency patients
10- Infection patients
11- Medical patients
12- Surgical patients
13- Intensive care patients

2 Search strategy for MEDLINE

1- ews OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8
10- MH "Cardiovascular Diseases") OR (MH "Cardiology")
11- (MH "Thoracic Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Musculoskeletal Diseases") OR (MH "Orthopedics")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney Diseases, Cystic") OR (MH "Kidney Failure, Chronic") OR (MH "Polycystic Kidney Diseases") OR (MH "Renal Insufficiency, Chronic")
20- renal disease* OR renal failure OR kidney disease*

21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- hematologic disorder* OR hematologic disease* OR hematology
25- 23 OR 24
26- 1 AND 25
27- (MH "Respiratory Tract Diseases")
28- respiratory disease* OR respiratory disorder*
29- 27 OR 28
30- 1 AND 29
31- (MH "Gastroenterology")
32- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
33- 31 OR 32
34- 1 AND 33
35- (MH "Medical Oncology") OR (MH "Surgical Oncology")
36- oncology OR cancer OR chemotherapy
37- 35 OR 36
38- 1 AND 37
39- (MH "Wounds and Injuries") OR (MH "Emergency Medicine")
40- emergency department* OR emergency OR emergency room* OR trauma*
41- 39 OR 40
42- 1 AND 41
43- (MH "Sepsis") OR (MH "Infection")
44- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
45- 43 OR 44
46- 1 AND 45
47- (MH "Obstetrics")
48- (obstetrics and gynecology) OR OBSTETRIC*
49- 47 OR 48
50- 1 AND 49
51- (MH "Allergy and Immunology")
52- immunological disease* OR immunological disorder*
53- 51 OR 52
54- 1 AND 53
55- (MH "Internal Medicine")
56- medical ward*
57- 55 OR 56
58- 1 AND 57
59- (MH "General Surgery")
60- surgical ward*
61- 59 OR 60
62- 1 AND 61
63- 5 OR 9 OR 14 OR 18 OR 22 OR 26 OR 30 OR 34 OR 38 OR 42 OR 46 OR 50 OR 54 OR 58 OR 62

3 Search strategy for CINAHL

1- ews OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8

10- (MH "Heart Diseases") OR (MH "Cardiovascular Diseases")
11- (MH "Heart Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Orthopedic Surgery") OR (MH "Musculoskeletal Diseases")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney, Cystic") OR (MH "Kidney Diseases")
20- renal disease* OR renal failure OR kidney disease*
21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- (MH "Lymphatic Diseases")
25- hematologic disorder* OR hematologic disease* OR hematology
26- 23 OR 24 O 25
27- 1 AND 26
28- (MH "Respiratory Tract Diseases")
29- respiratory disease* OR respiratory disorder*
30- 28 OR 29
31- 1 AND 30
32- (MH "Digestive System Diseases")
33- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
34- 32 OR 33
35- 1 AND 34
36- (MH "Cancer Patients") OR (MH "Oncology")
37- oncology OR cancer OR chemotherapy
38- 36 OR 37
39- 1 AND 38
40- (MH "Wounds and Injuries") OR (MH "Trauma")
41- emergency department* OR emergency OR emergency room* OR trauma*
42- 40 OR 41
43- 1 AND 42
44- (MH "Infection")
45- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
46- 44 OR 45
47- 1 AND 46
48- (MH "Obstetric Emergencies") OR (MH "Obstetric Patients")
49- (obstetrics and gynecology) OR OBSTETRIC*
50- 48 OR 49
51- 1 AND 50
52- (MH "Internal Medicine")
53- (MH "Allergy and Immunology")
54- medical ward
55- immunological disease* OR immunological disorder*
56- 52 OR 53 OR 54 OR 55
57- 1 AND 56
58- (MH "Surgical Patients")
59- surgical ward*
60- 58 OR 59
61- 1 AND 60
62- 5 OR 9 OR 14 OR 18 OR 22 OR 27 OR 31 OR 35 OR 39 OR 43 OR OR 47 OR 51 OR 57 OR 61

4 Early warning scores used in studies of patients' sub-populations and settings

Study	Score	Parameters								
		HR	SBP	RR	Temp	APVU/ LOC	O2 Sat	Supp O2	Urine OP	Other ()
Kellett, 2012	VIEWS	✓	✓	✓	✓	✗	✓	✓	✗	✗
Seak, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bozkurt, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kim, 2017	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hu, 2016	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Mulligan, 2010	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liljehult, 2016	EWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Cooksley, 2012	MEWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Cooksley, 2012	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Vaughn, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Von, 2007	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Young, 2014	MEWS	✓	✓	✓	✓	✗	✗	✗	✗	✓
Barlow, 2007	EWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Bilben, 2016	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Brabrand, 2017	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Forster, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✓	✗
Jo, 2016	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pedersen, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018	NEWS2	✓	✓	✓	✓	✓	✓	✓	✗	✓
Sbiti-rohr, 2016	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Henry, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Innocenti, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Garcea, 2006	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Qin, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Albur, 2016	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Asiimwe, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Brink 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Camm, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Chang, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Chiew, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chiew, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Churpek, Sokol 2017	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, Sokol 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cildir, 2013	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Corfrield, 2014	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Delahanty, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Delahanty, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Faisal, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Geier, 2013	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ghanem, 2011	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Goulden, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hung, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Khwannimit, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Khwannimit, 2019	SOS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Khwannimit, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Martino, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Pong, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pong, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Redfern, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Saeed, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Samsudin, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Samsudin, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Schmedding, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Siddiqui, 2017	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Tirotta, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Vorwerk, 2009	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Yoo, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Awad, 2017	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Baker, 2015	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Calvert 2016	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Gök, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chen, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗

Jo, 2013	HOTEL	✗	✓	✗	✓	✓	✓	✗	✗	✓
Jo, 2013	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Moseson, 2014	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Reini, 2012	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bulut, 2014	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2009	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2014	WORTHING	✓	✓	✓	✓	✓	✓	✗	✗	✗
Cattermole, 2014	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Cattermole, 2014	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Heitz, 2010	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Dundar, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eckart, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eick, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu F.Y, 2015	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Liu F.Y, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ho, 2013	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Jang, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kivipuro, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Kown, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu, 2014	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lee, 2019	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Naidoo, 2014	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Najafi, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Singer, 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Skitch, 2018	HEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Skitch, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
So, 2015	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Sirivilaithon, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lam, 2006	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Usman, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Wei, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Xie, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bartkowiak, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗

Bartkowiak, 2019	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Cuthbertson, 2007	EWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Cuthbertson, 2007	MEWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Garcea, 2010	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Gardner-Thorpe 2006	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Hollis, 2016	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Kovacs, 2016	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Plate, 2018	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Sarani, 2012	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Abbot, 2016	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Duckitt, 2007	WPC	✓	✓	✓	✓	✓	✓	✗	✗	✗
Duckitt, 2007	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
F., 2017	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Gosh, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Gosh, 2018	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Graziadio, 2019	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Prytherch, 2010	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Ramsussen, 2018	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Smith, 2013	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Wheeler, 2013	Hotel	✓	✗	✓	✗	✓	✓	✗	✗	✓
Wheeler, 2013	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Total	133

5 Quality assessment results

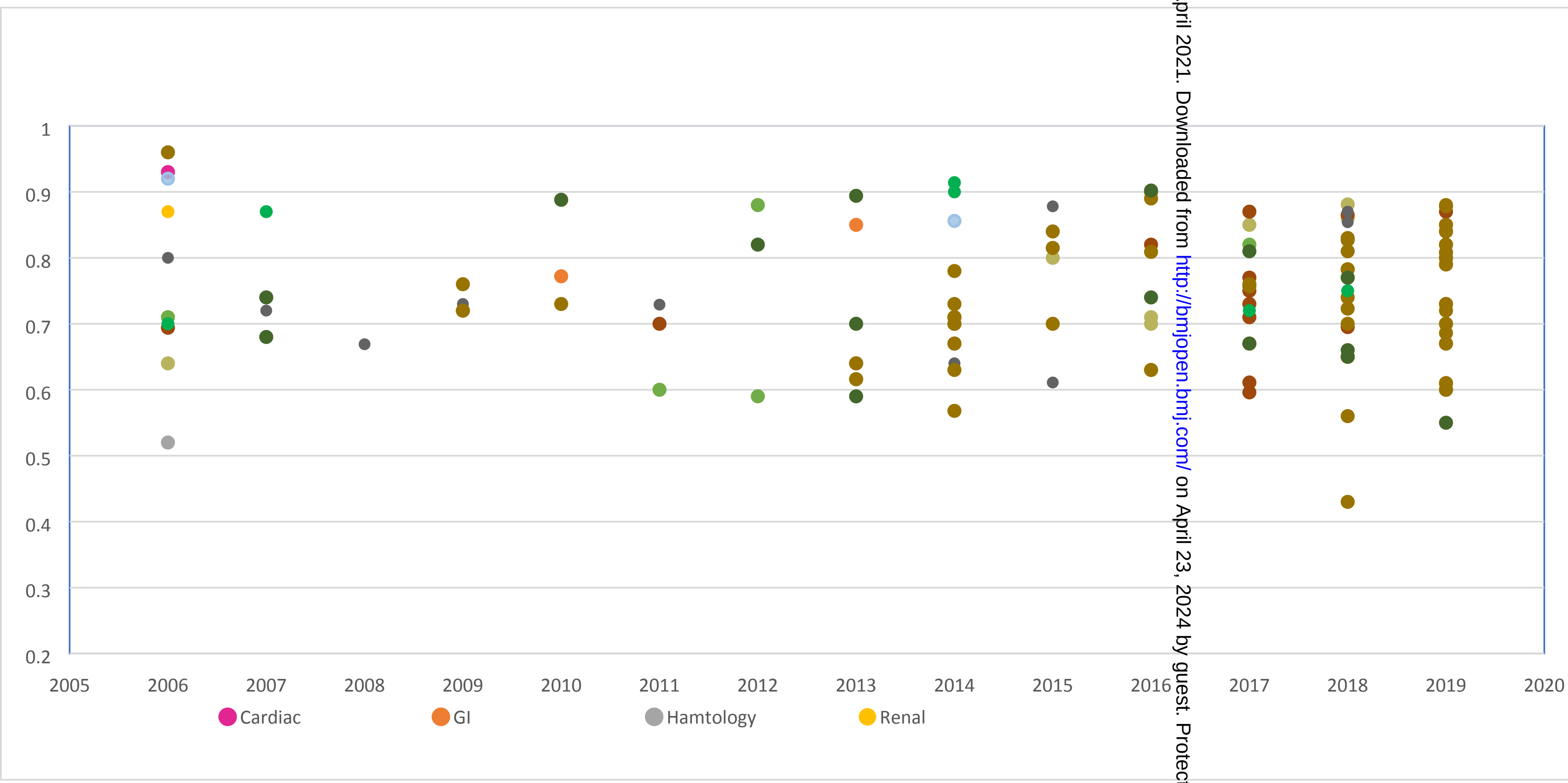
TOOL	Study	Validation	Judgment	
			Robust	Applicability
PROBAST	Kellett, 2012	Externally	low	low
	Seak, 2017	Externally	high	high
	Bozkurt, 2015	Externally	high	high
	Kim, 2017	Externally	unclear	unclear
	Hu, 2016	Internally	unclear	high
	Mulligan, 2010	Externally	high	high
	Liljehult, 2016	Externally	unclear	high
	Cooksley, 2012	Externally	unclear	unclear
	Vaughn, 2018	Externally	high	high
	Von, 2007	Externally	unclear	high
	Young, 2014	Externally	high	high
	Barlow, 2007	Externally	low	unclear
	BILBEN, 2016	Externally	unclear	unclear
	Brabrand, 2017	Externally	unclear	unclear
	Froster, 2018	Externally	low	low
	Jo, 2016	Externally	high	high
	Pedersen, 2018	Externally and Internally	low	low
	Pimentel, 2018	Externally	low	unclear
	Sbiti-rohr, 2016	Externally	unclear	high
	Henry, 2015	Internally	low	low
	Innocenti, 2018	Externally	unclear	unclear
	Garcea, 2006	Externally	unclear	high
	Qin, 2017	Externally	unclear	unclear
	Albur, 2016	Externally	unclear	unclear
	Asiimwe, 2015	Internally	unclear	unclear
	Brink 2019	Externally	unclear	unclear
	CAMM, 2018	Externally	unclear	unclear
	Chang, 2018	Externally	unclear	high
	Chiew, 2019	Externally	unclear	unclear
	Churpek, 2017	Externally	high	high
	Churpek, Sukul 2017	Externally	low	low
	Cildir, 2013	Externally	unclear	unclear
	Corfrield, 2014	Externally	low	low
	de Groot, 2014	Externally	unclear	unclear
	Delahanty, 2019	Internally	low	low
	Faisal, 2019	Externally	low	low
	Geier, 2013	Externally	unclear	unclear
	Gahnem, 2011	Externally	unclear	unclear

	Goulden, 2018	Externally	unclear	unclear
	Hung, 2017	Externally	unclear	high
	Khwannimit, 2019	Externally	unclear	unclear
	Martino, 2018	Externally	unclear	unclear
	Pong, 2019	Internally	unclear	unclear
	Parabhakar, 2019	Internally	unclear	unclear
	Redfern, 2018	Externally	low	low
	Saeed, 2019	Internally	unclear	unclear
	Samsudin, 2018	Internally	unclear	unclear
	Schmedding, 2019	Externally	unclear	unclear
	Siddiqui, 2017	Externally	unclear	unclear
	Tirotta, 2017	Externally	unclear	unclear
	Vorwerk, 2009	Externally	unclear	unclear
	Yoo, 2015	Externally	unclear	unclear
	Awad, 2017	Internally	low	low
	Baker, 2015	Externally	unclear	unclear
	Calvert 2016	Internally	low	unclear
	Gök, 2019	Externally	low	unclear
	Chen, 2019	Externally	unclear	high
	Jo, 2013	Externally	unclear	unclear
	Moseson, 2014	Externally	unclear	unclear
	Reini, 2012	Externally	unclear	unclear
	BULUT, 2014	Externally	unclear	unclear
	Cattermole, 2009	Internally	unclear	unclear
	Cattermole, 2014	Externally	unclear	unclear
	CR, 2010	Externally	high	unclear
	Dundar, 2016	Externally	unclear	high
	Dundar, 2019	Externally	unclear	high
	Eckart, 2019	Externally	unclear	unclear
	Eick, 2015	Externally	unclear	unclear
	F.Y, 2015	Externally	low	unclear
	Ho, 2013	Externally	unclear	unclear
	Jang, 2019	Internally	low	low
	Kivipuro, 2018	Externally	unclear	unclear
	Kown, 2018	Externally and Internally	unclear	unclear
	Liu, 2014	Internally	low	unclear
	Lee, 2019	Internally	low	low
	Naidoo, 2014	Externally	unclear	unclear
	Najafi, 2018	Externally	unclear	high
	Singer, 2017	Externally	unclear	unclear
	Skitch, 2018	Externally	unclear	unclear
	So, 2015	Externally	unclear	unclear
	Sirivilaithon, 2019	Internally	unclear	unclear
	T.S, 2006	Externally	unclear	unclear
	Usman, 2019	Externally	high	high
	W.C., 2018	Externally	unclear	high
	Wei, 2019	Externally	high	high

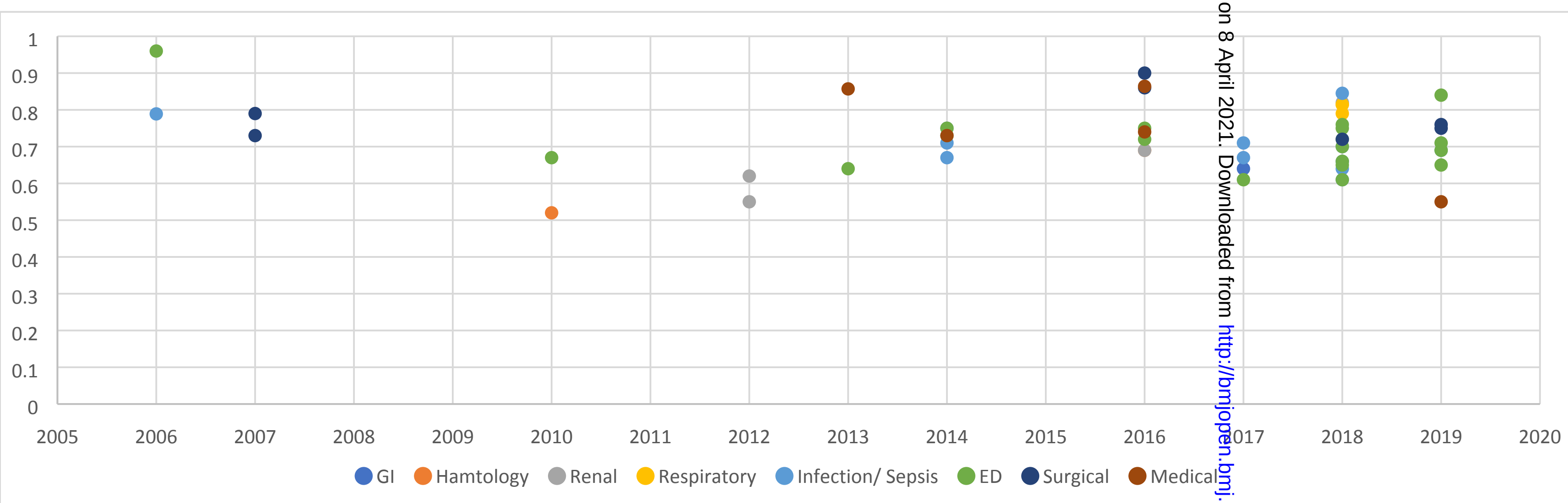
	Xie, 2018	Externally	unclear	unclear
	Bartkowiak, 2019	Externally	unclear	unclear
	Cuthbertson, 2007	Externally	high	unclear
	Garcea, 2010	Externally	high	high
	Gardner-Thorpe 2006	Externally	unclear	unclear
	Hollis, 2016	Externally	unclear	unclear
	Kovacs, 2016	Externally	low	low
	Plate, 2018	Externally	low	low
	Sarani, 2012	Externally	low	low
	Abbot, 2016	Externally	high	high
	Duckitt, 2007	Internally	low	low
	F., 2017	Externally	high	high
	Gosh, 2018	Internally	low	low
	Graziadio, 2019	Externally	unclear	unclear
	Prytherch, 2010	Internally	low	low
	Ramsussen, 2018	Externally	unclear	unclear
	Smith, 2013	Externally	low	low
	Wheeler, 2013	Externally	unclear	unclear
ROBINS-I		Overall bias assessment		
	Moon, 2011	Low		
	Subbe, 2003	Moderate		
	Dawes, 2014	Low		
	Sutherasan, 2018	Moderate		
	Heller, 2018	Low		

Total	108 studies

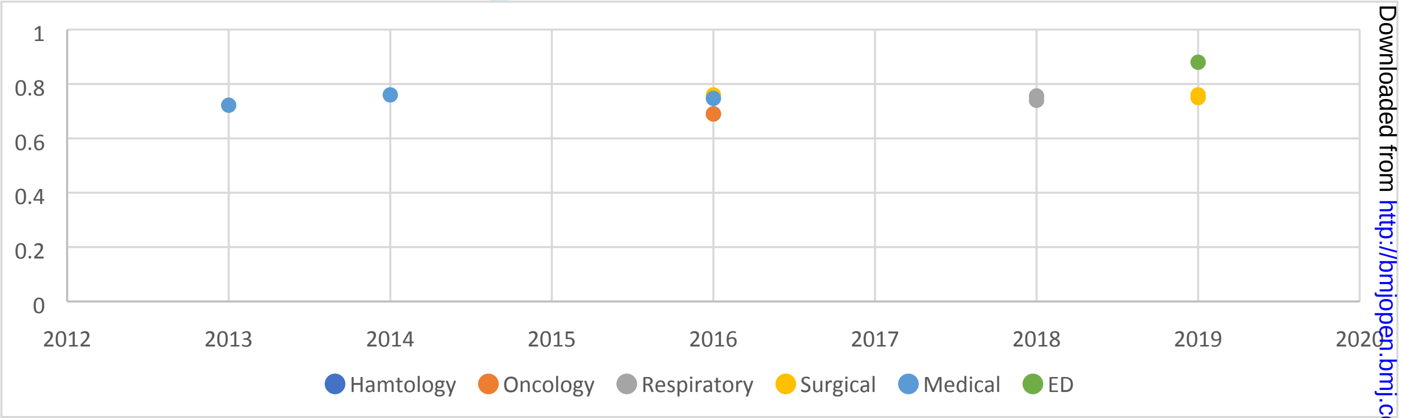
6 EWS' predictive performance (measured by AUC) for mortality in different subgroups and settings



7 EWS' predictive performance (measured by AUC) for ICU admission in different subgroups and settings



8 EWS' predictive performance (measured by AUC) for cardiac arrest in different subgroups and settings



PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	1
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4; Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	4
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	4
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5; figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	5
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	5-6
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	5; Table1; Table 2; Supplementary.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	5-6
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	5-6
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Supplementary
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	7
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	7-8
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	8
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	8

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045849.R1
Article Type:	Original research
Date Submitted by the Author:	29-Dec-2020
Complete List of Authors:	Alhmoud, Baneen; University College London Bonnici, Tim; University College London Patel, Riyaz; UCL, Farr Institute Melley, Daniel; Barts Health NHS Trust Williams, Bryan; University College London, Institute of Cardiovascular Science; Banerjee, Amitava; University College London, Farr Institute of Health Informatics Research
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Epidemiology, Health informatics, Intensive care, Medical management, Patient-centred medicine
Keywords:	Adult intensive & critical care < ANAESTHETICS, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, EPIDEMIOLOGY

SCHOLARONE™
Manuscripts

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review

Baneen Alhmoud[1,2] *research fellow/PhD student*
Timothy Bonnici[1,2] *consultant in intensive care medicine*
Riyaz Patel[1,2,3] *professor of cardiology and honorary consultant cardiologist*
Daniel Melley[3] *consultant in intensive care medicine and honorary senior lecturer*
Bryan Williams[1,2] *professor of medicine and consultant physician*
Amitava Banerjee[1,2,3] *associate professor in clinical data science and honorary consultant cardiologist*

Authors' Address:
[1]University College London, London
[2]University College London Hospitals NHS Trust, London
[3]Barts Health NHS Trust, London

Corresponding author: Dr Amitava Banerjee
Institute of Health Informatics, 222 Euston Road, London. NW1 2DA
ami.banerjee@ucl.ac.uk

Keywords: prediction, early warning score, prognosis, disease, clinical setting, systematic review
Abstract word count: 243
Total word count: 2701

Abstract

Objective:
To assess predictive performance of early warning scores (EWS) in disease subgroups and clinical settings.
Design:
Systematic review.
Data sources:
Medline, CINAHL, EMBASE and Cochrane database of systematic reviews from 1997 to 2019.
Inclusion criteria:
Randomised trials and observational studies of internal or external validation of EWS to predict deterioration (mortality, ICU transfer and cardiac arrest) in disease subgroups or clinical settings.
Results:
We identified 770 studies, of which 108 were included. Study designs and methods were inconsistent, with significant risk of bias (high: *n*=16 and unclear: *n*=64 and low risk: *n*=28). There were only two randomised trials. There was a high degree of heterogeneity in all subgroups and in NEWS (I^2=72-99%). Predictive accuracy (mean AUC; 95% CI) was highest in medical (0.74; 0.74–0.75) and surgical (0.77; 0.75–0.80) settings and respiratory diseases (0.77; 0.75–0.80). Few studies evaluated EWS in specific diseases, e.g. cardiology (*n* = 1), respiratory (*n* = 7). Mortality and ICU transfer were most frequently studied outcomes, and cardiac arrest was least examined (*n*=8). Integration with electronic health records was uncommon (n=9).
Conclusion:
Methodology and quality of validation studies of EWS are insufficient to recommend their use in all diseases and all clinical settings despite good performance of EWS in some subgroups. There is urgent need for consistency in methods and study design, following consensus guidelines for predictive risk scores. Further research should consider specific diseases and settings, utilising electronic health record data, prior to large-scale implementation.

Systematic review registration:
PROSPERO CRD42019143141

Strengths and limitations

- The first systematic review to investigate the performance of general early warning scores in different patient disease subgroups and clinical settings.
- Meta-analysis was performed for different EWS and NEWS validation studies in different disease and clinical setting subgroups
- This study is limited to use of general EWS in specific diseases and settings and does not consider the use of early warning scores in the general population.
- This study did not include EWS designed specifically for particular diseases or clinical settings
- Analysis of predictive accuracy of early warning scores includes area-under-the curve, not other validation measures.

Introduction

Across diseases, patient deterioration can range from critical care review and sepsis, to cardiorespiratory arrest and death, resulting in strain on healthcare resources(1,2). Delays or failures in timely detection of deterioration adversely affect prognosis, morbidity, mortality, and healthcare utilisation(3). For example, the 20000 in-hospital cardiac arrests per year in England are associated with costs of £50 million for resuscitation and post-arrest care(4).

Specific characteristics have long been known to be associated with deteriorating patient health(2, 5–8), including physiological parameters, such as heart rate and blood pressure(5, 9–11). Early warning scores (EWS), widely used in high-income countries, were borne out of the need for early detection of patient deterioration. EWS are tools derived from prediction models that assess patient characteristics and physiological parameters to stratify the risk of developing a worsening event or need for medical attention(12). The algorithms underlying EWS can be "aggregate-weighted" to sum up a set of parameters to produce a score, or use more advanced statistical modelling(13). EWS inform clinical decision-making, enabling escalation of attention and care when required. Standardised tools, such as the modified early warning score (MEWS)(14) were developed for use across different hospital settings, but specialised, non-standard EWS are also designed for particular subgroups, e.g. Rapid Emergency Medicine Score (REMS)(15) and Quick Sequential Organ Failure Assessment (qSOFA) (16) for patients with infections. In recognising different settings, EWS may have compromised simplicity and timeliness of assessment(12). For example, a number of EWS rely on parameters that do not exist in the first hours of assessment, such as blood investigations and imaging(1,17,18).

From fragmented implementation and inadequate early assessment via specialised tools, EWS have shifted back to standardised prediction models, particularly, the national early warning score(NEWS)(19), followed by NEWS2(20). NEWS was designed to produce a standardised assessment of acute illness severity across the NHS(21). While showing good discrimination compared with other EWS, especially in predicting mortality, there was a need to accommodate additional clinical parameters in the score. The updated NEWS2, emphasising appropriate scoring for type 2 respiratory failure, confusion and severe sepsis(20), was formally endorsed by NHS England(22) to be the EWS used in acute care. However, there have been concerns regarding excessive calls to clinicians, administrative workload, and variable symptoms across diseases and settings(23). The effectiveness of standard EWS in specific disease populations is not clear(24), and requires validation to estimate discrimination and calibration, like other clinical prediction models(25). While internal validation is useful, generalisability and reproducibility needs external validation(26).

Systematic reviews have evaluated EWS in pre-hospital, intensive care unit (ICU) and general settings (3,27,28), and sepsis(14), with narrow inclusion criteria and inadequate assessment of study quality. A recent systematic review evaluated development and validation of EWS in general patients, but did not include studies in specific disease subgroups or settings(29).

Objective

In a systematic review, we will assess performance of standardised EWS in particular diseases and clinical settings in predicting mortality, transfer to ICU and cardiac arrest.

Methods

Search strategy

The protocol adhered to PRISMA-P guidelines (30). Published articles were identified in MEDLINE, CINHAL and EMBASE, between 1997 (initial development of EWS) and 2019. The Cochrane database was searched for systematic reviews (CDSR) and trials (CENTRAL). For grey literature, Google Scholar was searched. During the screening procedure, studies were added from references in review articles and studies. Search strategies were developed by two authors (BA and AB) and reviewed by a third author (TB). Terms used for searching databases include terms for early warning or track and trigger scores and acronyms, identified subgroups and settings (e.g., MeSH) and free-text search terms (**Figure 1**; **Supplementary methods**).

Inclusion and exclusion criteria

Patient subgroups were identified according to disease categories and clinical settings (**Supplementary methods**). *Studies were included if:* (1) validation of EWS in adult patients was in a specific setting or disease; (2) the performance of the EWS, or the impact on all-cause mortality, transfer to ITU (admission of a patient to ITU from another clinical setting) and cardiac arrest (loss of cardiac output and function), was examined; (3) they were prospective or retrospective cohort, cross-sectional, case-control design or trials.

Studies were excluded if: (1) patients were less than 16 years of age; (2) EWS performance was only examined in derivation, not validation; (3) non-standard EWS were developed for a specific subgroup, e.g. Obstetric early warning score (OEWS) for obstetric patients or qSOFA for patients with infections; or (4) EWS validation was performed in a general patient dataset or setting, e.g. validation in a general hospital without consideration of hospital subgroups.

Data extraction

Articles were screened by title and abstract by one author (BA), then full-text screening was by two reviewers (BA and AB). Data was extracted independently by two reviewers (BA and AB) using a standardised and piloted data form. A third reviewer (TB) resolved any disagreements. Items for extraction for studies examining predictive accuracy were based on the CHARMS(31) checklist, except for tool derivation which was excluded. For studies addressing clinical outcomes, data extracted were adapted from Agency for Healthcare Research and Quality criteria(32).

Quality assessment

Risk of biases in validation studies was assessed using PROBAST(33), which classifies studies as low, unclear, or high risk of bias in four aspects: participant selection, predictors, outcomes and analysis within the overall risk of bias and the study applicability domains. For studies examining the clinical outcomes of EWS, ROBINS-I(34) was used.

Data analysis:

We analysis conducted using MS Excel and R programmes, and meta-analysis for EWS performance in different subgroups, using AUC (Area Under the Curve), identifying NEWS in studies. Due to missing effect sizes and normal distribution in some studies, we converted AUC to Fisher Z and performed a metanalysis. We evaluated study effect size and tested heterogeneity. Where applicable, we condicted a narrative synthesis.

Results

Study characteristics

Of the 16,181 articles identified by our search, we screened 1,355 articles by title and abstract, assessing 770 articles in full for eligibility. We included 108 studies, published between 2006 and 2019, in the final stage: 103 regarding predictive accuracy of EWS, and five regarding EWS in specific diseases and settings. These studies were predominantly observational (retrospective= 65, prospective= 36 and RCT=2). Emergency department (ED) (*n*=48) was the most common clinical setting, followed by medical (*n* = 12), ICU (*n* = 12), and surgical (*n*=9) settings. Sepsis (*n*=33) was the commonest disease subgroup. Other subgroups ranged from respiratory (n=8) to renal (n=1)(**Figures 1 and 2**).

Mortality was the main studied outcome. Cardiac arrest was infrequently studied (*n*=8). The effect of EWS on longer-term clinical outcomes was assessed in clinical settings (*n*=5): including ICU (*n*=1), surgical (*n*=1) and medical settings(*n*=3).

Quality assessment

There was a significant risk of bias found in majority of studies(high risk=16; unclear risk=64), and low risk in only 28 studies. In terms of applicability, narrow inclusion of conditions in a certain disease group was commonly related to risk of bias, while in general settings, biases were often due to low sample size or unspecified timing of EWS assessment. There was a wide variation in sample size (median: 551 and range: 43 - 920029). There was variation in defining study population by number of patients, hospital admissions or not specifying the particular study sample. Almost half of the studies (*n*=49; 48%) validated in <500 patients with either multiple observations or a single observation set (**Tables 1 and 2**). Articles investigating clinical outcomes in different settings were either of low risk (*n* = 2) or moderate risk of bias(*n* =3). External validation was more common(*n* = 83) than internal validation(*n* = 18) and two studies included internal and external validation(**Table S1**).

EWS validation in patient subgroups

- Subgroups and EWS

In the studies validating EWS, there was heterogeneity in subgroup definitions, tools, and methods of predictive accuracy. There was overlap commonly between studies of patients with infections receiving care in ED(35–37) and patients with sepsis admitted to ITU (38,39). EWS models that were integrated with electronic health records (EHR) were examined in recent studies (*n* = 9). Research on datasets utilising EWS-embedded EHRs had larger sample sizes, ranging from 504(40) to 13,014 patients(41)(**Tables 1 and 2**), with moderate to high predictive ability(AUC: 0.65–0.85). Several studies included comparison between different EWS in the same cohort(n=21)(36,39,42)(**Table S2**).

- Methodology

There was significant heterogeneity in methods across studies. The majority of studies were observational. Evaluation of predictive accuracy of different EWS in the same study was common(21,43–45). To measure accuracy of EWS, AUC was most commonly used(*n*=94), especially when comparing different EWS in the same study(21). Presentation of results was variable; for example, confidence intervals were missing in many studies. Other measures, such as analysing sensitivity and specificity, prognostic index and odds ratios, were found in only eight studies(**Tables 1 and 2**). Consequently, it was only feasible to analyse predictive accuracy in studies where AUC was the selected measure.

Timing from EWS assessment to endpoints was variable. Many studies included (*n* = 43) AUC within 24 to 48 hours, while 11 studies had endpoints more than 48 hours after EWS. However, the majority (*n*=65; 63%) did not specify time horizon or in-hospital outcome.

- Predictive performance of EWS

Outcomes were most commonly mortality, transfer to ICU, developing sepsis (in patients with infections), and cardiac arrest. Few studies examined other outcomes, e.g. respiratory arrest (*n* = 1) and organ failure (*n* = 4). Mortality, ICU admission and cardiac arrest were best predicted in medical (AUC mean: 0.74, 0.75 and 0.74)(46–48) and surgical settings (0.80, 0.79 and 0.75)(49,50), and respiratory diseases (0.75, 0.80 and 0.75) respectively. EWS prediction of sepsis had reasonable predictive performance in all subgroups (AUC: 0.71–0.79), and infectious diseases in particular (AUC: 0.79). Certain outcomes related to specific disease groups were not studied, e.g. cardiac arrest was not studied in cardiac patients(21); respiratory arrest was not investigated in respiratory patients(46,51,52).

The best predictive performance was found in studies examining cardiac(46), stroke(46,53) and renal(46) diseases (AUC: 0.93, 0.88 and 0.87 respectively). In emergency settings, predictive accuracy was variable (AUC: 0.56–0.91)(54–58). In haematology and oncology diseases, EWS predictive accuracy was suboptimal in mortality(**Figure S1**), cardiac arrest and ICU transfer (AUC: 0.52-0.69; **Figures 3 and 4**)(59–61). EWS prediction of ICU transfer was reasonable in ED(58,61), infectious diseases (62,63), and where both groups overlap(43,64), but not in gastroenterology and haematology(AUC: 0.64 and 0.60) (59,65)(**Figure S2**). Cardiac arrest was the least examined outcome among the three endpoints (*n*=8) and unstudied in cardiac diseases. (**Figures 3, 4 and S3**)

For mortality prediction, EWS showed high degree of statistical heterogeneity across subgroups (I^2 = 72% -99%)(**Figure 5**).. In validation studies of NEWS in different disease subgroups, there was also significant heterogeneity (I^2= 99%; **Figure 6**).

Longer-term outcomes following EWS implementation were assessed in five studies in ICU, medical and surgical settings. Results were mixed. Mortality rate was reduced in three of the studies: in ICU(8) and medical settings(66); and no improvement was observed in a medical setting. However, the study duration was likely to be inadequate, e.g. four months(67). The ICU transfer and cardiac arrest rates improved in surgical(68) and medical settings(66), but deteriorated in another study in a medical setting(67).

Discussion

In this comprehensive review of EWS across all diseases and settings, we had three main findings. First, EWS studies in different diseases and clinical settings were heterogeneous in methodology, predictive performance measures, and number of studies in each subgroup. Second, validation of EWS is limited in specialised settings, including cardiac disease. Third, despite widespread EHR and EWS integration, few studies have explored EHR-based EWS.

Inconsistency in evaluation and the lack of high-quality validation ultimately affects how EWS can and should be used in clinical practice, e.g. predicting risk of future deterioration versus actual deterioration(29). Heterogeneity across studies in all subgroups challenges implementation of EWS in all diseases and all settings. The role of multiple observations and change over time is poorly evaluated, e.g. a single observation is generally associated with high AUC compared to multiple observations(46,69). Moreover, AUC, the most commonly used measure of predictive performance, has limitations and other metrics, including positive predictive value, should also be assessed(70). Recording observations at an agreed threshold point before events in a standardised method is necessary to evaluate EWS effectively.

EWS were primarily designed for general patient populations in wards and emergency departments and remain under-evaluated in specific diseases and settings. In medical and ED contexts, EWS perform well, suggesting the role of EWS in general settings, or at the early stage of clinical assessment. Our positive findings in respiratory disease may indicate the emphasis of several EWS, such as NEWS2, on respiratory changes when patients are deteriorating. Specific disease areas may show unique alarm signs when critical events are anticipated, which may not be captured by standardised EWS, such as NEWS2, where prediction of deterioration is based on pre-defined thresholds in all patients(22). Critical events are commonly associated with CVD. With CVD being a leading cause of mortality globally, and the significant impact of morbidity on health and social care, early detection of deterioration is necessary(71). However, EWS are poorly validated in CVD, some of the parameters may not be applicable, and EWS may be unrepresentative(24). A recent study of NEWS2 in patients with coronavirus infection found poor performance in severity prediction (72), despite pre-existing conditions being common and predictive in patients with severe outcomes. EWS may need to take account of disease-specific risk factors and comorbidities.

Widespread uptake of EHR and digitisation of patient observations are expected to contribute to efficient use of EWS, by reducing human errors in documentation and calculation, as well as delays in escalation of care. However, relatively few studies have considered EHR-based EWS, and those studies have not analysed whether predictive performance of EWS is related to EHR use, diseases or settings. Investigating implementation and adoption of EWS is necessary to understand the application and performance of EWS. Predictive algorithms derived by machine learning have been successfully used in developing and validating EWS (42,73), but will require robust evaluation. Studying the implementation process of EWS within EHR will provide opportunities for qualitative and quantitative insights into escalation of care, as well as facilitators and barriers to use of EWS in routine practice.

There are several limitations in this review and in included studies. We aimed for a comprehensive investigation of all EWS developing since 1997, but this long study period may lead to bias in comparing studies with old and new validation approaches statistically and technically. We excluded EWS specifically derived and validated for particular disease populations or settings, and excluded studies considering a general patient population. Meta-analysis was only done for studies using AUC, excluding other methods for assessing performance of EWS. At the title and abstract screening stage, 1170 articles were excluded since they were Non-English, concerned the pre-hospital setting or paediatric populations or were reviews/reports. At the full text screening stage, a further 662 articles were excluded due to incomplete data, general patients rather than subgroups, supplemented EWS (e.g. EWS with blood biomarkers) and the prior reasons. The exclusion of these studies may have affected our findings, particularly the exclusion of non-English studies and those concerning paediatric patients and supplemented EWS. The distinction between general patient settings and specific disease or patient subgroups is dependent on hospital, healthcare system and country, and there is inevitably overlap between patients and settings at different stages in patient pathways. It was only feasible to include studies with a clear disease or setting identified to avoid confusion.

Validation of EWS in disease subgroups should consider similarities and differences across diseases, sample size, and include measures of model discrimination and calibration. Further research should adhere to established guidelines on clinical outcomes and predictive clinical scoring for decision-making, such as the PROGRESS framework(74).

Conclusion

Early warning scores developed for general patient populations require further validation of their performance for detecting worsening outcomes in specific disease subgroups and settings. Despite good performance in respiratory patients and medical and surgical settings in studies to-date, the predictive accuracy of EWS in all disease subgroups and all clinical settings remains unknown. The current evidence base does not necessarily support use of standard EWS in all patients in all settings. Future research should include validation of EWS in particular patient subgroups and settings with standardised methodology.

Contributorship statement

AB conceived the study. BA, AB and TB conducted the search, data extraction and data analysis. BA wrote the initial draft of the manuscript. All authors contributed to revisions of the manuscript.

Competing interests

AB has received research grants from Astra Zeneca. All other authors report no competing interests.

Funding

The Saudi Arabian Cultural Bureau provided PhD funding for BA. There was no further funding for this study.

Data sharing statement

All data relevant to the study are included in the article or uploaded as supplementary information

References

1. Cetınkaya HB, Koksal O, Sigirli D, Leylek EH, Karasu O. The predictive value of the modified early warning score with rapid lactate level (ViEWS-L) for mortality in patients of age 65 or older visiting the emergency department. Intern Emerg Med. 2017 Dec;12(8):1253–7.
2. Cei M, Bartolomei C, Mumoli N. In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: A prospective study. Int J Clin Pract. 2009 Apr;63(4):591-5.
3. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PWB. The impact of the use of the Early Warning Score (EWS) on patient outcomes: A systematic review. Resuscitation. 2014 May;85(5):587-94.
4. Hogan H, Hutchings A, Wulff J, Carver C, Holdsworth E, Welch J, et al. Interventions to reduce mortality from in-hospital cardiac arrest: a mixed-methods study. Heal Serv Deliv Res. 2019;7(2):1–110.
5. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: A retrospective case record review study. BMJ Qual Saf. 2012; Sep;21(9):737-45.
6. De Meester K, Das T, Hellemans K, Verbrugghe W, Jorens PG, Verpooten GA, et al. Impact of a standardized nurse observation protocol including MEWS after Intensive Care Unit discharge. Resuscitation. 2013; Feb;84(2):184-8.
7. Paterson R, MacLeod DC, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: Clinical audit. Clin Med J R Coll Physicians London. May-Jun 2006;6(3):281-4.
8. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted to a tertiary referral intensive care unit after CPR. Resuscitation. 2011; Feb;82(2):150-4.
9. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K. A comparison of Antecedents to Cardiac Arrests, Deaths and EMergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom - The ACADEMIA study. Resuscitation. 2004 Sep;62(3):275-82.
10. Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, et al. Duration of life-threatening antecedents prior to intensive care admission. Intensive Care Med. 2002 Nov;28(11):1629-34.
11. Wilkinson K, Martin IC, Gough MJ. National confidential enquiry into patient outcome and death. An age old problem. A review of the care received by elderly patients undergoing surgery. NCEPOD, London. 2011.
12. Morgan RJM, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. Clin Intensive Care. 1997;8(2):100.
13. Linnen DT, Escobar GJ, Hu X, Scruth E, Liu V, Stephens C. Statistical modeling and aggregate-weighted scoring systems in prediction of mortality and ICU transfer: A systematic review. J Hosp Med. 2019;14(3):161–9.
14. Hamilton F, Arnold D, Baird A, Albur M, Whiting P. Early Warning Scores do not accurately predict mortality in sepsis: A meta-analysis and systematic review of the literature. J Infect. 2018 Mar;76(3):241-248.
15. Wuytack F, Meskell P, Conway A, McDaid F, Santesso N, Hickey FG, et al. The effectiveness of physiologically based early warning or track and trigger systems after triage in adult patients presenting to emergency departments: A systematic review. BMC Emerg Med. 2017 Dec 6;17(1):38.
16. Plevin R, Callcut R. Update in sepsis guidelines: what is really new? Trauma Surg Acute Care Open [Internet]. 2017 Sep 7;2(1):e000088.
17. Mohammed MA, Rudge G, Watson D, Wood G, Smith GB, Prytherch DR, et al. Index blood tests and national early warning scores within 24 hours of emergency admission can predict the risk of in-hospital mortality: a model development and validation study. PLoS One. 2013 May 29;8(5):e64340–e64340.

18. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J. 2009 Apr;26(4):254–8.
19. Royal College of Physicians of London. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. R Coll Physician. 2012.
20. Royal College of Physicians of London. NHS England approves use of National Early Warning Score (NEWS) 2 to improve detection of acutely ill patients. R Coll Physician. 2017.
21. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013 Apr;84(4):465–70.
22. Inada-Kim M, Nsutebu E. NEWS 2: an opportunity to standardise the management of deterioration and sepsis. BMJ [Internet]. 2018 Mar 20;360:k1260.
23. Direkze S, Jain S. Time to intervene? lessons from the NCEPOD cardiopulmonary resuscitation report 2012. Br J Hosp Med [Internet]. 2012 Oct 16;73(10):585–7.
24. Badreldin AMA, Doerr F, Bender EM, Bayer O, Brehm BR, Wahlers T, et al. Rapid clinical evaluation: An early warning cardiac surgical scoring system for hand-held digital devices. Eur J Cardio-thoracic Surg. 2013 Dec;44(6):992-7.
25. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000 Feb 29;19(4):453–73.
26. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol. 2015;68(3):279–89.
27. Smith MEB, Chiovaro JC, O'Neil M, Kansagara D, Quiñones AR, Freeman M, et al. Early warning system scores for clinical deterioration in hospitalized patients: A systematic review. Annals of the American Thoracic Society. 2014 Nov;11(9):1454-65.
28. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: A systematic review. Resuscitation. 2016 May;102:35-43.
29. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ. 2020 May 20;369:m1501.
30. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009 Oct;62(10):1006-12.
31. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med. 2014 Oct 14;11(10):e1001744.
32. Chang SM. The Agency for Healthcare Research and Quality (AHRQ) effective health care (EHC) program methods guide for comparative effectiveness reviews: keeping up-to-date in a rapidly evolving field. J Clin Epidemiol. 2011 Nov;64(11):1166–7.
33. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med [Internet]. 2019 Jan 1;170(1):51.
34. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016 Oct 12;i4919.
35. Brink A, Alsma J, Verdonschot RJCG, Rood PPM, Zietse R, Lingsma HF, et al. Predicting mortality in patients with suspected sepsis at the Emergency Department; A retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One. 2019 Jan 25;14(1):e0211133–e0211133.
36. Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. Crit Care Med [Internet]. 2017 Nov;45(11):1805–12.
37. Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA, et al. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community

acquired sepsis in the emergency department. Intern Emerg Med. 2013 Apr;8(3):255–60.
38. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and q SOFA scores for predicting mortality and length of stay in ICU. J Crit Care. 2017 Oct;41:191–3.
39. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. Hear Lung J Crit Care. 2019 May;48(3):240–4.
40. Vaughn JL, Kline D, Denlinger NM, Andritsos LA, Exline MC, Walker AR. Predictive performance of early warning scores in acute leukemia patients receiving induction chemotherapy. Leuk Lymphoma. 2018 Jun;59(6):1498–500.
41. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015 Aug 5;7(299):299ra122-299ra122.
42. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2018 Oct;131:N.PAG-N.PAG.
43. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017 Apr 1;195(7):906–11.
44. de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scand J Trauma Resusc Emerg Med. 2017 Sep 11;25(1):91.
45. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted "track and trigger" systems. Resuscitation. 2008 May;77(2):170-9.
46. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012 Mar;83(3):297–302.
47. Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. BJA Br J Anaesth. 2007 May 22;98(6):769–74.
48. Abbott TEF, Torrance HDT, Cron N, Vaid N, Emmanuel J. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med. 2016 Nov;35:78–82.
49. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007 Feb;35(2):402–9.
50. Bartkowiak B, Snyder AM, Benjamin A, Schneider A, Twu NM, Churpek MM, et al. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–63.
51. Qin Q, Xia Y, Cao Y. Clinical study of a new Modified Early Warning System scoring system for rapidly evaluating shock in adults. J Crit Care. 2017 Feb;37:50–5.
52. Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open. 2016 Sep 28;6(9):e011021–e011021.
53. Liljehult J, Christensen T. Early warning score predicts acute mortality in stroke patients. Acta Neurol Scand. 2016 Apr;133(4):261–7.
54. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. Medicine (Baltimore). 2019 Feb;98(6):e14197–e14197.
55. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency

department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a prospective observational study. Scand J Trauma Resusc Emerg Med. 2016 Jun 2;24:80.
56. Goulden R, Hoyle M-C, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J EMJ. 2018 Jun;35(6):345–9.
57. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med Off J Eur Soc Emerg Med. 2016 Dec;23(6):406–12.
58. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J . 2014 Jun;31(6):476–81.
59. Mulligan A. Validation of a physiological track and trigger score to identify developing critical illness in haematology patients. Intensive Crit Care Nurs. 2010 Aug;26(4):196–206.
60. Cooksley T, Kitlowski E, Haji-Michael P. Effectiveness of Modified Early Warning Score in predicting outcomes in oncology patients. QJM Mon J Assoc Physicians. 2012 Nov;105(11):1083–8.
61. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. BMJ Open. 2019 Jan 17;9(1):e024636–e024636.
62. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care. 2011;15(2):R95–R95.
63. Albur M, Hamilton F, MacGowan AP. Early warning score: a dynamic marker of severity and prognosis in patients with Gram-negative bacteraemia and sepsis. Ann Clin Microbiol Antimicrob. 2016 Apr 12;15:23.
64. Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med. 2018 Apr;13(3):405–12. A
65. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. PLoS One. 2016 Aug 17;11(8):e0161401–e0161401.
66. Subbe CP, Davies RG, Williams E, Rutherford P, Gemmell L. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. Anaesthesia. 2003 Aug;58(8):797–802.
67. Sutherasan Y, Theerawit P, Suporn A, Nongnuch A, Phanachet P, Kositchaiwat C. The impact of introducing the early warning scoring system and protocol on clinical outcomes in tertiary referral university hospital. Ther Clin Risk Manag. 2018 Oct;14:2089–95.
68. Heller AR, Mees ST, Lauterwald B, Reeps C, Koch T, Weitz J. Detection of Deteriorating Patients on Surgical Wards Outside the ICU by an Automated MEWS-Based Early Warning System With Paging Functionality. Ann Surg. 2018 May 16;
69. Jarvis SW, Kovacs C, Briggs J, Meredith P, Schmidt PE, Featherstone PI, et al. Are observation selection methods important when comparing early warning score performance? Resuscitation. 2015 May 1;90:1–6.
70. Romero-Brufau S, Huddleston JM, Naessens JM, Johnson MG, Hickman J, Morlan BW, et al. Widely used track and trigger scores: Are they ready for automation in practice? Resuscitation. 2014 Apr 1;85(4):549–52.
71. Mozaffarian D. Global Scourge of Cardiovascular Disease. J Am Coll Cardiol. 2017 Jul 4;70(1):26 LP – 28.
72. Carr E, Bendayan R, Bean D, O'Gallagher K, Pickles A, Stahl D, et al. Supplementing the National Early Warning Score (NEWS2) for anticipating early deterioration among patients with COVID-19 infection. medRxiv. 2020 Jan 1;2020.04.24.20078006.
73. Churpek MM, Yuen TC, Park SY, Gibbons R, Edelson DP. Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards. Crit Care Med.

2014 Apr;42(4):841–8.

74. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. BMJ. 2013 Feb 5;346:e5595.

Table 1. Characteristics of included studies of predictive performance for early warning scores in patient subgroups and settings.

Author, year	Country	Subgroups								Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		CVD	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infect/sepsis	ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Kellett, 2012	Canada	●	○	○	●	●	●	○	○	●	○	●	●	●	○	○	○	10007	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kim, 2017	Korea	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	2172	✓	●	○	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Bozkurt, 2015	Turkey	○	●	○	○	○	○	○	○	○	○	○	○	○	●	○	○	202	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Seak, 2017	Taiwan	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	66	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Hu, 2016	USA	○	○	●	○	○	●	○	○	○	○	○	○	●	○	○	○	565	✓	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Liljehult, 2016	Denmark	○	○	○	○	●	○	○	○	○	○	○	○	●	○	○	○	274	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Mulligan,2010	UK	○	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	71	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cooksley, 2012	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	840	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Vaughn, 2018	USA	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	504	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Young, 2014	USA	○	○	●	○	○	●	○	○	○	○	○	○	●	○	○	○	61	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Von, 2007	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	43	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Pedersen, 2018	Denmark	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	11266	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Forster, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	8812	✓	○	○	○	●	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Pimentel, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	1394	✓	○	○	○	●	●	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Sbiti-rohr, 2016	Switzerland	○	○	○	○	○	○	●	○	○	●	○	○	○	○	●	○	925	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Brabrand, 2017	Denmark	○	○	○	○	○	○	●	○	○	○	○	●	○	●	○	○	570	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2016	Korea	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	553	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Barlow, 2007	UK	○	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	419	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bilben, 2016	Norway	○	○	○	○	○	○	●	○	○	●	○	○	○	●	○	○	246	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Delahanty, 2019	USA	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	920026	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Redfern, 2018	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	241996	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Churpek, Sokol 2017	USA	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	53849	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Faisal, 2019	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	36161		○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓

Author, year	Country	Subgroups								Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		CVD	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infect/sepsis	ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Churpek,2017	USA	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	18523	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Henry, 2015	USA	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	13014	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Brink,2019	Netherlands	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	8204	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
De Groot, 2017	Netherlands	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	2280	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Corfield, 2014	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	2003	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Goulden, 2018	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	1818	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Khwannimit, 2019	Thailand	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	1589	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Ghanem, 2011	Israel	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	1072	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Saeed, 2019	UK, France, Italy, Sweden & Spain	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	1058	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Innocenti, 2018	Italy	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	742	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Camm, 2018	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	533	✗	○	○	○	●	○	○	○	○	○	○	Sens &Spec	✓	✓	✗	✗	✗
Trotta, 2017	Italy	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	526	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Pong, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	364	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Prabhakar, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	343	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Martino, 2018	Italy	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	310	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Vorwerk, 2009	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	308	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Qin, 2017	China	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	292	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Schmedding, 2019	Gabon	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	277	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Albur, 2016	UK	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	245	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Cildir, 2013	Turkey	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	230	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Chiew, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	214	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Samsudin, 2018	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	214	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Zhang, 2018	China	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	152	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Geier, 2013	Germany	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	151	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Asiimwe, 2015	Uganda	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	150	✗	○	●	○	○	○	○	○	○	○	○	Prognostic index	✓	✗	✗	✗	✗
Hung, 2017	Taiwan	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	114	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Garcea, 2006	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	110	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yoo, 2015	Korea	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	100	✗	○	●	○	○	○	○	○	○	○	○	OR	✓	✓	✗	✗	✗

Siddiqui,2017	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	58	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗

Studies are ranked according to sample size from largest to smallest in each subgroup.

Abbreviations:

<u>Subgroup:</u> CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastro Intestinal diseases; ICU: Intensive Care Unit.

<u>EWS:</u> Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.

<u>EHR:</u> Electronic Health Records.

<u>Predictive measure:</u> AUC: Area Under the Curve; Sens and Spec: Sensitivity and Specificity; OR: Odds Ratio.

<u>Outcomes:</u> ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	CA	RA	Sepsis
Calvert 2016	Israel	●	○	○	○	●	○	○	○	29083	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Awad, 2017	UK	●	○	○	○	●	○	○	○	11722	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Reini, 2012	Sweden	●	○	○	○	○	●	○	○	518	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Chen, 2019	Taiwan	●	○	○	○	●	○	○	○	370	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✓	✗
Baker, 2015	Tanzania	●	○	○	○	○	●	○	○	269	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Gök, 2019	Turkey	●	○	○	○	●	○	○	○	250	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Moseson, 2014	USA	●	○	○	○	○	●	○	○	227	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2013	South Korea	●	○	○	○	●	○	○	○	151	✗	●	●	○	○	○	○	○	●	○	○	AUC	✓	✗	✗	✗	✗
Kown, 2018	Korea	○	●	○	○	●	○	○	○	1986334	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Usman, 2019	USA	○	●	○	○	●	○	○	○	115734	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Jang, 2019	Korea	○	●	○	○	●	○	○	○	56368	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✓	✗	✗
Wei, 2019	China	○	●	○	○	●	○	○	○	39977	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lee, 2019	Korea	○	●	○	○	●	○	○	○	27173	✗	○	●	○	●	○	○	○	○	●	○	AUC	✓	✗	✗	✗	✗
Singer, 2017	USA	○	●	○	○	●	○	○	○	22530	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Eick, 2015	Germany	○	●	○	○	○	●	○	○	5730	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bulut, 2014	Turkey	○	●	○	○	○	●	○	○	2000	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kivipuro, 2018	Finland	○	●	○	○	○	●	○	○	1354	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Eckart, 2019	USA	○	●	○	○	●	○	○	○	1303	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Ho, 2013	Malaysia	○	●	○	○	●	○	○	○	1024	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Skitch, 2018	Canada	○	●	○	○	●	○	○	○	845	✗	○	○	○	●	○	○	○	○	○	●	AUC	✗	✗	✗	✗	✓
Liu, 2014	Malaysia	○	●	○	○	○	●	○	○	702	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✓	✗	✗
Dundar, 2016	Turkey	○	●	○	○	○	●	○	○	671	✗	●	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yuan., 2018	China	○	●	○	○	○	●	○	○	621	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Naidoo, 2014	South Africa	○	●	○	○	●	○	○	○	590	✗	○	○	○	○	○	○	○	○	●	○	Sens & Spec	✓	✗	✗	✗	✗
Liu F.Y, 2015	China	○	●	○	○	○	●	○	○	551	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
So, 2015	China	○	●	○	○	○	●	○	○	544	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Dundar, 2019	Turkey	○	●	○	○	●	○	○	○	455	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lam, 2006	China	○	●	○	○	○	●	○	○	425	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Xie, 2018	China	○	●	○	○	○	●	○	○	383	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cattermole, 2009	China	○	●	○	○	○	●	○	○	330	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Heitz, 2010	USA	○	●	○	○	●	○	○	○	280	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Sirivilaithon, 2019	Thailand	○	●	○	○	○	○	○	●	250	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✗
Cattermole, 2014	China	○	●	○	○	○	●	○	○	230	✗	○	●	○	●	○	○	●	○	○	○	AUC	✓	✗	✗	✗	✗
Najafi, 2018	Iran	○	●	○	○	○	●	○	○	185	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bartkowiak, 2019	USA	○	○	●	○	●	○	○	○	32537	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Kovacs, 2016	UK	○	○	●	●	●	○	○	○	20626	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Plate, 2018	Netherlands	○	○	●	○	○	●	○	○	1782	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Sarani, 2012	Netherlands	○	○	●	○	○	●	○	○	572	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✓	✗	✗	✗
Hollis, 2016	USA	○	○	●	○	●	○	○	○	522	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Table 2. Characteristics of included studies of predictive performance for early warning scores in clinical settings.

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Gardner-Thorpe 2006	UK	○	○	●	○	○	●	○	○	334	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✓	✗	✗	✗
Garcea, 2010	UK	○	○	●	○	●	○	○	○	280	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Cuthbertson, 2007	UK	○	○	●	○	●	○	○	○	136	✗	○	●	●	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Prytherch, 2010	UK	○	○	○	●	●	○	○	○	35585	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Smith, 2013	UK	○	○	○	●	●	○	○	○	35585	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Rasmussen, 2018	Denmark	○	○	○	●	●	○	○	○	17312	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Ghosh, 2018	USA	○	○	○	●	●	○	○	○	2097	✓	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Duckitt, 2007	UK	○	○	○	●	○	●	○	○	1102	✗	○	○	●	○	○	○	●	○	○	○	AUC	✓	✓	✗	✗	✗
Colombo, 2017	Italy	○	○	○	●	●	○	○	○	471	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Abbot, 2016	UK	○	○	○	●	○	●	○	○	322	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Wheeler, 2013	Malawi	○	○	○	●	○	●	○	○	302	✗	○	●	○	○	○	○	○	●	○	○	AUC	✓	✗	✗	✗	✗
Graziadio, 2019	UK	○	○	○	●	○	●	○	○	292	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Studies are ranked according to sample size from largest to smallest in each subgroup.
Abbreviations:
Subgroup: CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastro Intestinal diseases; ICU: Intensive Care Unit.
EWS: Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.
EHR: Electronic Health Records.
Predictive measure: AUC: Area Under the Curve; Sens and Spec: Sensitivity and Specificity; OR: Odds Ratio
Outcomes: ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.

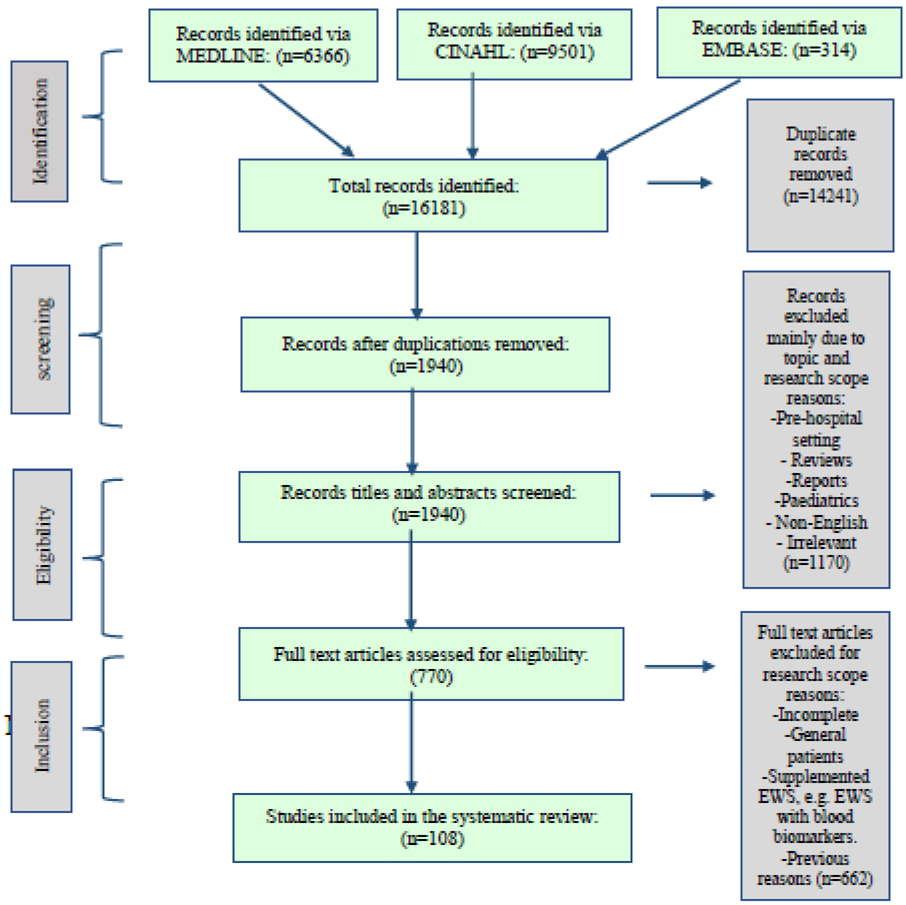


Figure 1. Search strategy and included studies regarding early warning scores in different disease subgroups and clinical settings.

85x82mm (144 x 144 DPI)

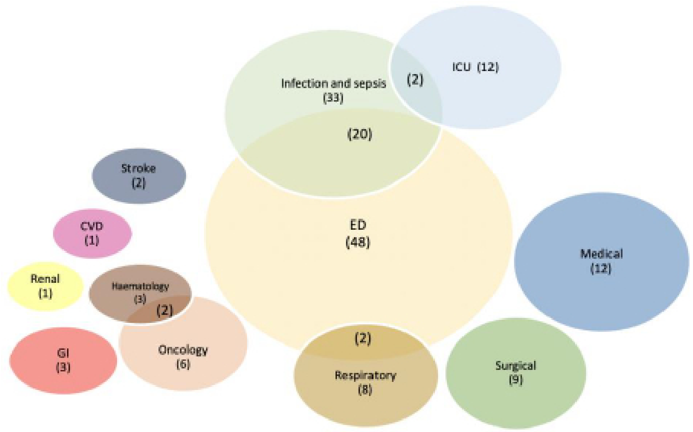


Figure 2. Number of studies regarding performance of early warning scores in different disease subgroups and clinical settings.

Each bubble represents the disease subgroup and/or setting where different early warning scores were examined. The size of the bubble represents the number of studies (n); and overlapping bubbles show studies where disease subgroup and settings overlap. Abbreviations: CVD: Cardiovascular Diseases; ED: Emergency Department; GI: Gastro Intestinal Diseases; ICU: Intensive Care Unit.

115x65mm (144 x 144 DPI)

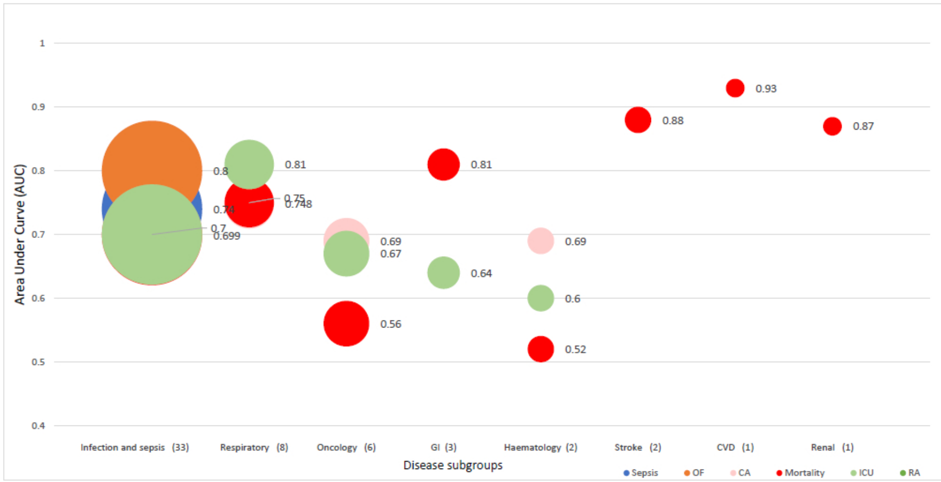


Figure 3. Early warning score performance in different disease subgroups.

Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: CA: cardiac arrest; CVD: cardiovascular diseases; GI: Gastro Intestinal Diseases; ICU: Transfer to Intensive Care Unit; OF: Organ Failure; RA: Respiratory Arrest.

187x94mm (144 x 144 DPI)

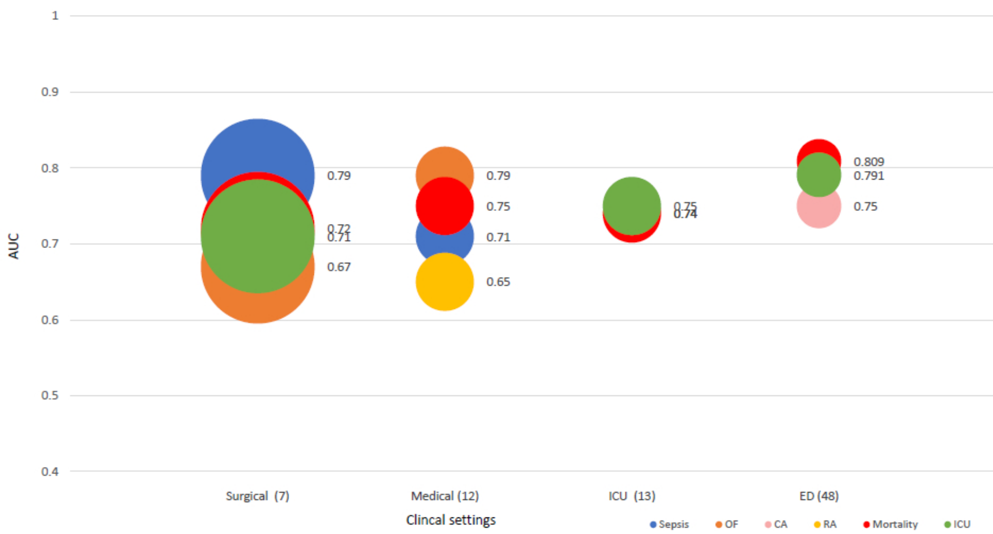


Figure 4. Early warning score performance in different clinical settings.

Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: ED: Emergency Department; ICU: Intensive Care Units; OF: organ failure; CA: Cardiac Arrest; ICU: Transfer to Intensive Care Units; RA: Respiratory Arrest.

166x88mm (144 x 144 DPI)

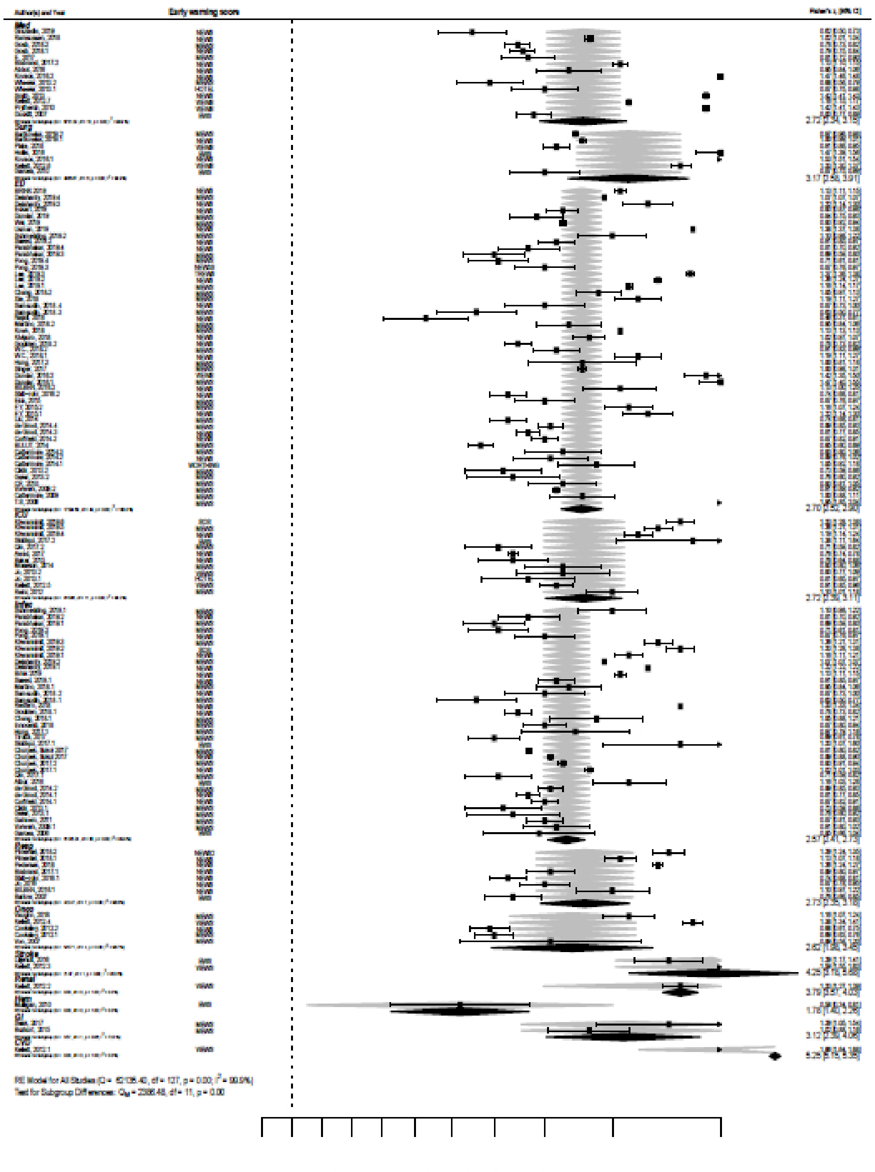


Figure 5: Forest plot of predictive accuracy of early warning scores for mortality in different disease subgroups and clinical settings.

Abbreviations: Med: medical settings, Surg: surgical settings, ED: Emergency Department, ICU: Intensive Care Units, Infec: Infectious Diseases, Resp: Respiratory Diseases, Onco: Oncology diseases, Stroke: Patients with stroke, Renal: Renal diseases, Hem: Haematological diseases, GI: Gastro Intestinal diseases, CVD: Cardiovascular Diseases.
Note: number following Author(s) and year indicate more than one EWS evaluated in the study.

92x127mm (144 x 144 DPI)

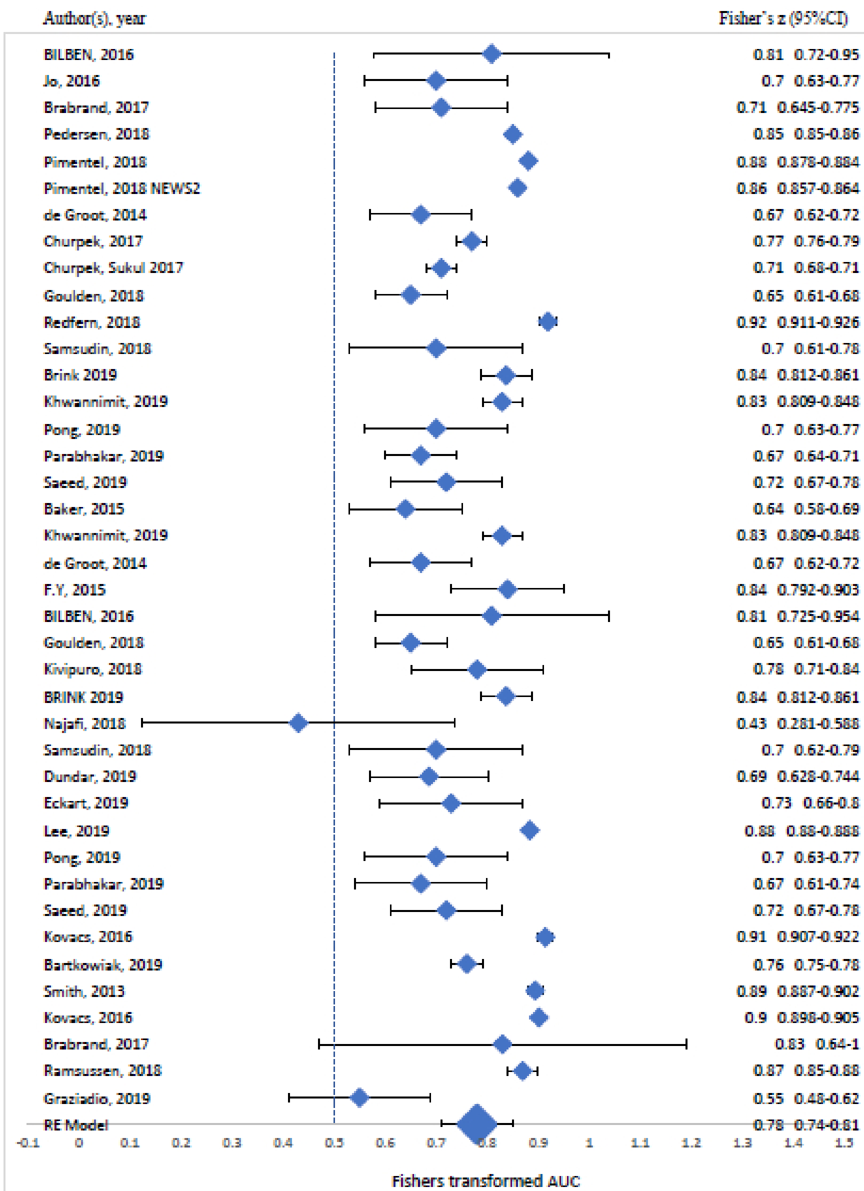


Figure 6: Forest plot predictive accuracy of NEWS for mortality.

RE model for all studies: Q (df = 39) = 37566.8345, p-val < .0001, I2 = 99.87%

84x112mm (144 x 144 DPI)

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review.

Supplementary methods: Search strategy for MEDLINE

1- ews OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8
10- MH "Cardiovascular Diseases") OR (MH "Cardiology")
11- (MH "Thoracic Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Musculoskeletal Diseases") OR (MH "Orthopedics")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney Diseases, Cystic") OR (MH "Kidney Failure, Chronic") OR (MH "Polycystic Kidney Diseases") OR (MH "Renal Insufficiency, Chronic")
20- renal disease* OR renal failure OR kidney disease*
21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- hematologic disorder* OR hematologic disease* OR hematology
25- 23 OR 24
26- 1 AND 25
27- (MH "Respiratory Tract Diseases")
28- respiratory disease* OR respiratory disorder*
29- 27 OR 28
30- 1 AND 29
31- (MH "Gastroenterology")
32- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
33- 31 OR 32
34- 1 AND 33
35- (MH "Medical Oncology") OR (MH "Surgical Oncology")
36- oncology OR cancer OR chemotherapy
37- 35 OR 36
38- 1 AND 37
39- (MH "Wounds and Injuries") OR (MH "Emergency Medicine")
40- emergency department* OR emergency OR emergency room* OR trauma*
41- 39 OR 40
42- 1 AND 41
43- (MH "Sepsis") OR (MH "Infection")
44- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
45- 43 OR 44
46- 1 AND 45
47- (MH "Obstetrics")
48- (obstetrics and gynecology) OR OBSTETRIC*
49- 47 OR 48
50- 1 AND 49
51- (MH "Allergy and Immunology")
52- immunological disease* OR immunological disorder*

53- 51 OR 52
54- 1 AND 53
55- (MH "Internal Medicine")
56- medical ward*
57- 55 OR 56
58- 1 AND 57
59- (MH "General Surgery")
60- surgical ward*
61- 59 OR 60
62- 1 AND 61
63- 5 OR 9 OR 14 OR 18 OR 22 OR 26 OR 30 OR 34 OR 38 OR 42 OR 46 OR 50 OR 54 OR 58 OR 62

Supplementary methods: Search strategy for CINAHL

1- ews OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8
10- (MH "Heart Diseases") OR (MH "Cardiovascular Diseases")
11- (MH "Heart Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Orthopedic Surgery") OR (MH "Musculoskeletal Diseases")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney, Cystic") OR (MH "Kidney Diseases")
20- renal disease* OR renal failure OR kidney disease*
21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- (MH "Lymphatic Diseases")
25- hematologic disorder* OR hematologic disease* OR hematology
26- 23 OR 24 O 25
27- 1 AND 26
28- (MH "Respiratory Tract Diseases")
29- respiratory disease* OR respiratory disorder*
30- 28 OR 29
31- 1 AND 30
32- (MH "Digestive System Diseases")
33- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
34- 32 OR 33
35- 1 AND 34
36- (MH "Cancer Patients") OR (MH "Oncology")
37- oncology OR cancer OR chemotherapy
38- 36 OR 37
39- 1 AND 38
40- (MH "Wounds and Injuries") OR (MH "Trauma")
41- emergency department* OR emergency OR emergency room* OR trauma*
42- 40 OR 41
43- 1 AND 42
44- (MH "Infection")
45- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
46- 44 OR 45
47- 1 AND 46
48- (MH "Obstetric Emergencies") OR (MH "Obstetric Patients")
49- (obstetrics and gynecology) OR OBSTETRIC*
50- 48 OR 49
51- 1 AND 50
52- (MH "Internal Medicine")
53- (MH "Allergy and Immunology")
54- medical ward
55- immunological disease* OR immunological disorder*

56- 52 OR 53 OR 54 OR 55
57- 1 AND 56
58- (MH "Surgical Patients")
59- surgical ward*
60- 58 OR 59
61- 1 AND 60
62- 5 OR 9 OR 14 OR 18 OR 22 OR 27 OR 31 OR 35 OR 39 OR 43 OR OR 47 OR 51 OR 57 OR 61

Supplementary methods: Patients' subgroups

1- Cardiology patients
2- Neurology patients
3- Orthopaedic patients
4- Renal patients
5- Haematology patients
6- Respiratory patients
7- Gastroenterology patients
8- Oncology patients
9- Emergency patients
10- Infection patients
11- Medical patients
12- Surgical patients
13- Intensive care patients

Supplementary References

1. Kellett J, Kim A. Validation of an abbreviated VitalpacTM Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012 Mar;83(3):297–302.
2. Kim W-Y, Lee J, Lee J-R, Jung YK, Kim HJ, Huh JW, et al. A risk scoring model based on vital signs and laboratory data predicting transfer to the intensive care unit of patients admitted to gastroenterology wards. J Crit Care. 2017 Aug;40:213–7.
3. Bozkurt S, Köse A, Arslan ED, Erdoğan S, Üçbilek E, Çevik İ, et al. Validity of modified early warning, Glasgow Blatchford, and pre-endoscopic Rockall scores in predicting prognosis of patients presenting to emergency department with upper gastrointestinal bleeding. Scand J Trauma Resusc Emerg Med. 2015 Dec 30;23:109.
4. Seak C-J, Yen DH-T, Ng C-J, Wong Y-C, Hsu K-H, Seak JC-Y, et al. Rapid Emergency Medicine Score: A novel prognostic tool for predicting the outcomes of adult patients with hepatic portal venous gas in the emergency department. PLoS One. 2017 Sep 15;12(9):e0184813–e0184813.
5. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. PLoS One. 2016 Aug 17;11(8):e0161401–e0161401.
6. Liljehult J, Christensen T. Early warning score predicts acute mortality in stroke patients. Acta Neurol Scand. 2016 Apr;133(4):261–7.
7. Mulligan A. Validation of a physiological track and trigger score to identify developing critical illness in haematology patients. Intensive Crit Care Nurs. 2010 Aug;26(4):196–206.
8. Cooksley T, Kitlowski E, Haji-Michael P. Effectiveness of Modified Early Warning Score in predicting outcomes in oncology patients. QJM Mon J Assoc Physicians. 2012 Nov;105(11):1083–8.
9. Vaughn JL, Kline D, Denlinger NM, Andritsos LA, Exline MC, Walker AR. Predictive performance of early warning scores in acute leukemia patients receiving induction chemotherapy. Leuk Lymphoma. 2018 Jun;59(6):1498–500.
10. Young RS, Gobel BH, Schumacher M, Lee J, Weaver C, Weitzman S. Use of the modified early warning score and serum lactate to prevent cardiopulmonary arrest in hematology-oncology patients: a quality improvement study. Am J Med Qual. 2014 Nov;29(6):530–7.
11. von Lilienfeld-Toal M, Midgley K, Lieberbach S, Barnard L, Glasmacher A, Gilleece M, et al. Observation-based early warning scores to detect impending critical illness predict in-hospital and overall survival in patients undergoing allogeneic stem cell transplantation. Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant. 2007 May;13(5):568–76.
12. Pedersen NE, Rasmussen LS, Petersen JA, Gerds TA, Østergaard D, Lippert A. Modifications of the National Early Warning Score for patients with chronic respiratory disease. Acta Anaesthesiol Scand. 2018 Feb;62(2):242–52.
13. Forster S, Housley G, McKeever TM, Shaw DE. Investigating the discriminative value of Early Warning Scores in patients with respiratory disease using a retrospective cohort analysis of admissions to Nottingham University Hospitals Trust over a 2-year period. BMJ Open. 2018 Jul 30;8(7):e020269–e020269.
14. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2018 Oct;131:N.PAG-N.PAG.

15. Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open 2016 Sep 28;6(9):e011021–e011021.
16. Brabrand M, Hallas P, Hansen SN, Jensen KM, Madsen JLB, Posth S. Using scores to identify patients at risk of short term mortality at arrival to the acute medical unit: A validation study of six existing scores. Eur J Intern Med. 2017 Nov;45:32–6.
17. Jo S, Jeong T, Lee JB, Jin Y, Yoon J, Park B, et al. Validation of modified early warning score using serum lactate level in community-acquired pneumonia patients. The National Early Warning Score-Lactate score. Am J Emerg Med. 2016 Mar;34(3):536–41.
18. Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia. Thorax. 2007 Mar;62(3):253–9.
19. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a prospective observational study. Scand J Trauma Resusc Emerg Med. 2016 Jun 2;24:80.
20. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med 2019 Apr;73(4):334–44.
21. Redfern OC, Smith GB, Prytherch DR, Meredith P, Inada-Kim M, Schmidt PE. A Comparison of the Quick Sequential (Sepsis-Related) Organ Failure Assessment Score and the National Early Warning Score in Non-ICU Patients With/Without Infection. Crit Care Med. 2018 Dec;46(12):1923–33.
22. Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. Crit Care Med. 2017 Nov;45(11):1805–12.
23. Faisal M, Richardson D, Scally AJ, Howes R, Beatson K, Speed K, et al. Computer-aided National Early Warning Score to predict the risk of sepsis following emergency medical admission to hospital: a model development and external validation study. C Can Med Assoc J = J L'association Medicale Can. 2019 Apr 8;191(14):E382–9.
24. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017 Apr 1;195(7):906–11.
25. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015 Aug 5;7(299):299ra122-299ra122.
26. Brink A, Alsma J, Verdonschot RJCG, Rood PPM, Zietse R, Lingsma HF, et al. Predicting mortality in patients with suspected sepsis at the Emergency Department; A retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One. 2019 Jan 25;14(1):e0211133–e0211133.
27. de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scand J Trauma Resusc Emerg Med. 2017 Sep 11;25(1):91.
28. Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emerg Med J. 2014 Jun;31(6):482–7.

29. Goulden R, Hoyle M-C, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J EMJ. 2018 Jun;35(6):345–9.
30. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. Hear Lung J Crit Care. 2019 May;48(3):240–4.
31. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care. 2011;15(2):R95–R95.
32. Saeed K, Wilson DC, Bloos F, Schuetz P, van der Does Y, Melander O, et al. The early identification of disease progression in patients with suspected infection presenting to the emergency department: a multi-centre derivation and validation study. Crit Care. 2019 Feb 8;23(1):40.
33. Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med. 2018 Apr;13(3):405–12.
34. Camm CF, Hayward G, Elias TCN, Bowen JST, Hassanzadeh R, Fanshawe T, et al. Sepsis recognition tools in acute ambulatory care: associations with process of care and clinical outcomes in a service evaluation of an Emergency Multidisciplinary Unit in Oxfordshire. BMJ Open. 2018 Apr 9;8(4):e020497–e020497.
35. Tirotta D, Gambacorta M, La Regina M, Attardo T, Lo Gullo A, Panzone F, et al. Evaluation of the threshold value for the modified early warning score (MEWS) in medical septic patients: a secondary analysis of an Italian multicentric prospective cohort (SNOOPII study). QJM Mon J Assoc Physicians. 2017 Jun 1;110(6):369–73.
36. Pong JZ, Fook-Chong S, Koh ZX, Samsudin MI, Tagami T, Chiew CJ, et al. Combining Heart Rate Variability with Disease Severity Score Variables for Mortality Risk Stratification in Septic Patients Presenting at the Emergency Department. Int J Environ Res Public Health. 2019 May 16;16(10).
37. Prabhakar SM, Tagami T, Liu N, Samsudin MI, Ng JCJ, Koh ZX, et al. Combining quick sequential organ failure assessment score with heart rate variability may improve predictive ability for mortality in septic patients at the emergency department. PLoS One. 2019 Mar 18;14(3):e0213445–e0213445.
38. Martino IF, Figgiaconi V, Seminari E, Muzzi A, Corbella M, Perlini S. The role of qSOFA compared to other prognostic scores in septic patients upon admission to the emergency department. Eur J Intern Med. 2018 Jul;53:e11–3.
39. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J EMJ. 2009 Apr;26(4):254–8.
40. Qin Q, Xia Y, Cao Y. Clinical study of a new Modified Early Warning System scoring system for rapidly evaluating shock in adults. J Crit Care. 2017 Feb;37:50–5.
41. Schmedding M, Adegbite BR, Gould S, Beyeme JO, Adegnika AA, Grobusch MP, et al. A Prospective Comparison of Quick Sequential Organ Failure Assessment, Systemic Inflammatory Response Syndrome Criteria, Universal Vital Assessment, and Modified Early Warning Score to Predict Mortality in Patients with Suspected Infection in Gabon. Am J Trop Med Hyg. 2019 Jan;100(1):202–8.
42. Albur M, Hamilton F, MacGowan AP. Early warning score: a dynamic marker of severity and prognosis in patients with Gram-negative bacteraemia and sepsis. Ann Clin Microbiol Antimicrob. 2016 Apr 12;15:23.
43. Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA, et al. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with

community acquired sepsis in the emergency department. Intern Emerg Med. 2013 Apr;8(3):255–60.
44. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. Medicine (Baltimore). 2019 Feb;98(6):e14197–e14197.
45. Samsudin MI, Liu N, Prabhakar SM, Chong S-L, Kit Lye W, Koh ZX, et al. A novel heart rate variability based risk prediction model for septic patients presenting to the emergency department. Medicine (Baltimore). 2018 Jun;97(22):e10866–e10866.
46. Chang S-H, Hsieh C-H, Weng Y-M, Hsieh M-S, Goh ZNL, Chen H-Y, et al. Performance Assessment of the Mortality in Emergency Department Sepsis Score, Modified Early Warning Score, Rapid Emergency Medicine Score, and Rapid Acute Physiology Score in Predicting Survival Outcomes of Adult Renal Abscess Patients in the Emergency D. Biomed Res Int. 2018 Sep 19;2018:6983568.
47. Geier F, Popp S, Greve Y, Achterberg A, Glöckner E, Ziegler R, et al. Severity illness scoring systems for early identification and prediction of in-hospital mortality in patients with suspected sepsis presenting to the emergency department. Wien Klin Wochenschr. 2013 Sep;125(17–18):508–15.
48. Asiimwe SB, Abdallah A, Ssekitoleko R. A simple prognostic index based on admission vital signs data among patients with sepsis in a resource-limited setting. Crit Care. 2015 Mar 16;19:86.
49. Hung S-K, Ng C-J, Kuo C-F, Goh ZNL, Huang L-H, Li C-H, et al. Comparison of the Mortality in Emergency Department Sepsis Score, Modified Early Warning Score, Rapid Emergency Medicine Score and Rapid Acute Physiology Score for predicting the outcomes of adult splenic abscess patients in the emergency department. PLoS One. 2017 Nov 1;12(11):e0187495–e0187495.
50. Garcea G, Jackson B, Pattenden CJ, Sutton CD, Neal CP, Dennison AR, et al. Early warning scores predict outcome in acute pancreatitis. J Gastrointest Surg Off J Soc Surg Aliment Tract. 2006 Jul;10(7):1008–15.
51. Yoo J-W, Lee JR, Jung YK, Choi SH, Son JS, Kang BJ, et al. A combination of early warning score and lactate to predict intensive care unit transfer of inpatients with severe sepsis/septic shock. Korean J Intern Med. 2015 Jul;30(4):471–7.
52. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and q SOFA scores for predicting mortality and length of stay in ICU. J Crit Care. 2017 Oct;41:191–3.
53. Calvert J, Desautels T, Chettipally U, Barton C, Hoffman J, Jay M, et al. High-performance detection and early prediction of septic shock for alcohol-use disorder patients. Ann Med Surg. 2016 May 10;8:50–5.
54. Awad A, Bader-El-Den M, McNicholas J, Briggs J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. Int J Med Inform. 2017 Dec;108:185–95.
55. Reini K, Fredrikson M, Oscarsson A. The prognostic value of the Modified Early Warning Score in critically ill patients: a prospective, observational study. Eur J Anaesthesiol. 2012 Mar;29(3):152–7.
56. Chen Y-C, Yu W-K, Ko H-K, Pan S-W, Chen Y-W, Ho L-I, et al. Post-intensive care unit respiratory failure in older patients liberated from intensive care unit and ventilator: The predictive value of the National Early Warning Score on intensive care unit discharge. Geriatr Gerontol Int. 2019 Apr;19(4):317–22.
57. Baker T, Blixt J, Lugazia E, Schell CO, Mulungu M, Milton A, et al. Single Deranged Physiologic Parameters Are Associated With Mortality in a Low-Income Country. Crit Care Med. 2015 Oct;43(10):2171–9.

58. Gök RGY, Gök A, Bulut M. Assessing prognosis with modified early warning score, rapid emergency medicine score and worthing physiological scoring system in patients admitted to intensive care unit from emergency department. Int Emerg Nurs. 2019 Mar;43:9–14.
59. Moseson EM, Zhuo H, Chu J, Stein JC, Matthay MA, Kangelaris KN, et al. Intensive care unit scoring systems outperform emergency department scoring systems for mortality prediction in critically ill patients: a prospective cohort study. J Intensive Care. 2014 Jul 1;2:40.
60. Jo S, JB L, YH J, TO J, JC Y, YK J, et al. Modified early warning score with rapid lactate level in critically ill medical patients: the ViEWS-L score. Emerg Med J. 2013 Feb;30(2):123–9.
61. Kwon J-M, Lee Y, Lee Y, Lee S, Park H, Park J. Validation of deep-learning-based triage and acuity score using a large national dataset. PLoS One. 2018 Oct 15;13(10):e0205836–e0205836.
62. Usman OA, Usman AA, Ward MA. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the Emergency Department. Am J Emerg Med. 2018 Nov 7.
63. Jang D-H, Kim J, Jo YH, Lee JH, Hwang JE, Park SM, et al. Developing neural network models for early detection of cardiac arrest in emergency department. Am J Emerg Med. 2019 Apr 7.
64. Wei X, Ma H, Liu R, Zhao Y. Comparing the effectiveness of three scoring systems in predicting adult patient outcomes in the emergency department. Medicine (Baltimore). 2019 Feb;98(5):e14289–e14289.
65. Lee SB, Kim DH, Kim T, Kang C, Lee SH, Jeong JH, et al. Triage in Emergency Department Early Warning Score (TREWS) is predicting in-hospital mortality in the emergency department. Am J Emerg Med. 2019 Feb 17.
66. Singer AJ, Ng J, Thode HC, Spiegel R, Weingart S, Thode HCJ. Quick SOFA Scores Predict Mortality in Adult Emergency Department Patients With and Without Suspected Infection. Ann Emerg Med.. 2017 Apr;69(4):475–9.
67. Eick C, Rizas KD, Meyer-Zürn CS, Groga-Bada P, Hamm W, Kreth F, et al. Autonomic nervous system activity as risk predictor in the medical emergency department: a prospective cohort study. Crit Care Med. 2015 May;43(5):1079–86.
68. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J. 2014 Jun;31(6):476–81.
69. Kivipuro M, Tirkkonen J, Kontula T, Solin J, Kalliomäki J, Pauniaho S-L, et al. National early warning score (NEWS) in a Finnish multidisciplinary emergency department and direct vs. late admission to intensive care. Resuscitation. 2018 Jul;128:164–9.
70. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. BMJ Open. 2019 Jan 17;9(1):e024636–e024636.
71. Ho LO, Li H, Shahidah N, Koh ZX, Sultana P, Hock Ong ME. Poor performance of the modified early warning score for predicting mortality in critically ill patients presenting to an emergency department. World J Emerg Med. 2013;4(4):273–8.
72. Skitch S, Tam B, Xu M, McInnis L, Vu A, Fox-Robichaud A. Examining the utility of the Hamilton early warning scores (HEWS) at triage: Retrospective pilot study in a Canadian emergency department. CJEM Can J Emerg Med. 2018 Mar;20(2):266–74.

73. Liu N, Koh ZX, Goh J, Lin Z, Haaland B, Ting BP, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. BMC Med Informatics Decis Mak. 2014 Jan;14(1):75.
74. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med Off J Eur Soc Emerg Med. 2016 Dec;23(6):406–12.
75. Yuan WC, Tao C, Dan ZD, Yi SC, Jing W, Jian Q. The significance of National Early Warning Score for predicting prognosis and evaluating conditions of patients in resuscitation room. Hong Kong J Emerg Med. 2018;25(6):324–30.
76. Naidoo DK, Rangiah S, Naidoo SS. An evaluation of the Triage Early Warning Score in an urban accident and emergency department in KwaZulu-Natal. South African Fam Pract. 2014 Jan;56(1):69–73.
77. Liu FY, Qin J, Wang RX, Fan XI, Wang J, Sun CY et al. A prospective validation of national early warning score in emergency intensive care unit patients at Beijing. Hong Kong J Emerg Med. 2015;22(3):137–44.
78. So S-N, Ong C-W, Wong L-Y, Chung JYM, Graham CA. Is the Modified Early Warning Score able to enhance clinical observation to detect deteriorating patients earlier in an Accident & Emergency Department? Australas Emerg Nurs J. 2015 Feb;18(1):24–32.
79. Dundar ZD, Kocak S, Girisgin AS. Lactate and NEWS-L are fair predictors of mortality in critically ill geriatric emergency department patients. Am J Emerg Med. 2019 Feb 7.
80. Lam TS, Mak PSK, Siu WS, Lam MY, Cheung TF, Rainer TH. Validation of a Modified Early Warning Score (MEWS) in emergency department observation ward patients. Hong Kong J Emerg Med. 2006;13(1):24–30.
81. Xie X, Huang W, Liu Q, Tan W, Pan L, Wang L, et al. Prognostic value of Modified Early Warning Score generated in a Chinese emergency department: a prospective cohort study. BMJ Open. 2018 Dec 14;8(12):e024120–e024120.
82. Cattermole GN, Mak SKP, Liow CHE, Ho MF, Hung KYG, Keung KM, et al. Derivation of a prognostic score for identifying critically ill patients in an emergency department resuscitation room. Resuscitation. 2009 Sep;80(9):1000–5.
83. Heitz CR, Gaillard JP, Blumstein H, Case D, Messick C, Miller CD. Performance of the maximum modified early warning score to predict the need for higher care utilization among admitted emergency department patients. J Hosp Med. 2010 Jan;5(1):E46-52.
84. Srivilaithon W, Amnuaypattanapon K, Limjindaporn C, Imsuwan I, Daorattanachai K, Dasanadeba I, et al. Predictors of in-hospital cardiac arrest within 24 h after emergency department triage: A case-control study in urban Thailand. Emerg Med Australas EMA. 2019 Mar 18.
85. Cattermole GN, Liow ECH, Graham CA, Rainer TH. THERM: the Resuscitation Management score. A prognostic tool to identify critically ill patients in the emergency department. Emerg Med J. 2014 Oct;31(10):803–7.
86. Najafi Z, Zakeri H, Mirhaghi A. The accuracy of acuity scoring tools to predict 24-h mortality in traumatic brain injury patients: A guide to triage criteria. Int Emerg Nurs. 2018 Jan;36:27–33.
87. Bartkowiak B, Snyder AM, Benjamin A, Schneider A, Twu NM, Churpek MM, et al. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–63.

88. Kovacs C. Outreach and early warning systems for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. Int J Nurs Pract. 2016;
89. Plate JD, Peelen LM, Leenen LP, Hietbrink F. Validation of the VitalPAC Early Warning Score at the Intermediate Care Unit. World J Crit Care Med. 2018 Aug 4;7(3):39–45.
90. Sarani B. Accuracy of an expanded early warning score for patients in general and trauma surgery wards (Br J Surg 2012; 99: 192-197). Br J Surg. 2012 Feb;99(2):197–8.
91. Hollis RH, Graham LA, Lazenby JP, Brown DM, Taylor BB, Heslin MJ, et al. A Role for the Early Warning Score in Early Identification of Critical Postoperative Complications. Ann Surg. 2016 May;263(5):918–23.
92. Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Ann R Coll Surg Engl. 2006 Oct;88(6):571–5.
93. Garcea G, Ganga R, Neal CP, Ong SL, Dennison AR, Berry DP. Preoperative early warning scores can predict in-hospital mortality and critical care admission following emergency surgery. J Surg Res. 2010 Apr;159(2):729–34.
94. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007 Feb;35(2):402–9.
95. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation. 2010 Aug;81(8):932–7.
96. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013 Apr;84(4):465–70.
97. Rasmussen LJH, Ladelund S, Haupt TH, Ellekilde GE, Eugen-Olsen J, Andersen O. Combining National Early Warning Score With Soluble Urokinase Plasminogen Activator Receptor (suPAR) Improves Risk Prediction in Acute Medical Patients: A Registry-Based Cohort Study. Crit Care Med. 2018 Dec;46(12):1961–8.
98. Ghosh E, Eshelman L, Yang L, Carlson E, Lord B. Early Deterioration Indicator: Data-driven approach to detecting deterioration in general ward. Resuscitation. 2018 Jan;122:99–105. =29122648&site=ehost-live&scope=site
99. Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. BJA Br J Anaesth. 2007 May 22;98(6):769–74.
100. Colombo F, Taurino L, Colombo G, Amato M, Rizzo S, Murolo M et al. The Niguarda MEWS, a new and refined tool to determine criticality and instability in Internal Medicine Ward and Emergency Medicine Unit. Ital J Med. 2017;11(3):310–7.
101. Abbott TEF, Torrance HDT, Cron N, Vaid N, Emmanuel J. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med. 2016 Nov;35:78–82.
102. Wheeler I, Price C, Sitch A, Banda P, Kellett J, Nyirenda M, et al. Early warning scores generated in developed healthcare settings are not sufficient at predicting early mortality in Blantyre, Malawi: a prospective cohort study. PLoS One. 2013;8(3):e59830–e59830.
103. Graziadio S, O'Leary RA, Stocken DD, Power M, Allen AJ, Simpson AJ, et al. Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS

in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study. BMJ Open. 2019 Feb 22;8(11):e020337–e020337.
104. Subbe CP, Davies RG, Williams E, Rutherford P, Gemmell L. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. Anaesthesia. 2003 Aug;58(8):797–802.
105. Dawes TR, Cheek E, Bewick V, Dennis M, Duckitt RW, Walker J, et al. Introduction of an electronic physiological early warning system: effects on mortality and length of stay. Br J Anaesth. 2014 Oct;113(4):603–9.
106. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted to a tertiary referral intensive care unit after CPR. Resuscitation. 2011.
107. Sutherasan Y, Theerawit P, Suporn A, Nongnuch A, Phanachet P, Kositchaiwat C. The impact of introducing the early warning scoring system and protocol on clinical outcomes in tertiary referral university hospital. Ther Clin Risk Manag. 2018 Oct;14:2089–95.
108. Heller AR, Mees ST, Lauterwald B, Reeps C, Koch T, Weitz J. Detection of Deteriorating Patients on Surgical Wards Outside the ICU by an Automated MEWS-Based Early Warning System With Paging Functionality. Ann Surg. 2018 May 16;

Table S1. Risk of bias assessment results

TOOL	Study	Validation	Quality	
			Risk of bias	Applicability
PROBAST	Kellett, 2012 (S1)	External	low	low
	Kim, 2017 (S2)	External	Unclear	unclear
	Bozkurt, 2015 (S3)	External	High	high
	Seak, 2017 (S4)	External	High	high
	Hu, 2016 (S5)	Internal	Unclear	high
	Liljehult, 2016 (S6)	External	Unclear	high
	Mulligan, 2010 (S7)	External	High	high
	Cooksley, 2012 (S8)	External	Unclear	unclear
	Vaughn, 2018 (S9)	External	High	high
	Young, 2014 (S10)	External	High	high
	von Lilienfeld-Toal, 2007 (S11)	External	Unclear	high
	Pedersen, 2018 (S12)	External and Internal	low	low
	Forster, 2018 (S13)	External	low	low
	Pimentel, 2018 (S14)	External	low	unclear
	Sbiti-rohr, 2016 (S15)		Unclear	high
	Brabrand, 2017 (S16)	External	Unclear	unclear
	Jo, 2016 (S17)	External	High	high
	Barlow, 2007 (S18)	External	low	unclear
	Bilben, 2016 (S19)	External	Unclear	unclear
	Delahanty, 2019 (S20)	Internal	low	low
	Redfern, 2018 (S21)	External	low	low
	Churpek, 2017 (S22)	External	High	high
	Faisal, 2019 (S23)	External	low	low
	Churpek 2017 (S24)	External	low	low
	Henry, 2015 (S25)	Internal	low	low
	Brink 2019 (S26)	External	Unclear	unclear
	de Groot, 2014 (S27)	External	Unclear	unclear
	Corfield, 2014 (S28)	External	low	low
	Goulden, 2018 (S29)	External	Unclear	unclear
	Khwannimit, 2019 (S30)	External	Unclear	unclear
	Ghanem-Zoubi, 2011 (S31)	External	Unclear	unclear
	Saeed, 2019 (S32)	Internal	Unclear	unclear
	Innocenti, 2018 (S33)	External	Unclear	unclear
	Camm, 2018 (S34)	External	Unclear	unclear
	Tirotta, 2017 (S35)	External	Unclear	unclear
	Pong, 2019 (S36)	Internal	Unclear	unclear
	Prabhakar, 2019 (S37)	Internal	Unclear	unclear
	Martino, 2018 (S38)	External	Unclear	unclear
	Vorwerk, 2009 (S39)	External	Unclear	unclear
	Qin, 2017 (S40)	External	Unclear	unclear
	Schmedding, 2019 (S41)	External	Unclear	unclear
	Albur, 2016 (S42)	External	Unclear	unclear
	Cildir, 2013 (S43)	External	Unclear	unclear
	Chiew, 2019 (S44)	External	Unclear	unclear

	Samsudin, 2018 (S45)	Internal	Unclear	unclear
	Chang, 2018 (S46)	External	Unclear	high
	Geier, 2013 (S47)	External	Unclear	unclear
	Asiimwe, 2015 (S48)	Internal	Unclear	unclear
	Hung, 2017 (S49)	External	Unclear	high
	Garcea, 2006 (S50)	External	Unclear	high
	Yoo, 2015 (S51)	External	Unclear	unclear
	Siddiqui, 2017 (S52)	External	Unclear	unclear
	Calvert, 2016 (S53)	Internal	low	unclear
	Awad, 2017 (S54)	Internal	low	low
	Reini, 2012 (S55)	External	Unclear	unclear
	Chen, 2019 (S56)	External	Unclear	high
	Baker, 2015 (S57)	External	Unclear	unclear
	Gök, 2019 (S58)	External	low	unclear
	Moseson, 2014 (S59)	External	Unclear	unclear
	Jo, 2013 (S60)	External	Unclear	unclear
	Kwon, 2018 (S61)	External and Internal	Unclear	unclear
	Usman, 2019 (S62)	External	High	high
	Jang, 2019 (S63)	Internal	low	low
	Wei, 2019 (S64)	External	High	high
	Lee, 2019 (S65)	Internal	low	low
	Singer, 2017 (S66)	External	Unclear	unclear
	Eick, 2015 (S67)	External	Unclear	unclear
	Bulut, 2014 (S68)	External	Unclear	unclear
	Kivipuro, 2018 (S69)	External	Unclear	unclear
	Eckart, 2019 (S70)	External	Unclear	unclear
	Ho, 2013 (S71)	External	Unclear	unclear
	Skitch, 2018 (S72)	External	Unclear	unclear
	Liu, 2014 (S73)	Internal	low	unclear
	Dundar, 2016 (S74)	External	Unclear	high
	Yuan, 2018 (S75)	External	Unclear	high
	Naidoo, 2014 (S76)	External	Unclear	unclear
	Liu, 2015 (S77)	External	low	unclear
	So, 2015 (S78)	External	Unclear	unclear
	Dundar, 2019 (S79)	External	Unclear	high
	Lam, 2006 (S80)	External	Unclear	unclear
	Xie, 2018 (S81)	External	Unclear	unclear
	Cattermole, 2009 (S82)	Internal	Unclear	unclear
	Heitz, 2010 (S83)	External	High	unclear
	Srivilaithon, 2019 (S84)	Internal	Unclear	unclear
	Cattermole, 2014 (S85)	External	Unclear	unclear
	Najafi, 2018 (S86)	External	Unclear	high
	Bartkowiak, 2019 (S87)	External	Unclear	unclear
	Kovacs, 2016 (S88)	External	low	low
	Plate, 2018 (S89)	External	low	low
	Sarani, 2012 (S90)	External	low	low
	Hollis, 2016 (S91)	External	Unclear	unclear

	Gardner-Thorpe 2006 (S92)	External	Unclear	unclear
	Garcea, 2010 (S93)	External	High	high
	Cuthbertson, 2007 (S94)	External	High	unclear
	Prytherch, 2010 (S95)	Internal	low	low
	Smith, 2013 (S96)	External	low	low
	Ramsussen, 2018 (S97)	External	Unclear	unclear
	Ghosh, 2018 (S98)	Internal	low	low
	Duckitt, 2007 (S99)	Internal	low	low
	Colombo, 2017 (S100)	External	High	high
	Abbot, 2016 (S101)	External	High	high
	Wheeler, 2013 (S102)	External	Unclear	unclear
	Graziadio, 2019 (S103)	External	Unclear	unclear
ROBINS-I		Overall bias assessment		
	Subbe, 2003 (S104)	Moderate		
	Dawes, 2014 (S105)	Low		
	Moon, 2011 (S106)	Low		
	Sutherasan, 2018 (S107)	Moderate		
	Heller, 2018 (S108)	Low		

Total	108 studies

Table S2. Early warning scores used in studies of patients' sub-populations and settings

		HR	SBP	RR	Temp	APVU/ LOC	O2 Sat	Supp O2	Urine OP	Other
Kellett, 2012 (S1)	VIEWS	✓	✓	✓	✓	✗	✓	✓	✗	✗
Seak, 2017 (S4)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bozkurt, 2015 (S3)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kim, 2017 (S2)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hu, 2016 (S5)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Mulligan, 2010 (S7)	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liljehult, 2016 (S6)	EWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Cooksley, 2012 (S8)	MEWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Cooksley, 2012 (S8)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Vaughn, 2018 (S9)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Von Lilienfeld-Toal, 2007 (S11)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Young, 2014 (S10)	MEWS	✓	✓	✓	✓	✗	✗	✗	✗	✓
Barlow, 2007 (S18)	EWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Bilben, 2016 (S19)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Brabrand, 2017 (S16)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Forster, 2018 (S13)	NEWS	✓	✓	✓	✓	✓	✓	✓	✓	✗
Jo, 2016 (S16)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pedersen, 2018 (S12)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018 (S14)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018 (S14)	NEWS2	✓	✓	✓	✓	✓	✓	✓	✗	✓
Sbiti-rohr, 2016 (S15)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Henry, 2015 (S25)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Innocenti, 2018 (S33)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Garcea, 2006 (S50)	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Qin, 2017 (S40)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Albur, 2016 (S42)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Asiimwe, 2015 (S48)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Brink 2019 (S26)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Camm, 2018 (S34)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Chang, 2018 (S46)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chiew, 2019 (S44)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chiew, 2019 (S44)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S22)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S22)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Churpek, 2017 (S24)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S24)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cildir, 2013 (S43)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Corfield, 2014 (S28)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014 (S27)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014 (S27)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Delahanty, 2019 (S20)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Delahanty, 2019 (S20)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Faisal, 2019 (S23)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Geier, 2013 (S47)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ghanem-Zoubi, 2011 (S31)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Goulden, 2018 (S29)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hung, 2017 (S49)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Khwannimit, 2019 (S30)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Khwannimit, 2019 (S30)	SOS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Khwannimit, 2019 (S30)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Martino, 2018 (S30)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Pong, 2019 (S36)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pong, 2019 (S36)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019 (S37)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019 (S37)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Redfern, 2018 (S21)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Saeed, 2019 (S32)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Samsudin, 2018 (S45)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Samsudin, 2018 (S45)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Schmedding, 2019 (S41)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Siddiqui, 2017 (S52)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Tirotta, 2017 (S35)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Vorwerk, 2009 (S39)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Yoo, 2015 (S51)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Awad, 2017 (S54)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Baker, 2015 (S57)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Calvert 2016 (S53)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Gök, 2019 (S58)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chen, 2019 (S56)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Jo, 2013 (S60)	HOTEL	✗	✓	✗	✓	✓	✓	✗	✗	✓
Jo, 2013 (S60)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Moseson, 2014 (S59)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Reini, 2012 (S55)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bulut, 2014 (S68)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2009 (S82)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2014 (S85)	WORTHING	✓	✓	✓	✓	✓	✓	✗	✗	✗
Cattermole, 2014 (S85)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗

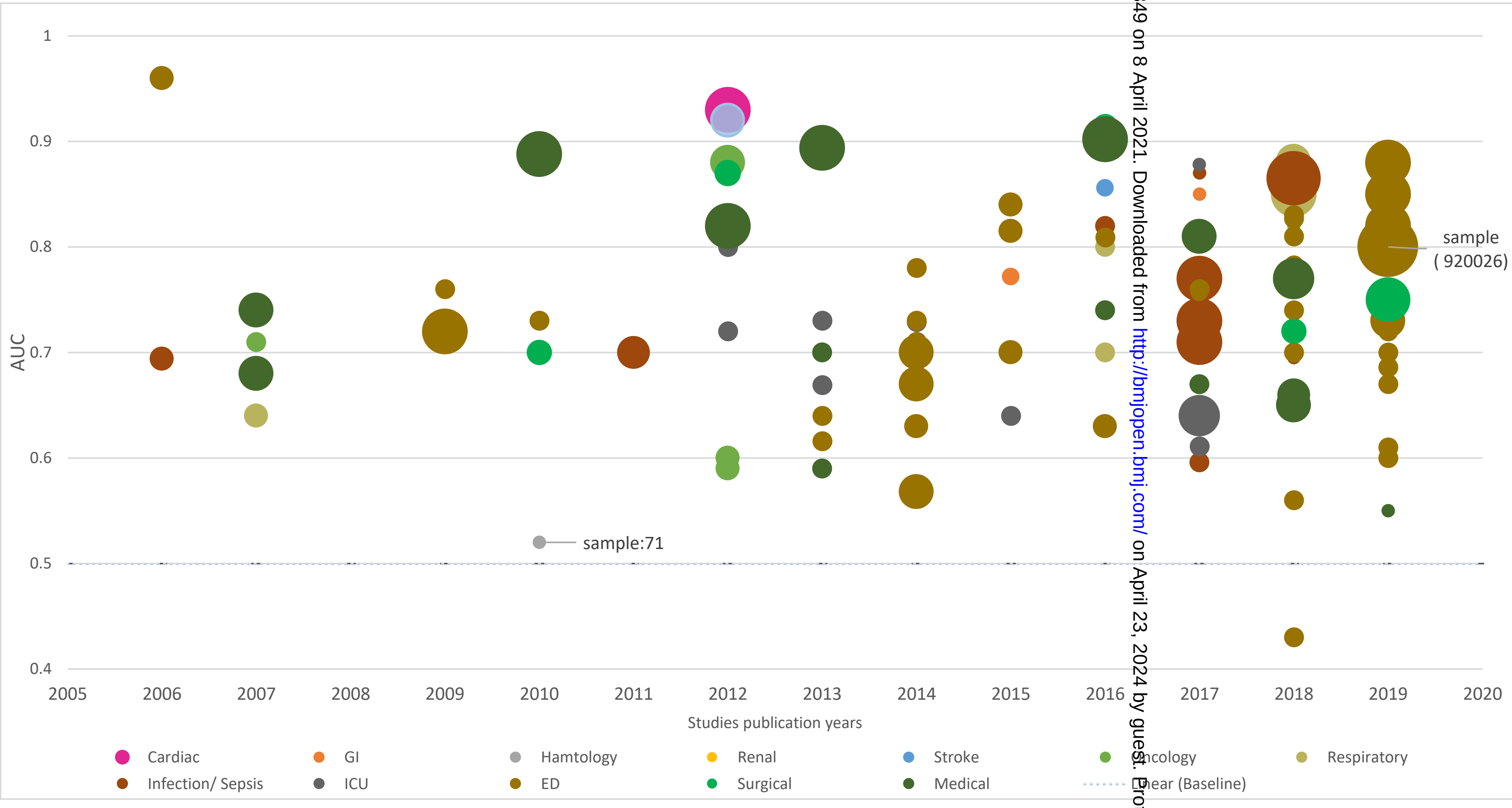
Cattermole, 2014 (S85)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Heitz, 2010 (S83)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016 (S74)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016 (S74)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Dundar, 2019 (S79)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eckart, 2019 (S70)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eick, 2015 (S67)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu, 2015 (S77)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Liu, 2015 (S77)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ho, 2013 (S71)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Jang, 2019 (S63)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kivipuro, 2018 (S69)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Kwon, 2018 (S61)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu, 2014 (S73)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019 (S65)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019 (S65)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lee, 2019 (S65)	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Naidoo, 2014 (S76)	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Najafi, 2018 (S86)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Singer, 2017 (S66)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Skitch, 2018 (S72)	HEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Skitch, 2018 (S72)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
So, 2015 (S78)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Srivilaithon, 2019 (S84)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lam, 2006 (S80)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Usman, 2019 (S62)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018 (S75)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018 (S75)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Wei, 2019 (S64)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Xie, 2018 (S81)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bartkowiak, 2019 (S87)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Bartkowiak, 2019 (S87)	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Cuthbertson, 2007 (S94)	EWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Cuthbertson, 2007 (S94)	MEWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Garcea, 2010 (S50)	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Gardner-Thorpe 2006 (S92)	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Hollis, 2016 (S91)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Kovacs, 2016 (S88)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Plate, 2018 (S89)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Sarani, 2012 (S90)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Abbott, 2016 (S101)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Duckitt, 2007 (S99)	WPC	✓	✓	✓	✓	✓	✓	✗	✗	✗
Duckitt, 2007 (S99)	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Colombo, 2017 (S100)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ghosh, 2018 (S98)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Ghosh, 2018 (S98)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Graziadio, 2019 (S103)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Prytherch, 2010 (S95)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Ramsussen, 2018 (S97)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Smith, 2013 (S96)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Wheeler, 2013 (S102)	Hotel	✓	✗	✓	✗	✓	✓	✗	✗	✓
Wheeler, 2013 (S102)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Total	133

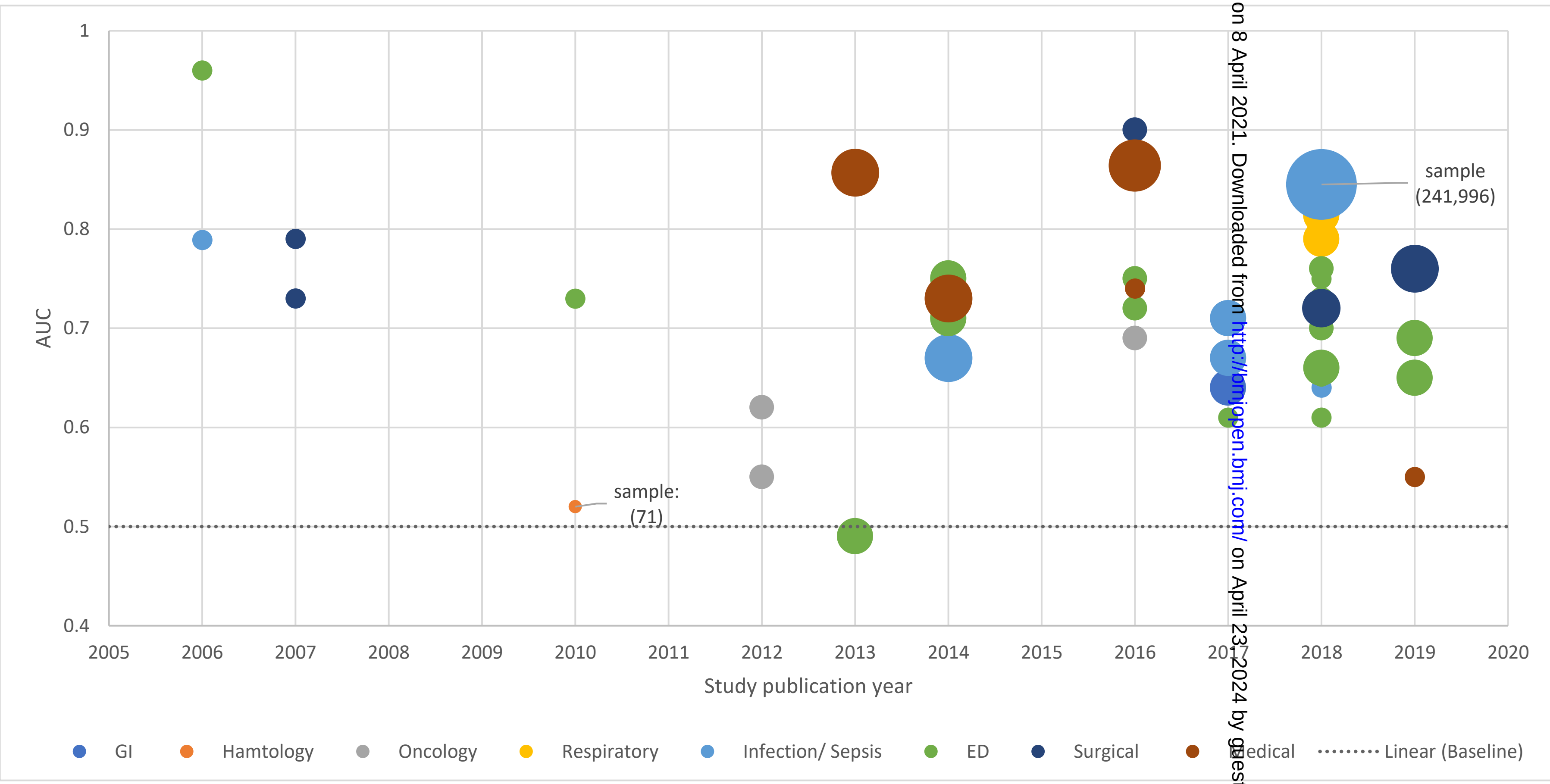
Abbreviations: HR: heart rate, SBP: systolic blood pressure, RR: respiratory rate, Temp: temperature, AVPU/LOC: alert, verbal response, physical response, unresponsive score or level of consciousness, O2 sat: Oxygen saturation, Supp O2: supplemental oxygen, Urine OP: urine output, Other: other parameters, i.e., blood biomarkers. VIEWS: Vitalpack early warning score, MEWS: modified early warning score, EWS: early warning score, NEWS: national early warning score, NEWS2: national early warning score 2, SOS: Search Out Severity score, Worthing: Worthing physiological scoring system, HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score, TREWS: Triage in Emergency department Early Warning Score, HEWS: Hamilton early warning score.

Figure S1. Predictive performance of early warning scores for mortality in studies from 2005 to 2020 for different disease subgroups and clinical settings



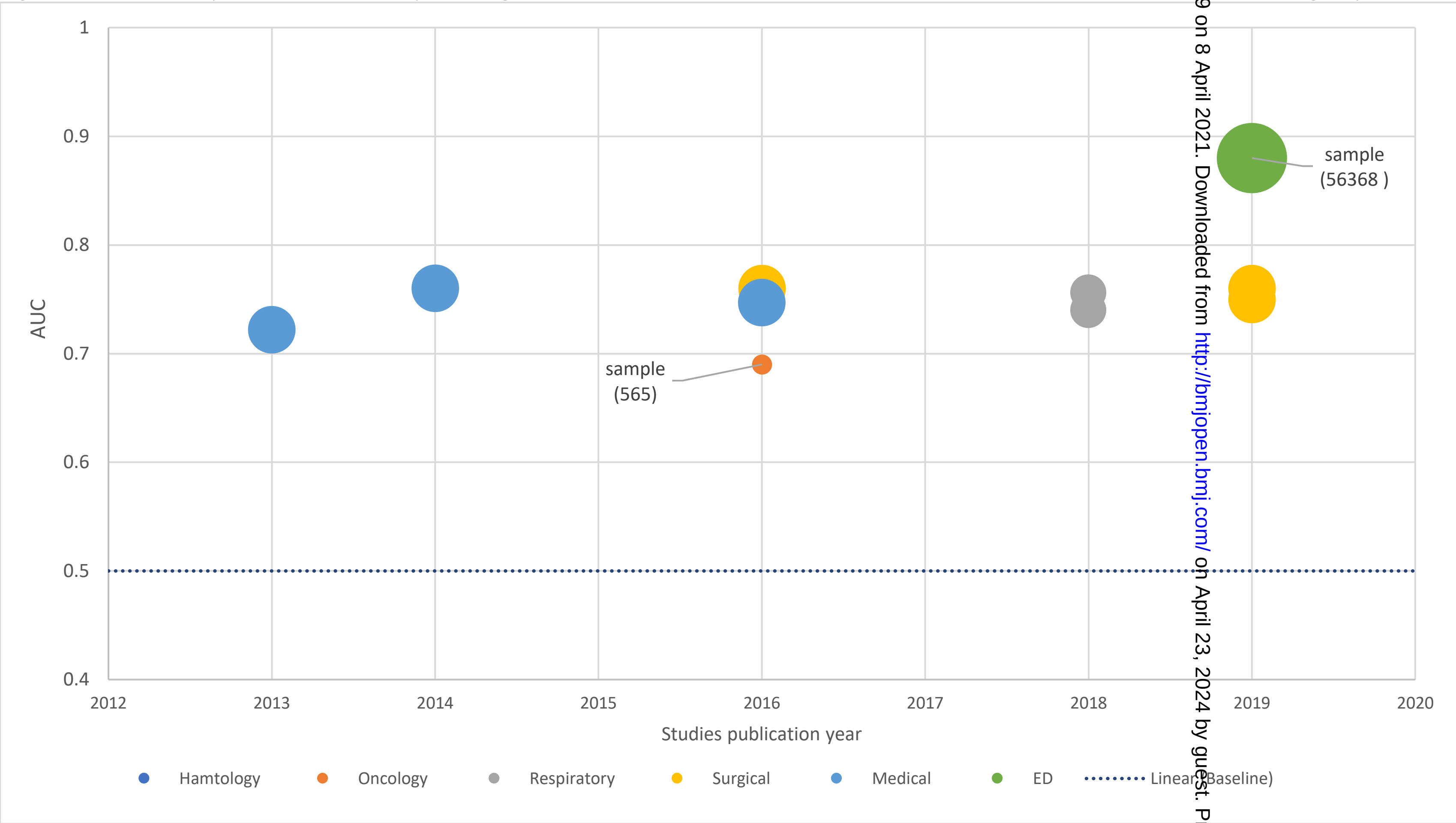
Abbreviations: AUC: Area Under the Curve; ED: Emergency Department; GI: Gastro Intestinal diseases; ICU: Intensive Care Unit. Note: Bubbles sizes represents the sample size in each study.

Figure S2. Predictive performance of early warning scores for intensive care admission in studies from 2005 to 2020 for different disease subgroups and clinical settings



Abbreviations: AUC: Area Under the Curve; ED: Emergency Department; GI: Gastro Intestinal diseases. Note: Bubbles sizes represents the sample size in each study.

Figure S3. Predictive performance of early warning scores for cardiac arrest in studies from 2005 to 2020 for different disease subgroups and clinical settings



Abbreviations: AUC: Area Under the Curve; ED: Emergency Department; GI: Gastro Intestinal diseases. Note: Bubbles sizes represents the sample size in each study.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	1
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4; Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	4
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	4
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5; figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	5
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	5-6
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	5; Table1; Table 2; Supplementary.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	5-6
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	5-6
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Supplementary
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	7
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	7-8
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	8
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	8

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

The performance of universal early warning scores in different patient subgroups and clinical settings: A systematic review

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045849.R2
Article Type:	Original research
Date Submitted by the Author:	01-Mar-2021
Complete List of Authors:	Alhmoud, Baneen; University College London Bonnici, Tim; University College London Patel, Riyaz; UCL, Farr Institute Melley, Daniel; Barts Health NHS Trust Williams, Bryan; University College London, Institute of Cardiovascular Science; Banerjee, Amitava; University College London, Farr Institute of Health Informatics Research
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Epidemiology, Health informatics, Intensive care, Medical management, Patient-centred medicine
Keywords:	Adult intensive & critical care < ANAESTHETICS, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, EPIDEMIOLOGY

SCHOLARONE™
Manuscripts

The performance of universal early warning scores in different patient subgroups and clinical settings: A systematic review

Baneen Alhmoud[1,2] *research fellow/PhD student*
Timothy Bonnici[1,2] *consultant in intensive care medicine*
Riyaz Patel[1,2,3] *professor of cardiology and honorary consultant cardiologist*
Daniel Melley[3] *consultant in intensive care medicine and honorary senior lecturer*
Bryan Williams[1,2] *professor of medicine and consultant physician*
Amitava Banerjee[1,2,3] *associate professor in clinical data science and honorary consultant cardiologist*

Authors' Address:
[1]University College London, London
[2]University College London Hospitals NHS Trust, London
[3]Barts Health NHS Trust, London

Corresponding author: Dr Amitava Banerjee
Institute of Health Informatics, 222 Euston Road, London. NW1 2DA
ami.banerjee@ucl.ac.uk

Keywords: prediction, early warning score, prognosis, disease, clinical setting, systematic review
Abstract word count: 245
Total word count: 2713

Abstract

Objective:
To assess predictive performance of universal early warning scores (EWS) in disease subgroups and clinical settings.
Design:
Systematic review.
Data sources:
Medline, CINAHL, EMBASE and Cochrane database of systematic reviews from 1997 to 2019.
Inclusion criteria:
Randomised trials and observational studies of internal or external validation of EWS to predict deterioration (mortality, ICU transfer and cardiac arrest) in disease subgroups or clinical settings.
Results:
We identified 770 studies, of which 103 were included. Study designs and methods were inconsistent, with significant risk of bias (high: *n*=16 and unclear: *n*=64 and low risk: *n*=28). There were only two randomised trials. There was a high degree of heterogeneity in all subgroups and in NEWS (I^2=72-99%). Predictive accuracy (mean AUC; 95% CI) was highest in medical (0.74; 0.74–0.75) and surgical (0.77; 0.75–0.80) settings and respiratory diseases (0.77; 0.75–0.80). Few studies evaluated EWS in specific diseases, e.g. cardiology (*n* = 1), respiratory (*n* = 7). Mortality and ICU transfer were most frequently studied outcomes, and cardiac arrest was least examined (*n*=8). Integration with electronic health records was uncommon (n=9).
Conclusion:
Methodology and quality of validation studies of EWS are insufficient to recommend their use in all diseases and all clinical settings despite good performance of EWS in some subgroups. There is urgent need for consistency in methods and study design, following consensus guidelines for predictive risk scores. Further research should consider specific diseases and settings, utilising electronic health record data, prior to large-scale implementation.

Systematic review registration:
PROSPERO CRD42019143141

Strengths and limitations

- The first systematic review to investigate the performance of early warning scores in different patient disease subgroups and clinical settings.
- Meta-analysis was performed for different EWS and NEWS validation studies in different disease and clinical setting subgroups
- This study is limited to specific diseases and settings and does not consider the use of early warning scores in the general population.
- Analysis of predictive accuracy of early warning scores is based on area-under-the curve, not other validation measures.
- During the study period 1997-2019, approaches to early warning scores and their validation have changed.

Introduction

Across diseases, patient deterioration can range from critical care review and sepsis, to cardiorespiratory arrest and death(1,2). Delays or failures in timely detection of deterioration adversely affect prognosis, morbidity, mortality, and healthcare utilisation(3). For example, the 20,000 in-hospital cardiac arrests per year in England are associated with costs of £50 million for resuscitation and post-arrest care(4). Around the world, earlier recognition and prevention of deterioration in unwell patients has far-reaching implications for reduction in mortality and morbidity, reduction in the cost of healthcare, and allocation of scarce high dependency and critical care resources. Preventive interventions are needed to overcome these challenges (5).

Specific characteristics have long been known to be associated with deteriorating patient health(2, 5–8), including physiological parameters, such as heart rate and blood pressure(5, 9–12). Early warning scores (EWS), widely used in high-income countries, were borne out of the need for early detection of patient deterioration. EWS are tools derived from prediction models that assess patient characteristics and physiological parameters to stratify the risk of developing a worsening event or need for medical attention(13). The algorithms underlying EWS can be "aggregate-weighted" to sum up a set of parameters to produce a score, or use more advanced statistical modelling(14). EWS inform clinical decision-making, enabling escalation of attention and care when required. Universal tools, such as the modified early warning score (MEWS)(15) were developed for use across different hospital settings, but specialised, non-standard EWS are also designed for particular subgroups, e.g. Rapid Emergency Medicine Score (REMS)(16) and Quick Sequential Organ Failure Assessment (qSOFA) (17) for patients with infections. In recognising different settings, EWS may have compromised simplicity and timeliness of assessment(13). For example, a number of EWS rely on parameters that do not exist in the first hours of assessment, such as blood investigations and imaging(1,18,19).

From fragmented implementation and inadequate early assessment via specialised tools, EWS have shifted back to universal prediction models, particularly, the national early warning score(NEWS)(20), followed by NEWS2(21). NEWS was designed to produce a universal assessment of acute illness severity across the NHS(22). While showing good discrimination compared with other EWS, especially in predicting mortality, there was a need to accommodate additional clinical parameters in the score. The updated NEWS2, emphasising appropriate scoring for type 2 respiratory failure, confusion and severe sepsis(21), was formally endorsed by NHS England(23) to be the EWS used in acute care. However, there have been concerns regarding excessive calls to clinicians, administrative workload, and variable symptoms across diseases and settings(24). The effectiveness of the universal EWS(**Box 1**) with standardised use across all settings is not clear in specific disease populations (25), and requires validation to estimate discrimination and calibration, like other clinical prediction models(26). While internal validation is useful, generalisability and reproducibility needs external validation(27).

Systematic reviews have evaluated EWS in pre-hospital, intensive care unit (ICU) and general settings (3,28,29), and sepsis(15), with narrow inclusion criteria and inadequate assessment of study quality. A recent systematic review evaluated development and validation of EWS in general patients, but did not include studies in specific disease subgroups or settings(30).

Objective

In a systematic review, we will assess performance of universal EWS in particular diseases and clinical settings in predicting mortality, transfer to ICU and cardiac arrest.

Methods

Search strategy

The protocol adhered to PRISMA-P guidelines (31). Published articles were identified in MEDLINE, CINHAL and EMBASE, between 1997 (initial development of EWS) and 2019. The Cochrane database was searched for systematic reviews (CDSR) and trials (CENTRAL). For grey literature, Google Scholar was searched. During the screening procedure, studies were added from references in review articles and studies. Search strategies were developed by two authors (BA and AB) and reviewed by a third author (TB). Terms used for searching databases include terms for early warning or track and trigger scores and acronyms, identified subgroups and settings (e.g., MeSH) and free-text search terms (**Figure 1**; **Supplementary methods**).

Inclusion and exclusion criteria

Patient subgroups were identified according to disease categories and clinical settings (**Supplementary methods**). *Studies were included if:* (1) validation of a universal EWS with standardised prediction model in adult patients; (2) EWS validation was in a specific setting or disease; (3) the performance of the EWS, or the impact on mortality, transfer to ITU and cardiac arrest, was examined; and (4) they were prospective or retrospective cohort, cross-sectional, case-control design or trials.

Studies were excluded if: (1) patients were less than 16 years of age; (2) EWS performance was only examined in derivation, not validation; (3) non-universal EWS was developed for a specific subgroup, e.g. Obstetric early warning score (OEWS) for obstetric patients or qSOFA for patients with infections; or (4) EWS validation was performed in a general patient dataset or setting, e.g. validation in a general hospital without consideration of hospital subgroups.

Data extraction

Articles were screened by title and abstract by one author (BA), then full-text screening was by two reviewers (BA and AB). Data was extracted independently by two reviewers (BA and AB) using a standardised and piloted data form. A third reviewer (TB) resolved any disagreements. Items for extraction for studies examining predictive accuracy were based on the CHARMS (32) checklist, except for tool derivation which was excluded.

Quality assessment

Risk of biases in validation studies was assessed using PROBAST(33), which classifies studies as low, unclear, or high risk of bias in four aspects: participant selection, predictors, outcomes and analysis within the overall risk of bias and the study applicability domains.

Evidence synthesis:

We conducted the analysis using MS Excel and R programmes. We summarised the results using descriptive statistics and graphical plots. Meta-analysis was performed, in different subgroups, using AUC (Area Under the Curve) for identified Universal EWS and for NEWS in studies. Fisher-Z transformation for correlation coefficients was conducted for AUC into normally distributed Z with 95% CI to evaluate the effect size and test for the heterogeneity. Where applicable, narrative synthesis was conducted.

Patient and public involvement:

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research

Box 1. Definitions:
Universal EWS: EWS that are globally adopted and applicable in every setting and for any disease sub-group. Standardised EWS: EWS model with a set of parameters used in a unified approach to predict deterioration in any patient subgroup (8,23) External validation: evaluation of the model's predictive accuracy with data different than the one sued for model development. (27) Internal validation: evaluation of a model's predictive accuracy with the same data set used for the development or in a population in which the model is intended for use.(27) Discrimination: the ability of a model to distinguish between the patients who will develop an outcome of interest and the ones who will not(26) Calibration: The accuracy of risk estimates in relation to the observed number of events (34)

Results

Study characteristics

Of the 16,181 articles identified by our search, we screened 1,355 articles by title and abstract, assessing 770 articles in full for eligibility. We included 103 studies, published between 2006 and 2019, in the final stage. These studies were predominantly observational (retrospective= 65, prospective= 36 and RCT=2). Emergency department (ED) (*n*=48) was the most common clinical setting, followed by medical (*n* = 12), ICU (*n* = 12), and surgical (*n*=9) settings. Sepsis (*n*=33) was the commonest disease subgroup. Other subgroups ranged from respiratory (n=8) to renal (n=1)(**Figures 1 and 2**). Mortality was the main studied outcome. Cardiac arrest was infrequently studied (*n*=8).

Quality assessment

There was a significant risk of bias found in majority of studies (high risk=16; unclear risk=64), and low risk in only 28 studies. In terms of applicability, narrow inclusion of conditions in a certain disease group was commonly related to risk of bias, while in general settings, biases were often due to low sample size or unspecified timing of EWS assessment. There was a wide variation in sample size (median: 551 and range: 43 - 920029). There was variation in defining study population by number of patients, hospital admissions or not specifying the particular study sample. Almost half of the studies (*n*=49; 48%) validated in <500 patients with either multiple observations or a single observation set (**Tables 1 and 2**). External validation was more common (*n* = 83) than internal validation (*n* = 18) and two studies included internal and external validation(**Table S1**).

EWS validation in patient subgroups

- Subgroups and EWS

In the studies validating EWS, there was heterogeneity in subgroup definitions, models, and methods of predictive accuracy. There was overlap between diseases and settings commonly between studies of patients with infections receiving care in ED(35–37) and patients with sepsis admitted to ICU (38,39). EWS models that were integrated with electronic health records (EHR) were examined in recent studies (*n* = 9). Research on datasets utilising EWS-embedded EHRs had larger sample sizes, ranging from 504(40) to 13,014 patients(41)(**Tables 1 and 2**), with moderate to high predictive ability(AUC: 0.65–0.85). Several studies included comparison between different EWS in the same cohort(n=21)(36,39,42)(**Table S2**).

- Methodology

There was significant heterogeneity in methods across studies. The majority of studies were observational. Evaluation of predictive accuracy of different EWS in the same study was common(22,43–45). To measure accuracy of EWS, AUC was most commonly used(*n*=94), especially when comparing different EWS in the same study(22,46). Presentation of results was variable; for example, confidence intervals were missing in many studies. Other measures, such as analysing sensitivity and specificity, prognostic index and odds ratios, were found in only eight studies (**Tables 1 and 2**). Consequently, it was only feasible to analyse predictive accuracy in studies where AUC was the selected measure.

Timing from EWS assessment to endpoints was variable. Many studies included (*n* = 43) AUC within 24 to 48 hours, while 11 studies had endpoints more than 48 hours after EWS. However, the majority (*n*=65; 63%) did not specify time horizon or in-hospital outcome.

- Predictive performance of EWS

Outcomes were most commonly mortality, transfer to ICU, developing sepsis (in patients with infections), and cardiac arrest. Few studies examined other outcomes, e.g. respiratory arrest (*n* = 1) and organ failure (*n* = 4). Mortality, ICU admission and cardiac arrest were best predicted in medical (AUC mean: 0.74, 0.75 and 0.74)(47–49) and surgical settings (0.80, 0.79 and 0.75)(50,51), and respiratory diseases (0.75, 0.80 and 0.75) respectively. EWS prediction of sepsis had reasonable predictive performance in all subgroups (AUC: 0.71–0.79), and infectious diseases in particular (AUC: 0.79). Certain outcomes related to specific disease groups were not studied, e.g. cardiac arrest was not studied in cardiac patients(22); respiratory arrest was not tested in respiratory patients(47-53).

The best predictive performance was found in studies examining cardiac(47), stroke(47,54) and renal(47) diseases (AUC: 0.93, 0.88 and 0.87 respectively). In emergency settings, predictive accuracy was variable (AUC: 0.56–0.91)(55–59). In haematology and oncology diseases, EWS predictive accuracy was suboptimal in mortality(**Figure S1**), cardiac arrest and ICU transfer (AUC: 0.52-0.69; **Figures 3 and 4**)(60–62). EWS prediction of ICU transfer was reasonable in ED(58,63), infectious diseases (64,65), and where both groups overlap(43,66), but not in gastroenterology and haematology(AUC: 0.64 and 0.60) (61,67)(**Figure S2**). Cardiac arrest was the least examined outcome among the three endpoints (*n*=8) and unstudied in cardiac diseases. (**Figures 3, 4 and S3**)

For mortality prediction, meta-analysis of included EWS showed high degree of statistical heterogeneity across all subgroups (I^2 = 72% -99%)(**Figure 5**). In validation studies of NEWS in different disease subgroups, there was also significant heterogeneity (I^2= 99%; **Figure 6**).

Discussion

In this comprehensive review of Universal EWS across all diseases and settings, we had three main findings. First, EWS studies in different diseases and clinical settings were heterogeneous in methodology, predictive performance measures, and number of studies in each subgroup. Second, validation of EWS is limited in specialised settings, including cardiac disease. Third, despite widespread EHR and EWS integration, few studies have explored EHR-based EWS.

Inconsistency in evaluation and the lack of high-quality validation makes the evidence of validity questionable, ultimately affects how EWS can and should be used in clinical practice as a risk score for deterioration prediction. Heterogeneity across studies in all subgroup's challenges implementation of EWS in all diseases and all settings. In methodology, observations selections method, time horizon between EWS score and event, and the metric used in assessment were inconsistent. Choosing multiple observations or a single observation prior the outcome may not significantly affect the ranking of EWS (68). Yet, selecting a single observation is generally associated with high AUC compared to multiple observations(47,68), supporting the use of multiple observations for each episode. Moreover, AUC, the most commonly used measure of predictive performance, has limitations and other metrics, including positive predictive value, should also be assessed(69).
Recording observations at an agreed threshold point before events in a standardised method is necessary to evaluate EWS effectively.

The Universal EWS with standardised models were primarily designed for general patient populations in wards and emergency departments and remain under-evaluated in specific diseases and settings. In medical and ED contexts, EWS perform well, suggesting the role of EWS in general settings, or at the early stage of clinical assessment. Our positive findings in respiratory disease may indicate the emphasis of several EWS, such as NEWS2, on respiratory changes when patients are deteriorating. Specific disease areas may show unique alarm signs when critical events are anticipated, which may not be captured by universal EWS, such as NEWS2, where prediction of deterioration is based on pre-defined thresholds in all patients(23). Critical events are commonly associated with CVD. With CVD being a leading cause of mortality globally, and the significant impact of morbidity on health and social care, early detection of deterioration is necessary(70). However, EWS are poorly validated in CVD, some of the parameters may not be applicable, and EWS may be unrepresentative(25). A recent study of NEWS2 in patients with coronavirus infection found poor performance in severity prediction (71), despite pre-existing conditions being common and predictive in patients with severe outcomes. EWS may need to take account of disease-specific risk factors and comorbidities.

Widespread uptake of EHR and digitisation of patient observations are expected to contribute to efficient use of EWS, by reducing human errors in documentation and calculation, as well as delays in escalation of care. However, relatively few studies have considered EHR-based EWS, and those studies have not analysed whether predictive performance of EWS is related to EHR use, diseases or settings. Investigating implementation and adoption of EWS is necessary to understand the application and performance of EWS. Predictive algorithms derived by machine learning have been successfully used in developing and validating EWS (42,72), but will require robust evaluation. Studying the implementation process of EWS within EHR will provide opportunities for qualitative and quantitative insights into escalation of care, as well as facilitators and barriers to use of EWS in routine practice.

There are several limitations in this review and in included studies. We aimed for a comprehensive investigation of all EWS developing since 1997, but this long study period may lead to bias in comparing studies with old and new validation approaches statistically and technically. We excluded EWS specifically derived and validated for particular disease populations or settings, and excluded studies considering a general patient population. Meta-analysis was only done for studies using AUC, excluding other methods for assessing performance of EWS. The distinction between general patient settings and specific disease or patient subgroups is dependent on hospital, healthcare system and country, and there is inevitably overlap between patients and settings at different stages in patient pathways. It was only feasible to include studies with a clear disease or setting identified to avoid confusion.

Validation of EWS in disease subgroups should consider similarities and differences across diseases, sample size, and include measures of model discrimination and calibration. Further research should adhere to established guidelines on clinical outcomes and predictive clinical scoring for decision-making, such as the PROGRESS framework (73).

Conclusion

Universal Early warning scores in specific disease subgroups and settings require further validation of their performance in detecting worsening outcomes. Despite good performance in respiratory patients and medical and surgical settings in studies to-date, the predictive accuracy of EWS in all disease subgroups and all clinical settings remains unknown. The current evidence base does not necessarily support use of standard EWS in all patients in all settings. Future research should include validation of EWS in particular patient subgroups and settings, with standardised methodology following established guidelines. Going toward the utilisation of EHR for EWS development, validation and implementation within EHR should be considered for improved early warning score systems.

Contributorship statement: AB conceived the study. BA, AB and TB conducted the search, data extraction and data analysis. BA wrote the initial draft of the manuscript. All authors (BA, AB, TB, DM, RP and BW) contributed to interpretation of findings, critical review and revisions of the manuscript.

Competing interests: AB has received research grants from Astra Zeneca. All other authors report no competing interests.

Funding: BA has received PhD funding from the Saudi Arabian Cultural Bureau.

Data sharing statement: All data relevant to the study are included in the article or uploaded as supplementary information

References

1. Cetınkaya HB, Koksal O, Sigirli D, Leylek EH, Karasu O. The predictive value of the modified early warning score with rapid lactate level (ViEWS-L) for mortality in patients of age 65 or older visiting the emergency department. Intern Emerg Med. 2017 Dec;12(8):1253–7.
2. Cei M, Bartolomei C, Mumoli N. In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: A prospective study. Int J Clin Pract. 2009;
3. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PWB. The impact of the use of the Early Warning Score (EWS) on patient outcomes: A systematic review. Resuscitation. 2014.
4. Hogan H, Hutchings A, Wulff J, Carver C, Holdsworth E, Welch J, et al. Interventions to reduce mortality from in-hospital cardiac arrest: a mixed-methods study. Heal Serv Deliv Res. 2019;7(2):1–110.
5. Adhikari NKJ, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. Lancet. 2010;376(9749):1339–46.
6. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: A retrospective case record review study. BMJ Qual Saf. 2012;
7. De Meester K, Das T, Hellemans K, Verbrugghe W, Jorens PG, Verpooten GA, et al. Impact of a standardized nurse observation protocol including MEWS after Intensive Care Unit discharge. Resuscitation. 2013;
8. Paterson R, MacLeod DC, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: Clinical audit. Clin Med J R Coll Physicians London. 2006;
9. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted to a tertiary referral intensive care unit after CPR. Resuscitation. 2011;
10. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K. A comparison of Antecedents to Cardiac Arrests, Deaths and EMergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom - The ACADEMIA study. Resuscitation. 2004;
11. Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, et al. Duration of life-threatening antecedents prior to intensive care admission. Intensive Care Med. 2002;
12. Wilkinson K, Martin IC, Gough MJ. National confidential enquiry into patient outcome and death. An age old problem. A review of the care received by elderly patients undergoing surgery. NCEPOD, London. 2011.
13. Morgan RJM, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. Clin Intensive Care. 1997;8(2):100.
14. Linnen DT, Escobar GJ, Hu X, Scruth E, Liu V, Stephens C. Statistical modeling and aggregate-weighted scoring systems in prediction of mortality and ICU transfer: A systematic review. J Hosp Med. 2019;14(3):161–9.
15. Hamilton F, Arnold D, Baird A, Albur M, Whiting P. Early Warning Scores do not accurately predict mortality in sepsis: A meta-analysis and systematic review of the literature. J Infect. 2018;
16. Wuytack F, Meskell P, Conway A, McDaid F, Santesso N, Hickey FG, et al. The effectiveness of physiologically based early warning or track and trigger systems after triage in adult patients presenting to emergency departments: A systematic review. BMC Emerg Med. 2017;
17. Plevin R, Callcut R. Update in sepsis guidelines: what is really new? Trauma Surg Acute Care Open. 2017 Sep 7;2(1):e000088.
18. Mohammed MA, Rudge G, Watson D, Wood G, Smith GB, Prytherch DR, et al. Index blood tests and national early warning scores within 24 hours of emergency admission can predict the risk of in-hospital mortality: a model development and validation study. PLoS One. 2013

May 29;8(5):e64340–e64340.
19. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J EMJ. 2009 Apr;26(4):254–8.
20. Royal College of Physicians of London. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. R Coll Physician. 2012;
21. Royal College of Physicians of London. NHS England approves use of National Early Warning Score (NEWS) 2 to improve detection of acutely ill patients. R Coll Physician. 2017;
22. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013 Apr;84(4):465–70.
23. Inada-Kim M, Nsutebu E. NEWS 2: an opportunity to standardise the management of deterioration and sepsis. BMJ. 2018 Mar 20; 360:k1260.
24. Direkze S, Jain S. Time to intervene? lessons from the NCEPOD cardiopulmonary resuscitation report 2012. Br J Hosp Med. 2012 Oct 16;73(10):585–7.
25. Badreldin AMA, Doerr F, Bender EM, Bayer O, Brehm BR, Wahlers T, et al. Rapid clinical evaluation: An early warning cardiac surgical scoring system for hand-held digital devices*. Eur J Cardio-thoracic Surg. 2013;
26. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000 Feb 29;19(4):453–73.
27. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol. 2015;68(3):279–89.
28. Smith MEB, Chiovaro JC, O'Neil M, Kansagara D, Quiñones AR, Freeman M, et al. Early warning system scores for clinical deterioration in hospitalized patients: A systematic review. Annals of the American Thoracic Society. 2014.
29. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: A systematic review. Resuscitation. 2016.
30. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. bmj. 2020;369.
31. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;
32. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med. 2014;
33. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019 Jan 1;170(1):51.
34. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: the Achilles heel of predictive analytics. BMC Med [Internet]. 2019;17(1):230.
35. Brink A, Alsma J, Verdonschot RJCG, Rood PPM, Zietse R, Lingsma HF, et al. Predicting mortality in patients with suspected sepsis at the Emergency Department; A retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One. 2019 Jan 25;14(1):e0211133–e0211133.
36. Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. Crit Care Med. 2017 Nov;45(11):1805–12.
37. Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA, et al. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department. Intern Emerg Med. 2013 Apr;8(3):255–60.
38. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and q SOFA scores for predicting mortality and length of stay in ICU. J Crit Care. 2017

Oct;41:191–3.
39. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. Hear Lung J Crit Care. 2019 May;48(3):240–4.
40. Vaughn JL, Kline D, Denlinger NM, Andritsos LA, Exline MC, Walker AR. Predictive performance of early warning scores in acute leukemia patients receiving induction chemotherapy. Leuk Lymphoma. 2018 Jun;59(6):1498–500.
41. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015 Aug 5;7(299):299ra122-299ra122.
42. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2018 Oct;131:N.PAG-N.PAG.
43. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017 Apr 1;195(7):906–11.
44. de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scand J Trauma Resusc Emerg Med. 2017 Sep 11;25(1):91.
45. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted “track and trigger” systems. Resuscitation. 2008.
46. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter “track and trigger” systems. Resuscitation. 2008;
47. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012 Mar;83(3):297–302.
48. Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. BJA Br J Anaesth. 2007 May 22;98(6):769–74.
49. Abbott TEF, Torrance HDT, Cron N, Vaid N, Emmanuel J. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med. 2016 Nov;35:78–82.
50. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007 Feb;35(2):402–9.
51. Bartkowiak B, Snyder AM, Benjamin A, Schneider A, Twu NM, Churpek MM, et al. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–63.
52. Qin Q, Xia Y, Cao Y. Clinical study of a new Modified Early Warning System scoring system for rapidly evaluating shock in adults. J Crit Care. 2017 Feb;37:50–5.
53. Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open. 2016 Sep 28;6(9):e011021–e011021.
54. Liljehult J, Christensen T. Early warning score predicts acute mortality in stroke patients. Acta Neurol Scand. 2016 Apr;133(4):261–7.
55. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. Medicine (Baltimore). 2019 Feb;98(6):e14197–e14197.
56. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a

prospective observational study. Scand J Trauma Resusc Emerg Med. 2016 Jun 2;24:80.
57. Goulden R, Hoyle M-C, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J EMJ. 2018 Jun;35(6):345–9.
58. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med Off J Eur Soc Emerg Med. 2016 Dec;23(6):406–12.
59. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J. 2014 Jun;31(6):476–81.
60. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted "track and trigger" systems. Resuscitation. 2008.
61. Mulligan A. Validation of a physiological track and trigger score to identify developing critical illness in haematology patients. Intensive Crit Care Nurs. 2010 Aug;26(4):196–206.
62. Cooksley T, Kitlowski E, Haji-Michael P. Effectiveness of Modified Early Warning Score in predicting outcomes in oncology patients. QJM Mon J Assoc Physicians. 2012 Nov;105(11):1083–8.
63. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. BMJ Open. 2019 Jan 17;9(1):e024636–e024636.
64. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care. 2011;15(2):R95–R95.
65. Albur M, Hamilton F, MacGowan AP. Early warning score: a dynamic marker of severity and prognosis in patients with Gram-negative bacteraemia and sepsis. Ann Clin Microbiol Antimicrob. 2016 Apr 12;15:23.
66. Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med. 2018 Apr;13(3):405–12.
67. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. PLoS One. 2016 Aug 17;11(8):e0161401–e0161401.
68. Jarvis SW, Kovacs C, Briggs J, Meredith P, Schmidt PE, Featherstone PI, et al. Are observation selection methods important when comparing early warning score performance? Resuscitation. 2015 May 1;90:1–6.
69. Romero-Brufau S, Huddleston JM, Naessens JM, Johnson MG, Hickman J, Morlan BW, et al. Widely used track and trigger scores: Are they ready for automation in practice? Resuscitation. 2014 Apr 1;85(4):549–52.
70. Mozaffarian D. Global Scourge of Cardiovascular Disease. J Am Coll Cardiol. 2017 Jul 4;70(1):26 LP – 28.
71. Carr E, Bendayan R, Bean D, O'Gallagher K, Pickles A, Stahl D, et al. Supplementing the National Early Warning Score (NEWS2) for anticipating early deterioration among patients with COVID-19 infection. medRxiv. 2020 Jan 1;2020.04.24.20078006.
72. Churpek MM, Yuen TC, Park SY, Gibbons R, Edelson DP. Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards*. Crit Care Med. 2014 Apr;42(4):841–8.
73. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. BMJ Br Med J. 2013 Feb 5;346:e5595.

Tables and Figures and Legends.

Figure 1. Search strategy and included studies regarding universal early warning scores in different disease subgroups and clinical settings.

Figure 2. Number of studies regarding performance of early warning scores in different disease subgroups and clinical settings.

Figure 2 Legend: Each bubble represents the disease subgroup and/or setting where different early warning scores were examined. The size of the bubble represents the number of studies (n); and overlapping bubbles show studies where disease subgroup and settings overlap. Abbreviations: CVD: Cardiovascular Diseases; ED: Emergency Department; GI: Gastro Intestinal Diseases; ICU: Intensive Care Unit.

Figure 3. Early warning score performance in different disease subgroups.

Figure 3 Legend: Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: CA: cardiac arrest; CVD: cardiovascular diseases; GI: Gastro Intestinal Diseases; ICU: Transfer to Intensive Care Unit; OF: Organ Failure; RA: Respiratory Arrest.

Figure 4. Early warning score performance in different clinical settings.

Figure 4 Legend: Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: ED: Emergency Department; ICU: Intensive Care Units; OF: organ failure; CA: Cardiac Arrest; ICU: Transfer to Intensive Care Units; RA: Respiratory Arrest.

Figure 5: Forest plot of predictive accuracy of universal early warning scores for mortality in different disease subgroups and clinical settings.

Figure 5 Legend: Abbreviations: Med: medical settings, Surg: surgical settings, ED: Emergency Department, ICU: Intensive Care Units, Infec: Infectious Diseases, Resp: Respiratory Diseases, Onco: Oncology diseases, Stroke: Patients with stroke, Renal: Renal diseases, Hem: Haematological diseases, GI: Gastro Intestinal diseases, CVD: Cardiovascular Diseases. Note: number following Author(s) and year indicate more than one EWS evaluated in the study.

Figure 6: Forest plot of predictive accuracy of NEWS for mortality.

Table 1. Characteristics of included studies of predictive performance for early warning scores in patient subgroups and settings.

Table 1 Legend: *Studies are ranked according to sample size from largest to smallest in each subgroup.* *Abbreviations:*

Subgroup: CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastrointestinal diseases; ICU: Intensive Care Unit.

EWS: Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.

EHR: Electronic Health Records.

Predictive measure: AUC: Area Under the Curve; Sens & Spec: Sensitivity and Specificity; OR: Odds Ratio.

Outcomes: ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.

Note: Black dots in the subgroup column represent the disease or the settings where the sample was studied and brown dots in the study by Kellet (2012) represent different samples for each subgroup.

Table 2. Characteristics of included studies of predictive performance for early warning scores in clinical settings.

Table 2 Legend: *Studies are ranked according to sample size from largest to smallest in each subgroup.*

Abbreviations:

Subgroup: CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastrointestinal diseases; ICU: Intensive Care Unit.

EWS: Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.

EHR: Electronic Health Records.

Predictive measure: AUC: Area Under the Curve; Sens and Spec: Sensitivity and Specificity; OR: Odds Ratio.

Outcomes: ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.

Table 1. Characteristics of included studies of predictive performance for early warning scores in patient subgroups and settings.

Author, year	Country	Subgroups								Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		CVD	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infection/sepsis	ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Kellett, 2012	Canada	●	○	○	●	●	●	○	○	●	○	●	●	●	○	○	○	10007	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kim, 2017	Korea	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	2172	✓	●	○	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Bozkurt, 2015	Turkey	○	●	○	○	○	○	○	○	○	○	○	○	○	●	○	○	202	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Seak, 2017	Taiwan	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	66	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Hu, 2016	USA	○	○	●	○	○	●	○	○	○	○	○	○	●	○	○	○	565	✓	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Liljehult, 2016	Denmark	○	○	○	○	●	○	○	○	○	○	○	○	●	○	○	○	274	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Mulligan,2010	UK	○	○	●	○	○	○	○	○	○	○	○	○	○	●	○	○	71	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cooksley, 2012	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	840	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Vaughn, 2018	USA	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	504	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Young, 2014	USA	○	○	●	○	○	●	○	○	○	○	○	○	●	○	○	○	61	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Von, 2007	UK	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	43	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Pedersen, 2018	Denmark	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	11266	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Forster, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	8812	✓	○	○	○	●	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Pimentel, 2018	UK	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	1394	✓	○	○	○	●	●	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Sbiti-rohr, 2016	Switzerland	○	○	○	○	○	○	●	○	○	●	○	○	○	○	●	○	925	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Brabrand, 2017	Denmark	○	○	○	○	○	○	●	○	○	○	○	●	○	●	○	○	570	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2016	Korea	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	553	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Barlow, 2007	UK	○	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	419	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bilben, 2016	Norway	○	○	○	○	○	○	●	○	○	●	○	○	○	●	○	○	246	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Delahanty, 2019	USA	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	920026	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Redfern, 2018	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	241996	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Churpek, Sokol 2017	USA	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	53849	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Faisal, 2019	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	36161		○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓

Author, year	Country	Subgroups								Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		CVD	GI	Haematology	Renal	Stroke	Oncology	Respiratory	Infection/sepsis	ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	[illegible]	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Churpek,2017	USA	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	18523	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Henry, 2015	USA	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	13014	✓	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Brink,2019	Netherlands	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	8204	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
De Groot, 2017	Netherlands	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	2280	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Corfield, 2014	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	2003	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Goulden, 2018	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	1818	✓	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Khwannimit, 2019	Thailand	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	1589	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Ghanem, 2011	Israel	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	1072	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Saeed, 2019	UK, France, Italy, Sweden & Spain	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	1058	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Innocenti, 2018	Italy	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	742	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Camm, 2018	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	533	✗	○	○	○	●	○	○	○	○	○	○	Sens &Spec	✓	✓	✗	✗	✗
Tirotta, 2017	Italy	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	526	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Pong, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	364	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Prabhakar, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	343	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Martino, 2018	Italy	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	310	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Vorwerk, 2009	UK	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	308	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Qin, 2017	China	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	292	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Schmedding, 2019	Gabon	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	277	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Albur, 2016	UK	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	245	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Cildir, 2013	Turkey	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	230	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Chiew, 2019	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	214	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Samsudin, 2018	Malaysia	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	214	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Chang, 2018	China	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	152	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Geier, 2013	Germany	○	○	○	○	○	○	○	●	○	●	○	○	○	●	○	○	151	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Asiimwe, 2015	Uganda	○	○	○	○	○	○	○	●	○	○	○	○	○	●	○	○	150	✗	○	●	○	○	○	○	○	○	○	○	Prognostic index	✓	✗	✗	✗	✗
Hung, 2017	Taiwan	○	○	○	○	○	○	○	●	○	●	○	○	●	○	○	○	114	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Garcea, 2006	UK	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	110	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yoo, 2015	Korea	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	100	✗	○	●	○	○	○	○	○	○	○	○	OR	✓	✓	✗	✗	✗
Siddiqui,2017	Malaysia	○	○	○	○	○	○	○	●	●	○	○	○	●	○	○	○	58	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗

Studies are ranked according to sample size from largest to smallest in each subgroup. *Abbreviations:*
Subgroup: *CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastrointestinal diseases; ICU: Intensive Care Unit.*
EWS: *Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.*
EHR: *Electronic Health Records.*
Predictive measure: *AUC: Area Under the Curve; Sens & Spec: Sensitivity and Specificity; OR: Odds Ratio.*
Outcomes: *ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.*
Note: Black dots in the subgroup column represent the disease or the settings where the sample was studied and brown dots in the study by Kellet (2012) represent different samples for each subgroup.

Table 2. Characteristics of included studies of predictive performance for early warning scores in clinical settings.

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	CA	RA	Sepsis
Calvert 2016	Israel	●	○	○	○	●	○	○	○	29083	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Awad, 2017	UK	●	○	○	○	●	○	○	○	11722	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Reini, 2012	Sweden	●	○	○	○	○	●	○	○	518	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Chen, 2019	Taiwan	●	○	○	○	●	○	○	○	370	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✓	✗
Baker, 2015	Tanzania	●	○	○	○	○	●	○	○	269	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Gök, 2019	Turkey	●	○	○	○	●	○	○	○	250	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✓
Moseson, 2014	USA	●	○	○	○	○	●	○	○	227	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Jo, 2013	South Korea	●	○	○	○	●	○	○	○	151	✗	●	●	○	○	○	○	○	●	○	○	AUC	✓	✗	✗	✗	✗
Kown, 2018	Korea	○	●	○	○	●	○	○	○	1986334	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Usman, 2019	USA	○	●	○	○	●	○	○	○	115734	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✓
Jang, 2019	Korea	○	●	○	○	●	○	○	○	56368	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✗	✓	✗	✗
Wei, 2019	China	○	●	○	○	●	○	○	○	39977	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lee, 2019	Korea	○	●	○	○	●	○	○	○	27173	✗	○	●	○	●	○	○	○	○	●	○	AUC	✓	✗	✗	✗	✗
Singer, 2017	USA	○	●	○	○	●	○	○	○	22530	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Eick, 2015	Germany	○	●	○	○	○	●	○	○	5730	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bulut, 2014	Turkey	○	●	○	○	○	●	○	○	2000	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Kivipuro, 2018	Finland	○	●	○	○	○	●	○	○	1354	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Eckart, 2019	USA	○	●	○	○	●	○	○	○	1303	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Ho, 2013	Malaysia	○	●	○	○	●	○	○	○	1024	✗	○	●	○	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Skitch, 2018	Canada	○	●	○	○	●	○	○	○	845	✗	○	○	○	●	○	○	○	○	○	●	AUC	✗	✗	✗	✗	✓
Liu, 2014	Malaysia	○	●	○	○	○	●	○	○	702	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✓	✗	✗
Dundar, 2016	Turkey	○	●	○	○	○	●	○	○	671	✗	●	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Yuan., 2018	China	○	●	○	○	○	●	○	○	621	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Naidoo, 2014	South Africa	○	●	○	○	●	○	○	○	590	✗	○	○	○	○	○	○	○	○	●	○	Sens & Spec	✓	✗	✗	✗	✗
Liu F.Y, 2015	China	○	●	○	○	○	●	○	○	551	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
So, 2015	China	○	●	○	○	○	●	○	○	544	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✗	✗	✗	✗
Dundar, 2019	Turkey	○	●	○	○	●	○	○	○	455	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Lam, 2006	China	○	●	○	○	○	●	○	○	425	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Xie, 2018	China	○	●	○	○	○	●	○	○	383	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Cattermole, 2009	China	○	●	○	○	○	●	○	○	330	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Heitz, 2010	USA	○	●	○	○	●	○	○	○	280	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Sirivilaithon, 2019	Thailand	○	●	○	○	○	○	○	●	250	✗	○	○	○	●	○	○	○	○	○	○	AUC	✗	✗	✗	✗	✗
Cattermole, 2014	China	○	●	○	○	○	●	○	○	230	✗	○	●	○	●	○	○	●	○	○	○	AUC	✓	✗	✗	✗	✗
Najafi, 2018	Iran	○	●	○	○	○	●	○	○	185	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Bartkowiak, 2019	USA	○	○	●	○	●	○	○	○	32537	✗	○	●	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Kovacs, 2016	UK	○	○	●	●	●	○	○	○	20626	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Plate, 2018	Netherlands	○	○	●	○	○	●	○	○	1782	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗
Sarani, 2012	Netherlands	○	○	●	○	○	●	○	○	572	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✓	✗	✗	✗
Hollis, 2016	USA	○	○	●	○	●	○	○	○	522	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Author, year	Country	Settings				Study design				Number of patients	EHR	EWS										Predictive measure	Outcomes studied				
		ICU	ED	Surgical	Medical	Retrospective	Prospective	RCT	Case Control			VIEWS	MEWS	EWS	NEWS	NEWS2	SOS	WORTHING	HOTEL	TREWS	HEWS		Mortality	ICU	C A	R A	Sepsis
Gardner-Thorpe 2006	UK	○	○	●	○	○	●	○	○	334	✗	○	●	○	○	○	○	○	○	○	○	Sens & Spec	✓	✓	✗	✗	✗
Garcea, 2010	UK	○	○	●	○	●	○	○	○	280	✗	○	○	●	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Cuthbertson, 2007	UK	○	○	●	○	●	○	○	○	136	✗	○	●	●	○	○	○	○	○	○	○	AUC	✗	✓	✗	✗	✗
Prytherch, 2010	UK	○	○	○	●	●	○	○	○	35585	✗	●	○	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Smith, 2013	UK	○	○	○	●	●	○	○	○	35585	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✓	✗	✗
Rasmussen, 2018	Denmark	○	○	○	●	●	○	○	○	17312	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Ghosh, 2018	USA	○	○	○	●	●	○	○	○	2097	✓	○	●	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Duckitt, 2007	UK	○	○	○	●	○	●	○	○	1102	✗	○	○	●	○	○	○	●	○	○	○	AUC	✓	✓	✗	✗	✗
Colombo, 2017	Italy	○	○	○	●	●	○	○	○	471	✗	○	●	○	○	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Abbot, 2016	UK	○	○	○	●	○	●	○	○	322	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✗	✗	✗	✗
Wheeler, 2013	Malawi	○	○	○	●	○	●	○	○	302	✗	○	●	○	○	○	○	○	●	○	○	AUC	✓	✗	✗	✗	✗
Graziadio, 2019	UK	○	○	○	●	○	●	○	○	292	✗	○	○	○	●	○	○	○	○	○	○	AUC	✓	✓	✗	✗	✗

Studies are ranked according to sample size from largest to smallest in each subgroup.

Abbreviations:

Subgroup: CVD: Cardiovascular Disease, ED: Emergency Department; GI: Gastrointestinal diseases; ICU: Intensive Care Unit.

EWS: Early warning score; VIEWS: Vital pack Early Warning Score, MEWS: Modified Early Warning Score; EWS: Early Warning Score; NEWS: National Early Warning Score; HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score; Worthing: Worthing physiological scoring system; TREWS: Triage in Emergency department Early Warning Score; SOS: Search Out Severity score, HEWS: Hamilton early warning score.

EHR: Electronic Health Records.

Predictive measure: AUC: Area Under the Curve; Sens and Spec: Sensitivity and Specificity; OR: Odds Ratio.

Outcomes: ICU: transfer to Intensive Care Unit; CA: Cardiac Arrest; RA: Respiratory Arrest.

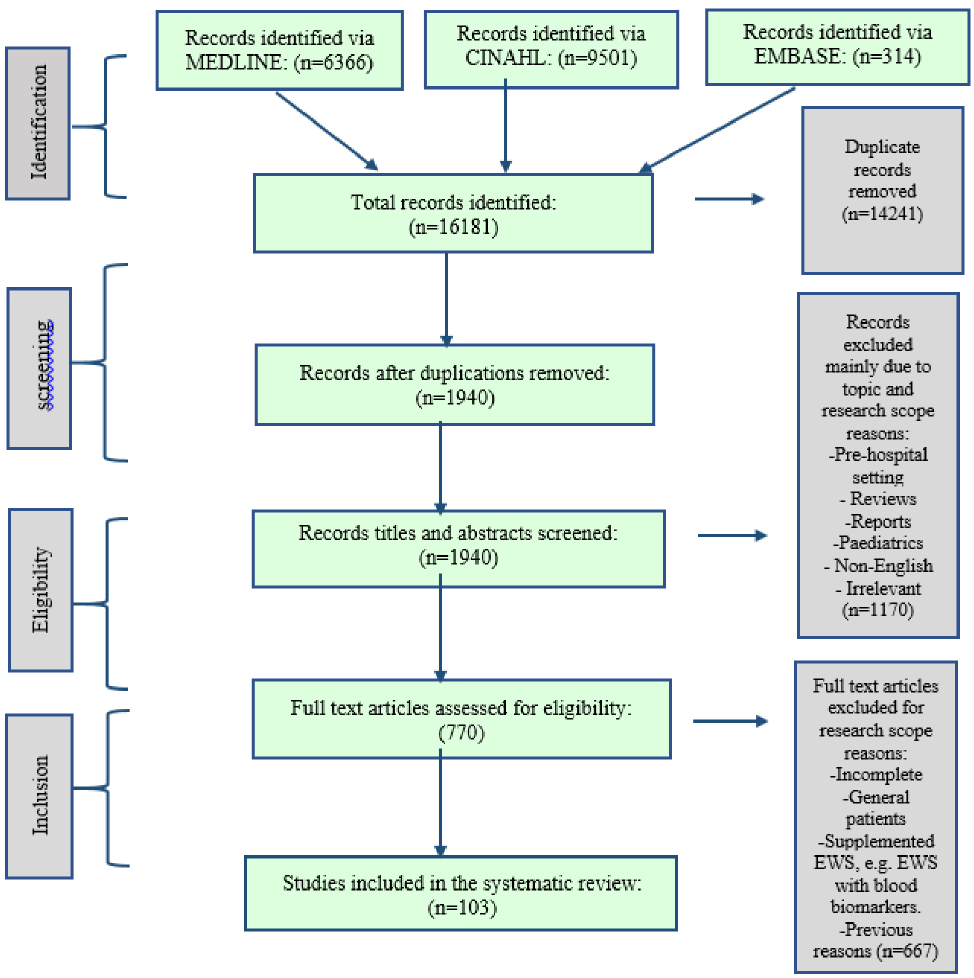


Figure 1. Search strategy and included studies regarding universal early warning scores in different disease subgroups and clinical settings.

115x115mm (144 x 144 DPI)

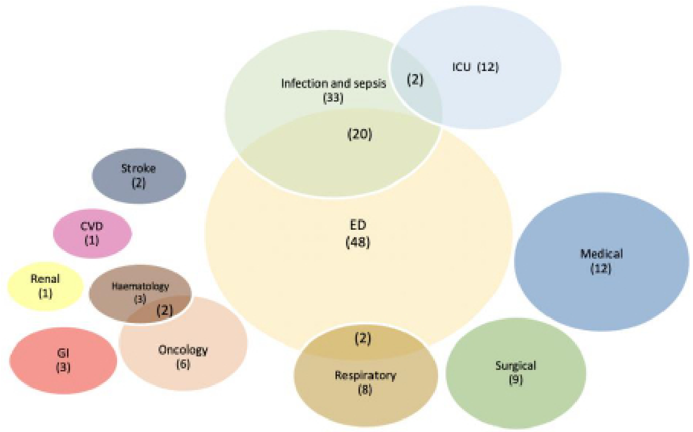


Figure 2. Number of studies regarding performance of early warning scores in different disease subgroups and clinical settings.

Legend: Each bubble represents the disease subgroup and/or setting where different early warning scores were examined. The size of the bubble represents the number of studies (n); and overlapping bubbles show studies where disease subgroup and settings overlap. Abbreviations: CVD: Cardiovascular Diseases; ED: Emergency Department; GI: Gastro Intestinal Diseases; ICU: Intensive Care Unit.

115x65mm (144 x 144 DPI)

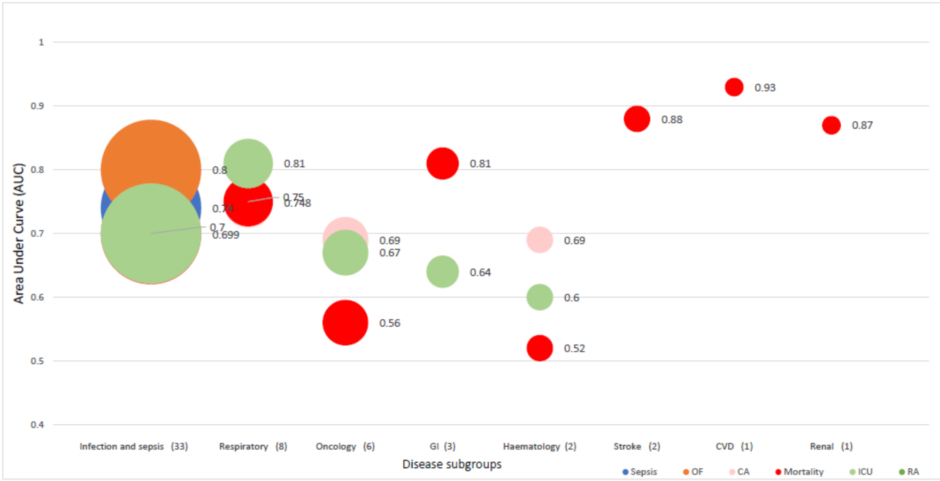


Figure 3. Early warning score performance in different disease subgroups.

Legend: Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: CA: cardiac arrest; CVD: cardiovascular diseases; GI: Gastro Intestinal Diseases; ICU: Transfer to Intensive Care Unit; OF: Organ Failure; RA: Respiratory Arrest.

187x94mm (144 x 144 DPI)

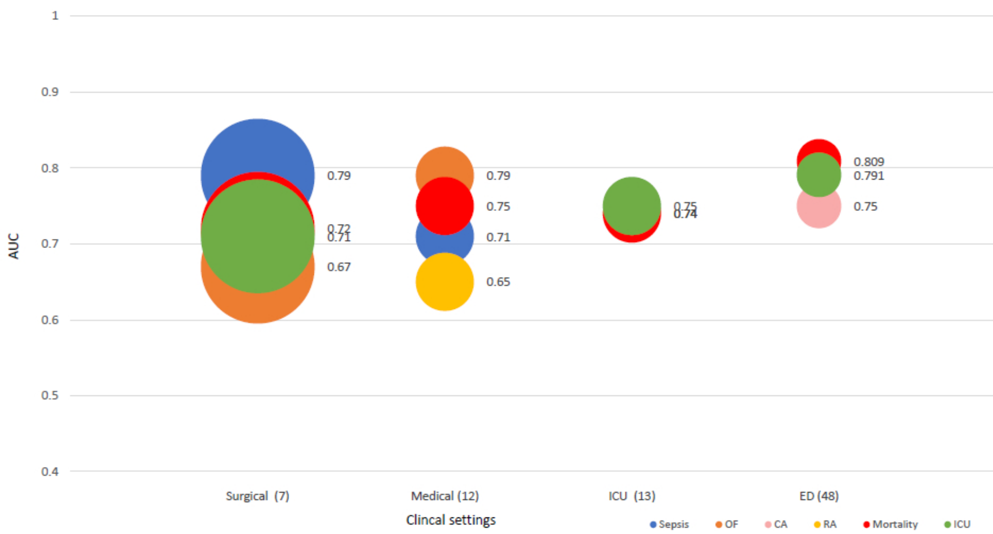


Figure 4. Early warning score performance in different clinical settings.

Legend: Each bubble represents critical events predicted by early warning scores for each disease subgroup with average AUC of studies beside each event type. The size of the bubble represents the number of studies in each subgroup. Abbreviations: ED: Emergency Department; ICU: Intensive Care Units; OF: organ failure; CA: Cardiac Arrest; ICU: Transfer to Intensive Care Units; RA: Respiratory Arrest.

166x88mm (144 x 144 DPI)

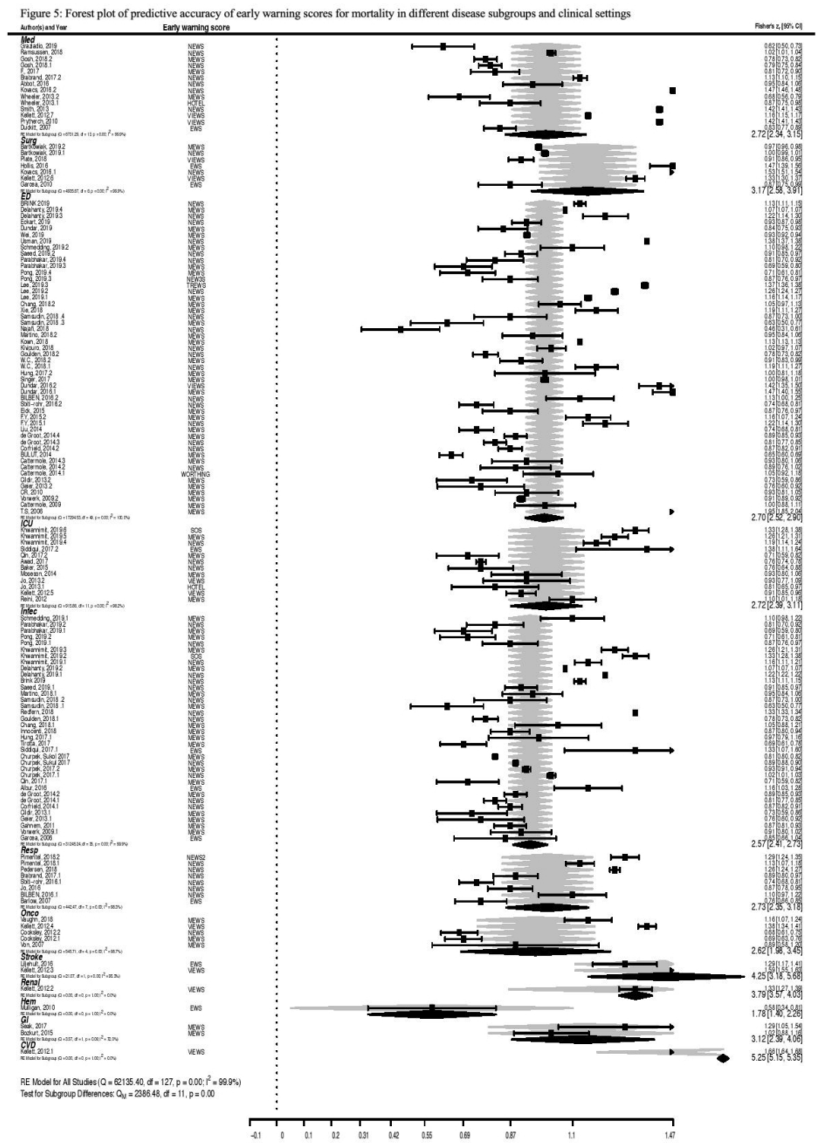


Figure 5: Forest plot of predictive accuracy of universal early warning scores for mortality in different disease subgroups and clinical settings.
Legend: Abbreviations: Med: medical settings, Surg: surgical settings, ED: Emergency Department, ICU: Intensive Care Units, Infec: Infectious Diseases, Resp: Respiratory Diseases, Onco: Oncology diseases, Stroke: Patients with stroke, Renal: Renal diseases, Hem: Haematological diseases, GI: Gastro Intestinal diseases, CVD: Cardiovascular Diseases. Note: number following Author(s) and year indicate more than one EWS evaluated in the study.

209x297mm (150 x 150 DPI)

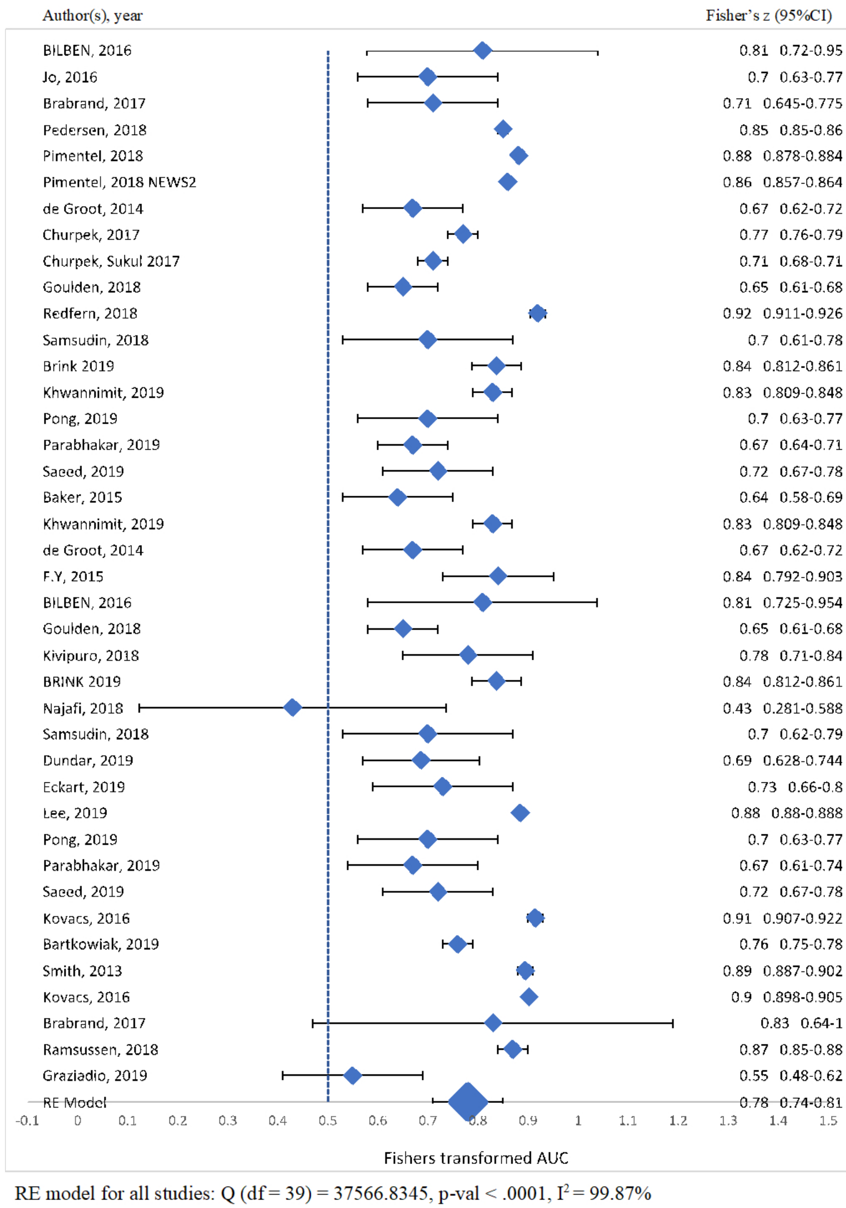


Figure 6. Forest plot of predictive accuracy of NEWS for mortality.

178x253mm (144 x 144 DPI)

The performance of early warning scores in different patient subgroups and clinical settings: A systematic review.

Supplementary methods: Search strategy for MEDLINE

1- EWS OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8
10- MH "Cardiovascular Diseases") OR (MH "Cardiology")
11- (MH "Thoracic Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Musculoskeletal Diseases") OR (MH "Orthopedics")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney Diseases, Cystic") OR (MH "Kidney Failure, Chronic") OR (MH "Polycystic Kidney Diseases") OR (MH "Renal Insufficiency, Chronic")
20- renal disease* OR renal failure OR kidney disease*
21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- hematologic disorder* OR hematologic disease* OR hematology
25- 23 OR 24
26- 1 AND 25
27- (MH "Respiratory Tract Diseases")
28- respiratory disease* OR respiratory disorder*
29- 27 OR 28
30- 1 AND 29
31- (MH "Gastroenterology")
32- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
33- 31 OR 32
34- 1 AND 33
35- (MH "Medical Oncology") OR (MH "Surgical Oncology")
36- oncology OR cancer OR chemotherapy
37- 35 OR 36
38- 1 AND 37
39- (MH "Wounds and Injuries") OR (MH "Emergency Medicine")
40- emergency department* OR emergency OR emergency room* OR trauma*
41- 39 OR 40
42- 1 AND 41
43- (MH "Sepsis") OR (MH "Infection")
44- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
45- 43 OR 44
46- 1 AND 45
47- (MH "Obstetrics")
48- (obstetrics and gynecology) OR OBSTETRIC*
49- 47 OR 48
50- 1 AND 49
51- (MH "Allergy and Immunology")
52- immunological disease* OR immunological disorder*

53- 51 OR 52
54- 1 AND 53
55- (MH "Internal Medicine")
56- medical ward*
57- 55 OR 56
58- 1 AND 57
59- (MH "General Surgery")
60- surgical ward*
61- 59 OR 60
62- 1 AND 61
63- 5 OR 9 OR 14 OR 18 OR 22 OR 26 OR 30 OR 34 OR 38 OR 42 OR 46 OR 50 OR 54 OR 58 OR 62

Supplementary methods: Search strategy for CINAHL

1- EWS OR early warning score* OR EARLY WARNING SYSTEM* OR RAPID RESPONSE SYSTEM* OR MEWS OR MODIFIED EARLY WARNING SCORE* OR MODIFIED EARLY WARNING SYSTEM* OR news OR national early warning score OR news2 OR national early warning score 2 OR (track and trigger system*)
2- (MH "Intensive Care Units")
3- ICU OR intensive care unit* OR CRITICAL CARE UNIT* OR CRITICAL CARE
4- 2 OR 3
5- 1 AND 4
6- (MH "Nervous System Diseases")
7- neurological disorder* OR neurological disease OR neurological condition*
8- 6 OR 7
9- 1 AND 8
10- (MH "Heart Diseases") OR (MH "Cardiovascular Diseases")
11- (MH "Heart Surgery")
12- cardiovascular disease* OR cardiovascular disorder* OR heart disease* OR cardiology* OR cardiac surgery OR thoracic surgery
13- 10 OR 11 OR 12
14- 1 AND 13
15- (MH "Orthopedic Surgery") OR (MH "Musculoskeletal Diseases")
16- orthopedic disease* OR orthopedic surgery
17- 15 OR 16
18- 1 AND 17
19- (MH "Kidney, Cystic") OR (MH "Kidney Diseases")
20- renal disease* OR renal failure OR kidney disease*
21- 19 OR 20
22- 1 AND 21
23- (MH "Hematologic Diseases")
24- (MH "Lymphatic Diseases")
25- hematologic disorder* OR hematologic disease* OR hematology
26- 23 OR 24 O 25
27- 1 AND 26
28- (MH "Respiratory Tract Diseases")
29- respiratory disease* OR respiratory disorder*
30- 28 OR 29
31- 1 AND 30
32- (MH "Digestive System Diseases")
33- gastrointestinal disorder* OR gastrointestinal disease* OR gastroenterology OR hepatology
34- 32 OR 33
35- 1 AND 34
36- (MH "Cancer Patients") OR (MH "Oncology")
37- oncology OR cancer OR chemotherapy
38- 36 OR 37
39- 1 AND 38
40- (MH "Wounds and Injuries") OR (MH "Trauma")
41- emergency department* OR emergency OR emergency room* OR trauma*
42- 40 OR 41
43- 1 AND 42
44- (MH "Infection")
45- INFECTION* OR INFECTIOUS DISEASE* OR SEPSIS
46- 44 OR 45
47- 1 AND 46
48- (MH "Obstetric Emergencies") OR (MH "Obstetric Patients")
49- (obstetrics and gynecology) OR OBSTETRIC*
50- 48 OR 49
51- 1 AND 50
52- (MH "Internal Medicine")
53- (MH "Allergy and Immunology")

54- medical ward
55- immunological disease* OR immunological disorder*
56- 52 OR 53 OR 54 OR 55
57- 1 AND 56
58- (MH "Surgical Patients")
59- surgical ward*
60- 58 OR 59
61- 1 AND 60
62- 5 OR 9 OR 14 OR 18 OR 22 OR 27 OR 31 OR 35 OR 39 OR 43 OR OR 47 OR 51 OR 57 OR 61

Supplementary methods: Patients' subgroups

1- Cardiology patients
2- Neurology patients
3- Orthopaedic patients
4- Renal patients
5- Haematology patients
6- Respiratory patients
7- Gastroenterology patients
8- Oncology patients
9- Emergency patients
10- Infection patients
11- Medical patients
12- Surgical patients
13- Intensive care patients

Supplementary References

1. Kellett J, Kim A. Validation of an abbreviated VitalpacTM Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012 Mar;83(3):297–302.
2. Kim W-Y, Lee J, Lee J-R, Jung YK, Kim HJ, Huh JW, et al. A risk scoring model based on vital signs and laboratory data predicting transfer to the intensive care unit of patients admitted to gastroenterology wards. J Crit Care. 2017 Aug;40:213–7.
3. Bozkurt S, Köse A, Arslan ED, Erdoğan S, Üçbilek E, Çevik İ, et al. Validity of modified early warning, Glasgow Blatchford, and pre-endoscopic Rockall scores in predicting prognosis of patients presenting to emergency department with upper gastrointestinal bleeding. Scand J Trauma Resusc Emerg Med. 2015 Dec 30;23:109.
4. Seak C-J, Yen DH-T, Ng C-J, Wong Y-C, Hsu K-H, Seak JC-Y, et al. Rapid Emergency Medicine Score: A novel prognostic tool for predicting the outcomes of adult patients with hepatic portal venous gas in the emergency department. PLoS One. 2017 Sep 15;12(9):e0184813–e0184813.
5. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. PLoS One. 2016 Aug 17;11(8):e0161401–e0161401.
6. Liljehult J, Christensen T. Early warning score predicts acute mortality in stroke patients. Acta Neurol Scand. 2016 Apr;133(4):261–7.
7. Mulligan A. Validation of a physiological track and trigger score to identify developing critical illness in haematology patients. Intensive Crit Care Nurs. 2010 Aug;26(4):196–206.
8. Cooksley T, Kitlowski E, Haji-Michael P. Effectiveness of Modified Early Warning Score in predicting outcomes in oncology patients. QJM Mon J Assoc Physicians. 2012 Nov;105(11):1083–8.
9. Vaughn JL, Kline D, Denlinger NM, Andritsos LA, Exline MC, Walker AR. Predictive performance of early warning scores in acute leukemia patients receiving induction chemotherapy. Leuk Lymphoma. 2018 Jun;59(6):1498–500.
10. Young RS, Gobel BH, Schumacher M, Lee J, Weaver C, Weitzman S. Use of the modified early warning score and serum lactate to prevent cardiopulmonary arrest in hematology-oncology patients: a quality improvement study. Am J Med Qual. 2014 Nov;29(6):530–7.
11. von Lilienfeld-Toal M, Midgley K, Lieberbach S, Barnard L, Glasmacher A, Gilleece M, et al. Observation-based early warning scores to detect impending critical illness predict in-hospital and overall survival in patients undergoing allogeneic stem cell transplantation. Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant. 2007 May;13(5):568–76.
12. Pedersen NE, Rasmussen LS, Petersen JA, Gerds TA, Østergaard D, Lippert A. Modifications of the National Early Warning Score for patients with chronic respiratory disease. Acta Anaesthesiol Scand. 2018 Feb;62(2):242–52.
13. Forster S, Housley G, McKeever TM, Shaw DE. Investigating the discriminative value of Early Warning Scores in patients with respiratory disease using a retrospective cohort analysis of admissions to Nottingham University Hospitals Trust over a 2-year period. BMJ Open. 2018 Jul 30;8(7):e020269–e020269.
14. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early

Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2018 Oct;131:N.PAG-N.PAG.
15. Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open 2016 Sep 28;6(9):e011021–e011021.
16. Brabrand M, Hallas P, Hansen SN, Jensen KM, Madsen JLB, Posth S. Using scores to identify patients at risk of short term mortality at arrival to the acute medical unit: A validation study of six existing scores. Eur J Intern Med. 2017 Nov;45:32–6.
17. Jo S, Jeong T, Lee JB, Jin Y, Yoon J, Park B, et al. Validation of modified early warning score using serum lactate level in community-acquired pneumonia patients. The National Early Warning Score-Lactate score. Am J Emerg Med. 2016 Mar;34(3):536–41.
18. Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia. Thorax. 2007 Mar;62(3):253–9.
19. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a prospective observational study. Scand J Trauma Resusc Emerg Med. 2016 Jun 2;24:80.
20. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med 2019 Apr;73(4):334–44.
21. Redfern OC, Smith GB, Prytherch DR, Meredith P, Inada-Kim M, Schmidt PE. A Comparison of the Quick Sequential (Sepsis-Related) Organ Failure Assessment Score and the National Early Warning Score in Non-ICU Patients With/Without Infection. Crit Care Med. 2018 Dec;46(12):1923–33.
22. Churpek MM, Snyder A, Sokol S, Pettit NN, Edelson DP. Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores. Crit Care Med. 2017 Nov;45(11):1805–12.
23. Faisal M, Richardson D, Scally AJ, Howes R, Beatson K, Speed K, et al. Computer-aided National Early Warning Score to predict the risk of sepsis following emergency medical admission to hospital: a model development and external validation study. C Can Med Assoc J = J L'association Medicale Can. 2019 Apr 8;191(14):E382–9.
24. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017 Apr 1;195(7):906–11.
25. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015 Aug 5;7(299):299ra122-299ra122.
26. Brink A, Alsma J, Verdonschot RJCG, Rood PPM, Zietse R, Lingsma HF, et al. Predicting mortality in patients with suspected sepsis at the Emergency Department; A retrospective cohort study comparing qSOFA, SIRS and National Early Warning Score. PLoS One. 2019 Jan 25;14(1):e0211133–e0211133.
27. de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scand J Trauma Resusc Emerg Med. 2017 Sep 11;25(1):91.

28. Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emerg Med J. 2014 Jun;31(6):482–7.
29. Goulden R, Hoyle M-C, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J EMJ. 2018 Jun;35(6):345–9.
30. Khwannimit B, Bhurayanontachai R, Vattanavanit V. Comparison of the accuracy of three early warning scores with SOFA score for predicting mortality in adult sepsis and septic shock patients admitted to intensive care unit. Hear Lung J Crit Care. 2019 May;48(3):240–4.
31. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care. 2011;15(2):R95–R95.
32. Saeed K, Wilson DC, Bloos F, Schuetz P, van der Does Y, Melander O, et al. The early identification of disease progression in patients with suspected infection presenting to the emergency department: a multi-centre derivation and validation study. Crit Care. 2019 Feb 8;23(1):40.
33. Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med. 2018 Apr;13(3):405–12.
34. Camm CF, Hayward G, Elias TCN, Bowen JST, Hassanzadeh R, Fanshawe T, et al. Sepsis recognition tools in acute ambulatory care: associations with process of care and clinical outcomes in a service evaluation of an Emergency Multidisciplinary Unit in Oxfordshire. BMJ Open. 2018 Apr 9;8(4):e020497–e020497.
35. Tirotta D, Gambacorta M, La Regina M, Attardo T, Lo Gullo A, Panzone F, et al. Evaluation of the threshold value for the modified early warning score (MEWS) in medical septic patients: a secondary analysis of an Italian multicentric prospective cohort (SNOOPII study). QJM Mon J Assoc Physicians. 2017 Jun 1;110(6):369–73.
36. Pong JZ, Fook-Chong S, Koh ZX, Samsudin MI, Tagami T, Chiew CJ, et al. Combining Heart Rate Variability with Disease Severity Score Variables for Mortality Risk Stratification in Septic Patients Presenting at the Emergency Department. Int J Environ Res Public Health. 2019 May 16;16(10).
37. Prabhakar SM, Tagami T, Liu N, Samsudin MI, Ng JCJ, Koh ZX, et al. Combining quick sequential organ failure assessment score with heart rate variability may improve predictive ability for mortality in septic patients at the emergency department. PLoS One. 2019 Mar 18;14(3):e0213445–e0213445.
38. Martino IF, Figgiaconi V, Seminari E, Muzzi A, Corbella M, Perlini S. The role of qSOFA compared to other prognostic scores in septic patients upon admission to the emergency department. Eur J Intern Med. 2018 Jul;53:e11–3.
39. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J EMJ. 2009 Apr;26(4):254–8.
40. Qin Q, Xia Y, Cao Y. Clinical study of a new Modified Early Warning System scoring system for rapidly evaluating shock in adults. J Crit Care. 2017 Feb;37:50–5.
41. Schmedding M, Adegbite BR, Gould S, Beyeme JO, Adegnika AA, Grobusch MP, et al. A Prospective Comparison of Quick Sequential Organ Failure Assessment, Systemic Inflammatory Response Syndrome Criteria, Universal Vital Assessment, and Modified Early Warning Score to Predict Mortality in Patients with Suspected Infection in Gabon. Am J Trop Med Hyg. 2019 Jan;100(1):202–8.

42. Albur M, Hamilton F, MacGowan AP. Early warning score: a dynamic marker of severity and prognosis in patients with Gram-negative bacteraemia and sepsis. Ann Clin Microbiol Antimicrob. 2016 Apr 12;15:23.
43. Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA, et al. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department. Intern Emerg Med. 2013 Apr;8(3):255–60.
44. Chiew CJ, Liu N, Tagami T, Wong TH, Koh ZX, Ong MEH. Heart rate variability based machine learning models for risk prediction of suspected sepsis patients in the emergency department. Medicine (Baltimore). 2019 Feb;98(6):e14197–e14197.
45. Samsudin MI, Liu N, Prabhakar SM, Chong S-L, Kit Lye W, Koh ZX, et al. A novel heart rate variability based risk prediction model for septic patients presenting to the emergency department. Medicine (Baltimore). 2018 Jun;97(22):e10866–e10866.
46. Chang S-H, Hsieh C-H, Weng Y-M, Hsieh M-S, Goh ZNL, Chen H-Y, et al. Performance Assessment of the Mortality in Emergency Department Sepsis Score, Modified Early Warning Score, Rapid Emergency Medicine Score, and Rapid Acute Physiology Score in Predicting Survival Outcomes of Adult Renal Abscess Patients in the Emergency D. Biomed Res Int. 2018 Sep 19;2018:6983568.
47. Geier F, Popp S, Greve Y, Achterberg A, Glöckner E, Ziegler R, et al. Severity illness scoring systems for early identification and prediction of in-hospital mortality in patients with suspected sepsis presenting to the emergency department. Wien Klin Wochenschr. 2013 Sep;125(17–18):508–15.
48. Asiimwe SB, Abdallah A, Ssekitoleko R. A simple prognostic index based on admission vital signs data among patients with sepsis in a resource-limited setting. Crit Care. 2015 Mar 16;19:86.
49. Hung S-K, Ng C-J, Kuo C-F, Goh ZNL, Huang L-H, Li C-H, et al. Comparison of the Mortality in Emergency Department Sepsis Score, Modified Early Warning Score, Rapid Emergency Medicine Score and Rapid Acute Physiology Score for predicting the outcomes of adult splenic abscess patients in the emergency department. PLoS One. 2017 Nov 1;12(11):e0187495–e0187495.
50. Garcea G, Jackson B, Pattenden CJ, Sutton CD, Neal CP, Dennison AR, et al. Early warning scores predict outcome in acute pancreatitis. J Gastrointest Surg Off J Soc Surg Aliment Tract. 2006 Jul;10(7):1008–15.
51. Yoo J-W, Lee JR, Jung YK, Choi SH, Son JS, Kang BJ, et al. A combination of early warning score and lactate to predict intensive care unit transfer of inpatients with severe sepsis/septic shock. Korean J Intern Med. 2015 Jul;30(4):471–7.
52. Siddiqui S, Chua M, Kumaresh V, Choo R. A comparison of pre ICU admission SIRS, EWS and q SOFA scores for predicting mortality and length of stay in ICU. J Crit Care. 2017 Oct;41:191–3.
53. Calvert J, Desautels T, Chettipally U, Barton C, Hoffman J, Jay M, et al. High-performance detection and early prediction of septic shock for alcohol-use disorder patients. Ann Med Surg. 2016 May 10;8:50–5.
54. Awad A, Bader-El-Den M, McNicholas J, Briggs J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. Int J Med Inform. 2017 Dec;108:185–95.
55. Reini K, Fredrikson M, Oscarsson A. The prognostic value of the Modified Early Warning Score in critically ill patients: a prospective, observational study. Eur J Anaesthesiol. 2012 Mar;29(3):152–7.
56. Chen Y-C, Yu W-K, Ko H-K, Pan S-W, Chen Y-W, Ho L-I, et al. Post-intensive care unit respiratory failure in older patients liberated from intensive care unit and ventilator: The

predictive value of the National Early Warning Score on intensive care unit discharge. Geriatr Gerontol Int. 2019 Apr;19(4):317–22.
57. Baker T, Blixt J, Lugazia E, Schell CO, Mulungu M, Milton A, et al. Single Deranged Physiologic Parameters Are Associated With Mortality in a Low-Income Country. Crit Care Med. 2015 Oct;43(10):2171–9.
58. Gök RGY, Gök A, Bulut M. Assessing prognosis with modified early warning score, rapid emergency medicine score and worthing physiological scoring system in patients admitted to intensive care unit from emergency department. Int Emerg Nurs. 2019 Mar;43:9–14.
59. Moseson EM, Zhuo H, Chu J, Stein JC, Matthay MA, Kangelaris KN, et al. Intensive care unit scoring systems outperform emergency department scoring systems for mortality prediction in critically ill patients: a prospective cohort study. J Intensive Care. 2014 Jul 1;2:40.
60. Jo S, JB L, YH J, TO J, JC Y, YK J, et al. Modified early warning score with rapid lactate level in critically ill medical patients: the ViEWS-L score. Emerg Med J. 2013 Feb;30(2):123–9.
61. Kwon J-M, Lee Y, Lee Y, Lee S, Park H, Park J. Validation of deep-learning-based triage and acuity score using a large national dataset. PLoS One. 2018 Oct 15;13(10):e0205836–e0205836.
62. Usman OA, Usman AA, Ward MA. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the Emergency Department. Am J Emerg Med. 2018 Nov 7.
63. Jang D-H, Kim J, Jo YH, Lee JH, Hwang JE, Park SM, et al. Developing neural network models for early detection of cardiac arrest in emergency department. Am J Emerg Med. 2019 Apr 7.
64. Wei X, Ma H, Liu R, Zhao Y. Comparing the effectiveness of three scoring systems in predicting adult patient outcomes in the emergency department. Medicine (Baltimore). 2019 Feb;98(5):e14289–e14289.
65. Lee SB, Kim DH, Kim T, Kang C, Lee SH, Jeong JH, et al. Triage in Emergency Department Early Warning Score (TREWS) is predicting in-hospital mortality in the emergency department. Am J Emerg Med. 2019 Feb 17.
66. Singer AJ, Ng J, Thode HC, Spiegel R, Weingart S, Thode HCJ. Quick SOFA Scores Predict Mortality in Adult Emergency Department Patients With and Without Suspected Infection. Ann Emerg Med.. 2017 Apr;69(4):475–9.
67. Eick C, Rizas KD, Meyer-Zürn CS, Groga-Bada P, Hamm W, Kreth F, et al. Autonomic nervous system activity as risk predictor in the medical emergency department: a prospective cohort study. Crit Care Med. 2015 May;43(5):1079–86.
68. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J. 2014 Jun;31(6):476–81.
69. Kivipuro M, Tirkkonen J, Kontula T, Solin J, Kalliomäki J, Pauniaho S-L, et al. National early warning score (NEWS) in a Finnish multidisciplinary emergency department and direct vs. late admission to intensive care. Resuscitation. 2018 Jul;128:164–9.
70. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study. BMJ Open. 2019 Jan 17;9(1):e024636–e024636.
71. Ho LO, Li H, Shahidah N, Koh ZX, Sultana P, Hock Ong ME. Poor performance of the modified early warning score for predicting mortality in critically ill patients presenting to an emergency department. World J Emerg Med. 2013;4(4):273–8.

72. Skitch S, Tam B, Xu M, McInnis L, Vu A, Fox-Robichaud A. Examining the utility of the Hamilton early warning scores (HEWS) at triage: Retrospective pilot study in a Canadian emergency department. CJEM Can J Emerg Med. 2018 Mar;20(2):266–74.
73. Liu N, Koh ZX, Goh J, Lin Z, Haaland B, Ting BP, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. BMC Med Informatics Decis Mak. 2014 Jan;14(1):75.
74. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med Off J Eur Soc Emerg Med. 2016 Dec;23(6):406–12.
75. Yuan WC, Tao C, Dan ZD, Yi SC, Jing W, Jian Q. The significance of National Early Warning Score for predicting prognosis and evaluating conditions of patients in resuscitation room. Hong Kong J Emerg Med. 2018;25(6):324–30.
76. Naidoo DK, Rangiah S, Naidoo SS. An evaluation of the Triage Early Warning Score in an urban accident and emergency department in KwaZulu-Natal. South African Fam Pract. 2014 Jan;56(1):69–73.
77. Liu FY, Qin J, Wang RX, Fan XI, Wang J, Sun CY et al. A prospective validation of national early warning score in emergency intensive care unit patients at Beijing. Hong Kong J Emerg Med. 2015;22(3):137–44.
78. So S-N, Ong C-W, Wong L-Y, Chung JYM, Graham CA. Is the Modified Early Warning Score able to enhance clinical observation to detect deteriorating patients earlier in an Accident & Emergency Department? Australas Emerg Nurs J. 2015 Feb;18(1):24–32.
79. Dundar ZD, Kocak S, Girisgin AS. Lactate and NEWS-L are fair predictors of mortality in critically ill geriatric emergency department patients. Am J Emerg Med. 2019 Feb 7.
80. Lam TS, Mak PSK, Siu WS, Lam MY, Cheung TF, Rainer TH. Validation of a Modified Early Warning Score (MEWS) in emergency department observation ward patients. Hong Kong J Emerg Med. 2006;13(1):24–30.
81. Xie X, Huang W, Liu Q, Tan W, Pan L, Wang L, et al. Prognostic value of Modified Early Warning Score generated in a Chinese emergency department: a prospective cohort study. BMJ Open. 2018 Dec 14;8(12):e024120–e024120.
82. Cattermole GN, Mak SKP, Liow CHE, Ho MF, Hung KYG, Keung KM, et al. Derivation of a prognostic score for identifying critically ill patients in an emergency department resuscitation room. Resuscitation. 2009 Sep;80(9):1000–5.
83. Heitz CR, Gaillard JP, Blumstein H, Case D, Messick C, Miller CD. Performance of the maximum modified early warning score to predict the need for higher care utilization among admitted emergency department patients. J Hosp Med. 2010 Jan;5(1):E46-52.
84. Srivilaithon W, Amnuaypattanapon K, Limjindaporn C, Imsuwan I, Daorattanachai K, Dasanadeba I, et al. Predictors of in-hospital cardiac arrest within 24 h after emergency department triage: A case-control study in urban Thailand. Emerg Med Australas EMA. 2019 Mar 18.
85. Cattermole GN, Liow ECH, Graham CA, Rainer TH. THERM: the Resuscitation Management score. A prognostic tool to identify critically ill patients in the emergency department. Emerg Med J. 2014 Oct;31(10):803–7.
86. Najafi Z, Zakeri H, Mirhaghi A. The accuracy of acuity scoring tools to predict 24-h mortality in traumatic brain injury patients: A guide to triage criteria. Int Emerg Nurs. 2018 Jan;36:27–33.
87. Bartkowiak B, Snyder AM, Benjamin A, Schneider A, Twu NM, Churpek MM, et al. Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification

of Surgical Inpatients in the Postoperative Setting: Retrospective Cohort Study. Ann Surg. 2019 Jun;269(6):1059–63.
88. Kovacs C. Outreach and early warning systems for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. Int J Nurs Pract. 2016;
89. Plate JD, Peelen LM, Leenen LP, Hietbrink F. Validation of the VitalPAC Early Warning Score at the Intermediate Care Unit. World J Crit Care Med. 2018 Aug 4;7(3):39–45.
90. Sarani B. Accuracy of an expanded early warning score for patients in general and trauma surgery wards (Br J Surg 2012; 99: 192-197). Br J Surg. 2012 Feb;99(2):197–8.
91. Hollis RH, Graham LA, Lazenby JP, Brown DM, Taylor BB, Heslin MJ, et al. A Role for the Early Warning Score in Early Identification of Critical Postoperative Complications. Ann Surg. 2016 May;263(5):918–23.
92. Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Ann R Coll Surg Engl. 2006 Oct;88(6):571–5.
93. Garcea G, Ganga R, Neal CP, Ong SL, Dennison AR, Berry DP. Preoperative early warning scores can predict in-hospital mortality and critical care admission following emergency surgery. J Surg Res. 2010 Apr;159(2):729–34.
94. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007 Feb;35(2):402–9.
95. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation. 2010 Aug;81(8):932–7.
96. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013 Apr;84(4):465–70.
97. Rasmussen LJH, Ladelund S, Haupt TH, Ellekilde GE, Eugen-Olsen J, Andersen O. Combining National Early Warning Score With Soluble Urokinase Plasminogen Activator Receptor (suPAR) Improves Risk Prediction in Acute Medical Patients: A Registry-Based Cohort Study. Crit Care Med. 2018 Dec;46(12):1961–8.
98. Ghosh E, Eshelman L, Yang L, Carlson E, Lord B. Early Deterioration Indicator: Data-driven approach to detecting deterioration in general ward. Resuscitation. 2018 Jan;122:99–105. =29122648&site=ehost-live&scope=site
99. Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. BJA Br J Anaesth. 2007 May 22;98(6):769–74.
100. Colombo F, Taurino L, Colombo G, Amato M, Rizzo S, Murolo M et al. The Niguarda MEWS, a new and refined tool to determine criticality and instability in Internal Medicine Ward and Emergency Medicine Unit. Ital J Med. 2017;11(3):310–7.
101. Abbott TEF, Torrance HDT, Cron N, Vaid N, Emmanuel J. A single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med. 2016 Nov;35:78–82.
102. Wheeler I, Price C, Sitch A, Banda P, Kellett J, Nyirenda M, et al. Early warning scores generated in developed healthcare settings are not sufficient at predicting early mortality in Blantyre, Malawi: a prospective cohort study. PLoS One. 2013;8(3):e59830–e59830.

103. Graziadio S, O'Leary RA, Stocken DD, Power M, Allen AJ, Simpson AJ, et al. Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study. BMJ Open. 2019 Feb 22;8(11):e020337–e020337.

Table S1. Risk of bias assessment results

TOOL	Study	Validation	Quality: Risk of bias	Quality: Applicability
PROBAST	Kellett, 2012 (S1)	External	low	low
	Kim, 2017 (S2)	External	Unclear	unclear
	Bozkurt, 2015 (S3)	External	High	high
	Seak, 2017 (S4)	External	High	high
	Hu, 2016 (S5)	Internal	Unclear	high
	Liljehult, 2016 (S6)	External	Unclear	high
	Mulligan, 2010 (S7)	External	High	high
	Cooksley, 2012 (S8)	External	Unclear	unclear
	Vaughn, 2018 (S9)	External	High	high
	Young, 2014 (S10)	External	High	high
	von Lilienfeld-Toal, 2007 (S11)	External	Unclear	high
	Pedersen, 2018 (S12)	External and Internal	low	low
	Forster, 2018 (S13)	External	low	low
	Pimentel, 2018 (S14)	External	low	unclear
	Sbiti-rohr, 2016 (S15)		Unclear	high
	Brabrand, 2017 (S16)	External	Unclear	unclear
	Jo, 2016 (S17)	External	High	high
	Barlow, 2007 (S18)	External	low	unclear
	Bilben, 2016 (S19)	External	Unclear	unclear
	Delahanty, 2019 (S20)	Internal	low	low
	Redfern, 2018 (S21)	External	low	low
	Churpek, 2017 (S22)	External	High	high
	Faisal, 2019 (S23)	External	low	low
	Churpek 2017 (S24)	External	low	low
	Henry, 2015 (S25)	Internal	low	low
	Brink 2019 (S26)	External	Unclear	unclear
	de Groot, 2014 (S27)	External	Unclear	unclear
	Corfield, 2014 (S28)	External	low	low
	Goulden, 2018 (S29)	External	Unclear	unclear
	Khwannimit, 2019 (S30)	External	Unclear	unclear
	Ghanem-Zoubi, 2011 (S31)	External	Unclear	unclear
	Saeed, 2019 (S32)	Internal	Unclear	unclear
	Innocenti, 2018 (S33)	External	Unclear	unclear
	Camm, 2018 (S34)	External	Unclear	unclear
	Tirotta, 2017 (S35)	External	Unclear	unclear
	Pong, 2019 (S36)	Internal	Unclear	unclear
	Prabhakar, 2019 (S37)	Internal	Unclear	unclear
	Martino, 2018 (S38)	External	Unclear	unclear
	Vorwerk, 2009 (S39)	External	Unclear	unclear
	Qin, 2017 (S40)	External	Unclear	unclear
	Schmedding, 2019 (S41)	External	Unclear	unclear
	Albur, 2016 (S42)	External	Unclear	unclear
	Cildir, 2013 (S43)	External	Unclear	unclear
	Chiew, 2019 (S44)	External	Unclear	unclear

	Samsudin, 2018 (S45)	Internal	Unclear	unclear
	Chang, 2018 (S46)	External	Unclear	high
	Geier, 2013 (S47)	External	Unclear	unclear
	Asiimwe, 2015 (S48)	Internal	Unclear	unclear
	Hung, 2017 (S49)	External	Unclear	high
	Garcea, 2006 (S50)	External	Unclear	high
	Yoo, 2015 (S51)	External	Unclear	unclear
	Siddiqui, 2017 (S52)	External	Unclear	unclear
	Calvert, 2016 (S53)	Internal	low	unclear
	Awad, 2017 (S54)	Internal	low	low
	Reini, 2012 (S55)	External	Unclear	unclear
	Chen, 2019 (S56)	External	Unclear	high
	Baker, 2015 (S57)	External	Unclear	unclear
	Gök, 2019 (S58)	External	low	unclear
	Moseson, 2014 (S59)	External	Unclear	unclear
	Jo, 2013 (S60)	External	Unclear	unclear
	Kwon, 2018 (S61)	External and Internal	Unclear	unclear
	Usman, 2019 (S62)	External	High	high
	Jang, 2019 (S63)	Internal	low	low
	Wei, 2019 (S64)	External	High	high
	Lee, 2019 (S65)	Internal	low	low
	Singer, 2017 (S66)	External	Unclear	unclear
	Eick, 2015 (S67)	External	Unclear	unclear
	Bulut, 2014 (S68)	External	Unclear	unclear
	Kivipuro, 2018 (S69)	External	Unclear	unclear
	Eckart, 2019 (S70)	External	Unclear	unclear
	Ho, 2013 (S71)	External	Unclear	unclear
	Skitch, 2018 (S72)	External	Unclear	unclear
	Liu, 2014 (S73)	Internal	low	unclear
	Dundar, 2016 (S74)	External	Unclear	high
	Yuan, 2018 (S75)	External	Unclear	high
	Naidoo, 2014 (S76)	External	Unclear	unclear
	Liu, 2015 (S77)	External	low	unclear
	So, 2015 (S78)	External	Unclear	unclear
	Dundar, 2019 (S79)	External	Unclear	high
	Lam, 2006 (S80)	External	Unclear	unclear
	Xie, 2018 (S81)	External	Unclear	unclear
	Cattermole, 2009 (S82)	Internal	Unclear	unclear
	Heitz, 2010 (S83)	External	High	unclear
	Srivilaithon, 2019 (S84)	Internal	Unclear	unclear
	Cattermole, 2014 (S85)	External	Unclear	unclear
	Najafi, 2018 (S86)	External	Unclear	high
	Bartkowiak, 2019 (S87)	External	Unclear	unclear
	Kovacs, 2016 (S88)	External	low	low
	Plate, 2018 (S89)	External	low	low
	Sarani, 2012 (S90)	External	low	low
	Hollis, 2016 (S91)	External	Unclear	unclear

	Gardner-Thorpe 2006 (S92)	External	Unclear	unclear
	Garcea, 2010 (S93)	External	High	high
	Cuthbertson, 2007 (S94)	External	High	unclear
	Prytherch, 2010 (S95)	Internal	low	low
	Smith, 2013 (S96)	External	low	low
	Ramsussen, 2018 (S97)	External	Unclear	unclear
	Ghosh, 2018 (S98)	Internal	low	low
	Duckitt, 2007 (S99)	Internal	low	low
	Colombo, 2017 (S100)	External	High	high
	Abbot, 2016 (S101)	External	High	high
	Wheeler, 2013 (S102)	External	Unclear	unclear
	Graziadio, 2019 (S103)	External	Unclear	unclear

Table S2. Early warning scores used in studies of patients' sub-populations and settings

		HR	SBP	RR	Temp	APVU/ LOC	O2 Sat	Supp O2	Urine OP	Other
Kellett, 2012 (S1)	VIEWS	✓	✓	✓	✓	✗	✓	✓	✗	✗
Seak, 2017 (S4)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bozkurt, 2015 (S3)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kim, 2017 (S2)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hu, 2016 (S5)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Mulligan, 2010 (S7)	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liljehult, 2016 (S6)	EWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Cooksley, 2012 (S8)	MEWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Cooksley, 2012 (S8)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Vaughn, 2018 (S9)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Von Lilienfeld-Toal, 2007 (S11)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Young, 2014 (S10)	MEWS	✓	✓	✓	✓	✗	✗	✗	✗	✓
Barlow, 2007 (S18)	EWS	✓	✓	✓	✓	✓	✓	✗	✓	✗
Bilben, 2016 (S19)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Brabrand, 2017 (S16)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Forster, 2018 (S13)	NEWS	✓	✓	✓	✓	✓	✓	✓	✓	✗
Jo, 2016 (S16)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pedersen, 2018 (S12)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018 (S14)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pimentel, 2018 (S14)	NEWS2	✓	✓	✓	✓	✓	✓	✓	✗	✓
Sbiti-rohr, 2016 (S15)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Henry, 2015 (S25)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Innocenti, 2018 (S33)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Garcea, 2006 (S50)	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Qin, 2017 (S40)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Albur, 2016 (S42)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Asiimwe, 2015 (S48)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Brink 2019 (S26)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Camm, 2018 (S34)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Chang, 2018 (S46)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chiew, 2019 (S44)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chiew, 2019 (S44)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S22)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S22)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Churpek, 2017 (S24)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Churpek, 2017 (S24)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Cildir, 2013 (S43)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Corfield, 2014 (S28)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014 (S27)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
De Groot, 2014 (S27)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Delahanty, 2019 (S20)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Delahanty, 2019 (S20)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Faisal, 2019 (S23)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Geier, 2013 (S47)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ghanem-Zoubi, 2011 (S31)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Goulden, 2018 (S29)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Hung, 2017 (S49)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Khwannimit, 2019 (S30)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Khwannimit, 2019 (S30)	SOS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Khwannimit, 2019 (S30)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Martino, 2018 (S30)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Pong, 2019 (S36)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Pong, 2019 (S36)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019 (S37)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Prabhakar, 2019 (S37)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Redfern, 2018 (S21)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Saeed, 2019 (S32)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Samsudin, 2018 (S45)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Samsudin, 2018 (S45)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Schmedding, 2019 (S41)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Siddiqui, 2017 (S52)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Tirotta, 2017 (S35)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Vorwerk, 2009 (S39)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Yoo, 2015 (S51)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Awad, 2017 (S54)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Baker, 2015 (S57)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Calvert 2016 (S53)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Gök, 2019 (S58)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Chen, 2019 (S56)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Jo, 2013 (S60)	HOTEL	✗	✓	✗	✓	✓	✓	✗	✗	✓
Jo, 2013 (S60)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Moseson, 2014 (S59)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Reini, 2012 (S55)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bulut, 2014 (S68)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2009 (S82)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Cattermole, 2014 (S85)	WORTHING	✓	✓	✓	✓	✓	✓	✗	✗	✗

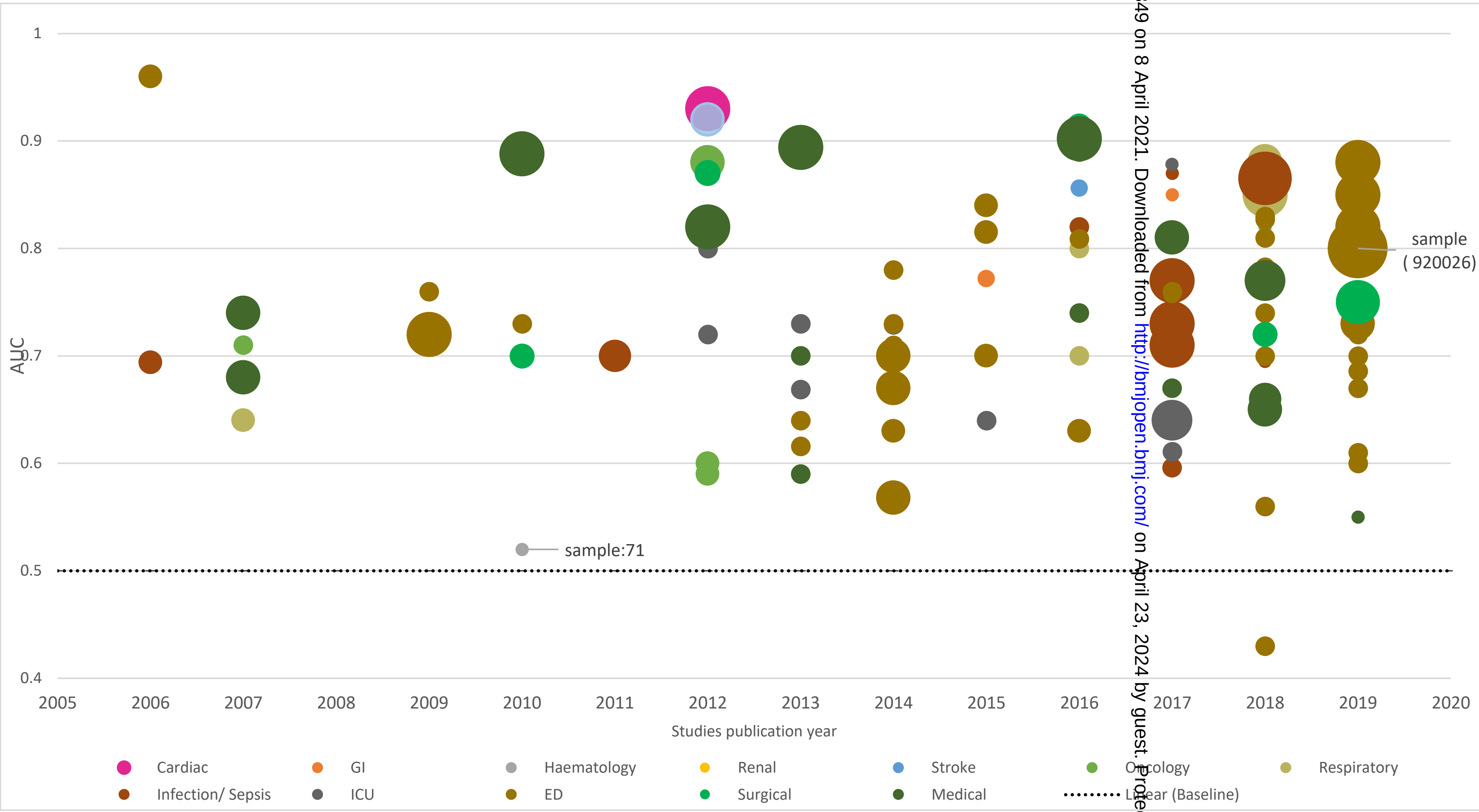
Cattermole, 2014 (S85)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Cattermole, 2014 (S85)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Heitz, 2010 (S83)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016 (S74)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Dundar, 2016 (S74)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Dundar, 2019 (S79)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eckart, 2019 (S70)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Eick, 2015 (S67)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu, 2015 (S77)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Liu, 2015 (S77)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ho, 2013 (S71)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Jang, 2019 (S63)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Kivipuro, 2018 (S69)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Kwon, 2018 (S61)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Liu, 2014 (S73)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019 (S65)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Lee, 2019 (S65)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lee, 2019 (S65)	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Naidoo, 2014 (S76)	TREWS	✓	✓	✓	✓	✓	✗	✗	✗	✓
Najafi, 2018 (S86)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Singer, 2017 (S66)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Skitch, 2018 (S72)	HEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Skitch, 2018 (S72)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
So, 2015 (S78)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Srivilaithon, 2019 (S84)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Lam, 2006 (S80)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Usman, 2019 (S62)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018 (S75)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Yuan, 2018 (S75)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Wei, 2019 (S64)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Xie, 2018 (S81)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Bartkowiak, 2019 (S87)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Bartkowiak, 2019 (S87)	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Cuthbertson, 2007 (S94)	EWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Cuthbertson, 2007 (S94)	MEWS	✓	✓	✓	✓	✗	✓	✗	✗	✗
Garcea, 2010 (S50)	EWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Gardner-Thorpe 2006 (S92)	MEWS	✓	✓	✓	✓	✓	✗	✗	✓	✗
Hollis, 2016 (S91)	EWS	✓	✓	✓	✓	✓	✓	✗	✗	✗
Kovacs, 2016 (S88)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Plate, 2018 (S89)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗

Sarani, 2012 (S90)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Abbott, 2016 (S101)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Duckitt, 2007 (S99)	WPC	✓	✓	✓	✓	✓	✓	✗	✗	✗
Duckitt, 2007 (S99)	EWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Colombo, 2017 (S100)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Ghosh, 2018 (S98)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Ghosh, 2018 (S98)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗
Graziadio, 2019 (S103)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Prytherch, 2010 (S95)	VIEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Ramsussen, 2018 (S97)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Smith, 2013 (S96)	NEWS	✓	✓	✓	✓	✓	✓	✓	✗	✗
Wheeler, 2013 (S102)	Hotel	✓	✗	✓	✗	✓	✓	✗	✗	✓
Wheeler, 2013 (S102)	MEWS	✓	✓	✓	✓	✓	✗	✗	✗	✗

Total	133

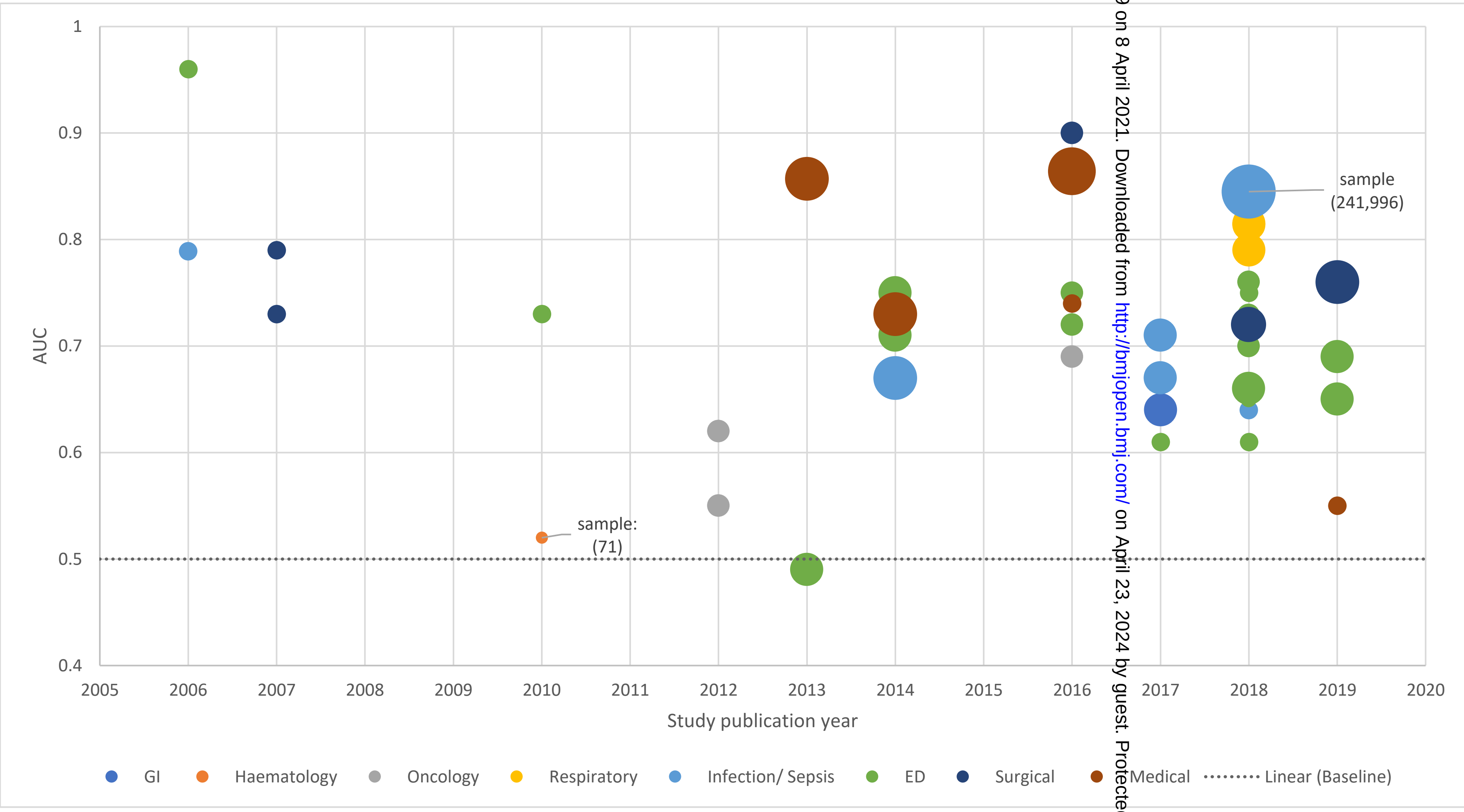
<u>Abbreviations:</u> HR: heart rate, SBP: systolic blood pressure, RR: respiratory rate, Temp: temperature, AVPU/LOC: alert, verbal response, physical response, unresponsive score or level of consciousness, O2 sat: Oxygen saturation, Supp O2: supplemental oxygen, Urine OP: urine output, Other: other parameters, i.e., blood biomarkers. VIEWS: Vitalpack early warning score, MEWS: modified early warning score, EWS: early warning score, NEWS: national early warning score, NEWS2: national early warning score 2, SOS: Search Out Severity score, Worthing: Worthing physiological scoring system, HOTEL: Hypotension, Oxygen saturation, Temperature, ECG abnormality, Loss of independence score, TREWS: Triage in Emergency department Early Warning Score, HEWS: Hamilton early warning score.

Figure S1. Predictive performance of early warning scores for mortality in studies from 2005 to 2020 for different disease subgroups and clinical settings



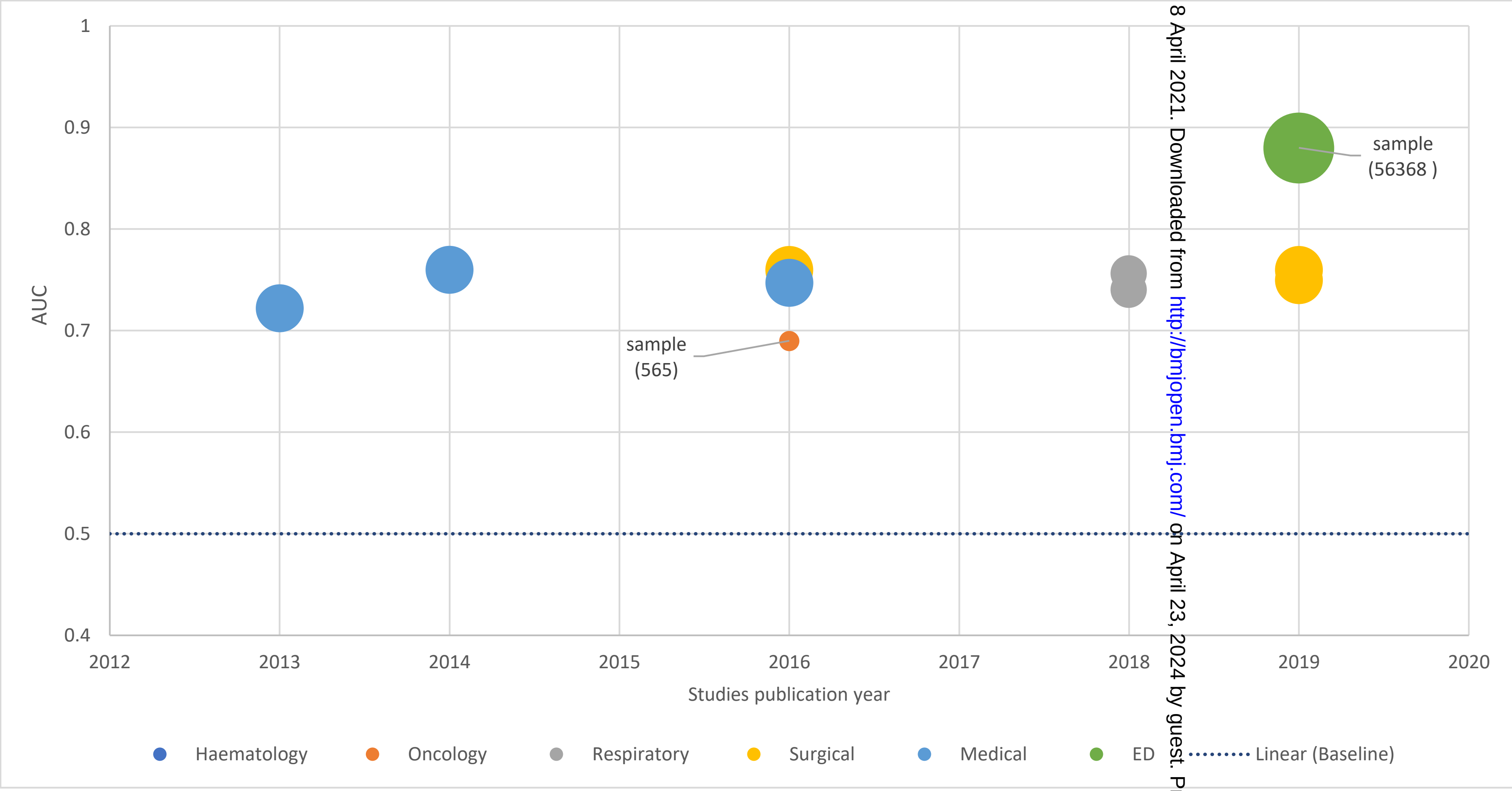
Abbreviations: *AUC: Area Under the Curve; ED: Emergency Department; GI: Gastrointestinal diseases; ICU: Intensive Care Unit. Note: Bubbles sizes represents the sample size in each study.*

Figure S2. Predictive performance of early warning scores for intensive care admission in studies from 2005 to 2020 for different disease subgroups and clinical settings



Abbreviations: *AUC: Area Under the Curve; ED: Emergency Department; GI: Gastrointestinal diseases. Note: Bubbles sizes represents the sample size in each study.*

Figure S3. Predictive performance of early warning scores for cardiac arrest in studies from 2012 to 2020 for different disease subgroups and clinical settings



<u>Abbreviations:</u> AUC: Area Under the Curve; ED: Emergency Department; GI: Gastrointestinal diseases. Note: Bubbles sizes represents the sample size in each study.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	1
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4; Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	4
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	4
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5; figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	5
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	5-6
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	5; Table1; Table 2; Supplementary.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	5-6
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	5-6
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Supplementary
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	7
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	7-8
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	8
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	8

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.