The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: a systematic review.

Teresa Williams, Hideo Tohira, Judith Finn, Gavin D Perkins, Kwok M. Ho.

ABSTRACT

Aim: To examine whether early warning scores (EWS) can accurately predict critical illness in the prehospital setting and affect patient outcomes.

Methods: We searched bibliographic databases for comparative studies that examined prehospital EWS for patients transported by ambulance in the prehospital setting. The ability of the different EWS, including pre-alert protocols and physiological-based EWS, to predict critical illness (sensitivity, odds ratio [OR], area under receiver operating characteristic [AUROC] curves) and hospital mortality was summarized. Study quality was assessed using the Newcastle–Ottawa Scale.

Results: Eight studies were identified. Two studies compared the use of EWS to standard practice using clinical judgement alone to identify critical illness: the pooled diagnostic OR and summary AUROC for EWS were 10.9 (95%CI 4.2-27.9) and 0.78 (95%CI 0.74-0.82), respectively. A study of 144,913 patients reported age and physiological variables predictive of critical illness: AUROC in the independent validation sample was 0.77, 95% CI 0.76-0.78. The high-risk patients stratified by the national early warning score (NEWS) were significantly associated with a higher risk of both mortality and intensive care admission. Data on comparing between different EWS were limited; the Prehospital Early Sepsis Detection (PRESEP) score predicted occurrence of sepsis better than the Modified EWS (AUROC 0.93 versus 0.77, respectively).

Conclusion: EWS in the prehospital setting appeared useful in predicting clinically important outcomes, but the significant heterogeneity between different EWS suggests that these positive promising findings may not be generalizable. Adequately powered prospective studies are needed to identify the EWS best suited to the prehospital setting.

Introduction

The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: a systematic review.

Teresa Williams, Hideo Tohira, Judith Finn, Gavin D Perkins, Kwok M. Ho.

ABSTRACT

Aim: To examine whether early warning scores (EWS) can accurately predict critical illness in the prehospital setting and affect patient outcomes.

Methods: We searched bibliographic databases for comparative studies that examined prehospital EWS for patients transported by ambulance in the prehospital setting. The ability of the different EWS, including pre-alert protocols and physiological-based EWS, to predict critical illness (sensitivity, odds ratio [OR], area under receiver operating characteristic [AUROC] curves) and hospital mortality was summarized. Study quality was assessed using the Newcastle–Ottawa Scale.

Results: Eight studies were identified. Two studies compared the use of EWS to standard practice using clinical judgement alone to identify critical illness: the pooled diagnostic OR and summary AUROC for EWS were 10.9 (95%CI 4.2-27.9) and 0.78 (95%CI 0.74-0.82), respectively. A study of 144,913 patients reported age and physiological variables predictive of critical illness: AUROC in the independent validation sample was 0.77, 95% CI 0.76-0.78. The high-risk patients stratified by the national early warning score (NEWS) were significantly associated with a higher risk of both mortality and intensive care admission. Data on comparing between different EWS were limited; the Prehospital Early Sepsis Detection (PRESEP) score predicted occurrence of sepsis better than the Modified EWS (AUROC 0.93 versus 0.77, respectively).

Conclusion: EWS in the prehospital setting appeared useful in predicting clinically important outcomes, but the significant heterogeneity between different EWS suggests that these positive

promising findings may not be generalizable. Adequately powered prospective studies are needed to identify the EWS best suited to the prehospital setting.

Introduction

Early warning scores (EWS), also known as track and trigger systems, have been developed to facilitate early recognition of the deteriorating hospitalized patient.[1] The EWS may be a single parameter or multiple parameters but often take the form of a composite score weighted by the severity of derangement of physiological variables[2,3] such as systolic blood pressure (SBP), heart rate (HR), respiratory rate (RR) and oxygen saturation (SpO_2). Some EWS also include results from laboratory tests and therapeutic variables such as the requirement for use of supplemental oxygen therapy.[4,5] The composite score is then linked to predefined triggers for review by a critical care team and / or escalation to different levels of care.

While EWS in the hospital setting, including the emergency department (ED), are now considered a standard of care in many parts of the world[6-10], use of EWS by paramedics in the prehospital setting is much less established.[7,11] However, there is interest in the potential for a prehospital EWS to improve patient outcomes – especially for those with a time-critical illness – through earlier access to definitive care.[6,12]

The initial prehospital EWS - the Rapid Acute Physiology Score (RAPS)[4] – is an abbreviated version of the Acute Physiology and Chronic Health Evaluation (APACHE-II),[13] and was developed and tested for air transport of the critically ill. There are now several EWS in use in-hospital (e.g. Modified EWS [MEWS],[14] VitalPAC Early Warning Score [VIEWS],[5] physiological-social EWS (PMEWS),[15] National EWS (NEWS);[6] some of which have also been used in the prehospital setting.[6,14,15] Applying EWS developed in the hospital setting to the prehospital setting to assist early identification of critically ill patients, including those with severe sepsis, acute respiratory failure, or improve triage decisions, may not be appropriate without validation.[16] In this systematic review, we examined the evidence for the use of EWS in the prehospital setting.

Specifically, we sought to assess whether EWS can be used to identify a critically ill patient, predict the likelihood of adverse outcome and whether their implementation into pre-hospital practice has an influence on patient outcomes.

Methods

The review protocol of this systematic review was registered with PROSPERO (CRD42015016818).

Search strategy

We defined EWS as pre-alert protocols and numerical EWS. Four bibliographic databases were searched: MEDLINE (1966- Aug 2015), EMBASE (1980- Aug 2015), CINAHL (1982-Aug 2015) and the Cochrane Library (2004- Aug 2015), using the following MeSH/EMTREE subject headings: ("early warning score" OR "risk score") AND ("ambulance" OR "paramedic" OR ["emergency medical services" and "prehospital"] OR ["emergency medical services" and "out of hospital"]). The reference lists of the relevant or potential papers were also reviewed. The MEDLINE search strategy is shown in Supplementary Table S1.

Study selection

Studies were included if they examined the effect of EWS on identification of a patient condition, prognosis or outcomes for patients transported by road ambulance by paramedics and/or emergency medical technicians in the prehospital setting. The outcomes of interest were paramedic identification of a patient's critical illness: admission to ICU, in-hospital mortality, sepsis. Only randomised controlled trials, case control, cross-sectional or cohort studies were included in this systematic review. Case series or studies involving paediatric patients, rural settings, air

transport,[17,18] or inter-facility transfers[19,20] were excluded. Helicopter emergency services (HEMS) were excluded because these patients are attended by intensive care paramedics and/or critical care physicians and the patients are known to be critically ill and requiring urgent transfer to hospital. We also excluded studies that assessed trauma scores and stroke scales. If a study was reported in multiple publications, we cited the most complete or recent publication and included information from all the reports related to the same study.

Papers identified during the initial literature search were assessed for relevance to this review based on the information contained in the title, abstract and subject descriptor/ MeSH heading (authors TW and HT). Full text articles were obtained if the study was considered relevant or if the information contained in the title and abstract of the study were inconclusive. Any disagreement regarding eligibility was resolved by discussion and consensus involving a third author (JF).

Data extraction

Data on study design, patient characteristics, and patient outcomes were retrieved from the eligible studies. Methodological quality was assessed by the two reviewers independently (authors TW and HT) using the GRADE system for randomised controlled trials[21] and the Newcastle–Ottawa tool (NOS) for cohort and case control studies.[22] The eight-item tool categorised studies into three domains: selection of the study groups (four items), comparability of the groups (one item) and ascertainment of the outcome of interest for cohort studies (three items): a series of response options are provided for each item.[22] A star system for assessment of each item provided a visual semi-quantitative assessment of study quality: the highest quality studies were awarded a maximum of one star for each item within the selection and outcome categories and a maximum of two stars for comparability.[22]

Data synthesis

Study characteristics, methods and results were described according to recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Supplementary Table S2).[23] We proposed to assess heterogeneity first, using the Higgins I^2 test,[24] and only estimate a pooled effect if the statistical heterogeneity was not high risk. The risk of heterogeneity is considered low if I^2 values are less than 25%, moderate for values 25-50% and high if greater than 50%.[24] In the event of significant heterogeneity, forest plots were simply used to provide a graphical representation of the data. *A priori* sensitivity analyses were proposed to explore sources of heterogeneity. For factors associated with critical illness we estimated the odds ratio (OR) and 95% confidence intervals (CIs). We used sensitivity and specificity to assess the diagnostic accuracy of EWS. A funnel plot was used to assess publication bias, using mortality as an end-point.[25] Pre-planned subgroup analyses included studies examining prehospital factors associated with critical illness and pre-alerting the emergency department of the patients impending arrival. Data were analysed using Review Manager (*RevMan*) version 5.3 (Cochrane Collaboration, Oxford, UK), STATA (Release 13: StataCorp LP, College Station, TX, USA), and Meta-Disc (version 1.4, Madrid, Spain). Statistical significance was defined by a two-sided alpha of 0.05.

Results

Study characteristics

The initial search identified 293 papers plus 77 studies specific for sepsis, but 358 were excluded after deleting duplicates and reviewing the title and abstract, and four excluded after reviewing the full paper (Figure 1). One of these excluded studies[15] used a prehospital physiological-social EWS (PMEWS) to assist paramedic decision-making for the need to transfer patients with a presenting

complaint of "shortness of breath" or "difficulty breathing" but not to identify critical illness. The MEDLINE search is shown in Supplementary Table S1. Eight studies[7,16,26-31] met the selection criteria and were included in this systematic review: three from the United States of America (US),[26,27,31] three from the United Kingdom (UK),[7,16,28] one from Sweden[29] and one from Germany.[30] Study characteristics are shown in Table 1. The number of patients totalled 150,797 (range 112 to 144,913) but 96% of these were from one study.[31] We summarised the studies based on (1) diagnosis – the ability of EWS to identify a patient who has a critical illness; (2) predict the risk of an adverse outcome (e.g. admission to ICU, need for ventilation, in-hospital mortality) and (3) determine if introduction of an EWS system improved patients outcomes. One study assessed both diagnosis and outcome.[26] Agreement on the decision on which group a study was assigned was by consensus of the three authors (TW, JF and HT)

Figure 1. PRISMA flow chart of the study selection process[23]

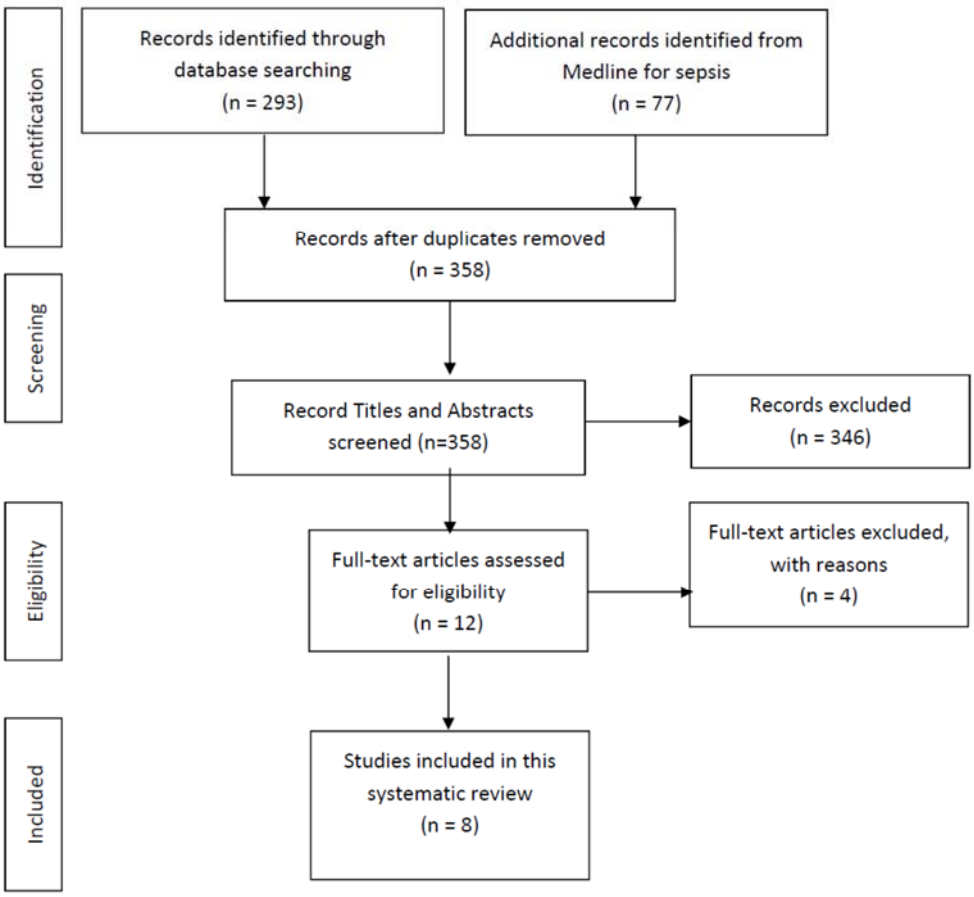


Table 1 Characteristics, identification of critical illness/sepsis, need for pre-alert to hospital and outcomes adverse event/mortality of the nine studies included in this systematic review, grouped by diagnosis, prognosis and outcomes

Study/country/ EMS service	Population	Tool assessed	Comparator	Outcome	Findings
Identification of critical illness including sepsis					
Guerra et al. (2013)[26] US Ambulance service: 2 paramedics or paramedic-EMT pair	112 patients with severe sepsis transported to 3 tertiary hospitals, 2009 Prospective cohort study to identify sepsis; retrospective case control study to assess in-hospital mortality Included: age18+ years, not pregnant, 2+ SIRS criteria, suspected or documented infection, hypoperfusion (SBP < 90 mmHg/MAP < 65 mmHg /lactate level >= 4 mmol/L) Excluded: scheduled transfers	Sepsis Alert Protocol	Clinician judgement	Identification of severe sepsis	32/67 (48%) Sepsis Alert Protocol patients identified sepsis correctly versus 5/45 (11%) patients treated by EMS providers not trained in use of Sepsis Alert Protocol
Suffoletto (2011)[27] Pennsylvania, US 84% trained EMS (EMT)-paramedics, 33 (16%) trained as EMT basics	199 patients transported to single teaching tertiary-care ED Included: age >=18 years transported to single tertiary care ED Excluded: trauma and stroke patients transported with prehospital alerts Convenience sample of EMS providers and ED clinicians blinded to prehospital assessments.	Abnormal prehospital physiologic variables prehospital physiology: HR>90 beats/min, SBP<100 mmHg, RR >20 breaths/min, SpO_2<95%, history of fever, altered mental status	Clinician judgement	Identification of serious infection, i.e. presence of ED report of acute infection plus patient admission	Serious infection: 32/199 (16%) patients, 50% septic (2+ abnormal ED vital signs), 16% admitted to ICU 39% of patients with serious infection had no abnormal prehospital vital signs Prehospital factors associated with serious infection: SBP <100 mmHg, EMS-elicited history or suspicion of fever, and prehospital judgment of infection Model 1 (prehospital physiology only) discrimination AUC 0.66, sensitivity=0.50 (95% CI 0.32–0.68), specificity=0.84 (95% CI 0.77–0.89), PLR =0.22 (95% CI 0.16–0.28), and NLR =0.78 (0.72–0.84) Model 2 (prehospital physiology plus prehospital

					impression of infection), discrimination AUC 0.71, sensitivity=0.59 (95% CI 0.40–0.76), specificity= 0.81 (95% CI (0.74–0.86), PLR =0.26 (95% CI 0.20–0.32), NLR =0.74
Booth et al. (2013)[28] Aberdeen Royal Infirmary Scotland Ambulance crew not described	Prospective study, 7 weeks 104 patients transported by ambulance to ED resuscitation area and reviewed when investigator was on duty Excluded procedural monitoring	Pragmatic alert requirement determined by consultant physician, blinded to outcome	Pragmatic alert requirement by ambulance crew	Pre-alert sensitivity, specificity, PPV, NPV	90 pre-alert, 14 no pre-alert required Ambulance crew decisions to alert 72/104 Sensitivity 72% (CI 62% to 80%), specificity 50% (CI 27% to 73%), PPV 90% and NPV 22% Pre-alert guidance alert prompt: sensitivity 99% (CI 94%-100%), specificity 64% (CI 39%-84%), PPV 95% and NPV 22% 28% of patients under-alerted by ambulance crews, mostly patients with chest pain
Wallgren et al. (2014)[29] Stockholm, Sweden Ambulances staffed with a specialist nurse and an EMT	Retrospective cross-sectional study, 1 January 2007 to 18 May 2008 (17 months) 353 adult patients transported by the EMS, with a hospital discharge ICD code consistent with sepsis. Severe sepsis 148/333 (44%)	[a]Robson screening tool:[32] [b]BAS 90-30-90: SpO_2, RR, SBP	Clinical judgement	Identification of infection, sensitivity	Clinical judgement suspected sepsis in 42/353 (12%) patients and 25/148 (17%) patients with severe sepsis Robson screening tool: sensitivity 93% (13/14 patients with all Robson score parameters) BAS 90-30-90 sensitivity 70% (57/81 patients with all parameters to calculate BAS 90-30-90 score) Robson score ($p=0.004$) and BAS 90-30-90 ($p<0.001$) better predictors of severe sepsis compared to clinical judgment alone
Bayer (2015)[30] Jena University Hospital, Germany	Cohort study-retrospective analysis of 375 patients transported to ED May 2010-April 2013:	Prehospital Early Sepsis Detection (PRESEP) score	[c]MEWS, [a]Robson screening tool; [b]BAS 90-30-90	Predictive validity sensitivity, specificity, positive predictive value (PPV)	PRESP score sensitivity 0.85, specificity 0.86, PPV 0.66, NPV 0.95

	93 (24.8%) patients with sepsis: 60 patients severe sepsis, 12 septic shock Included: age 18+ years, transported to ED by EMS, complete ePCR, i.e. documentation of at least RR, HR, and temperature		Sepsis diagnosis verified by intensivist and emergency physician using ePCR data and clinical records		MEWS sensitivity 0.74, specificity 0.75, PPV 0.45, NPV 0.91 BAS 90-60-90 sensitivity 0.62, specificity 0.83, PPV 0.51, NPV 0.89 Robson screening tool sensitivity 0.95 specificity 0.43, PPV 0.32, NPV 0.9 AUROC = 0.93 (95% CI 0.89 to 0.96) versus AUROC of MEWS = 0.77, p < 0.001)
Prognosis					
Seymour (2010)[31] Greater King County, Washington (excluded metropolitan Seattle), US 16 receiving facilities 2-tier response EMS (1) EMT–fire fighters with BLS skills (2) paramedics with ALS skills	Population-based cohort study, 2002-2006 144,913 patients: Development cohort n=87,266 Validation cohort n=57,647 Included: non-trauma, non-cardiac arrest adult patients Critical illness defined as severe sepsis, received mechanical ventilation, or death during hospitalisation	Patients likely to develop critical illness	Patients unlikely to develop critical illness	Hospital mortality, severe sepsis, mechanical ventilation administered	Critical illness during hospitalization: development cohort n=4,835 (5.5%) and validation cohort n=3,121 (5.4%) 61% of patients severe sepsis Independent factors associated with critical illness: age, SBP, RR, GCS score, SpO_2, nursing home residence. Sex, nursing home not included in final score AUROC in independent validation sample 0.77 (95% CI 0.76-0.78) Outcome components: hospital mortality, 0.78 (95% CI, 0.77-0.79); severe sepsis, 0.76 (95% CI, 0.75-0.77); mechanical ventilation 0.81 (95% CI, 0.80-0.82) Score threshold for critical illness 4+ sensitivity 0.22 (95% CI 0.20-0.23), specificity 0.98 (95% CI 0.98-0.98), positive likelihood ratio 9.8 (95% CI 8.9-10.6), negative likelihood ratio 0.80 (95% CI 0.79- 0.82) Score threshold 1+ sensitivity 0.98 (95% CI 0.97-0.98),

					specificity 0.17 (95% CI 0.17-0.17)
Silcock et al. (2015)[16] Royal Alexandra Hospital Paisley, Scotland	Retrospective cohort study, 1 Oct to 30 Nov 2012 (2-months) 1,684 patients Included all emergency ambulances dispatched with intention to transfer to hospital matched to patients presenting to the hospital's ED Excluded patients <16 years, known to be pregnant, transfers from other hospitals, STEMI patients diverted to other hospital	NEWS[6] Combined score of 4+ considered critically ill	Clinician judgement	48-hour and 30-day mortality, ICU admission, combined endpoint of 48 hour mortality or ICU admission	All 3 primary endpoints and the combined endpoint associated with higher NEWS scores (p =< 0.01 for each) Medium-risk NEWS group associated with a statistically significant increase in ICU admission (RR = 2.466, 95% CI 1.0–6.09), but not hospital mortality relative to the low risk group High risk NEWS group increased 48-hour mortality (RR 35.32 [10.08–123.7]), 30 day mortality (RR 6.7 [3.79–11.88]), and ICU admission (5.43 [2.29–12.89]) Similar results when trauma and non-trauma patients analysed separately
Fullerton et al. (2012)[7] Birmingham Heartlands Hospital, UK Paramedic-led 2,082 (68%) cases, EMT-led 854 (28%) cases Ambulance crew type missing in 121 (4%) cases	Retrospective observational cohort study, single centre, April - June 2010 (2 months) Included 3057/3504 adult ED attendances >=16 years Missing observation data range 1.2% (AVPU) to 36% (temperature), missing values imputed Excluded 26 (0.7%) cases with missing outcome data First record retained, other records excluded (n=421)	MEWS scores using pre-hospital observations	Clinician judgement	Adverse events within 24 hours of admission	Paramedics pre-alerted hospital in 224 cases (7.3%) 76 (2.5%) suffered an adverse event (death, critical care/CCU admission, medical emergency, cardiac arrest, emergency surgery, urgent transfer) Ambulance clinical judgement; identified 47/67 adverse events: sensitivity 62% (95% CI 51-73%), specificity 94% (95% CI 93-95%) MEWS AUC 0.80 (95% CI 0.74-0.86) Combination of MEWS >=4 and clinical judgement: sensitivity 72% (95% CI 62-83%), specificity 85% (95% CI 84-86%)
Outcome					
Guerra et al. (2013)[26] US	112 patients with severe sepsis from 3 tertiary hospitals, 2009	Sepsis Alert Protocol	Clinician judgement	In-hospital mortality	Mortality for Sepsis Alert Protocol patients 14% (5/37) versus no Sepsis

Ambulance service: 2 paramedics or paramedic-EMT pair	Prospective cohort study to identify sepsis; retrospective case control study to assess in-hospital mortality Included: age18+ years, not pregnant, 2+ SIRS criteria, suspected or documented infection, hypoperfusion (SBP < 90 mmHg or MAP < 65 mmHg or lactate level >= 4 mmol/L) Excluded: scheduled transfers				Alert Protocol 33% (25/75) Unadjusted in-hospital survival OR 3.19, 95% CI 1.14-8.88; p = 0.04

ALS=advanced life support, ATLS=advanced trauma life support, AUROC=area under the receiver operating characteristic curve, AVPU - level of consciousness= alert, verbal, pain, or unresponsive). BLS=basic life support, CI=confidence interval, EMS=Emergency Medical Service, ePCR=electronic Patient Care Record, GCS=Glasgow Coma Score, HR=heart rate, ICD-9-CM=International Classification of Diseases version 9 Clinical Modification, MAP=mean arterial pressure, MEWS=Modified Early Warning Score, MTS=Manchester Triage System, NLR= negative likelihood ratio OR=odds ratio, PMEWS=Physiological-social EWS; PLR=positive likelihood ratio RR=respiratory rate, SaO_2=arterial oxygen saturation, SBP=systolic blood pressure, SD=standard deviation, SI=shock index, SIRS=systematic inflammatory response syndrome SpO_2=peripheral oxygen saturation, STEMI ST elevation myocardial infarction, US=United States of America

[a] Robson screening tool:[32] any 2 of these criteria - temperature, HR, RR, altered mental status, plasma glucose, history suggestive of new infection

[b] BAS 90-30-90: SpO_2<90%, RR>30 breaths per minute, SBP<90 mmHg

[c] MEWS[14] uses 5 physiological variables (SBP, HR, RR, temperature, AVPU) rated 0 to 3 to form an aggregated weighted EWS score. AVPU may be substituted with GCS alert=15 verbal=12 pain=8 unresponsive=3[33]

[d] PMEWS[34] Physiological-social EWS RR, SpO_2, HR, SBP, temperature, AVPU, age>65 and (social isolation or chronic disease or performance status)

[e] NEWS[6] National EWS HR SBP RR SpO_2, level of consciousness (AVPU), temperature, supplemental oxygen

Methodological quality

Overall the level of evidence was low – there were no randomised controlled trials. No study was excluded because of methodological quality. The Newcastle Ottawa Scale[22] for cohort and case control studies ranged from 5 to 9 stars, as shown in Supplementary Table S2. In the four cohort studies[26-28,30] and a cross sectional study[29] to identify sepsis, there were small sample sizes and different tools were used. There was potential selection bias in Suffoletto et al.[27] because the data were collected during five- to 10-hour blocks chosen randomly according to research assistant availability. There were no details on randomisation or balancing processes.

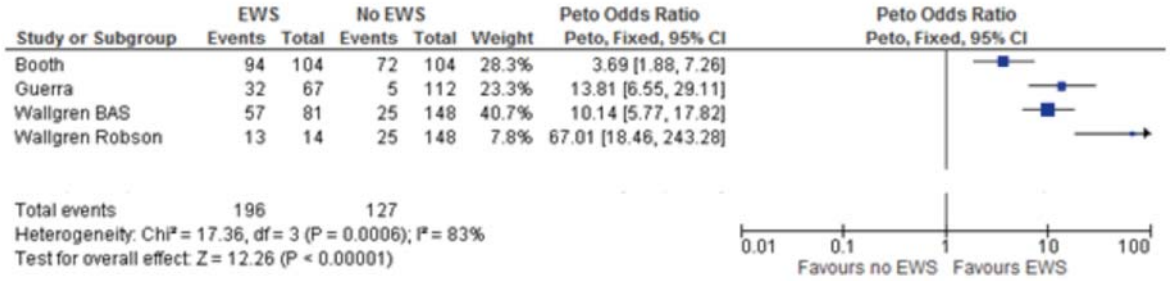
Missing data for the Robson screening tool for severe sepsis was reported in 91% of the septic patients in Wallgren et al.'s study,[29] which was problematic and seriously challenges interpretation of the results. Missing observation data, range 1.2% (AVPU) to 36% (temperature), were imputed.[7]

In Bayer et al.'s single centre study[30] examining independent effects of factors associated with sepsis, important predictor variables were adjusted for in the analyses but the differences in age between the groups, proportion of patients with medical diagnoses, and incidence of sepsis could influence generalisability of the results.[30] Only one other study adjusted for important predictor variables.[31] All studies stated that Human Research Ethics approval had been obtained.

Heterogeneity and Publication bias

Heterogeneity was high: studies used different study designs, selection criteria, definitions of critical illness, tools and outcome measures. In three studies[26,28,29] with four comparisons of EWS versus clinical judgement to identify critical illness (Figure 2), statistical heterogeneity was very high (I^2=83%). However, restricting the comparison to clinical judgement versus a sepsis alert protocol by Guerra et al.[26] and the Swedish BAS 90-30-90, an acronym for SBP <90mmHg, respiratory rate >30 breaths per minute and oxygen saturation <90%,[35] by Wallgren et al.,[29] heterogeneity was substantially reduced (I^2 =0%). We could not assess publication bias in the funnel plot because there were only four studies in the meta-analysis (Supplementary Figure S1).

Figure 2. Graphical representation of the association of early warning scores (including pre-alerts) on identification of critical illness in the prehospital setting



(1) Do EWS assist with the identification of patients with a critical illness such as sepsis?

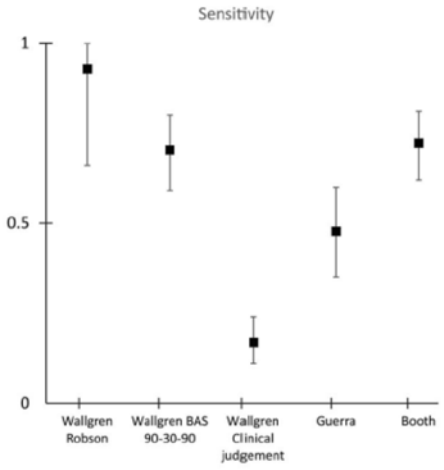
The review identified five low quality studies which addressed this question, all in patients with suspected infection.[26-30] The studies used different methods to identify sepsis and assess outcomes. Guerra et al.[26] used a Sepsis Alert Protocol screening tool to assess the Emergency Medical Service (EMS) ability to identify patients with severe sepsis. Forty-eight percent of patients were correctly identified as having severe sepsis by EMS providers trained to use the Sepsis Alert Protocol compared to 4% identified by EMS who did not receive the sepsis protocol training.[26] Booth et al.[28] also used a pre-alert guidance tool and compared it to ambulance crew decisions and a prehospital EWS to pre-alert EDs of their impending arrival with potentially critically ill patients. The pre-alert guidance prompts had a high sensitivity (99%, 95% CI 94-100%) (95%) as shown in Figure 3 compared to ambulance crew decisions without the alert prompts, although the specificity was modest (64%, 95% CI 39-84%).[28]

In a third study of sepsis, Suffoletto et al.[27] compared the agreement between paramedic judgment and prehospital physiologic variables to the emergency physician diagnosis of acute infection.

Sampling was balanced between weekdays and weekends, between daytime and evening over a two–month period. Prehospital SBP <100 mmHg, EMS-elicited history or suspicion of fever, and prehospital judgment of infection were factors associated with serious infection. The model's overall predictive ability of identifying serious infection was, however, only moderate (the area under the receiver operating characteristic curve [AUROC] 0.71). Sensitivity was 0.59 (95% CI 0.40-0.76) and specificity of 0.81 (95% CI 0.74-0.86).[27]

Comparing two prehospital sepsis screening tools, the Robson screening tool[32]and the BAS 90-30-90,[35] with EMS clinical judgment in predicting sepsis, Wallgren et al.[29] found that both the Robson screening tool and BAS 90-30-90 performed better than clinical judgement to identify sepsis. The Robson screening tool[32] had better sensitivity in the 14 of 148 (9%) patients with severe sepsis who had the data for the score to be calculated. (Figure 3). All four comparisons of EWS to clinical judgement to identify sepsis favoured EWS as shown in Figure 3. The OR in the meta-analysis ranged from 3.7 in Booth et al.'s study[28] to 67.0 (95% CI 18.5-243) for the Robson screening score in Wallgren et al.'s study.[29]

Figure 3. Sensitivity with 95% confidence intervals for studies of identification of critical illness in the prehospital setting.



Bayer et al.[30] reported the development and validation of a PRESEP score, an EWS combining temperature, RR, HR and SBP (GCS and blood sugar were not significant). Physiological variable cut-points were refined by Bayer et al.[30] from those defined by the American College of Chest Physicians/Society of Critical Care Medicine (ACCP/SCCM) criteria[36] and the Surviving Sepsis Campaign Guidelines.[37] The PRESEP score was highly predictive for sepsis (AUROC 0.93 95%CI 0.89-0.96). The PRESEP >=4 sensitivity was 0.85 (95% CI = 0.77 to 0.92), higher than MEWS>=4 (0.77), BAS 90-30-90 (0.62) but lower than the Modified Robson score (0.95).[30] The specificity for the PRESEP was 0.86 (95% CI = 0.82 to 0.90).

(2) Do EWS predict the risk of an adverse outcome (e.g. admission to ICU, need for ventilation, in-hospital mortality)?

Three studies examined the prognostic effect of EWS. [7,16,31] A large population-based cohort study using prehospital data linked to hospital discharge data[31] and two smaller cohort studies [7,16] assessed prognosis using different methods to accomplish this. The quality of the studies varied. Seymour et al.[31] large population-based study developed a prediction score to identify critical illness, defined as either having severe sepsis, requiring mechanical ventilation, or death after hospitalisation. More than half the patients were severe sepsis (61%) and trauma patients were excluded. Data were randomly split into development (n = 87,266 [60%]) and validation (n = 57,647 [40%]) cohorts. The *ICD-9-CM* codes used for sepsis and organ failure were 995.91, 995.92, 785.52 and the procedure code 96.7x for mechanical ventilation.[38] Only the initial prehospital vital signs, documented by the first arriving EMS personnel were used.[31] Candidate variables were selected by (1) clinical relevance, (2) generalizability (3) timing of prehospital care exposure. The independent factors associated with critical illness reported by Seymour et al.[31] included age >=45 years, RR <12 or >=24, SBP =<90 mmHg, HR >=120 beats per minute, SpO_2 <88% and Glasgow Coma Score <15. Being a nursing home resident was also significant but was not included in the

regression models. The predictive ability of the model to identify critical illness was also only moderate (AUROC in an independent validation sample 0.77, 95% CI 0.76-0.78).[31]

A second study of prognosis, a retrospective cohort study of 1,684 patients, was the only study to assess the ability of NEWS, proposed for implementation throughout the UK's National Health Service, to predict patient outcomes in the prehospital setting.[16] Silcock et al.[16] reported higher NEWS were associated with three primary endpoints (survival to admission or 30 days, death within 48-hours of admission, ICU admission, all p =< 0.01) and a combined endpoint (48 hour mortality or ICU admission) but the results were inconsistent across risk groups. Thirty-day mortality was 6/251 (2%) for medium (scores 5-6) versus 19/146 (13%) for high-risk (scores 7+) NEWS categories, ICU admission 7/251 (3%) for medium versus 8/146 (5%) high-risk NEWS categories and 48-hour mortality 1/251 (0.4%) for moderate versus 12/146 (8%) high risk NEWS categories.[16] The high-risk NEWS group was associated with an increased risk of 48-hour mortality (risk ratio 35.32, 95%CI 10.08–123.7]), 30-day mortality (RR 6.7, 95%CI 3.79–11.88]) and ICU admission (5.43, 95%CI 2.29–12.89]); medium-risk NEWS group was associated with an increased risk of ICU admission (risk ratio = 2.47, 95% CI 1.0–6.09), but not hospital mortality relative to the low risk group.[16] These results were similar when trauma and non-trauma patients were analysed separately.[16]

Fullerton et al.,[7] in a single centre study of 3,504 patients conducted over two months, compared the accuracy of a pre-alerting system to the modified EWS (MEWS) in the third study of EWS and prognosis.[14] The study used prehospital observations to detect critical illness, defined as the occurrence of adverse events within 24 hours of hospital admission. Missing data were: outcomes 0.7%, RR (2.3%), HR (1.9%), temperature (36%), SBP (6.0%), SpO2 (5.4%) and AVPU ([Alert, Verbal, Pain, Unresponsive] 1.2%). The sensitivity and specificity of clinical judgement to detect critical illness were 61.8% (95% CI 51.0-72.8%) and 94.1% (95% CI 93.2-94.9%) respectively.[7]

The MEWS was a better predictor of adverse outcomes such as ICU admission, cardiac arrest and death than clinical judgement (AUROC 0.799, 95% CI 0.738-0.856). Comparing the MEWS category >=4 and clinical judgement improved the sensitivity (72.4%, 95% CI 62.5-82.7%) and specificity (84.8%, 95% CI 83.52-86.1%).[7]

(3) Determine if introduction of an EWS system improved patient outcomes

One study[26] used a retrospective case control study to assess the effect of the Sepsis Alert Protocol on survival to hospital discharge, with all the inherent weaknesses of no randomisation of patients to the control and intervention groups and the use of a retrospective design. Guerra et al.[26] reported hospital mortality was 14% (37/112) for patients with severe sepsis for whom a Sepsis Alert Protocol was initiated compared to 33% (75/112) for those without a Sepsis Alert Protocol initiated (unadjusted OR=3.19, 95% CI 1.14-8.88; p = 0.04).[26] There was no adjustment for potential confounders. None of the studies included in this systematic review assessed whether using an EWS in the prehospital setting was effective in improving outcomes compared to clinical judgement alone.

Discussion

Despite the plethora of publications relating to use of EWS in the in-hospital setting – there are relatively few studies that have examined the use of EWS in the prehospital emergency ambulance setting. In the eight studies[7,16,26-31] examining the use of EWS in the prehospital setting, it appeared that EWS were helpful in assisting ambulance services in identifying critically ill patients,[26,28-30] prognosis[7] and outcomes.[16] However we noted that there was substantial heterogeneity between studies, in terms of the populations, how the EWS were constructed as well as the definitions of adverse outcomes that were predicted by different EWS. One recent study[39] that did not meet our review inclusion criteria also suggested that EWS may be useful in assisting clinicians' triage

decision at the ED. These results of the use of EWS in the prehospital setting are clinically relevant and require further discussion.

First, identifying time-critical conditions such as sepsis early may benefit prehospital patients by delivering timely pre-alert to ED, resuscitation and antibiotics.[40-42] A recent study found that pre-alerting before arrival to ED almost halved the time for in-hospital treatment.[43] This finding is not specific for EWS but supports the strategy of a structured pre-alerting for critically ill patients. Early warming scores are used to trigger ED pre-alerting but the pre-alert may not be required. However, a high sensitivity of an EWS is essential to avoid missing seriously ill patients not treated urgently resulting in adverse outcomes. In line with this clinical concern, most EWS included in this review did have a reasonably high sensitivity in identifying critically ill patients in the prehospital setting.

Second, an ideal EWS should have both a high sensitivity and specificity. Our results found the existing EWS appeared not to perform as well as EWS in a hospital setting. In the prehospital setting, identifying critically ill patients is extremely challenging because patients often present with non-specific signs and symptoms with limited clinical history and laboratory tests are unavailable. It is possible that the trend in how the prehospital EWS score changes within the same patient while on route to the ED may improve the specificity of the EWS, but this has not been assessed thoroughly.[44] Nevertheless, in the prehospital setting, paramedics have a much shorter time to re-evaluate their patients' response to treatment and hence an ideal prehospital EWS can be very difficult to achieve.[45] While EWS may be useful in the prehospital setting, the focus solely on the "number" of a score in clinical decision-making should not replace clinical judgement but rather complement EWS.[6,45]

Third, methods used by paramedics to calculate the EWS in the prehospital setting have received little attention. Depending on the particular ambulance service practice, paramedics use either paper-based or electronic patient care records (ePCR) to record patient observations. However, observations may not be documented until the end of the job, e.g. paramedics may record observations on a note pad (or the back of their glove) until time permits for entering the data onto the ePCR. An EWS needs to be generated automatically by the ePCR or similar portable devices (e.g. smartphone app) in real time to have any value. Automatic calculation of EWS improves speed and accuracy[46-48] and allows integration of physiological variables with patient characteristics from the patient record.[49] Ultimately tablet-computer solutions integrating machine learning algorithms[50] linked to the monitor-defibrillator unit to produce automatic score generation will facilitate EWS to be used to inform appropriate and timely care decisions.

Finally, we would like to acknowledge the limitations of this study. Despite an exhaustive literature search and the inclusion of studies, based on our pre-determined selection criteria, we may have missed some studies. The low number of studies in this systematic review may be due to the fact that the importance of EWS in the prehospital setting became apparent only recently and hence more studies are needed before we can recommend widespread adoption of EWS in all ambulance services. Perhaps, a consensus meeting between stakeholders from different ambulance services is needed before an adequately powered studies can be conducted. None of the included studies in this review had assessed whether using an EWS in the prehospital setting was cost-effective in improving patient-centred outcomes compared to clinical judgement alone.

Conclusion

Using EWS in a prehospital setting is an important emerging theme in emergency and critical care medicine. Despite promising results from a limited number of studies, the predictive accuracy, clinical utility and generalizability of many prehospital EWS, particularly in conjunction with clinical judgement, remain uncertain. Adequately powered prospective studies are definitely needed to identify the best EWS for use in the prehospital setting.

Acknowledgement

Vale-Ian Jacobs for his support and substantial contribution to this study.

Dr Ho is funded by the Raine Medical Research Foundation and Western Australian Department of Health through the Raine Clinical Research Fellowship

Conflict of interest statement

No author has a conflict of interest

Funding

Nil

Figure Legends

Figure 1. Flow chart of the study selection process.

Figure 2. Summary of the association of early warning scores including pre-alerts on identification of critical illness in the prehospital setting.

Figure 3. Sensitivity with 95% confidence intervals for studies of identification of critical illness in the prehospital setting.

Supplementary Figure S1. Publication bias: EWS/pre-alert on identification of critical illness

References

References

1. Morgan R, Williams F, Wright M. An early warning scoring system for detecting developing critical illness. Clin Intensive Care 1997;8:100.

2. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. Resuscitation 2008;79:11-21.

3. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. Resuscitation 2008;77:170-9.

4. Rhee KJ, Fisher CJ, Jr., Willitis NH. The Rapid Acute Physiology Score. Am J Emerg Med 1987;5:278-82.

5. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010;81:932-7.

6. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Report of a working party. London: RCP. 2012.

7. 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;83:557-62.

8. Corfield AR, Lees F, Zealley I, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emerg Med J 2014;31:482-7.

9. Groarke JD, Gallagher J, Stack J, et al. Use of an admission early warning score to predict patient morbidity and mortality and treatment success. Emerg Med J 2008;25:803-6.

10. Griffiths JR, Kidney EM. Current use of early warning scores in UK emergency departments. Emerg Med J 2012;29:65-6.

11. 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;27:943-7.

12. Australian Commission on Safety and Quality in Health Care. Safety and Quality Improvement Guide Standard 9: Recognising and Responding to Clinical Deterioration in Acute Health Care (October 2012). Sydney. ACSQHC2012.

13. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13:818-29.

14. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. Q J Med 2001;94:521-6.

15. Challen K, Walter D. Physiological scoring: an aid to emergency medical services transport decisions? Prehosp Disaster Med 2010;25:320-3.

16. Silcock DJ, Corfield AR, Gowens PA, Rooney KD. Validation of the National Early Warning Score in the prehospital setting. Resuscitation 2015;89:31-5.

17. Bertollo S, Rodenberg H. Correlation of the RTS (Revised Trauma Score) and RAPS (Rapid Acute Physiology Score) in rotor-wing prehospital care. Air Med J 1994;13:91-5.

18. Rhee KJ, Baxt WG, Mackenzie JR, et al. Differences in air ambulance patient mix demonstrated by physiologic scoring. Ann Emerg Med 1990;19:552-6.

19. Lee LL, Yeung KL, Lo WY, Lau YS, Tang SY, Chan JT. Evaluation of a simplified therapeutic intervention scoring system (TISS-28) and the modified early warning score (MEWS) in predicting physiological deterioration during inter-facility transport. Resuscitation 2008;76:47-51.

20. Markakis C, Dalezios M, Chatzicostas C, et al. Evaluation of a risk score for interhospital transport of critically ill patients. Emerg Med J 2006;23:313-7.

21. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6.

22. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses, 1998 (Accessed 23 Jan 2012, at.http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm).

23. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264-9.

24. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.

25. Higgins J, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011 (Accessed 29 March 2011, at www.cochrane-handbook.org).

26. Guerra WF, Mayfield TR, Meyers MS, Clouatre AE, Riccio JC. Early detection and treatment of patients with severe sepsis by prehospital personnel. J Emerg Med 2013;44:1116-25.

27. Suffoletto B, Frisch A, Prabhu A, Kristan J, Guyette FX, Callaway CW. Prediction of serious infection during prehospital emergency care. Prehosp Emerg Care 2011;15:325-30.

28. Booth SM, Bloch M. An evaluation of a new prehospital pre-alert guidance tool. Emerg Med J 2013;30:820-3.

29. Wallgren UM, Castren M, Svensson AE, Kurland L. Identification of adult septic patients in the prehospital setting: a comparison of two screening tools and clinical judgment. Eur J Emerg Med 2014;21:260-5.

30. Bayer O, Schwarzkopf D, Stumme C, et al. An Early Warning Scoring System to Identify Septic Patients in the Prehospital Setting: The PRESEP Score. Acad Emerg Med 2015;22:868-71.

31. Seymour CW, Kahn JM, Cooke CR, Watkins TR, Heckbert SR, Rea TD. Prediction of critical illness during out-of-hospital emergency care. JAMA 2010;304:747-54.

32. Gardlund B, Cronqvist J, Follin P, Furebring M, Gille-Johnson P, Sjolin J. [Severe sepsis and septic shock require immediate care]. Lakartidningen 2011;108:271-5.

33. Robson W, Nutbeam T, Daniels R. Sepsis: a need for prehospital intervention? Emerg Med J 2009;26:535-8.

34. Bone RC, Balk RA, Cerra FB, et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. Chest 1992;101:1644-55.

35. Dellinger RP, Levy MM, Rhodes A, et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012. Intensive Care Med 2013;39:165-228.

36. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Crit Care Med 2001;29:1303-10.

37. 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? 2015.

38. Kumar A, Roberts D, Wood KE, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med 2006;34:1589-96.

39. Ferrer RMDP, Martin-Loeches IMDP, Phillips GMAS, et al. Empiric Antibiotic Treatment Reduces Mortality in Severe Sepsis and Septic Shock From the First Hour: Results From a Guideline-Based Performance Improvement Program. Crit Care Med 2014;42:1749-55.

40. Gaieski DF, Mikkelsen ME, Band RA, et al. Impact of time to antibiotics on survival in patients with severe sepsis or septic shock in whom early goal-directed therapy was initiated in the emergency department. Crit Care Med 2010;38:1045-53.

41. Studnek JR, Artho MR, Garner CL, Jr., Jones AE. The impact of emergency medical services on the ED care of severe sepsis. Am J Emerg Med 2012;30:51-6.

42. Kellett J, Wang F, Woodworth S, Huang W. Changes and their prognostic implications in the abbreviated VitalPACTM Early Warning Score (ViEWS) after admission to hospital of 18,827 surgical patients. Resuscitation 2013;84:471-6.

43. Subbe CP, Hillman K. "To live and let die" - the search for the best way to identify at-risk patients? Resuscitation 2013;84:407-8.

44. Mohammed M, Hayton R, Clements G, Smith G, Prytherch D. Improving accuracy and efficiency of early warning scores in acute care. Br J Nurs 2009;18:18-24.

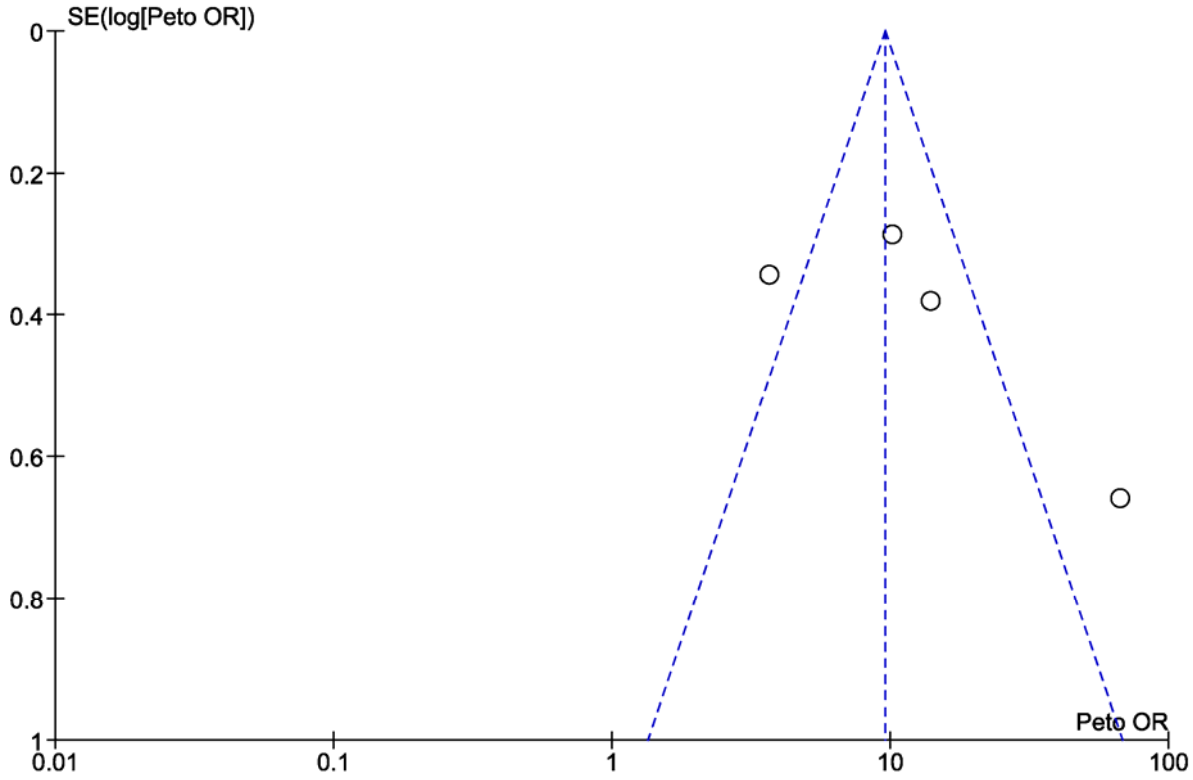
45. Prytherch DR, Smith GB, Schmidt P, et al. Calculating early warning scores--a classroom comparison of pen and paper and hand-held computer methods. Resuscitation 2006;70:173-8.

46. Jones S, Mullally M, Ingleby S, Buist M, Bailey M, Eddleston JM. Bedside electronic capture of clinical observations and automated clinical alerts to improve compliance with an Early Warning Score protocol. Crit Care Resusc 2011;13:83-8.

47. 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;42:841-8.

48. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. Crit Care Med 2016;44:368-74.

Supplementary Figure S1. Publication bias: EWS/pre-alert on identification of critical illness



Supplementary Data Table S1 Medline search

No	Search	
1	early warning score.mp.	183
2	exp Triage/mt [Methods]	2036
3	pre-alert.mp.	15
4	risk score.mp.	6825
5	1 or 2 or 3 or 4	9019
6	pre-hospital.mp.	2390
7	prehospital.mp.	7600
8	out of hospital.mp.	6261
9	6 or 7 or 8	15270
10	emergency medical services.mp. or exp Emergency Medical Services/	103682
11	paramedic.mp.	1621
12	ambulance.mp. or exp Ambulances/	10602
13	emergency medical technician.mp. or exp Emergency Medical Technicians/	5235
14	11 or 12 or 13	15593
15	9 or 14	27541
16	10 and 15	16976
17	5 and 15	293
	Sepsis	
18	exp Sepsis/ or sepsis.mp.	135288
19	16 and 18	77

Supplementary Data Table S2. Assessment of methodological quality using the Newcastle-Ottawa Scale[22] for cohort studies

	Guerra[26]	Suffoleto[27]	Booth[28]	Wallgren[29]	Bayer[30]	Seymour[31]	Fullerton[7]	Silcock[16]
Intervention/ exposure	Sepsis Alert Protocol	Provider judgement plus physiology	Pre-alert	Robson score, BAS 90-30-90 Provider judgement	PRESEP EWS MEWS	Predictors of critical illness	Pre-alert	NEWS validation
Selection								
1. Representativeness of intervention cohort	⚙	⚙		⚙	⚙	⚙	⚙	⚙
2. Selection of non-intervention cohort	⚙	⚙		⚙	⚙	⚙	⚙	⚙
3. Ascertainment of intervention	⚙	⚙	⚙	⚙	⚙	⚙	⚙	⚙
4. Outcome of interest not present at start of study	⚙	⚙	⚙	⚙	⚙	⚙	⚙	⚙
Comparability								
a) study controls for physiological values					⚙	⚙		
b) study controls for additional factors					⚙	⚙		
Outcome								
1. Assessment of outcome	⚙	⚙	⚙	⚙	⚙	⚙	⚙	⚙
2. Follow up long enough for outcomes to occur	⚙	⚙	⚙	⚙	⚙	⚙	⚙	⚙
3. Adequacy of follow up of cohorts	⚙	⚙	⚙	⚙	⚙	⚙	⚙	⚙
Total Score	**7**	**7**	**5**	**7**	**9**	**9**	**7**	**7**