Resuscitation 132 (2018) 101–111

Contents lists available at ScienceDirect

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Review

Can early warning scores identify deteriorating patients in pre-hospital settings? A systematic review

Rita Patel[a,b,*], Manjula D. Nugawela[a,b], Hannah B. Edwards[a,b], Alison Richards[a,b], Hein Le Roux[c,d], Anne Pullyblank[d,e], Penny Whiting[a,b]

[a] The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West), University Hospitals Bristol NHS Foundation Trust, Whitefriars, Lewins Mead, Bristol, BS1 2NT, UK
[b] Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK
[c] Gloucestershire Clinical Commissioning Group, Sanger House, 5220 Valiant Court, Gloucester Business Park, Brockworth, Gloucester, GL3 4FE, UK
[d] West of England Academic Health Science Network (WEAHSN), South Plaza, Marlborough Street, Bristol, BS1 3NX, UK
[e] North Bristol NHS Trust, Trust Headquarters, Southmead Hospital, Southmead Road, Westbury-on-Trym, Bristol, BS10 5NB, UK

ARTICLE INFO

Keywords:
Deteriorating patients
Critical care
Pre hospital setting
Early warning score
Track and trigger system

ABSTRACT

Objective: To evaluate the effectiveness and predictive accuracy of early warning scores (EWS) to predict deteriorating patients in pre-hospital settings.
Methods: Systematic review. Seven databases searched to August 2017. Study quality was assessed using QUADAS-2. A narrative synthesis is presented.
Eligibility: Studies that evaluated EWS predictive accuracy or that compared outcomes in populations that did or did not use EWS, in any pre-hospital setting were eligible for inclusion. EWS were included if they aggregated three or more physiological parameters.
Results: Seventeen studies (157,878 participants) of predictive accuracy were included (16 in ambulance service and 1 in nursing home). AUCs ranged from 0.50 (CI not reported) to 0.89 (95%CI 0.82, 0.96). AUCs were generally higher (>0.80) for prediction of mortality within short time frames or for combination outcomes that included mortality and ICU admission. Few patients with low scores died at any time point. Patients with high scores were at risk of deterioration. Results were less clear for intermediate thresholds (≥ 4 or 5). Five studies were judged at low or unclear risk of bias, all others were judged at high risk of bias.
Conclusions: Very low and high EWS are able to discriminate between patients who are not likely and those who are likely to deteriorate in the pre-hospital setting. No study compared outcomes pre- and post-implementation of EWS so there is no evidence on whether patient outcomes differ between pre-hospital settings that do and do not use EWS. Further studies are required to address this question and to evaluate EWS in pre-hospital settings.

Introduction

When patients are acutely ill, it can be challenging to identify those likely to deteriorate who need urgent intervention. Acute changes in physiological parameters (heart rate, systolic blood pressure, respiratory rate, oxygen saturation, level of consciousness and temperature) occur before deterioration. Early detection of changes in these parameters provides an opportunity to initiate a prioritised clinical response and prevent serious health outcomes [1–3]. Early Warning Scores (EWS) were developed to reliably and systematically identify patients who are deteriorating. In the UK there are different EWS systems in use, mostly in hospital settings [4–6]. In 2012 the Royal College of Physicians (RCP) developed a National Early Warning Score (NEWS) to facilitate a standardised approach [3].

Recently, NEWS has been advocated for use in pre-hospital settings in the UK, such as general practice, mental health services, and ambulance services [3,7]. Use of pre-hospital EWS is controversial due to lack of evidence of effectiveness in these settings [8–10]. EWS systems were developed using observations from hospitalised patients where EWS are recorded over a period of time, allowing tracking of a patient's progress or deterioration. This leads to a pre-defined response designed to ensure prompt recognition and treatment of deterioration [5,11,12]. In pre-hospital settings, EWS should be used as an adjunct to clinical decision making rather than replacing it.

We conducted a systematic review on the use of EWS, comprised of at least three physiological parameters, in pre-hospital settings. We

* Corresponding author at: NIHR CLAHRC West, University Hospitals Bristol NHS Foundation Trust, Whitefriars, Lewins Mead, Bristol, BS1 2NT, UK.
E-mail address: rita.patel@bristol.ac.uk (R. Patel).

https://doi.org/10.1016/j.resuscitation.2018.08.028
Received 3 April 2018; Accepted 26 August 2018

aimed to summarise current evidence of effectiveness and predictive accuracy of EWS in these settings.

Methods

The review followed guidance published by the Centre for Reviews and Dissemination and the Cochrane Collaboration [13,14], and was registered on the PROSPERO database (id CRD42017059305) [15]. We followed the new PRISMA guidelines for DTA reviews when reporting results [16].

Study identification

Embase (OvidSP), Medline (OvidSP), Medline In-Process Citations & Daily Update (OvidSP), PsycINFO (OvidSP), CINAHL (Cumulative Index to Nursing and Allied Health Literature) (EBSCO), Science Citation Index (SCI) (Web of Science), Cochrane Database of Systematic Reviews (Wiley) and the Database of Abstracts of Reviews of Effects (DARE) (CRD, York) were searched from inception to August 2017. Search strategies combined terms for early warning scores with terms for pre-hospital setting (Medline search strategy available as a Web Appendix). A Google search for 'grey literature' was carried out in November 2016 using keywords 'early warning scor*, 'early warning system* and 'EWS'. Websites of organisations identified as likely to have further relevant material were then also searched. Searches were not limited by language, date or publication status (unpublished or published). References of included studies were screened to identify additional relevant studies.

Study selection

We included studies that provided information on accuracy of EWS for predicting outcomes or compared outcomes in a population that used EWS with one that did not, in any pre-hospital setting. The review was restricted to studies of adults over 16; studies of pregnant women were excluded. EWS were included if they aggregated three or more physiological parameters to produce a single score. Scores designed for specific conditions such as sepsis or acute kidney injury were excluded. Studies had to report data on serious health outcomes such as mortality, cardiac arrest, intensive care unit (ICU) admission or the length of hospital/ICU stay. Accuracy studies were required to report sufficient data to construct 2 × 2 tables of predictive performance or a measure of accuracy (e.g. area under curve, sensitivity or specificity). Search results and full text articles were independently assessed by two reviewers; disagreements were resolved through consensus or referral to a third reviewer.

Data extraction and quality assessment

We extracted data on inclusion/exclusion criteria, type of pre-hospital setting, patient demographics, EWS, EWS threshold, outcome(s) assessed, 2 × 2 tables of EWS performance and reported estimates of accuracy. Area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and positive and negative likelihood ratios were either extracted directly from the papers or calculated from 2 × 2 tables, where available. To minimise bias and errors, data extraction was performed by one reviewer and checked by a second.

Study quality was assessed using the QUADAS-2 tool designed to assess quality of primary studies for risk of bias and concerns regarding applicability [17]. To be judged at low concerns regarding applicability of the patient population, studies had to enrol a general population of patients but exclude trauma patients. We were most interested in the NEWS score and so to be judged at low concerns regarding applicability of the index test the study had to have evaluated NEWS. As we were interested in a broad range of outcomes (reference standards) we did not assess the applicability of the reference standard.

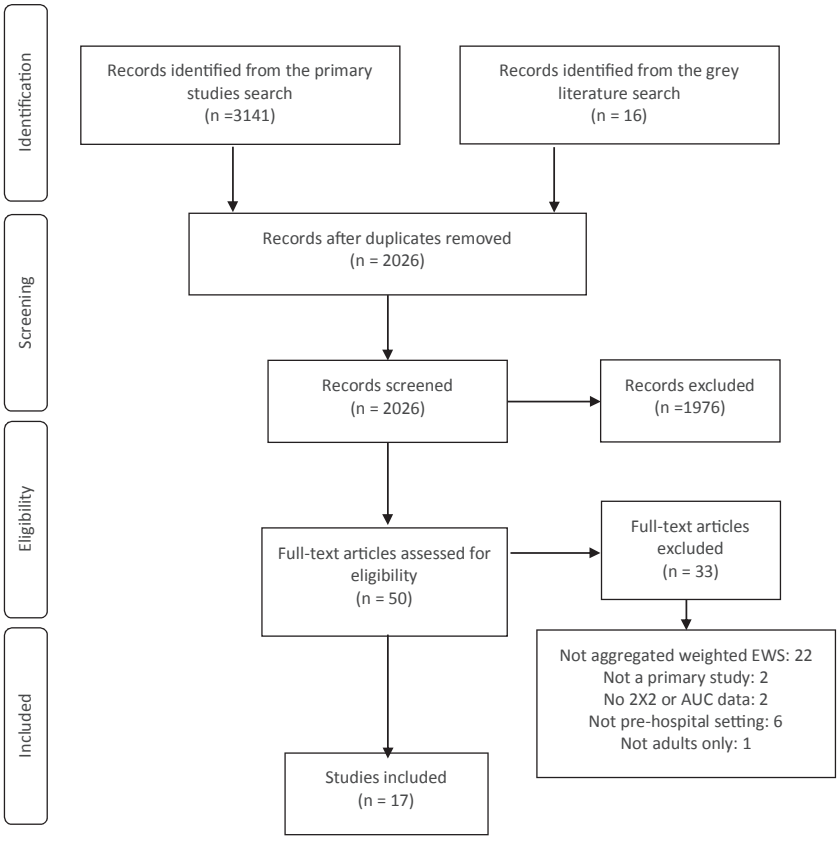

Fig. 1. PRISMA Flow Diagram for systematic review of Early Warning Scores in pre-hospital settings.

Data synthesis

Due to differences in outcomes and EWS assessed it was not possible to conduct a meta-analysis, hence a narrative synthesis is presented. We grouped studies by outcome assessed and EWS evaluated. We used forest plots to display estimates of AUC and of sensitivity and specificity together with 95% confidence intervals, separately for low, medium and high-risk thresholds – lower EWS scores suggest lower risk. Major methodological problems or biases that affected the studies were considered. All analyses were performed in Stata 14 (StataCorp LP, College Station, Texas, USA).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing this report.

Patient involvement

This review was part of a larger project to evaluate early warning scores in pre-hospital settings. The NIHR CLAHRC West PPI health systems panel provided advice on early findings of the project and will help with implementation of findings.

Results

The searches identified 2026 hits of which 50 were considered potentially relevant and screened as full text articles (Fig. 1). No studies compared pre-hospital settings that used EWS with settings that did not. Seventeen studies (157,878 participants) using EWS for predicting outcomes were included.

One study was conducted in a community nursing home; [18] all other studies evaluated EWS based on readings taken by ambulance staff. Five studies were available only as abstracts [18–22]. Thirteen studies were retrospective chart reviews/record linkage where EWS were calculated specifically for the study based on routinely collected data. One of these used a nested case-control design where EWS were only calculated for a subset of patients based on whether they had experienced an outcome of interest (ward, or ICU admission, death in ED or discharge) [23]. Three studies were prospective cohort studies and one used a case-control design. Studies were conducted in the UK [8,9,18,23–25], USA, [19,21,26,27], Ireland [28], Germany [29], Hong Kong [30], China [31], Finland [20], Iran [32], and New Zealand [22]. Sample sizes ranged from 101 [19] to over 57,647 [26]. EWS evaluated included NEWS (5 studies) [19–21,23,24], MEWS (5 studies) [8,18,29–31], PMEWS (3 studies) [9,25,32], prehospital risk score (2 studies) [26,27], nzPHEWS (1 study) [22], and Ph-ViEWS (1 study) [28]. Table 1 summarises the parameters included in each EWS, the maximum score and recommended thresholds for action. Mean or median age, where reported, ranged from 50 to 83 years and the proportion of men from 37 to 63%. Timing of EWS was not well reported. Several studies reported only that this was based on ambulance or pre-hospital service data. Six studies reported that the first set of pre-hospital values was used [8,22,25–27,30], and one that data were recorded "on the scene" [31]. A further study took two sets of EWS data, one calculated based on the first set of observations recorded by the ambulance personnel and one based on the last set of data before admission to ED [23]. An overview of included studies is provided in Table 2.

All studies had limitations in study quality (Table 3). Overall, one study was judged low risk of bias [8] and four as unclear risk of bias; [20,22,25,30], all others were judged at high risk of bias. All studies had some concerns regarding applicability.

Six studies were judged at high risk of bias for patient selection [9,18,21,23,31,32]. Four excluded certain patients potentially introducing bias [9,18,31,32]. Restriction included those who had a final working diagnosis of a respiratory condition [9], patients transferred to hospital (outcome was whether transfer to hospital was appropriate) [32], patients who accepted treatment [31] and patients with single

Table 1
Parameters used in studies of predictive ability of early warning scores in pre-hospital settings.

Name of scoring system	Parameters used in the scoring system							Maximum possible score	Threshold for action
	Heart rate	Respiratory rate	Systolic blood pressure	Temperature	Oxygen saturation	Any supplemental O_2	Level of consciousness[a]	Other, specify	
NEWS	√	√	√	√	√	√	√		20	5–6 medium clinical risk ≥7 high clinical risk
MEWS	√	√	√	√			√		14	≥4 or ≥3
PMEWS	√	√	√	√	√		√	Age, social isolation, chronic disease, performance status	19	≥4
Ph-ViEWS	√	√	√	√	√	√	√		21	Review & alert: 3 (4-hourly) 5-6 (hourly) ≥7 (½-hourly)
Prehospital risk score [26]	√	√	√		√		√	Gender, age, location e.g. nursing home	8	3 moderate risk 4-8 high risk
nzPHEWS[b]	√	√	√	√	√	√	√		20	NR

MEWS = Modified Early Warning Score; NEWS = National Early Warning Score; PMEWS = Pandemic Medical Early Warning Score [9] or physiological social score [25] or physiological-social modified early warning score [32]; Ph-ViEWS = Prehospital applied VitalPAC™ Early Warning Score; nzPHEWS = New Zealand Prehospital Early Warning Score.

[a] For example, level of consciousness AVPU (alert, verbal, pain, or unresponsive status); consciousness (clear, responded to sound, responded to pain, no respond); or neurological (alert, confused/agitated, voice, pain, unconscious); Glasgow Coma Scale score.

[b] nzPHEWS is based on New Zealand Early Warning Score but no details are reported in this abstract. We report parameters for New Zealand Early Warning Score from Health Quality & Safety Commission New Zealand [35].

Table 2
Characteristics of studies of the predictive ability of early warning scores (EWS) in pre-hospital settings.

Study and Country	EWS type	Design	Setting	Inclusion and exclusion criteria	EWS timing	n	Mean age, years	% male
Silcock [24] UK	NEWS	Retrospective chart review	Ambulance service	Consecutive patients transported to a single hospital including trauma patients. 737 had incomplete observations and were excluded	Ambulance data; no further details	1684	NR	NR
#Infinger [19] USA	NEWS	Retrospective chart review	Ambulance service	Patients transported to a level 1 trauma centre suspected of severe sepsis	NEWS applied retrospectively; timing unclear	101	NR	NR
#Pirneskoski [20] Finland	NEWS	Retrospective record linkage	Emergency Medical Services	Electronic patient record with sufficient prehospital data, including trauma patients	NEWS applied retrospectively; timing unclear	35,845	66	48
Shaw [23] UK	NEWS	Nested case-control	Ambulance service	Ambulance patients with data extracted for sample approx. 100 with each outcome, including trauma patients.	1) on ambulance arrival 2) before emergency dept	287	63	52
#Studnek [21] USA	NEWS	Retrospective case control record review	Emergency Medical Services	215 from sepsis registry and 100 controls from the Emergency Medical Services reports, unclear if includes trauma patients.	NEWS applied retrospectively; timing unclear	315	NR	NR
#Pattison [18] UK	MEWS	Retrospective chart review	Community nursing home	Residents triaged during 4-month period repeated measures excluded	MEWS used routinely by nursing home case management service, to triage [medical] visit requests.	178	83	37
Fullerton [8] England	MEWS	Retrospective chart review	Ambulance service	Consecutive adults with clinical observations recorded and not in cardiac arrest (includes trauma patients)	First set of pre-hospital values	3057	55	50
Bayer [29] Germany	MEWS	Retrospective chart review	Emergency medical services	Consecutive patients including trauma patients admitted to emergency department.	Emergency medical services data; no further details	375	71	54
Leung [30] Hong Kong	MEWS	Prospective cohort	Ambulance service	Non-trauma patients who arrived by ambulance, excluding those in cardiac arrest	First set of data recorded by ambulance crews	1493	Median 78	48
Ruan [31] China	MEWS	Prospective cohort	Pre-hospital emergency care	All patients visited and treated 'on the scene' including trauma; those with poor compliance and cooperation excluded	Recorded 'on the scene' by pre-hospital emergency care staff	10,517	53	63
Challen [25] England	PMEWS	Retrospective chart review	Ambulance service	Cases presenting with 'shortness of breath' or 'difficulty in breathing'	First set of data recorded by ambulance crews	213[a]	NR	NR
Gray [9] England	PMEWS	Retrospective chart review	Ambulance service	Cases with diagnosis of respiratory condition	Ambulance data; no further details	387	56% aged 65 +	NR
Ebrahimian [32] Iran	PMEWS	Prospective cohort	Ambulance service	Consecutive patients with internal diseases transferred to hospital, trauma and mental illness excluded	Prehospital on transfer	2157	50	45
Seymour [26] USA	Prehospital risk score	Retrospective chart review	Emergency Medical Services	Nontrauma, non–cardiac arrest patients transported to a hospital with complete data	First set of prehospital vital signs	57,647	60	47
Kievlan [27] USA	Prehospital risk score	Retrospective record linkage	Emergency Medical Services	Scene-to-hospital transports excluded cardiac arrest, trauma, burn, falls or duplicate encounters	First set of prehospital vital signs; timing unclear	42,550	62	40
#Swain [22] New Zealand	nzPHEWS	Retrospective chart review	Ambulance service	Consecutive patients in each status level admitted to emergency dept likely including trauma patients.	First set of vital signs recorded	800	NR	NR
Gaumont [28] Ireland	Ph-ViEWS	Retrospective chart review	Ambulance service	High acuity (Manchester Triage System 1&2) non-traumatic patients	Ph-ViEWS applied retrospectively; no further details on timing	272	Median 66	54

EMS = Emergency Medical Services.
MEWS = Modified Early Warning Score.
NEWS = National Early Warning Score.
NR = not reported.
PMEWS = Pandemic medical early warning score [9] or physiological social score [25] or physiological-social modified early warning score [32].
Ph-ViEWS = prehospital applied VitalPAC™ Early Warning Score.
[a]In paper N = 215.
[b]Calculated from critically/non-critically ill patient characteristics given in study.
[c]Kievlan [27] and Seymour [26] use the same parameters but scored differently.
[#] = Abstract only.

Table 3
QUADAS-2 Evaluation of the risk of bias and applicability of studies of the predictive performance of EWS scores in pre-hospital settings.

Study	EWS type	Risk of Bias				Applicability Concerns	
		Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient population	Index Test
Silcock [24]	NEWS	☺	☺	☺	☹	☹	☺
#Infinger [19]	NEWS	☺	?	☺	☹	☹	☺
#Pirneskoski [20]	NEWS	?	☺	☺	?	☹	☺
Shaw [23]	NEWS	☹	☺	☺	☹	☹	☺
#Studnek [21]	NEWS	☹	?	☹	☺	☹	☺
#Pattison [18]	MEWS	☹	☺	☺	☹	☹	☹
Fullerton [8]	MEWS	☺	☺	☺	☺	☹	☹
Bayer [29]	MEWS	☺	☺	☺	☹	☹	☹
Leung [30]	MEWS	☺	☺	☺	?	☹	☹
Ruan [31]	MEWS	☹	☺	☺	☹	☹	☹
Challen [25]	PMEWS	☺	?	☺	☺	☹	☹
Gray [9]	PMEWS	☹	?	☺	☹	☹	☹
Ebrahimian [32]	PMEWS	☹	☺	☹	?	☹	☹
Seymour [26]	Prehospital risk score	☺	☺	☺	☹	☹	☹
Kievlan [27]	Prehospital risk score	☺	☺	☺	☹	☹	☹
#Swain [22]	nzPHEWS	?	☺	☺	?	☹	☹
Gaumont [28]	Ph-ViEWS	☺	☺	☺	☹	☹	☹

☺Low Risk ☹High Risk ?Unclear Risk.
=Abstract only.

EWS measures only (i.e. repeated measures excluded) [18]. One study used a case-control design [21] and one selected a sample dependant on outcome [23].

All studies had high concerns regarding applicability of the patient population. Three studies were restricted to patients with respiratory problems [9,25], or sepsis registration [21]. Seven included trauma patients. [8,20,22–24,29,31]. One was restricted to patients with "internal pathology" (not further defined) and excluded those suspected to be mentally ill [32], three excluded cardiac arrest and other conditions [26,27,30]. Two studies were restricted to high risk patients –with suspected severe sepsis [19], or with Manchester triage system gradings of level 1 or 2- who would be expected to have higher EWS [28]. One was a nursing home population rather than the general population [18].

Thirteen studies were judged at low risk of bias for interpretation of the index test. The other four were judged at unclear risk of bias as no information was available on whether the person interpreting the EWS was blinded to the outcome [9,19,21,25]. Only five studies evaluated NEWS and were considered to have low concerns regarding applicability of the index test; [19–21,23,24] all other studies were judged as high concerns regarding applicability.

Most studies were judged at low risk of bias for how the reference standard (outcome) was assessed. One study was judged as high risk of bias as the outcome was subjective (need for hospital transfer) and assessed by a single specialist [32]. In the case-control study in which sepsis was the target condition, it was not clear that all the control group were sepsis free [21].

Nine studies were judged at high risk of bias for the flow and timing domain because a large proportion (18%–31%) of eligible patients were not included in the analysis. An additional study was judged as high risk of bias as data were collected prospectively as part of routine care and so it was considered likely that any EWS would be acted on before outcomes were assessed [18]. One was judged as unclear risk of bias as the number of patients excluded due to incomplete records was not reported [30]. Three were judged as unclear as there was no information on whether the EWS score was acted on before the outcome was assessed.

Outcomes assessed were: death within various time frames; ICU admission; adverse event within 24 h or requiring lifesaving intervention; critical illness; appropriate emergency department attendance; hospital admission; need for hospital transfer; or sepsis (Table 4). Overall AUCs ranged from 0.5 (no CI reported) suggesting poor discriminatory performance to 0.89 (95% CI 0.82, 0.96) suggesting very good discriminatory performance (Fig. 2). AUCs were generally higher (> 0.80) for prediction of mortality within short time frames (< 24 or < 48 h) or for combination outcomes that included both mortality and admission to ICU. However, one study that evaluated the predictive accuracy of NEWS for mortality within 24 h reported a much lower AUC of 0.5 [20]. This study was only available as an abstract but there were no clear differences between this study and other studies to explain this discrepancy in findings. Estimates of the AUC were also lower (0.60–0.80) for prediction of outcomes such as admission to hospital or ICU, and appropriate ED attendance. The study in the community nursing home reported lower AUCs than studies conducted in the ambulance service. This study evaluated mortality at 7 days, 30 days and 90 days with AUCs ranging from 0.53 to 0.63 [18].

Studies that evaluated EWS assessed by ambulance service staff reported estimates of sensitivity and specificity (or data from which these could be calculated) for different thresholds. We grouped these into low (for ruling out deterioration), intermediate, and high (for ruling in deterioration). Seven studies reported estimates of sensitivity and specificity for very low thresholds (≥1) (Fig. 3a). Estimates of sensitivity from these studies ranged from 94 to 100%, the two studies that assessed death within 48 h both reported estimates of sensitivity of 100%. This means that none of the patients that died within 48 h had an EWS of 0. A very small number of patients with an EWS of 0 died within 30 days, were admitted to ICU or had a critical illness. However, estimates of specificity were extremely low ranging from 0 to 13% meaning that many of the patients who were not likely to deteriorate had scores ≥1.

Estimates of sensitivity and specificity for the ability of EWS based on a "medium risk" threshold were available for ten studies (Fig. 3b). Estimates of sensitivity ranged from 37% (prediction of critical illness; specificity 92%) to 87% (death in ED; specificity 60%). Estimates of

Table 4
Sensitivity and specificity of EWS scores in pre-hospital settings.

Study	EWS type	N	Mortality timing	AUC (95% CI)	Threshold	% Sensitivity (95% CI)	% Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)
Death									
Shaw [23]	NEWS	287	In ED	NR	> = 5	87 (69, 96)	60 (54, 66)	2.16 (1.76, 2.65)	0.22 (0.09, 0.56)
					> = 7	63 (44, 80)	83 (78, 88)	3.79 (2.57, 5.57)	0.44 (0.27, 0.71)
‡Silcock [24]	NEWS	1684	24 hours	0.86 (0.69, 1)*					
Pirneskoski [20]	NEWS	35,845	24 hours	0.5*					
‡Leung [30]	MEWS	1493	24 hours	0.81 (0.72, 0.9)	> = 1	100 (75, 100)	0 (0, 1)	1.00 (1.00, 1.00)	–
					> = 4	62 (32, 86)	84 (82, 86)	3.81 (2.44, 5.95)	0.46 (0.23, 0.91)
Gaumont [28]	Ph-Views	272	24 hours		> = 7	78 (40, 97)	59 (53, 65)	1.89 (1.30, 2.76)	0.38 (0.11, 1.29)
‡Silcock [24]	NEWS	1684	48 hours	0.87 (0.75, 0.98)*	> = 1	100 (77, 100)*	13 (11, 15)*		
					> = 5	79 (49, 95)*	74 (71, 76)*		
					> = 7	71 (42, 92)*	90 (88, 92)*		
‡Silcock [24]	NEWS	1684	7 days	0.8 (0.7, 0.89)*					
‡Pattison [18]	MEWS	178	7 days	0.63 (0.4, 0.87)*					
‡Silcock [24]	NEWS	1684	14 days	0.79 (0.71, 0.86)*					
‡Silcock [24]	NEWS	1684	30 days	0.74 (0.66, 0.82)*	> = 1	98 (87, 100)*	13 (12, 15)*		
					> = 5	55 (39, 70)*	74 (72, 77)*		
					> = 7	40 (26, 57)*	91 (89, 92)*		
‡Pattison [18]	MEWS	178	30 days	0.57 (0.43, 0.7)*					
Ruan [31]	MEWS	10,517	90 days	0.88 (0.87, 0.9)	> = 1	100 (100, 100)	1 (0, 1)	1.01 (1.00, 1.01)	–
					> = 4	80 (77, 82)	79 (79, 80)	3.86 (3.66, 4.06)	0.26 (0.22, 0.29)
‡Pattison [18]	MEWS	178	90 days	0.53 (0.45, 0.66)*					
Death in ED/Admitted to ICU									
‡Silcock [24]	NEWS	1684		0.89 (0.82, 0.96)*					
Death within 48 hours/Admitted to ICU									
‡Silcock [24]	NEWS	1684		0.82 (0.73, 0.99)*	> = 1	97 (83,100)*	13 (12, 15)*		
					> = 5	77 (58, 90)*	75 (72, 77)*		
					> = 7	53 (34, 72)*	91 (89, 92)*		
Death/Admitted to ICU									
Shaw [23]	NEWS	287		NR	> = 5	86 (77, 93)	73 (66, 79)	3.19 (2.5, 4.07)	0.19 (0.11, 0.32)
					> = 7	56 (45, 67)	94 (89, 96)	8.66 (4.96, 15.13)	0.47 (0.37, 0.59)
Death/Admitted to ICU/ward									
Shaw [23]	NEWS	287		NR	> = 5	59 (51, 66)	81 (72, 88)	3.1 (2.03, 4.72)	0.51 (0.42, 0.62)
					> = 7	32 (25, 39)	97 (91, 99)	10.52 (3.38, 32.7)	0.71 (0.64, 0.78)
Admitted to ICU									
‡Silcock [24]	NEWS	1684		0.77 (0.66, 0.89)*	> = 1	94 (71, 1)*	13 (11, 15)*		
					> = 5	76 (50, 93)*	74 (72, 76)*		
					> = 7	41 (18, 67)*	90 (88, 92)*		
Adverse event within 24 h									
Fullerton [8]	MEWS	3057			> = 4	57 (45, 68)	89 (87, 90)	4.93 (3.95, 6.15)	0.49 (0.38, 0.63)
Adverse event requiring life-saving intervention									
‡Leung [30]	MEWS	1493		0.72 (0.69, 0.75)	> = 1	100 (99, 100)	0 (0, 1)	1.00 (1.00, 1.01)	–
					> = 4	38 (33, 44)	89 (88, 91)	3.62 (2.92, 4.5)	0.69 (0.63, 0.75)
Critical illness #									
Seymour [26]	Prehospital risk score	57,647		0.77 (0.76, 0.78)	> = 1	98 (97, 98)	17 (17, 17)	1.18 (1.17, 1.19)	0.12 (0.10, 0.16)
					> = 2	73 (72, 75)	70 (69, 70)	2.42 (2.36, 2.48)	0.38 (0.36, 0.41)
					> = 3	45 (43, 47)	91 (91, 91)	5.05 (4.82, 5.29)	0.61 (0.59, 0.62)
Kievlan [27]	Prehospital risk score	42,550		0.73 (0.72, 0.74)	> = 1	98 (97, 98)	17 (17, 17)	1.18 (1.17, 1.19)	0.14 (0.11, 0.19)
					> = 2	63 (61, 65)	73 (72, 73)	2.31 (2.22, 2.39)	0.51 (0.48, 0.54)
					> = 3	37 (35, 39)	92 (92, 92)	4.71 (4.40, 5.04)	0.68 (0.66, 0.71)
Appropriate ED attendance									
‡Challen [25]	PMEWS	213[a]		0.71 (0.64, 0.78)	> = 1	100 (97, 100)	1 (0, 7)	1.01 (0.99, 1.04)	–

(*continued on next page*)

Table 4 (*continued*)

Study	EWS type	N	Mortality timing	AUC (95% CI)	Threshold	% Sensitivity (95% CI)	% Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)
					> = 4	81 (74, 87)	43 (31, 55)	1.42 (1.15, 1.75)	0.44 (0.29, 0.68)
Hospital admission									
Gray [9]	PMEWS	300		0.8 (0.74, 0.85)	> = 1	100 (96, 100)	0 (0, 3)	1.00 (1.00, 1.01)	–
					> = 4	86 (77, 92)	41 (34, 48)	1.45 (1.26, 1.67)	0.35 (0.21, 0.58)
‡Challen [25]	PMEWS	213[a]		0.7 (0.63, 0.77)*					
#Swain [22]	nzPHEWS	200		0.6*					
Need for hospital transfer									
Ebrahimian [32]	PMEWS	2157		0.74 (0.71. 0.77)*					
Sepsis									
Infinger [19]	NEWS	101			NR	90 (77, 97)	25 (15, 38)	1.21 (1.01, 1.45)	0.37 (0.13, 1.05)
Studnek [21]	NEWS	315			NR	76 (70, 82)*	95 (88, 98)*		
Bayer [29]	MEWS	375		0.77*	> = 4	74*	75*		

All values are calculated from reported data, except for those marked with an asterix [*] which denotes insufficient data in paper to calculate and shows values extracted from paper.
‡More than one outcome reported in this paper.
[a]In paper N = 215.
#Critical illness during hospitalization, defined as severe sepsis, delivery of mechanical ventilation, or death during hospitalization [26] or defined as an intensive care unit stay with delivery of organ support (mechanical ventilation or vasopressor use) [27].
AUC = Area under the curve.
ED = Emergency Department.
ICU = Intensive Care Unit.
MEWS = Modified Early Warning Score.
NEWS = National Early Warning Score.
NR = not reported.
PMEWS = Pandemic medical early warning score [9] or physiological social score [25] or physiological-social modified early warning score [32].
Ph-ViEWS = prehospital applied VitalPAC™ Early Warning Score.

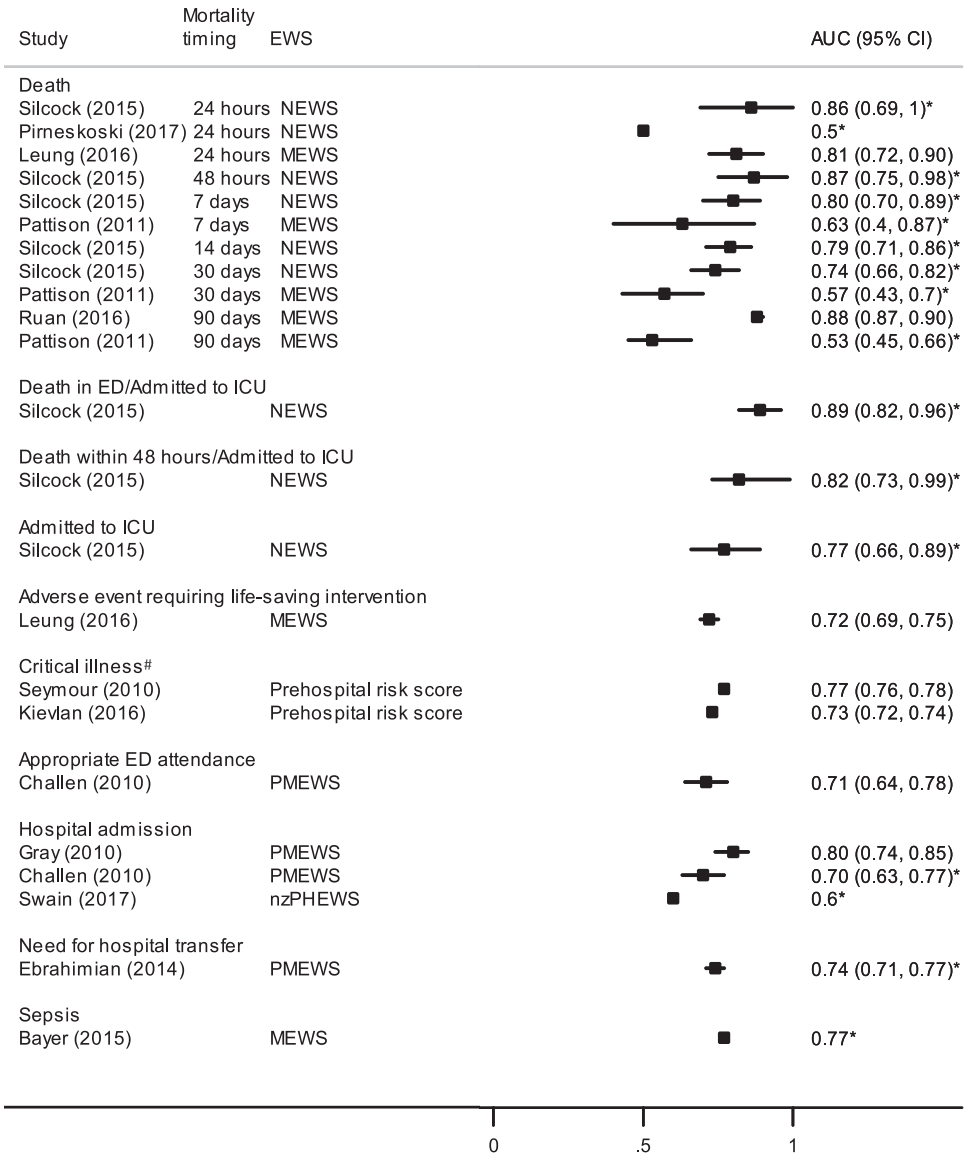

Fig. 2. Area under the curve, AUC (95% CI). All values are calculated from reported data, except for those marked with an asterix [*] which denotes insufficient data in paper to calculate and shows values extracted from paper.
[#]Critical illness during hospitalization, defined as severe sepsis, delivery of mechanical ventilation, or death during hospitalization [26] or defined as an intensive care unit stay with delivery of organ support (mechanical ventilation or vasopressor use) [27].
ED = Emergency Department.
ICU = Intensive Care Unit.
MEWS = Modified Early Warning Score.
NEWS = National Early Warning Score.
PMEWS = Pandemic medical early warning score [9] or physiological social score [25] or physiological-social modified early warning score [32].
Ph-ViEWS = prehospital applied VitalPAC™ Early Warning Score.

specificity ranged from 41% (prediction of hospital admission sensitivity 86%) to 92% (prediction of critical illness; sensitivity 37%). There was a clear trade-off between sensitivity and specificity with studies reporting higher estimates of sensitivity reporting lower estimates of specificity. These data are difficult to interpret but suggest that at this threshold it is difficult to rule in or rule out deterioration.

Nine studies reported estimates of sensitivity and specificity for a "high risk" threshold (Fig. 3c). Two studies evaluated NEWS where scores ≥7 indicate that a patient is at high clinical risk [23,24]. Estimates of specificity were high (≥83%) with some studies reporting 100% specificity. Estimates of sensitivity were lower (0%–78%), with the lowest values for the studies reporting 100% specificity.

Discussion

Summary of findings

We conducted a comprehensive systematic review on the evidence for use of EWS in pre-hospital settings. All included studies were conducted in the ambulance service apart from one conducted in a community nursing home. We did not find any studies in general practice, community or mental health settings. All studies used or reported data that could be adapted to an "accuracy" framework, looking at the number of people with each EWS who did or did not experience the outcome of interest and using these data to generate an ROC curve. Some studies also provided data on sensitivity and specificity at EWS thresholds. We did not find any studies that compared a pre-hospital setting using EWS with one that did not. Although there was variation across studies, AUCs generally suggested reasonable discriminatory performance for mortality at 48 h or combination outcomes that included mortality. Predictive performance was lower for longer term mortality and for other outcomes such as admission to ICU/hospital. The study conducted in a community nursing home suggested EWS do not perform as well in this setting as in the ambulance setting. This could suggest that EWS are not appropriate for use in this setting, or that EWS are less predictive of longer term outcome measures such as 7, 30 and 90-day mortality.

The studies suggested that patients with a score of 0 are very unlikely to deteriorate, and patients with high scores (NEWS ≥ 7) were more likely to deteriorate. No patient with a score of 0 died within 48 h of admission. Intermediate scores were harder to interpret. Overall the data suggests that EWS do distinguish between patients who are and are not likely to deteriorate, but only at more extreme values.

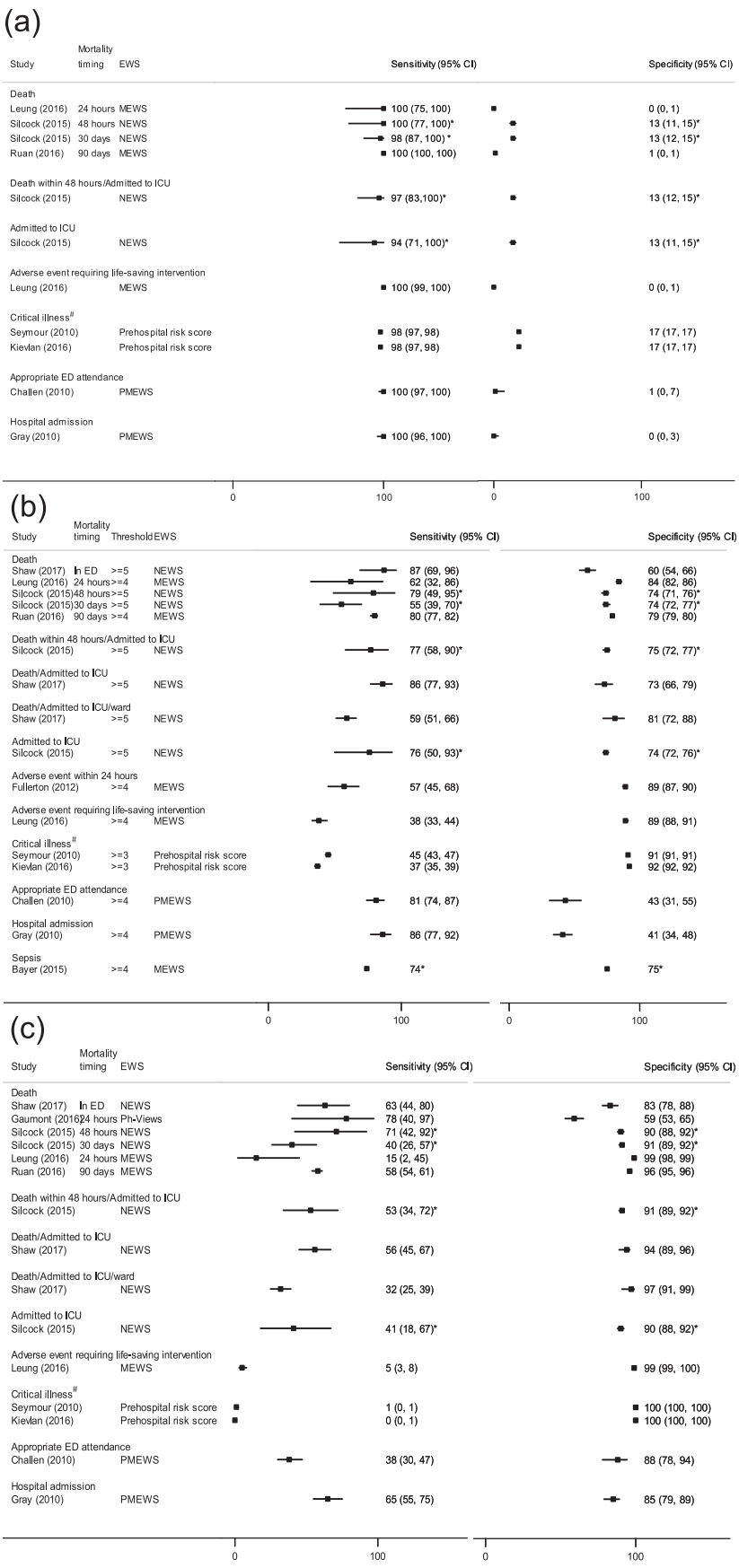

Fig. 3. (a) Sensitivity and specificity (95% CI) at low (≥1) threshold levels. (b) Sensitivity and specificity (95% CI) at medium threshold levels. (c) Sensitivity and specificity (95% CI) at high threshold levels.

All values are calculated from reported data, except for those marked with an asterix [*] which denotes insufficient data in paper to calculate and shows values extracted from paper.

#Critical illness during hospitalization, defined as severe sepsis, delivery of mechanical ventilation, or death during hospitalization [26] or defined as an intensive care unit stay with delivery of organ support (mechanical ventilation or vasopressor use) [27].

ED = Emergency Department.

ICU = Intensive Care Unit.

MEWS = Modified Early Warning Score.

NEWS = National Early Warning Score.

PMEWS = Pandemic medical early warning score [9] or physiological social score [25] or physiological-social modified early warning score [32].

Ph-ViEWS = prehospital applied VitalPAC™ Early Warning Score.

Strengths and weaknesses of this review

This review followed accepted guidance for the conduct of robust systematic reviews [14]. In order to identify as many relevant studies as possible and reduce the risk of publication bias, we used a very sensitive search strategy and searched across an extensive range of resources. Both published and unpublished studies were eligible for inclusion and we included five studies available only as abstracts. We did not statistically assess publication bias as formal assessment of publication bias in systematic reviews of predictive accuracy studies remains problematic and reliability is limited [13]. The extent to which publication bias occurs in predictive accuracy studies is unclear; however, simulation studies have indicated that the effect of publication bias on meta-analytic estimates of test accuracy is minimal [33]. We therefore consider it unlikely that the results of our review have been substantially influenced by publication bias. All stages of the review process involved two reviewers minimising bias and errors. We used the QUADAS-2 tool to assess the risk of bias and applicability of included studies. This highlighted a number of methodological weaknesses in the included studies such as exclusion of certain patients from the study or the analysis, lack of information on blinding, and potential acting on EWS which would be expected to impact patient outcomes and therefore bias estimates of accuracy. Potential limitations of the applicability of studies included restriction to certain patient groups such as those with sepsis, respiratory problems or high-risk patients, and evaluating different EWS. This variation in patient study groups will particularly affect the ability to identify the predictive value of EWS with intermediate scores.

Comparison with previous literature

We are aware of one systematic review that evaluated EWS in pre-hospital settings [34], restricted to studies conducted in the ambulance/emergency medical settings but used a broader definition of EWS than our review, including studies on disease-specific EWS and single item EWS. The review included eight studies, five of these would not have been eligible for inclusion in our review as they assessed sepsis specific screening tools (n = 3), individual physiologic variables rather than aggregate weighted track and trigger systems (n = 1), or a "pragmatic" alert (n = 1). A further study was included in our review as it was possible to extract data on the predictive performance of MEWS but was included in the other review as its primary objective was to develop a new EWS for identification of sepsis [29]. Only two papers were included in both reviews [8,24]. There is therefore little overlap between our systematic review and the previous review.

Unanswered questions and future research

This review has provided information on the predictive accuracy of EWS in pre-hospital settings. However, it is questionable whether an accuracy framework is the most appropriate method for evaluating EWS. EWS are "track and trigger" systems – they encourage action to be taken based on different EWS thresholds. In order to produce unbiased estimates, accuracy studies require that the index test (in this case EWS) are not acted upon prior to application of the reference standard (in this case the outcome of interest, usually mortality). If EWS are acted upon then we would expect this to impact on the outcome, such that the risk of experiencing the outcome is reduced. This would result in a biased estimate of the predictive accuracy of EWS. We considered this as part of our quality assessment – as most studies calculated EWS retrospectively this was not generally a problem in the studies included in our review. Our key question was "Can EWS correctly identify patients who are likely to deteriorate?" In out of hospital settings this is likely to be linked to whether using EWS is better at prioritising care for patients than clinical judgement alone. If EWS are not prioritising appropriate patients for more urgent assessment and intervention, there is a risk that resources may be targeted at patients unlikely to deteriorate whilst those at risk of deterioration may not receive appropriate care.

The optimum design to evaluate EWS is likely to be a cluster randomised trial where settings such as GP surgeries, out of hours providers or ambulance services, are randomised to either use an EWS combined with clinical judgement or clinical judgement alone with comparison of outcomes. If appropriate, settings could use EWS at multiple time points from first assessment until discharge with control settings not using EWS. The optimum outcome measure is unclear. Evidence from our review suggest that short term mortality may be appropriate, but since only a very small proportion of patients assessed in out of hospital settings are likely to die, this would require a very large sample size. Other potential outcomes include incidence of sepsis, admission to hospital or escalation of care. In the out of hospital setting 'time to appropriate care' could be a new measure which would reflect escalation of care for a sick patient. The difficulty in conducting such a study is that EWS are now used routinely in almost all hospital settings and in many out of hospital settings. It is therefore likely to be very difficult if not impossible to conduct a randomised trial. Alternatives rely on comparing routinely collected data in areas that do and do not use EWS and to conduct qualitative research on health professionals experience of using EWS. We are currently undertaking work in this area.

Conclusions

The aim of using pre-hospital EWS is to aid clinical decision making and to enable standardised communication. EWS should help identify patients at risk of deterioration who need referral to secondary care, and patients at low risk of deterioration who can be safely managed at home. Our review suggests that, based on data from ambulance settings, a very low EWS score (0) means patients are unlikely to deteriorate. This adds confidence to clinical judgement that such a patient can be safely managed outside hospital. Patients with very high scores (> = 7) are more likely to deteriorate and should receive appropriate intervention. In practice this means urgent referral to secondary care as the patient requires urgent assessment and treatment. There is insufficient data available to draw strong conclusions regarding the effectiveness or accuracy of EWS in patients with intermediate scores (1–6).

There is no evidence on whether patient outcomes differ between pre-hospital settings that do and do not use EWS. Further studies are required to address this question and to evaluate EWS in other out of hospital settings.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Author contributions

PW, HLR and AP designed this study. AR searched the literature. RP, MDN, HBE and PW extracted data. RP analysed data. RP and PW wrote the first draft of the manuscript. All authors contributed to revisions of the manuscript. PW is the guarantor.

Funding

This work was jointly funded by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West) at University Hospitals Bristol NHS

Foundation Trust and the West of England Academic Health Science Network. The views expressed are those of the author(s) and not necessarily those of NHS England, NHS Improvement, the NIHR or the Department of Health and Social Care.

Data sharing

All data are included in the paper or supplementary appendix. No additional data are available.

Transparency

PW affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.resuscitation.2018.08.028.

References

[1] 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;21(9):737–45. https://doi.org/10.1136/bmjqs-2011-001159. [published Online First: 2012/08/29].

[2] 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;28(11):1629–34. https://doi.org/10.1007/s00134-002-1496-y. [published Online First: 2002/11/05].

[3] Royal College of Physicians. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. London Royal College of Physicians; 2012.

[4] Griffiths JR, Kidney EM. Current use of early warning scores in UK emergency departments. Emerg Med J 2012;29(1):65–6. https://doi.org/10.1136/emermed-2011-200508. [published Online First: 2011/10/04].

[5] Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS—towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010;81(8):932–7. https://doi.org/10.1016/j.resuscitation.2010.04.014. [published Online First: 2010/07/20].

[6] Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. QJM 2001;94(10):521–6. [published Online First: 2001/10/06].

[7] NCEPOD. The National confidential enquiry into patient outcome and death. London: Inspiring Change; 2017http://www.ncepod.org.uk/2017report2/downloads/InspiringChange_FullReport.pdf.

[8] Fullerton JN, Price CL, Silvey NE, Brace SJ, Perkins GD. Is the Modified Early Warning Score (MEWS) superior to clinician judgement in detecting critical illness in the pre-hospital environment? Resuscitation 2012;83(5):557–62. https://doi.org/10.1016/j.resuscitation.2012.01.004. [published Online First: 2012/01/18].

[9] Gray JT, Challen K, Oughton L. Does the pandemic medical early warning score system correlate with disposition decisions made at patient contact by emergency care practitioners? Emerg Med J 2010;27(12):943–7. https://doi.org/10.1136/emj.2009.072959. [published Online First: 2010/11/16].

[10] Roland D, Jahn H. Are early warning scores too early for paramedic practice? J Paramed Pract 2012;4(1):16–21.

[11] Morgan R, Williams F, Wright M. An early warning scoring system for detecting developing critical illness. Clin Intensive Care 1997;8(100).

[12] 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;84(4):465–70. https://doi.org/10.1016/j.resuscitation.2012.12.016. [published Online First: 2013/01/09].

[13] J. Deeks (Ed.), Cochrane handbook for systematic reviews of diagnostic test accuracy Chapter 11 Interpreting results and drawing conclusions, 2013.

[14] Centre for Reviews and Dissemination. Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]. York: University of York; 2009 (Accessed 23 March 2011).

[15] PROSPERO International prospective register of systematic reviews. [Available from: https://www.crd.york.ac.uk/PROSPERO/ (Accessed 26 January 2017).

[16] McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018;319(4):388–96.

[17] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009. [published Online First: 2011/10/19].

[18] Pattison AT, Vernon MJ. The modified early warning score (MEWS) does not predict mortality in community dwelling nursing home residents. Age Ageing 2011;40:ii32. https://doi.org/10.1093/ageing/afr099.

[19] Infinger A, Robert J. Assessing the validity of prehospital identification of severe sepsis using two decision aids. Acad Emerg Med 2016;23:S274. https://doi.org/10.1111/acem.12974.

[20] Pirneskoski J, Nurmi J, Olkkola K, Kuisma M. Prehospital National Early Warning Score (NEWS) does not predict one day mortality. BMJ Open 2017;7:A8–9. https://doi.org/10.1136/bmjopen-2017-EMSabstracts.21.

[21] Studnek JR, Infinger A. Assessing and comparing the validity of three screening tools to identify sepsis in the prehospital setting. Acad Emerg Med 2017;24:S197–8. https://doi.org/10.1111/acem.13203.

[22] Swain A. Does a New Zealand early warning score contribute more to the prehospital assessment of patient acuity than ambulance status codes? Aust J Paramed 2017;14(1). 12–12.

[23] Shaw J, Fothergill RT, Clark S, Moore F. Can the prehospital National Early Warning Score identify patients most at risk from subsequent deterioration? Emerg Med J 2017;34(8):533–7. https://doi.org/10.1136/emermed-2016-206115.

[24] Silcock DJ, Corfield AR, Gowens PA, Rooney KD. Validation of the National Early Warning Score in the prehospital setting. Resuscitation 2015;89:31–5. https://doi.org/10.1016/j.resuscitation.2014.12.029.

[25] Challen K, Walter D. Physiological scoring: an aid to emergency medical services transport decisions? Prehosp Disaster Med 2010;25(4):320–3.

[26] Seymour CW, Kahn JM, Cooke CR, Watkins TR, Heckbert SR, Rea TD. Prediction of critical illness during out-of-hospital emergency care. JAMA 2010;304(7):747–54.

[27] Kievlan D, Martin-Gill C, Kahn J, Callaway C, Yealy D, Angus D, et al. External validation of a prehospital risk score for critical illness. Crit Care 2016;20(1):255. https://doi.org/10.1186/s13054-016-1408-0.

[28] Gaumont D, Cummins N, Hannigan A, Ryan D. VIEWS from the prehospital perspective: a comparison with a prehospital score to triage categorisation in the emergency department. Ir Med J 2016;109(6):424.

[29] Bayer O, Schwarzkopf D, Stumme C, Stacke A, Hartog CS, Hohenstein C, et al. An early warning scoring system to identify septic patients in the prehospital setting: the PRESEP score. Acad Emerg Med 2015;22(7):868–71. https://doi.org/10.1111/acem.12707.

[30] Leung SC, Leung LP, Fan KL, Yip WL. Can prehospital modified early warning score identify non-trauma patients requiring life-saving intervention in the emergency department? Emerg Med Australas 2016;28(1):84–9. https://doi.org/10.1111/1742-6723.12501.

[31] Ruan H, Zhu Y, Tang Z, Li B. Modified early warning score in assessing disease conditions and prognosis of 10,517 pre-hospital emergency cases. Int J Clin Exp Med 2016;9(7):14554–8.

[32] Ebrahimian A, Seyedin H, Jamshidi-Orak R, Masoumi G. Physiological-social scores in predicting outcomes of prehospital internal patients. Emerg Med Int 2014;2014:312189. https://doi.org/10.1155/2014/312189.

[33] Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58(9):882–93.

[34] 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;102:35–43. https://doi.org/10.1016/j.resuscitation.2016.02.011. [published Online First: 2016/02/26].

[35] Health Quality & Safety Commission New Zealand. New Zealand Early Warning Score vital sign chart user guide. 2017. July https://wwwhqscgovtnz/assets/Deteriorating-Patient/PR/Vital_sign_chart_user_guide_July_2017_pdf.