BMJ Open

Performance of seven different paediatric early warning scores to predict critical care admission in febrile children presenting to the emergency department: a retrospective cohort study

Sam T Romaine [ID] ,[1] Gerri Sefton,[2] Emma Lim,[3] Ruud G Nijman,[4] Jolanta Bernatoniene,[5] Simon Clark,[6] Luregn J Schlapbach,[7,8] Philip Pallmann [ID] ,[9] Enitan D Carrol[1,10]

Correspondence to
Professor Enitan D Carrol;
edcarrol@liverpool.ac.uk

ABSTRACT

Objective Paediatric Early Warning Scores (PEWS) are widely used in the UK, but the heterogeneity across tools and the limited data on their predictive performance represent obstacles to improving best practice. The standardisation of practice through the proposed National PEWS will rely on robust validation. Therefore, we compared the performance of the National PEWS with six other PEWS currently used in NHS hospitals, for their ability to predict critical care (CC) admission in febrile children attending the emergency department (ED).

Design Retrospective single-centre cohort study.

Setting Tertiary hospital paediatric ED.

Participants A total of 11 449 eligible febrile ED attendances were identified from the electronic patient record over a 2-year period. Seven PEWS scores were calculated (Alder Hey, Bedside, Bristol, National, Newcastle and Scotland PEWS, and the Paediatric Observation Priority Score, using the worst observations recorded during their ED stay.

Outcomes The primary outcome was CC admission within 48 hours, the secondary outcomes were hospital length of stay (LOS) >48 hours and sepsis-related mortality.

Results Of 11 449 febrile children, 134 (1.2%) were admitted to CC within 48 hours of ED presentation, 606 (5.3%) had a hospital LOS >48 hours. 10 (0.09%) children died, 5 (0.04%) were sepsis-related. All seven PEWS demonstrated excellent discrimination for CC admission (range area under the receiver operating characteristic curves (AUC) 0.91–0.95) and sepsis-related mortality (range AUC 0.95–0.99), most demonstrated moderate discrimination for hospital LOS (range AUC 0.69–0.75). In CC admission threshold analyses, bedside PEWS (AUC 0.90; 95% CI 0.86 to 0.93) and National PEWS (AUC 0.90; 0.87–0.93) were the most discriminative, both at a threshold of ≥6.

Conclusions Our results support the use of the proposed National PEWS in the paediatric ED for the recognition of suspected sepsis to improve outcomes, but further validation is required in other settings and presentations.

Strengths and limitations of this study

► We compared the performance of six PaediatricEarly Warning Scores (PEWS) in-use across the UK with the proposed National PEWS, in a large cohort of febrile children attending the emergencydepartment.

► Our cohort represents a group of children in whom those at risk of poor outcomes can be difficult to identify, meaning effective detection tools are required.

► We used retrospective data from a single-centre, in which a proportion of eligible cases had to be excluded due to missing observation data.

► Minor adaptations to five of the PEWS were necessary for inclusion in the study.

► Only the detection aspect (afferent limb) of the overall PEWS system could be assessed in the study.

INTRODUCTION

Early Warning Scores

The use of Early Warning Scores (EWS) has been recommended for optimising the recognition and management of serious illness to prevent avoidable deaths.[1 2] Many different EWS have been used across the UK,[3] initially for the purpose of identifying deterioration in adult inpatients.[4] The Royal College of Physicians (RCP) standardised UK adult practice in 2012 with the National Early Warning Score (NEWS)[3] and subsequent NEWS in 2017.[5] NEWS demonstrated good discriminative ability in predicting intensive care unit (ICU) admission and death in adult Medical Admission Unit patients[6] and has subsequently been successfully validated in both the Emergency Department (ED)[7–9] and prehospital settings.[10–12] In order to standardise systems of care, it was, therefore, suggested that NEWS should be used across the entire

journey of an acutely unwell patient, including primary care, ambulance service, the ED and inpatient ward.[5 13 14] This system-wide introduction of NEWS was associated with decreased mortality in adults with suspected sepsis.[15]

Paediatric Early Warning Scores

Paediatric EWS (PEWS) have become widely used across the UK, with many different systems in use.[16] Despite substantial overlap, currently used tools often differ in their included variables, thresholds, weighting and escalation algorithms.[17] Several PEWS have demonstrated good discrimination in predicting ICU admission and death,[17 18] and implementing PEWS alongside a rapid response system (RRS) has been shown to reduce in-hospital mortality.[19] However, significant variations in performance have been reported with different PEWS, and it is unclear to what extent this relates to tool performance or clinical context such as patient complexity.[17 18] The large number of contrasting PEWS, coupled with a lack of internal and external validation,[20 21] creates uncertainty over the optimal PEWS. In addition, the use of different PEWS may complicate the transfer of patients between settings, the experience of staff working in different settings and training requirements. Both Scotland and Northern Ireland use a single PEWS across their hospitals.[22 23] In England, a standardised National PEWS is being developed with plans for system-wide introduction in April 2021, including primary care, ambulance services, the ED and inpatient wards.[23] It is hoped that using a common language across these settings will improve outcomes in children, as has been achieved in adults with NEWS.[15]

Acute febrile illness is among the most common reasons for children to attend the paediatric ED,[24] but only a minority have sepsis or serious illness, making their detection challenging. This difficulty is exacerbated by the tendency of a fever to alter other vital signs. Outcomes are poor for those with sepsis that is recognised late;[25] therefore, it is essential that PEWS are able to effectively identify these children.

Objective

The aim of the present study was to compare the performance of the proposed National PEWS with six other PEWS currently in-use across the UK for their ability to predict poor outcomes in a large cohort of febrile children presenting to the ED.

METHODS

Study population and definitions

The study was conducted at a single large tertiary hospital, Alder Hey (AH) Children's NHS Foundation Trust, Liverpool, UK, which manages around 60 000 ED attendances annually.[26] We included children (<16 years) attending the ED with a fever over 38°C, or a history of fever within the previous 3 days, who presented between 1 September 2015 and 31 August 2017. Details of the cohort have been previously published.[27] Patients were identified retrospectively from the electronic patient record (EPR) by reviewing eligible ED attendances and critical care (CC) admissions across this 2-year period. Cases were excluded if there was missing observation data, age ≥16 years, no history of fever or patients transferred from another hospital. Cases with missing observation data were defined as those with two or more components of any of the included PEWS not being recorded during their ED stay. If only one component was not recorded, this was deemed to be normal and the case was included. The primary outcome was CC admission (high-dependency unit (HDU) or ICU) within 48 hours of ED presentation. CC admission within 48 hours was used to capture both those children with sepsis recognition in the ED, and a significant proportion admitted to a ward, who later deteriorate and are transferred to CC.[28] The secondary outcomes were hospital length of stay (LOS) >48 hours, and sepsis-related mortality, defined as in-hospital death within 28 days, determined to be sepsis related by a review of the medical notes. The seven PEWS evaluated were: AH PEWS,[29] Bedside PEWS,[30] Bristol PEWS (personal communication, Jolanta Bernatoniene, 2020), National PEWS (personal communication, Simon J Clark, 2020), Newcastle PEWS (personal communication, Emma Lim, 2020), Paediatric Observation Priority Score (POPS)[31] and Scotland PEWS[22] (online supplemental figures 1–7).

Score calculations

PEWS scores were calculated using the worst observations recorded during the ED stay. AH PEWS was in routine clinical use and was calculated electronically in real time, whereas the other six PEWS were calculated retrospectively for study purposes only.

Score adaptations

The routinely recorded EPR data are based on, and limited to, the information required to calculate a score for the PEWS in-use at our hospital (AH PEWS). The full information required to calculate scores for five PEWS was not available from the EPR data. It was, therefore, necessary to adapt these five PEWS for the purposes of the study. Care was taken to ensure that any changes were minor and as close to the original PEWS as possible (table 1).

Respiratory distress

The AH PEWS uses a different system for 'respiratory distress' than the five other PEWS that use this component. In the AH PEWS, severity is determined by totalling the number of 'effort of breathing' features present (grunting, head bobbing, marked subcostal recession, nasal flaring, stridor, tracheal tug; one feature=one point, two or more features=two points), whereas National PEWS, Newcastle PEWS and POPS grade the severity of each individual feature (eg, stridor=2 points) and the highest scoring feature determines the score (online supplemental figures 1, 4–6). Furthermore, the three

Table 1 Summary of parameters used by each PEWS including adaptations

Scoring system	AVPU	CRT	Co-morbidities	HR	Nurse/parent concern	O₂ delivery	Respiratory distress	RR	O₂ sats	SBP	Temperature
Alder Hey PEWS[29]	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	○
Bedside PEWS[30]	✗	✓	✗	✓	✗	✓	▲ Scored the same as adapted National and Newcastle PEWS	✓	✓	✓	✗
Bristol PEWS	✓	✓	✗	✓	○	✓	▲ Scored the same as adapted National and Newcastle PEWS	✓	✓	✓	✓
National PEWS	○	✓	○	✓	○	✓	▲ The three original 'severe' features not scored ▲ 'Mild recession' added to the mild category ▲ 'Marked subcostal recession' added to the severe category	✓	✓	✓	○
Newcastle PEWS	✓	✓	○	✓	✓	✓	▲ Three original 'severe' features not scored ▲ 'Mildrecession' added to the mild category ▲ 'Marked subcostal recession' added to the severe category	✓	✓	✓	✓
POPS[31]	✓	✗	▲ Category not scored	✓	✗ ▲ Nurse and parent concern scored as two points ▲ Nurse or parent concern	✗	▲ 'Marked subcostal recession' used as a surrogate for 'severe recession'	✓	✓	✗	✓
Scotland PEWS[22]	✓	✓	○	✓	○	✓	✗	✓	✓	✓	✓

Continued

Table 1 Continued

Scoring system	AVPU	CRT	Co-morbidities	HR	Nurse/parent concern	O₂ delivery	Respiratory distress	RR	O₂ sats	SBP	Temperature

✓=included and, as per the original PEWS, incorporated in the score calculation.
○=included but, as per the original PEWS, not incorporated in the score calculation.
X=not included in the original PEWS.
●= adaptation to parameter.
AVPU, Alert/Voice/Pain/Unresponsive scale; CRT, capillary refill time; HR, heart rate; O2, oxygen; PEWS, Paediatric Early Warning Score; POPS, Paediatric Observation Priority Score; RR, respiratory rate; sats, saturations; SBP, systolic blood pressure.

'severe' features in the National and Newcastle PEWS are not routinely recorded on the EPR so could not be scored. Conversely, 'mild recession' and 'marked subcostal recession' *are* routinely recorded on the EPR but are not explicitly categorised in the 'respiratory distress' severity grades described in the National and Newcastle PEWS (online supplemental figures 4,5). Therefore, the study authors took a pragmatic decision to include 'mild recession' in the mild category and 'marked subcostal recession' in the severe category (table 1). The latter decision was based partly on two previously published paediatric respiratory scoring systems that use 'marked' features in their severe categories only.[32 33] Similarly, for POPS, 'marked subcostal recession' was deemed a surrogate for 'severe recession' as the latter is not routinely recorded on the EPR (online supplemental figure 6). In the Bristol and Bedside PEWS, overall 'respiratory distress' is simply categorised as 'none', 'mild', 'moderate', or 'severe' but only individual 'respiratory effort' features are recorded on the EPR rather than this overall summary (online supplemental figures 2, 3). For Bristol and Bedside PEWS, the 'respiratory distress' category was scored identically to the National and Newcastle PEWS described above. This represents a less significant change to the original scores than other adaptations (eg, using the number of features present as in AH PEWS).

Other adaptations

The 'gut feeling' category in POPS (eg, low-level concern=1 point, child looks unwell=2 points) was also modified as there is no grading of the nurse and/or parent concern on the EPR (ie, it is either 'yes' or 'no'). To adapt the EPR information to POPS, nurse *and* parent concern was scored as 2 points, while nurse *or* parent concern was scored as 1 point. In addition, comorbidity information was unavailable, so this category in POPS could not be scored (table 1 and online supplemental figure 6). Finally, Newcastle PEWS is currently a 'trigger-based' system, whereas the other six systems are 'score-based'. To allow a more direct comparison, Newcastle PEWS was converted to a score-based system, with mild features scored as 1, moderate as 2, and severe as 4, as is the convention across Bedside, Bristol and National PEWS (online supplemental figures 2–5).

Statistical analysis

Area under the receiver operating characteristic curves (AUC) were calculated for each PEWS. The three optimal cut-offs within each PEWS for predicting CC admission were identified using Youden's J statistic.[34] Sensitivity, specificity, positive and negative predictive value, positive and negative likelihood ratio, OR and accuracy were calculated for each PEWS at each threshold, alongside an asymptotic 95% CI. Statistical analyses were performed using SPSS V.25.

Patient and public involvement

No patients or the public were involved in the present study.

RESULTS

Study population

Of 14 121 cases identified, 2672 were excluded: 2041 (14%) due to missing observation data, 454 (3.2%) due to no history of fever, 96 (0.7%) due to duplicate cases, 77 (0.5%) due to age ≥16 and 4 (0.03%) due to transfer from another hospital. A total of 11 449 cases remained in the final cohort (table 2 and online supplemental figure 8).

Missing data

A total of 2041 (14%) cases were excluded due to missing observation data. Blood pressure (BP) was by far the most frequently missed parameter, with 7569 (66%) cases in the final cohort missing a BP measurement, including 26% of all those admitted to CC. In comparison, oxygen delivery was absent in 118 (1.0%) final cohort cases, and all other parameters were missing in fewer than 0.2% of final cohort cases.

Outcomes

A total of 134 (1.2%) children were admitted to CC within 48 hours of ED attendance. A total of 114 were admitted to HDU, 28 to ICU, and 8 to both. Of those admitted to CC, 117 (87%) met Sepsis-3 criteria before CC admission,[35] while 99 (74%) met 2005 Goldstein criteria.[36] Hospital LOS was >48 hours in 606 (5.3%) children. A total of 10 (0.09%) children died and 5 (0.04%) of which were sepsis-related.

PEWS comparison

All seven PEWS demonstrated excellent discrimination (range AUC 0.91–0.95) for predicting CC admission within 48 hours of ED attendance (table 3). The greatest discrimination was seen with the Bedside PEWS (AUC 0.95; 95% CI 0.93 to 0.97), Bristol PEWS (AUC 0.95; 95% CI 0.92 to 0.97) and National PEWS (AUC 0.95; 95% CI 0.92 to 0.97). In threshold analyses, Bedside PEWS (AUC 0.90; 95% CI 0.86 to 0.93) and National PEWS (AUC 0.90; 0.87–0.93) demonstrated the best discrimination, both at a cut-off of ≥6. At this threshold, National PEWS achieved a sensitivity of 89.6% (95% CI 83.1 to 94.2) and a specificity of 90.6% (95% CI 90.1 to 91.2), whereas Bedside PEWS demonstrated a sensitivity of 87.3% (95% CI 80.5 to 92.4) and a specificity of 91.7% (95% CI 91.2 to 92.2). The remaining five PEWS had at least one threshold, whereby both sensitivity and specificity were over 80% (table 4).

Most PEWS demonstrated moderate discrimination for hospital LOS >48 hours (range AUC 0.69–0.75), with Bristol PEWS (AUC 0.75; 95% CI 0.73 to 0.77) and National PEWS (AUC 0.75; 95% CI 0.72 to 0.77) the most discriminative (table 3). For sepsis-related mortality, all seven PEWS demonstrated excellent discrimination (range AUC 0.95–0.99), with Bedside PEWS and Newcastle PEWS (both AUC 0.99; 95% CI 0.98 to 1) performing best (table 3). For both secondary outcomes,

Table 2	Summary of demographics, clinical outcomes and scores for each of the PEWS	
Demographics		
Number of cases	11 449	
Age in years, median (IQR)	2.5 (3.9)	
Male, n (%)	6112 (53)	
Outcomes		
Discharged home from ED, n (%)	9977 (87)	
Admitted to hospital, n (%)	1472 (13)	
Hospital LOS>48 hours, n (%)	606 (5.3)	
Admitted to CC within 48 hours (%)	134 (1.2)	
In-hospital mortality, n (%)	10 (0.09)	
Sepsis-related mortality, n (%)	5 (0.04)	
Hospital LOS in days, median (IQR)	0.12 (0.11)	
Inpatient LOS in days, median (IQR)	1.04 (2.73)	
CC LOS in days, median (IQR)	2.19 (4.28)	
Scoring systems		
Alder Hey PEWS[29]		
Range	0–29	
Median, (IQR)	1 (2)	
≥3, n (%)	1647 (14)	
≥4, n (%)	975 (8.5)	
≥5, n (%)	511 (4.5)	
Bedside PEWS[30]		
Range	0–21	
Median, (IQR)	2 (2)	
≥5, n (%)	1632 (14)	
≥6, n (%)	1057 (9.2)	
≥7, n (%)	726 (6.3)	
Bristol PEWS		
Range	0–29	
Median, (IQR)	2 (3)	
≥6, n (%)	1435 (13)	
≥7, n (%)	1005 (8.8)	
≥8, n (%)	752 (6.6)	
National PEWS		
Range	0–21	
Median, (IQR)	2 (3)	
≥5, n (%)	1857 (16)	
≥6, n (%)	1182 (10)	
≥7, n (%)	763 (6.7)	
Newcastle PEWS		
Range	0–35	
Median, (IQR)	5 (5)	
≥9, n (%)	1194 (10)	
≥10, n (%)	805 (7.0)	
≥11, n (%)	516 (4.5)	

Continued

Table 2	Continued
POPS[31]	
Range	0–13
Median, (IQR)	2 (2)
≥4, n (%)	2804 (25)
≥5, n (%)	1436 (13)
≥6, n (%)	725 (6.3)
Scotland PEWS[22]	
Range	0–17
Median, (IQR)	2 (3)
≥5, n (%)	1958 (17)
≥6, n (%)	783 (6.8)
≥7, n (%)	493 (4.3)

CC, critical care; ED, emergency department; LOS, length of stay; PEWS, Paediatric Early Warning Score; POPS, Paediatric Observation Priority Score.

the detailed performance of the seven PEWS as dichotomised scores is displayed in online supplemental tables 1 and 2).

DISCUSSION

Our results demonstrate excellent and relatively comparable performance across six different PEWS currently used in the UK, and the proposed National PEWS, in predicting CC admission in febrile children presenting to the ED. For predicting hospital LOS >48 hours, modest discrimination was seen. Given the proposed standardised system performs similarly to several systems currently in use, our findings provide evidence to support the use of the National PEWS in the ED in order to improve standardisation and reduce variability in escalation of care (as the same thresholds for escalation will be used nationally). Before applying the score to prehospital and inpatient settings, further validation in different cohorts across multiple presentations (eg, non-febrile and trauma) and settings will be required.

The development of National PEWS follows the development and introduction of an effective and standardised system across the entire adult patient journey in England.[5 15] There are clear theoretical and practical advantages in using a consistent 'language' to detect deterioration in patients across their entire journey. These benefits include (1) standardisation of care when transferring patients between hospitals, (2) consistency in handovers between primary care, ambulance services, ED, inpatient wards and ICU, (3) facilitated benchmarking and quality control across services and institutions, (4) improved standardised education of nursing and medical staff. Using a standardised system will allow for consistency of monitoring and make deterioration easier to recognise. Importantly, for staff members working in multiple departments or trusts (eg, bank/locum staff,

trainee doctors on rotation, staff moving between trusts), using the same system will increase the opportunity for standardised training and familiarity. This will reduce the risk of confusion, inconsistency and miscommunication, consequently improving early recognition of deterioration.

A critically ill child may present to the ED, and it is usually clear to the clinician that the child needs resuscitation, stabilisation and transfer to CC. Other children presenting earlier in the sepsis trajectory may not appear critically unwell in the ED and may be admitted to a ward, then deteriorate within the next 48 hours. PEWS are an ideal tool for identifying and tracking physiological changes, and the introduction of a standardised score would potentially allow the deterioration to be tracked from ED to ward to CC.

Further validation of the National PEWS is required in other presentations and other contexts, such as primary care and ambulance services. Studies of PEWS in these settings are limited. However, the Scotland PEWS has demonstrated promising performance in the prehospital setting,[22] and the system-wide introduction of NEWS has led to improved outcomes in adults with suspicion of sepsis.[15] Importantly, in the all-comer cohort studied by Corfield et al,[22] the performance of Scotland PEWS was inferior to all seven PEWS in the present study, including Scotland PEWS, suggesting the performance of these PEWS may be less discriminatory in presentations other than suspected sepsis. Furthermore, 79% of eligible cases were excluded from Corfield et al[22] due to missing data, highlighting a potential problem with the use of PEWS in the prehospital setting. A follow-up study provided a potential solution by demonstrating that mortality and ICU admission can be predicted as accurately using only four components as opposed to eight.[37] Even in the ED setting, missing vital sign data pose a significant problem. In the present study, 14% of cases were excluded due to missing two or more components. This would have increased to 69% if those missing one component had also been excluded. In our cohort, BP was by far the most commonly missed parameter, and similarly low rates of BP measurement in the paediatric ED have been reported consistently in the previous research.[38–40] This affects the development of scoring systems that include BP, as they are using cohorts with significant amounts of missing data. Furthermore, the 'real-world' performance of a score may be greatly affected if one parameter is consistently not measured, reducing opportunities for the reliable detection of serious illness.[41] One solution would be a quality improvement drive to ensure a minimum set of observations is done, including BP. However, current barriers to regular BP measurement, including inappropriate cuff-size, patient distress or lack of co-operation, and the concern over falsely elevated readings, may be difficult to overcome. An alternative solution would be to use simpler scoring systems, with fewer parameters, that do not include BP, and instead include parameters that are more consistently measured in the prehospital and

Table 3 AUCs for primary and secondary outcomes for each PEWS

Scoring system	Threshold	Critical care admission within 48 hours AUC (95% CI)	Hospital length of stay >48 hours AUC (95% CI)	Sepsis-related mortality AUC (95% CI)
Alder Hey PEWS[29]		0.94 (0.92 to 0.97)	0.71 (0.69 to 0.74)	0.98 (0.95 to 1)
	≥3	0.88 (0.85 to 0.91)	0.67 (0.64 to 0.69)	0.93 (0.89 to 0.97)
	≥4	0.88 (0.85 to 0.92)	0.65 (0.62 to 0.67)	0.96 (0.94 to 0.98)
	≥5	0.87 (0.83 to 0.91)	0.63 (0.60 to 0.65)	0.88 (0.67 to 1)
Bedside PEWS[30]		0.95 (0.93 to 0.97)	0.74 (0.72 to 0.77)	0.99 (0.98 to 1)
	≥5	0.88 (0.85 to 0.91)	0.67 (0.65 to 0.70)	0.93 (0.89 to 0.97)
	≥6	0.90 (0.86 to 0.93)	0.66 (0.64 to 0.69)	0.95 (0.93 to 0.98)
	≥7	0.88 (0.84 to 0.92)	0.64 (0.62 to 0.67)	0.97 (0.95 to 0.99)
Bristol PEWS		0.95 (0.92 to 0.97)	0.75 (0.73 to 0.77)	0.97 (0.93 to 1)
	≥6	0.89 (0.86 to 0.92)	0.67 (0.65 to 0.70)	0.94 (0.91 to 0.97)
	≥7	0.88 (0.85 to 0.92)	0.65 (0.63 to 0.68)	0.86 (0.65 to 1)
	≥8	0.87 (0.83 to 0.91)	0.64 (0.62 to 0.67)	0.87 (0.66 to 1)
National PEWS		0.95 (0.93 to 0.97)	0.75 (0.72 to 0.77)	0.97 (0.94 to 1)
	≥5	0.88 (0.84 to 0.90)	0.67 (0.65 to 0.70)	0.92 (0.88 to 0.96)
	≥6	0.90 (0.87 to 0.93)	0.67 (0.64 to 0.70)	0.95 (0.92 to 0.98)
	≥7	0.89 (0.85 to 0.93)	0.65 (0.63 to 0.68)	0.87 (0.66 to 1)
Newcastle PEWS		0.92 (0.89 to 0.95)	0.69 (0.67 to 0.72)	0.99 (0.98 to 1)
	≥9	0.87 (0.83 to 0.90)	0.65 (0.63 to 0.68)	0.95 (0.92 to 0.98)
	≥10	0.87 (0.82 to 0.91)	0.63 (0.61 to 0.66)	0.97 (0.95 to 0.98)
	≥11	0.85 (0.81 to 0.90)	0.62 (0.59 to 0.65)	0.98 (0.97 to 0.99)
POPS[31]		0.91 (0.88 to 0.94)	0.71 (0.68 to 0.73)	0.95 (0.91 to 0.99)
	≥4	0.83 (0.79 to 0.86)	0.66 (0.64 to 0.68)	0.88 (0.82 to 0.94)
	≥5	0.85 (0.81 to 0.89)	0.65 (0.62 to 0.67)	0.94 (0.91 to 0.97)
	≥6	0.81 (0.76 to 0.86)	0.62 (0.59 to 0.64)	0.77 (0.51 to 1)
Scotland PEWS[22]		0.91 (0.88 to 0.94)	0.71 (0.69 to 0.74)	0.97 (0.95 to 0.99)
	≥5	0.83 (0.79 to 0.86)	0.66 (0.64 to 0.69)	0.92 (0.87 to 0.96)
	≥6	0.82 (0.78 to 0.87)	0.64 (0.62 to 0.67)	0.97 (0.95 to 0.98)
	≥7	0.78 (0.73 to 0.83)	0.61 (0.59 to 0.64)	0.78 (0.52 to 1)

AUC, area under receiver operating characteristic curve; LOS, length of stay; PEWS, Paediatric Early Warning Score; POPS, Paediatric Observation Priority Score.

ED settings. We have previously developed such a score, the Liverpool quick Sequential Organ Failure Assessment (LqSOFA) and demonstrated its ability to identify those children at risk of poorer outcomes within a febrile ED cohort.[27] Compared with the AH PEWS, the LqSOFA was less discriminative (AUC 0.93 vs 0.81).[27] However, the LqSOFA has the advantage of using fewer parameters (four vs eight), which are measured easily and consistently, while still achieving good discrimination. This means it may be ideally suited for settings where BP measurement is inconsistent or impossible, such as prehospital, and resource-poor settings, but the LqSOFA requires further validation in these settings.

In addition to the missing data, the present study has some other important limitations. Our study evaluated only the 'detection' or afferent aspect of the PEWS system. The rest of the PEWS system, including the escalation plan and the efferent arm (eg, RRS), could not be assessed in this retrospective study, but are crucial components in preventing critical deterioration and improving patient outcomes.[19] Furthermore, the escalation plan of the AH PEWS may have influenced the performance of the scores in predicting CC admission, as urgent medical review is required for those with scores of 6 or greater (online supplemental figure 1). Second, it was necessary to adapt some of the included PEWS to allow a comparison to be made, which may have affected the performance of individual scores. In particular, comorbidity information was unavailable, possibly reducing performance of POPS, as this is the only PEWS in our study that

Table 4 Prognostic performance of each PEWS in predicting critical care admission within 48 hours

Scoring system and threshold		OR (95% CI)	Sensitivity; % (95% CI)	Specificity; % (95% CI)	Positive predictive value; % (95% CI)	Negative predictive value; % (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)	Accuracy; % (95% CI)
Alder Hey PEWS[29]	≥3	55 (31.5 to 95)	89.6 (83.1 to 94.2)	86.5 (85.9 to 87.1)	7.3 (6.8 to 7.8)	99.9 (99.8 to 99.9)	6.6 (6.2 to 7.2)	0.12 (0.07 to 0.20)	86.5 (85.9 to 87.2)
	≥4	65 (40 to 104)	84.3 (77.1 to 90.0)	92.4 (91.9 to 92.9)	11.6 (10.6 to 12.6)	99.8 (99.7 to 99.9)	11.1 (10.0 to 12.2)	0.17 (0.11 to 0.25)	92.3 (91.8 to 92.8)
	≥5	93 (61 to 141)	77.6 (69.6 to 84.4)	96.4 (96.0 to 96.7)	20.4 (18.3 to 22.6)	99.7 (99.6 to 99.8)	21.6 (18.9 to 24.6)	0.23 (0.17 to 0.32)	96.2 (95.8 to 96.5)
Bedside PEWS[30]	≥5	51 (30.0 to 88)	88.8 (82.2 to 93.6)	86.6 (86.0 to 87.3)	7.3 (6.8 to 7.8)	99.9 (99.8 to 99.9)	6.6 (6.2 to 7.2)	0.13 (0.08 to 0.21)	86.7 (86.0 to 87.3)
	≥6	76 (46 to 127)	87.3 (80.5 to 92.4)	91.7 (91.2 to 92.2)	11.1 (10.2 to 12.0)	99.8 (99.8 to 99.9)	10.5 (9.6 to 11.5)	0.14 (0.09 to 0.22)	91.6 (91.1 to 92.1)
	≥7	76 (49 to 118)	81.3 (73.7 to 87.6)	94.6 (94.1 to 95.0)	15.0 (13.6 to 16.5)	99.8 (99.7 to 99.8)	14.9 (13.3 to 16.7)	0.20 (0.14 to 0.28)	94.4 (94.0 to 94.8)
Bristol PEWS	≥6	60 (35.1 to 103)	88.8 (82.2 to 93.6)	88.4 (87.8 to 89.0)	8.3 (7.7 to 8.9)	99.9 (99.8 to 99.9)	7.6 (7.1 to 8.3)	0.13 (0.08 to 0.20)	88.4 (87.8 to 89.0)
	≥7	63 (39.3 to 101)	84.3 (77.1 to 90.0)	92.1 (91.6 to 92.6)	11.2 (10.3 to 12.2)	99.8 (99.7 to 99.9)	10.7 (9.7 to 11.8)	0.17 (0.11 to 0.25)	92.0 (91.5 to 92.5)
	≥8	63 (41 to 96)	79.1 (71.2 to 85.6)	94.3 (93.9 to 94.7)	14.1 (12.8 to 15.5)	99.7 (99.6 to 99.8)	13.9 (12.4 to 15.5)	0.22 (0.16 to 0.31)	94.1 (93.7 to 94.5)
National PEWS	≥5	47 (27.1 to 82)	89.6 (83.1 to 94.2)	84.7 (84.0 to 85.3)	6.5 (6.0 to 6.9)	99.9 (99.8 to 99.9)	5.8 (5.4 to 6.3)	0.12 (0.08 to 0.20)	84.7 (84.0 to 85.4)
	≥6	83 (47 to 145)	89.6 (83.1 to 94.2)	90.6 (90.1 to 91.2)	10.2 (9.4 to 10.9)	99.9 (99.8 to 99.9)	9.5 (8.8 to 10.4)	0.12 (0.07 to 0.19)	90.6 (90.1 to 91.1)
	≥7	83 (52 to 133)	83.6 (76.2 to 89.4)	94.3 (93.8 to 94.7)	14.7 (13.4 to 16.1)	99.8 (99.7 to 99.9)	14.5 (13.1 to 16.2)	0.17 (0.12 to 0.26)	94.1 (93.7 to 94.6)
Newcastle PEWS	≥9	46 (29.0 to 72)	82.8 (75.4 to 88.8)	90.4 (89.9 to 91.0)	9.3 (8.5 to 10.1)	99.8 (99.7 to 99.9)	8.7 (7.9 to 9.5)	0.19 (0.13 to 0.28)	90.3 (89.8 to 90.9)
	≥10	58 (37.6 to 88)	79.1 (71.2 to 85.6)	93.8 (93.4 to 94.3)	13.2 (11.9 to 14.5)	99.7 (99.6 to 99.8)	12.8 (11.4 to 14.3)	0.22 (0.16 to 0.31)	93.7 (93.2 to 94.1)
	≥11	74 (50 to 110)	73.9 (65.6 to 81.1)	96.3 (96.0 to 96.7)	19.1 (17.1 to 21.4)	99.7 (99.6 to 99.8)	20.1 (17.5 to 23.0)	0.27 (0.20 to 0.36)	96.1 (95.7 to 96.4)
POPS[31]	≥4	25.5 (14.9 to 44)	88.8 (82.2 to 93.6)	76.3 (75.5 to 77.1)	4.2 (4.0 to 4.5)	99.8 (99.7 to 99.9)	3.74 (3.49 to 4.0)	0.15 (0.09 to 0.24)	76.4 (75.6 to 77.2)
	≥5	32.8 (21.2 to 51)	81.3 (73.7 to 87.6)	88.3 (87.7 to 88.9)	7.6 (7.0 to 8.3)	99.8 (99.6 to 99.8)	6.9 (6.3 to 7.6)	0.21 (0.15 to 0.30)	88.2 (87.6 to 88.8)
	≥6	34.4 (23.8 to 50)	67.2 (58.5 to 75.0)	94.4 (94.0 to 94.8)	12.4 (11.0 to 14.0)	99.6 (99.5 to 99.7)	12.0 (10.4 to 13.8)	0.35 (0.27 to 0.44)	94.1 (93.6 to 94.5)
Scotland PEWS[22]	≥5	22.3 (14.4 to 34.6)	81.3 (73.7 to 87.6)	83.7 (83.0 to 84.3)	5.6 (5.1 to 6.1)	99.7 (99.6 to 99.8)	5.0 (4.5 to 5.5)	0.22 (0.16 to 0.32)	83.6 (82.9 to 84.3)
	≥6	37.6 (25.7 to 55)	70.9 (62.4 to 78.4)	93.9 (93.5 to 94.4)	12.1 (10.8 to 13.6)	99.6 (99.5 to 99.7)	11.7 (10.2 to 13.3)	0.31 (0.24 to 0.40)	93.7 (93.2 to 94.1)
	≥7	39.1 (27.3 to 56)	59.7 (50.9 to 68.1)	96.4 (96.0 to 96.7)	16.2 (14.1 to 18.7)	99.5 (99.4 to 99.6)	16.4 (13.8 to 19.4)	0.42 (0.34 to 0.51)	95.9 (95.5 to 96.3)

PEWS, Paediatric Early Warning Score; POPS, Paediatric Observation Priority Score.

incorporates this variable into the overall score. Third, some PEWS include features that are not accounted for in the overall score, such as 'staff or carer concern' in the Scotland PEWS (table 1 and online supplemental figure 7). Similarly, other practical factors such as acceptability to staff, ease-of-use and compliance could not be considered but are likely to affect overall performance. Other limitations include the use of retrospective data from a single centre, the criteria for CC admission which may depend on local practice, and the low mortality rate.

CONCLUSION

Our results demonstrate highly comparable performance of seven different PEWS. The proposed National PEWS demonstrated comparable or slightly improved performance compared with the other six PEWS, supporting the use of National PEWS in the paediatric ED. However, further validation is required in other settings (primary care and prehospital) and other presentations, such as trauma, asthma and diabetic ketoacidosis, to allow the proposed National PEWS to be used as a common language of illness severity across the health system.

Author affiliations
[1]Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK
[2]Paediatric Intensive Care Unit, Alder Hey Children's NHS Foundation Trust, Liverpool, UK
[3]Paediatric Immunology, Infectious Diseases & Allergy, Great North Children's Hospital, Newcastle Upon Tyne, UK
[4]Section of Paediatrics, Division of Infectious Diseases, Department of Medicine, Imperial College London, London, UK
[5]Paediatric Infectious Disease Department, Bristol Royal Hospital for Children, Bristol, UK
[6]The Jessop Wing Neonatal Unit, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[7]Child Health Research Centre, The University of Queensland, Saint Lucia, Queensland, Australia
[8]Children's Research Center, Neonatal and Pediatric Intensive Care Unit, University Children's Hospital Zürich, Zurich, Switzerland
[9]College of Biomedical and Life Sciences, Centre for Trials Research, Cardiff University, Cardiff, UK
[10]Department of Infectious Diseases, Alder Hey Children's NHS Foundation Trust, Liverpool, Merseyside, UK

Twitter Enitan D Carrol @CarrolEnitan

ORCID iDs
Sam T Romaine http://orcid.org/0000-0001-8411-7967
Philip Pallmann http://orcid.org/0000-0001-8274-9696

REFERENCES

1 Hébert A, Boucher M-P, Guimont C, National Institute for Health and Care Excellence. Acutely ill adults in hospital: recognising and responding to deterioration. Available: https://www.nice.org.uk/guidance/cg50 [Accessed 2 Jul 2020].
2 Pearson G. *Why children die: a pilot study*. London: CEMACH, 2008.
3 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. https://www.rcplondon.ac.uk/file/8504/download
4 Gao H, McDonnell A, Harrison DA, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive Care Med* 2007;33:667–79.
5 Royal College of Physicians. National early warning score (news) 2, 2017. Available: https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2 [Accessed 2 Jul 2020].
6 Smith GB, Prytherch DR, Meredith P, et al. 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;84:465–70.
7 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.
8 Churpek MM, Snyder A, Han X, et al. Quick sepsis-related organ failure assessment, systemic inflammatory response syndrome, and early warning scores for detecting clinical deterioration in infected patients outside the intensive care unit. *Am J Respir Crit Care Med* 2017;195:906–11.
9 Alam N, Vegting IL, Houben E, et al. Exploring the performance of the National early warning score (news) in a European emergency department. *Resuscitation* 2015;90:111–5.
10 Patel R, Nugawela MD, Edwards HB, et al. Can early warning scores identify deteriorating patients in pre-hospital settings? A systematic review. *Resuscitation* 2018;132:101–11.
11 Abbott TEF, Cron N, Vaid N, et al. Pre-Hospital national early warning score (news) is associated with in-hospital mortality and critical care unit admission: a cohort study. *Ann Med Surg* 2018;27:17–21.
12 Inada-Kim M, Knight T, Sullivan M, et al. The prognostic value of national early warning scores (news) during transfer of care from

community settings to hospital: a retrospective service evaluation. *BJGP Open* 2020;4:bjgpopen20X101071.

13 The Royal College of Emergency Medicine. Position statement: National early warning score (news) for adult patients attending emergency departments, 2016. Available: https://www.rcem.ac.uk/docs/News/CEM10125-Position%20statement%20-%20NEWS%20for%20adult%20patients%20attending%20EDs%20-%20June%202016.pdf [Accessed 2 Jul 2020].

14 NHS England. Sepsis guidance implementation advice for adults, 2017. Available: https://www.england.nhs.uk/wp-content/uploads/2017/09/sepsis-guidance- implementation-advice-for-adults.pdf [Accessed 2 Jul 2020].

15 Pullyblank A, Tavaré A, Little H, *et al*. Implementation of the National early warning score in patients with suspicion of sepsis: evaluation of a system-wide quality improvement project. *Br J Gen Pract* 2020;70:e381–8.

16 Roland D, Oliver A, Edwards ED, *et al*. Use of paediatric early warning systems in Great Britain: has there been a change of practice in the last 7 years? *Arch Dis Child* 2014;99:26–9.

17 Chapman SM, Wray J, Oulton K, *et al*. 'The score matters': wide variations in predictive performance of 18 paediatric track and trigger systems. *Arch Dis Child* 2017;102:487–95.

18 Seiger N, Maconochie I, Oostenbrink R, *et al*. Validity of different pediatric early warning scores in the emergency department. *Pediatrics* 2013;132:e841–50.

19 Chapman SM, Wray J, Oulton K, *et al*. Systematic review of paediatric track and trigger systems for hospitalised children. *Resuscitation* 2016;109:87–109.

20 Chapman SM, Maconochie IK. Early warning scores in paediatrics: an overview. *Arch Dis Child* 2019;104:395–9.

21 Trubey R, Huang C, Lugg-Widger FV, *et al*. Validity and effectiveness of paediatric early warning systems and track and trigger tools for identifying and reducing clinical deterioration in hospitalised children: a systematic review. *BMJ Open* 2019;9:e022105.

22 Corfield AR, Silcock D, Clerihew L, *et al*. Paediatric early warning scores are predictors of adverse outcome in the pre-hospital setting: a national cohort study. *Resuscitation* 2018;133:153–9.

23 Royal College of Paediatrics and Child Health. Paediatric Early Warning System (PEWSystem) - developing a standardised tool for England, 2019. Available: https://www.rcpch.ac.uk/resources/paediatric-early-warning-system-pewsystem- developing-standardised-tool-england [Accessed 2 Jul 2020].

24 Sands R, Shanmugavadivel D, Stephenson T, *et al*. Medical problems presenting to paediatric emergency departments: 10 years on. *Emerg Med J* 2012;29:379–82.

25 Liu L, Johnson HL, Cousens S, *et al*. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet* 2012;379:2151–61.

26 Irwin AD, Grant A, Williams R, *et al*. Predicting risk of serious bacterial infections in febrile children in the emergency department. *Pediatrics* 2017;140:e201162853.

27 Romaine ST, Potter J, Khanijau A, *et al*. Accuracy of a modified qSOFA score for predicting critical care admission in febrile children. *Pediatrics* 2020;146:e20200782.

28 Weiss SL, Balamuth F, Hensley J, *et al*. The epidemiology of hospital death following pediatric severe sepsis: when, why, and how children with sepsis die. *Pediatr Crit Care Med* 2017;18:823–30.

29 Sefton G, Lane S, Killen R, *et al*. Accuracy and efficiency of recording pediatric early warning scores using an electronic physiological surveillance system compared with traditional Paper-Based documentation. *Comput Inform Nurs* 2017;35:228–36.

30 Parshuram CS, Bayliss A, Reimer J, *et al*. Implementing the bedside paediatric early warning system in a community hospital: a prospective observational study. *Paediatr Child Health* 2011;16:e18–22.

31 Roland D, Arshad F, Coats T, *et al*. Baseline characteristics of the paediatric observation priority score in emergency departments outside its centre of derivation. *Biomed Res Int* 2017;2017:1–5.

32 Brashers V, Huether S. Alterations of pulmonary function in children. In: McCance K, Huether S, eds. *Pathophysiology: the biological basis for disease in adults and children*. 8th ed. St. Louis, MO: Elsevier, 2019: 1202–27.

33 Nayani K, Naeem R, Munir O, *et al*. The clinical respiratory score predicts paediatric critical care disposition in children with respiratory distress presenting to the emergency department. *BMC Pediatr* 2018;18:339.

34 Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–5.

35 Matics TJ, Sanchez-Pinto LN. Adaptation and validation of a pediatric sequential organ failure assessment score and evaluation of the Sepsis-3 definitions in critically ill children. *JAMA Pediatr* 2017;171:e172352:171.

36 Goldstein B, Giroir B, Randolph A, *et al*. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005;6:2–8.

37 Corfield AR, Clerihew L, Stewart E, *et al*. The discrimination of quick paediatric early warning scores in the pre-hospital setting. *Anaesthesia* 2020;75:353–8.

38 Gilhotra Y, Willis F. Blood pressure measurements on children in the emergency department. *Emerg Med Australas* 2006;18:148–54.

39 Hagedoorn NN, Zachariasse JM, Moll HA. Association between hypotension and serious illness in the emergency department: an observational study. *Arch Dis Child* 2020;105:545–51.

40 Burchell C, Michie CA. Blood pressure measurement in a district general paediatric A and E department. *Arch Dis Child* 2005;90:1097.

41 Hébert A, Boucher M-P, Guimont C, *et al*. Effect of measuring vital signs on recognition and treatment of septic children. *Paediatr Child Health* 2017;22:13–16.