Use of paediatric early warning scores in intermediate care units

Marie Emilie Lampin,[1,2] Alain Duhamel,[2,3] Hélène Behal,[3] Morgan Recher,[4] Francis Leclerc,[2,4] Stéphane Leteurtre[2,4]

ABSTRACT

Objective Paediatric early warning scores (EWS) were developed to detect deterioration in paediatric wards or emergency departments. The aim of this study was to assess the relationship between three paediatric EWS and clinical deterioration detected by the nurse in paediatric intermediate care units (PImCU).

Methods This was a prospective, observational, multicentre study at seven French regional hospitals that included all children <18 years of age. Clinical parameters included in three EWS (Paediatric Advanced Warning Score, Paediatric Early Warning Score and Bedside Paediatric Early Warning System) were prospectively recorded every 8 hours or in case of deterioration. The outcome was a call to physician by the nurse when a clinical deterioration was observed. The cohort was divided into derivation and validation cohorts. An updated methodology for repeated measures was used and discrimination was estimated by the area under the receiver-operating curve.

Results A total of 2636 children were included for 14 708 observations to compute a posteriori the EWS. The discrimination of the three EWS for predicting calls to physicians by nurses was good (range: 0.87–0.91) for the derivation cohort and moderate (range: 0.71–0.76) for the validation cohort. Equations for probability thresholds of calls to physicians, taking into account the time t, the score at time t and the score at admission, are available.

Conclusion These three EWS developed for children in paediatric wards or emergency departments can be used in PImCU to detect a clinical deterioration and predict the need for medical intervention.

INTRODUCTION

Intermediate care units (ImCU) or high dependency care units (HDC) are units between regular wards and intensive care units (ICUs) for patients who require monitoring due to potential organ failure.[1] A review on utilisation of ImCU reported diversity in their formats but common denominators are continuous monitoring and respiratory support, without mechanical ventilation and multiple vasoactive medications.[2] Children hospitalised in paediatric ImCU (PImCU) are at high risk of deterioration. Many paediatric early warning scores (EWS), aimed at the detection of deterioration, have been developed for patients in paediatric wards or emergency departments, using transfer to PICU, request for emergency assistance, or cardiac arrest as outcome variables.[3 4] These physiology-based scoring systems should alert staff to detect deterioration and accelerate access to

What is already known on this topic?

► No early warning scores (EWS) have been tested in paediatric intermediate care units (PImCU) while these children are at high risk of clinical deterioration.

What this study adds?

► It is the first study to focus on EWS used in PImCU and proposes call to physician by nurse as outcome for deterioration in patients.

appropriate intervention. No EWS has been developed for the population of PImCU to predict clinical deterioration. The aim of this study was to assess the relationship between three paediatric EWS (Paediatric Advanced Warning Score (PAWS),[5] Paediatric Early Warning Score (PEWS)[6] and Bedside Paediatric Early Warning System (Bedside PEWS)[7]) and the deterioration of a child's condition detected by the nurse in PImCU and predict the need for medical intervention.

METHODS
Setting

We performed a prospective observational study in seven PImCU of regional hospitals in northern France, each comprising four to six beds. Recruitments took place between 7 September 2012 and 7 January 2014. All patients admitted were assessed for eligibility.

Data collection

Collected data included demographic parameters, medical background, course of the care, primary reason for admission and primary disease at admission.

Clinical parameters (referred to as 'observations' in this paper) were collected by nurses on patient day sheets (standardised for the study in the seven hospitals) for each patient at PImCU admission (=H0), every 8 hours, and at each time the nurse detected deterioration in the child's condition (collection data form provided in online supplementary appendix A, table S1). These parameters included temperature, heart rate, blood pressure, capillary refill time, oxygen saturation, respiratory rate, work of breathing (0–3), apnoea, oxygen therapy (room air/cannula/mask, flow and

fractional inspired oxygen), level of consciousness (conscious/voice response/pain response/unconscious). The three EWS, easy to use and routinely feasible at the bedside (PAWS coted 0–21; PEWS 0–9 and Bedside PEWS 0–26; detailed in online supplementary appendix A, table S2-S4), were computed a posteriori from these clinical parameters. Note that, the value of the scores was not available to nurses and physicians, and, thus not used to inform the decision-making at the bedside.

Outcome

A call to physician by the nurse (yes or no) was the outcome variable. It was reported 'yes' only when the physician was called because the nurse was worried about the child's condition.

Reproducibility

One day per week and for one observation per day and per patient (day and time of the collection left to the choice of the teams), the clinical parameters were collected simultaneously by the nurse in charge of the child and by a second observer (the attending physician), to determine reproducibility. The collection time was chosen by the team of the different centres but once chosen, remained the same throughout the study.

Sample size

We estimated the number of calls to physician by nurses in hospitalised patients during their entire stay around 10%, as proposed by Tucker et al.[6] Calculation of the sample size was based on the method described by Hanley and McNeil.[8] We assumed an expected area under the receiver-operating curve (AUC) of 0.80. To show that the AUC was strictly better than 0.75, with a one-sided test (type I error=0.05, power=0.8) required 2000 patients. We, therefore, sought to recruit 2500 patients (corresponding to a power >0.85).

Statistical analysis

When one or two of the clinical parameters were missing, the score was calculated by considering the missing values as normal. The scores with more than two missing parameters were not analysed. The missing data for each score are presented in online supplementary appendix A, table S5. A sensitivity analysis was conducted to determine the impact of default to normal imputation compared with a complete-case analysis (these results are presented in online supplementary appendix B, table S6. The method of default to normal imputation was used up to two missing items per score. The interobserver agreement (nurse-physician) was assessed by calculating the intraclass correlation coefficient for each score.

The cohort was divided into derivation (the first 70% included patients) and validation (the remaining 30%) cohorts. In the derivation cohort, a logistic regression model with a random subject effect (general linear mixed model) to account for multiple observations per subject was performed. In this model, the fixed effects were the time t, the score at time t and the score at admission. This model allowed estimation of the predicted probability at each time t. The discriminant power of the model was estimated by the AUC associated with the predicted probabilities. The OR of each score adjusted for the score at admission and time was calculated with their 95% CI. For each score, we computed the threshold for the predicted probability, maximising the specificity for a sensitivity fixed to 90% allowing establishment of a prediction rule for calls to physician by nurses.
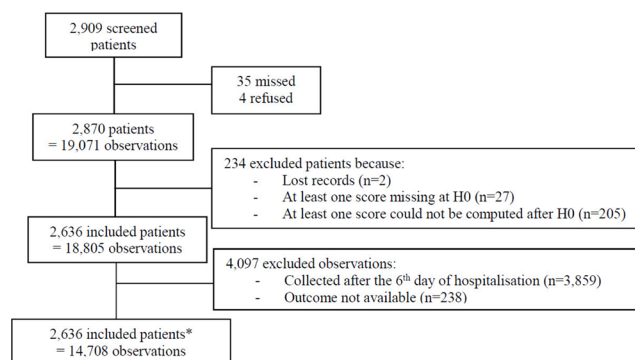

Figure 1 Flow chart of the inclusions. *The distribution of patient inclusions in the seven centres were 369 for centre 1, 333 for centre 2, 477 for centre 3, 246 for centre 4, 272 for centre 5, 465 for centre 6 and 474 for centre 7.

In the validation cohort, the AUC were calculated using the coefficients of the fixed effects estimated on the derivation cohort as done by Foulkes et al.[9] We assessed the stability of the prediction rules by computing the sensitivity and specificity corresponding to the thresholds determined in the derivation cohort. While we choose to exclude a random effect for centre from our prediction model, as it is not possible to factor this into a prediction rule for calls to the physician, we provide information on the centre differences, and the impact of centre variation on the AUC in online supplementary appendix B, table S7 and table S8. Analyses were performed using the observations from admission to 24 hours (as median length of stay), to 36 hours, to 48 hours and to 6 days (6 days corresponding to the 95th percentiles of the duration of the stay in PImCU).

All analyses were performed using a two-tailed test with an alpha level of 0.05. Statistical analyses were performed using SAS software V.9.4 (SAS Institute, Cary, North Carolina, USA).

Exclusion criteria of the statistical analysis: patients who did not have scores at H0 or for whom no score could be computed after H0 and observations recorded after 6 days were excluded from analysis.

Ethical considerations

This was an observational study which required no intervention; therefore, the institutional review board 'Société de Réanimation de Langue Française' waived the need for informed consent. All patients and their parents who were able to provide consent received written and oral information prior to the study and had the option for their data to be excluded from the study (NCT 02304341).

RESULTS

Of the 2909 patients assessed for eligibility, 2636 were included. A flow chart is presented in figure 1 and patients' characteristics are listed in table 1.

The results of the comparison between excluded and included patients (234 vs 2636 patients) are presented in table 1. Of the excluded patients, 88% (205/234) were excluded because the three scores could not be calculated at least once after H0, which implies a very short duration of stay (<8 hours). These patients were older, more frequently at home before admission to the PImCU (20.9%), had a lower incidence of respiratory failure (34.6%) at admission and a lower rate of infection (41.0%) than the included patients (all p<0.05).

Reproducibility of the three scores was good: the intraclass correlation coefficients (95% CI) for PAWS, PEWS and Bedside

Table 1 Baseline characteristics of the paediatric intermediate care unit population

Baseline characteristics	Total, n=2870	Included patients, n=2636	Excluded patients*, n=234	P value
Sex ratio	1.26	1.26	1.32	ns
Median age months (IQR)	29 (5–103)	28 (5–98)	42 (7–162)	0.01
Median length of stay days (IQR)	1.0 (1.0–3.0)	1.5 (0.8–2.7)	0.2 (0.1–0.3)	<10⁻⁴
Comorbidities, n (%)†	945 (32.9)	876 (33.2)	69 (29.5)	ns
Location prior to admission, n (%)†				0.001
Emergency	2006 (69.9)	1851 (70.2)	155 (66.2)	
Paediatric ward	386 (13.4)	361 (13.7)	25 (10.7)	
Operating room	71 (2.5)	68 (2.6)	3 (1.3)	
PICU	25 (0.9)	25 (0.9)	0 (0.0)	
Home	377 (13.1)	328 (12.4)	49 (20.9)	
Type of stay, n (%)†				ns
Medical	2714 (94.6)	2490 (94.5)	224 (95.7)	
Surgical	119 (4.1)	113 (4.3)	6 (2.6)	
Primary reason for admission, n (%)†				0.002
Respiratory	1272 (44.3)	1191 (45.2)	81 (34.6)	
Neurological	627 (21.8)	580 (22.0)	47 (20.0)	
Cardiovascular	218 (7.6)	198 (7.5)	20 (8.5)	
Gastrointestinal	222 (8.2)	202 (7.7)	20 (8.5)	
Others	529 (18.4)	465 (17.6)	64 (27.3)	
Cause of primary diseases, n (%)†				<10⁻⁴
Infectious	1546 (53.9)	1450 (55.0)	96 (41.0)	
Toxicological	201 (7.0)	176 (6.7)	25 (10.7)	
Congenital	136 (4.7)	127 (4.8)	9 (3.8)	
Trauma	126 (4.4)	114 (4.3)	12 (5.1)	
Oncological	65 (2.3)	42 (1.6)	23 (9.8)	
Others	794 (27.7)	727 (27.6)	67 (28.6)	
Destination, n (%)†				
Home	1024 (35.7)	937 (35.5)	87 (37.2)	
Paediatric ward	1586 (55.3)	1487 (56.4)	99 (42.3)	
Others	246 (8.6)	201 (7.6)	45 (19.2)	
Transfer to PICU n (%)†	85 (3.0)	68 (2.6)	17 (7.3)	<10⁻⁴
Death, n	1	1	0	
Median score at admission (IQR)				
PAWS	2 (0–4)	3 (2–4)	2 (1–4)	0.001
PEWS	0 (0–2)	1 (0–3)	0 (0–3)	0.005
Bedside PEWS	3 (0–5)	3 (2–6)	3 (1–5)	0.004

*Numbers and percentages take into account missing data.
†234 excluded patients: lost records (n=2), at least one score missing at H0 (n=27), at least one score could not be computed after H0 (n=205).
Bedside PEWS, Bedside Paediatric Early Warning System; ns, not significant; PICU, paediatric intensive care unit; PAWS, Paediatric Advanced Warning Score; PEWS, Paediatric Early Warning Score.

PEWS were 0.807 (0.802 to 0.812) (n=372 observations), 0.857 (0.853 to 0.861) (n=375) and 0.806 (0.801 to 0.811) (n=367), respectively.

Of 19 071 observations collected, 14 708 were included into analysis allowing calculating 12 668 (86.1%) PAWS, 11 756 (79.9%) PEWS and 12 191 (82.9%) Bedside PEWS complete scores and 1710 (11.6%) PAWS, 2916 (19.8%) PEWS and 2078 (14.1%) Bedside PEWS additional scores, after considering one or two missing parameters as normal.

A call to physician by nurse occurred in 1064 (7%) observations. The median scores for calls to physician by nurse were 3 (2–5) for the PAWS, 2 (0–3) for the PEWS and 5 (2–8) for the Bedside PEWS.

Performances of the three scores for predicting calls to physician by nurses on the derivation cohort are presented in table 2. These results regarded the complete scores and imputed scores with default to normal up to two missing items: sensitivity analysis results were comparable (online supplementary appendix B, table S6). AUC of the three scores were between 0.87 (≤24 hours) and 0.91 (≤6 days). Equations of probability thresholds of calls to

physician, taking into account the time t, the score at t time and the score at admission are provided in table 2.

Specificities were 48% and 49% (≤24 hours) and increased to 75% and 76% (≤6 days) (table 2). Negative predictive values were excellent (98% and 99%) for all scores and at all time points (table 2).

Each one-point increase in the PAWS, PEWS and Bedside PEWS significantly increased the risk of a physician being called (OR (95% CI): 1.37 (1.30 to 1.44), 1.57 (1.47 to 1.67) and 1.30 (1.26 to 1.35), respectively.

The prediction rule built from the derivation cohort (if the predicted probability was greater than the threshold) was applied to the validation cohort (table 3).

AUC of the three scores were between 0.71 (PAWS≤6 days) and 0.76 (Bedside PEWS≤24 hours). Sensitivities of the three scores were maximal at H24 (76% to 81%) and decreased until day 6 (55% to 66%), while the prevalence of calls to physician also decreased from 11% to 7%. Negative predictive values were excellent (95% and 96%) for all scores and at all-time points.

Table 2 Performances of the PAWS, PEWS and Bedside PEWS scores on the derivation sample on the first 24 hours, first 36 hours, first 48 hours and first 6 days

Total scores analysed (complete and imputed with 'default to normal' method)

Scores	AUC (95% CI)	Probability threshold (%)[*]	Specificity (%)	PPV (%)	NPV (%)
≤24 hours (n=4693 observations; call to physician=494 (11%) observations)					
PAWS[†]	0.87 (0.85 to 0.89)	5.9	49	17	98
PEWS[‡]	0.87 (0.86 to 0.89)	6.0	48	17	98
Bedside PEWS[§]	0.87 (0.85 to 0.89)	5.2	48	16	98
≤36 hours (n=5892 observations; call to physician=559 (9%) observations)					
PAWS[†]	0.90 (0.88 to 0.91)	7.3	64	20	98
PEWS[‡]	0.89 (0.88 to 0.91)	6.7	55	17	98
Bedside PEWS[§]	0.89 (0.87 to 0.91)	6.2	58	17	98
≤48 hours (n=7196 observations; call to physician=629 (9%) observations)					
PAWS[†]	0.91 (0.90 to 0.92)	7.5	69	22	99
PEWS[‡]	0.91 (0.89 to 0.92)	7.4	71	23	99
Bedside PEWS[§]	0.91 (0.89 to 0.92)	6.7	70	21	99
≤6 day (n=10 152 observations; call to physician=749 (7%) observations)					
PAWS[†]	0.91 (0.90 to 0.92)	7.1	77	23	99
PEWS[‡]	0.91 (0.90 to 0.92)	7.1	76	23	99
Bedside PEWS[§]	0.91 (0.90 to 0.92)	6.1	75	21	99

If PAWS at admission=3, time=48 hours, PAWS (H48)=4.
Predictive equation for *call to physician:* Logit=$-2.867+0.315\times(4)-0.072\times(3)-0.016\times(48)=-2.591$.
Probability of call to physician=$1/(1+\exp(-\text{Logit}))=7.0\%$.
This probability is <7.5% (probability threshold for PAWS from admission to 48 hours): so nurse must not call physician.
*Probability threshold maximising specificity for sensitivity fixed at 90%: probability computed from the general linear mixed model included the following effects: the score at time t, the score at admission and the time t.
[†]Predictive equation for call to physician: Logit=$-2.867+0.315\times$(score at the time t)$-0.072\times$(score at admission)$-0.016\times$(time t); probability of call to physician=$1/(1+\exp(-\text{Logit}))$.
[‡]Predictive equation for call to physician: Logit=$-2.630+0.448\times$(score at the time t)$-0.101\times$(score at admission)$-0.017\times$(time t); probability of call to physician=$1/(1+\exp(-\text{Logit}))$.
[§]Predictive equation for call to physician: Logit=$-3.033+0.263\times$(score at the time t)$-0.058\times$(score at admission)$-0.015\times$(time t); probability of call to physician=$1/(1+\exp(-\text{Logit}))$.
AUC, area under the receiver-operating curve; Bedside PEWS, Bedside Paediatric Early Warning System; NPV, negative predictive value; PAWS, Paediatric Advanced Warning Score; PEWS, Paediatric Early Warning Score; PPV, positive predictive value.

DISCUSSION

This study is the first to assess the use of paediatric EWS in PImCU in detecting deterioration of children. The discriminative ability estimated by the AUCs of the three scores for predicting calls to physician were good to excellent (range: 0.87–0.91) on the derivation cohort and were moderate (range: 0.71–0.76) on the validation cohort. Equations of probabilities thresholds of calls to physician, taking into account the score at time t, the score at admission and time t, are available. The three EWS had good interobserver reproducibilities.

PImCU are recent units in which children at high risk of deterioration are hospitalised.[1] Four systematic reviews of paediatric EWS have been published.[3 4 10 11] To develop or validate paediatric EWS, a gold standard that establishes clinical deterioration is necessary, but there is no consensus about the gold standard for this event.[3] There were multiple outcome measures: death,[12] cardiac arrest or code blue,[13–15] unplanned transfer to PICU or requirement for PICU,[6 7 15–21] a call for urgent medical assistance or rapid response system (RRS) activation[22–25] and length of hospital stay[22]. However,

Table 3 Performances of the PAWS, PEWS and Bedside PEWS scores on the validation sample on the first 24 hours, first 36 hours, first 48 hours and first 6 days

	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
≤24 hours (n=2028 observations; call to physician=219 (11%) observations)					
PAWS	0.74 (0.70 to 0.78)	76	50	15	95
PEWS	0.74 (0.70 to 0.78)	77	48	15	95
Bedside PEWS	0.76 (0.72 to 0.80)	81	46	15	95
≤36 hours (n=2550 observations; physician call=245 (10%) observations)					
PAWS	0.74 (0.70 to 0.77)	72	60	16	95
PEWS	0.75 (0.71 to 0.78)	78	53	15	96
Bedside PEWS	0.75 (0.72 to 0.79)	79	53	15	96
≤48 hours (n=3156 observations; call to physician=267 (8%) observations)					
PAWS	0.73 (0.70 to 0.77)	64	71	17	96
PEWS	0.75 (0.72 to 0.79)	73	63	15	96
Bedside PEWS	0.75 (0.72 to 0.79)	71	65	15	96
≤6 day (n=4556 observations; call to physician=315 (7%) observations)					
PAWS	0.71 (0.68 to 0.74)	55	76	14	96
PEWS	0.75 (0.72 to 0.78)	58	75	15	96
Bedside PEWS	0.75 (0.72 to 078)	66	69	13	96

Performances of scores were calculated from a predictive equation derived on the development sample; sensitivity and specificity were calculated using thresholds measured on the development sample.
AUC, area under the receiver-operating curve; Bedside PEWS, Bedside Paediatric Early Warning System; PPV, positive predictive value; NPV, negative predictive value; PAWS, Paediatric Advanced Warning Score; PEWS, Paediatric Early Warning Score.

the aim of these EWS was to identify deterioration before respiratory or cardiac arrest or transfer to a PICU. Moreover, most of these outcomes are rare events, and this affects the methodology used for the validation of EWS. In our study, the outcome variable used was call to physician by nurses in case of deterioration. This outcome is not ideal as it is influenced by numerous factors: experience of the nurse, knowledge of the patient, relationship between the nurse and the physician, ease of calling and workload. Call to physician by nurses may be a confounding factor because the collection of clinical parameters at the time of 'deterioration' depends on subjective clinical judgement by the nurse. In our study, it was not verified if call to physician by the nurse was justified. Bonafide *et al* in a qualitative study to identify mechanisms beyond the statistical ability of use of EWS by nurses and physician suggested that combination of EWS and clinical judgement could be a better system for detecting deterioration.[26] Jensen *et al* in a study interested in factors that may compromise the use of EWS in clinical practice reported the lack of clinical judgement in EWS.[27]

Many EWS were retrospectively developed and their validity was evaluated using a case/control type of methodology with a number of cases fewer than 120.[7 14 15 19 20 28 29]

Our study was large, prospective and multicentric, with clinical parameters of the scores (computed a posteriori) recorded for each patient at admission, every 8 hours and in case of deterioration and outcome (call to physician by the nurse) was evaluated for all collected scores. So there is a dependent relationship between regular observations evaluated by the nurse either every 8 hours or at any time of the stay. In current practice, a score to detect deterioration that is repeated at regular intervals is more relevant than a single score. We have used an update methodology for repeated measures. However, a score threshold could not be identified in this mixed model due to repeated measures in the same patient. The thresholds that have been proposed for EWS are variable and cannot be compared with each other.[4] Using the outcome *cardiac arrest*, *RRS activation* or *PICU admission*, in case/control studies, thresholds proposed for the Bedside PEWS were ≥7 and ≥ 8.[7 14 16 17 28] For different modified PEWS, the proposed thresholds were between ≥2 and ≥5[6 13 14 18 19 29] and for PAWS ≥3.[5] Two prospective studies used *PICU admission* as the outcome: the threshold was ≥1 for PEWS[30] and ≥3 for the PAWS.[31] Sieger *et al* analysed the validity of different EWS in 17 943 children and observed that the optimal threshold to calculate sensitivity and specificity for PICU admission was low (threshold at 1), except for the Bedside PEWS, which had a threshold of 3: AUCs were 0.77 for the PAWS, 0.79 for the PEWS and 0.82 for the Bedside PEWS.[32] In our study, AUCs of the three scores were good to excellent (0.87–0.91) for the derivation cohorts and moderate (0.71–0.76) for the validation cohorts. On the validation cohort, sensitivity decreased (from 76%–81% to 55%–66%) and specificity increased over time for the three scores (from 46%–50% to 69–76). Mulherin *et al* indicated that sensitivity and specificity subgroup variations were not a bias but was clinically relevant information to be identified and reported.[33] These authors suggested replacing the term 'spectral bias' by 'spectrum effect' that reflects the heterogeneity in the test performance when applied to different subgroups, like in our study that included different subgroups regarding length of PImCU stay. These three scores seem to better detect deterioration at the beginning of the stay and their performances decreased over time; this may be related to the stabilisation of patients, thus requiring fewer calls to physician throughout the stay.

The reproducibility of the three scores was good (coefficient range: 0.81–0.86). Few studies have analysed the interobserver reproducibility of these EWS: Chaiyakulsil and Pandee[31] reported a good inter-rater reliability (kappa=0.75) for the PAWS and Gold *et al*[30] reported an excellent inter-rater reliability (intraclass coefficient=0.91) for the PEWS.

None of the three scores appeared to be better than the others for detecting deterioration. However, the PAWS may be preferred for quick initial assessment in emergency care or in the ward because blood pressure measurement is not necessary. The Bedside PEWS has the advantage of considering oxygen saturation and oxygen supply, contrary to the PAWS, which only considers oxygen saturation, but it does not include any neurological assessment. The PEWS is divided into three categories—respiratory, circulatory and neurological, corresponding to the different types of organ failure and, thus, seems easier to use.

Sambeeck *et al* in a cross-sectional survey revealed that the use of 45 different EWS scores in Dutch hospitals can lead to a false sense of security and recommended to establish a national working group to coordinate implementation of an EWS usable for both general and university hospitals.[34] A recent review emphasised that despite widespread use, the evidence base for EWS remains limited because there is no consensus of the most effective EWS and there exists a lack of robust, valid and clinically meaningful outcomes.[35] In the evaluating processes of care and outcomes of children in hospital (EPOCH) Randomised Clinical Trial, including 144 539 patients, the effect of the Bedside PEWS intervention was assessed on all-cause hospital mortality and late admission to PICU, cardiac arrest and ICU resource use. Despite this large number of patients, to compensate the low number of events/outcome, responsible for the lack of robustness on the validity of the EWS, the Bedside PEWS intervention did not significantly decrease all-cause mortality, compared with usual care.[36] Some studies have proposed other approaches to test paediatrics EWS: Jensen *et al* conducted a multicentre, randomised controlled trial, comparing two different EWS models to predict deterioration requiring transfer to a higher level of care.[37] In this study, despite a large number of enrolled patients (n=16 213), 22 unplanned transfers to a higher level of care were identified. No significant difference between the two scores (Bedside PEWS and CDR PEWS) was identified but CDR PEWS seemed more acceptable to staff.[37] Thomas-Jones *et al* proposed a prospective, mixed-methods, before and after study-based approach to improvement. This study is still ongoing.[38] We propose another approach: the use of a prediction rule taking into account the EWS on admission and at the time *t* when the deterioration occurs, rather than a score threshold or an escalation algorithm indicating the care-team action.[36]

Limitations

First, there are missing data. The score was calculated considering the missing values (1 or 2) as normal. How are managed missing data not often explained. For the Bedside PEWS, Parshuram *et al*[16] took the most recent recorded data when the corresponding data were missing[16]. Some authors attributed normal values whatever the number of missing parameters.[5 14 15] Others used a multiple imputation model.[31 32] In our study, multiple imputation was not used because this would have added repetition in a model in which there were already repeated measurements. Second, although this study was multicentric, it included seven PImCU of regional hospitals from the same region, which may represent a recruitment bias. Patients between centres could be different. However, the potential centre effect could not be added as random effect in the model because it would not allow to obtain a rule of predictive decision for the call to physician by the nurse. Furthermore, when

we adjusted to the centre, the discriminant power evaluated by the AUC was similar. Moreover, the patients included in our study were admitted to PImCU not attached in PICU and these patients may be different (maybe less severe) from the patients hospitalised in PImCU attached to PICU. As reported by Plate *et al*, there is a great diversity of these ImCU.[2] A focus on HDC for children in the UK recommends to separate three care level for activity of these HDC (level 1: enhanced care unit, level 2: critical care unit, level 3: intensive care unit). Our population would correspond to level 1 or 2.[39] This could explain our transfer rate to PICU (3%), barely higher than that reported by Sieger *et al* in an emergency department (2%).[32] Third, because we used a methodology for repeated measures, we do not provide an absolute threshold value but equations of probability thresholds for calls to a physician, taking into account the score at time t, the score at admission and time t are available.

CONCLUSION

This prospective multicentre observational study indicates that three EWS (PAWS, PEWS and Bedside PEWS) initially developed for children admitted to paediatric wards or presenting to the emergency departments can be used in PImCU to predict the need for medical intervention. Further studies are needed in different contexts and different countries to validate the usefulness of EWS with a prediction rule for call to physician in case of deterioration.

REFERENCES

1 Jaimovich DG, American Academy of Pediatrics Committee on Hospital Care and Section on Critical Care. Admission and discharge guidelines for the pediatric patient requiring intermediate care. *Pediatrics* 2004;113:1430–3.
2 Plate JDJ, Leenen LPH, Houwert M, *et al*. Utilisation of intermediate care units: a systematic review. *Crit Care Res Pract* 2017;2017:1–10.
3 Lambert V, Matthews A, MacDonell R, *et al*. Paediatric early warning systems for detecting and responding to clinical deterioration in children: a systematic review. *BMJ Open* 2017;7:e014497.
4 Chapman SM, Wray J, Oulton K, *et al*. Systematic review of paediatric track and trigger systems for hospitalised children. *Resuscitation* 2016;109:87–109.
5 Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J* 2008;25:745–9.
6 Tucker KM, Brewer TL, Baker RB, *et al*. Prospective evaluation of a pediatric inpatient early warning scoring system. *J Spec Pediatr Nurs* 2009;14:79–85.
7 Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the bedside paediatric early warning system score. *Crit Care* 2009;13:R135.
8 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
9 Foulkes AS, Azzoni L, Li X, *et al*. Prediction based classification for longitudinal biomarkers. *Ann Appl Stat* 2010;4:1476–97.
10 Chapman SM, Grocott MP, Franck LS. Systematic review of paediatric alert criteria for identifying hospitalised children at risk of critical deterioration. *Intensive Care Med* 2010;36:600–11.
11 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.
12 Olson D, Davis NL, Milazi R, *et al*. Development of a severity of illness scoring system (inpatient triage, assessment and treatment) for resource-constrained hospitals in developing countries. *Trop Med Int Health* 2013;18:871–8.
13 Akre M, Finkelstein M, Erickson M, *et al*. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics* 2010;125:e763–9.
14 Robson MA, Cooper CL, Medicus LA, *et al*. Comparison of three acute care pediatric early warning scoring tools. *J Pediatr Nurs* 2013;28:e33–41.
15 Duncan H, Hutchison J, Parshuram CS. The pediatric early warning system score: a severity of illness score to predict urgent medical need in hospitalized children. *J Crit Care* 2006;21:271–8.
16 Parshuram CS, Duncan HP, Joffe AR, *et al*. Multicentre validation of the bedside paediatric early warning system score: a severity of illness score to detect evolving critical illness in hospitalised children. *Crit Care* 2011;15:R184.
17 Fuijkschot J, Vernhout B, Lemson J, *et al*. Validation of a paediatric early warning score: first results and implications of usage. *Eur J Pediatr* 2015;174:15–21.
18 Mandell IM, Bynum F, Marshall L, *et al*. Pediatric early warning score and unplanned readmission to the pediatric intensive care unit. *J Crit Care* 2015;30:1090–5.
19 Agulnik A, Forbes PW, Stenquist N, *et al*. Validation of a pediatric early warning score in hospitalized pediatric oncology and hematopoietic stem cell transplant patients. *Pediatr Crit Care Med* 2016;17:e146–53.
20 Skaletzky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. *Clin Pediatr* 2012;51:431–5.
21 Tume L. The deterioration of children in ward areas in a specialist children's hospital. *Nurs Crit Care* 2007;12:12–9.
22 Bell D, Mac A, Ochoa Y, *et al*. The Texas children's hospital pediatric advanced warning score as a predictor of clinical deterioration in hospitalized infants and children: a modification of the PEWS tool. *J Pediatr Nurs* 2013;28:e2–9.
23 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–e22.
24 McKay H, Mitchell IA, Sinn K, *et al*. Effect of a multifaceted intervention on documentation of vital signs and staff communication regarding deteriorating paediatric patients. *J Paediatr Child Health* 2013;49:48–56.
25 Randhawa S, Roberts-Turner R, Woronick K, *et al*. Implementing and sustaining evidence-based nursing practice to reduce pediatric cardiopulmonary arrest. *West J Nurs Res* 2011;33:443–56.
26 Bonafide CP, Roberts KE, Weirich CM, *et al*. Beyond statistical prediction: qualitative evaluation of the mechanisms by which pediatric early warning scores impact patient safety. *J Hosp Med* 2013;8:248–53.
27 Jensen CS, Nielsen PB, Olesen HV, *et al*. Pediatric early warning score systems, nurses perspective - a focus group study. *J Pediatr Nurs* 2018;41:e16–e22.
28 Gawronski O, Ciofi Degli Atti ML, Di Ciommo V, *et al*. Accuracy of Bedside Paediatric Early Warning System (BedsidePEWS) in a pediatric stem cell transplant unit. *J Pediatr Oncol Nurs* 2016;33:249–56.
29 McLellan MC, Gauvreau K, Connor JA. Validation of the Cardiac Children's Hospital early warning score: an early warning scoring tool to prevent cardiopulmonary arrests in children with heart disease. *Congenit Heart Dis* 2014;9:194–202.
30 Gold DL, Mihalov LK, Cohen DM. Evaluating the Pediatric Early Warning Score (PEWS) system for admitted patients in the pediatric emergency department. *Acad Emerg Med* 2014;21:1249–56.
31 Chaiyakulsil C, Pandee U. Validation of pediatric early warning score in pediatric emergency department. *Pediatr Int* 2015;57:694–8.
32 Seiger N, Maconochie I, Oostenbrink R, *et al*. Validity of different pediatric early warning scores in the emergency department. *Pediatrics* 2013;132:e841–50.
33 Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med* 2002;137:598–602.
34 Sambeeck SJV, Fuijkschot J, Kramer BW, *et al*. Pediatric early warning system scores: lessons to be learned. *J Pediatr Intensive Care* 2018;7:27–32.
35 Chapman SM, Maconochie IK. Early warning scores in paediatrics: an overview. *Arch Dis Child* 2019;104:395–9.

36 Parshuram CS, Dryden-Palmer K, Farrell C, *et al*. Effect of a pediatric early warning system on all-cause mortality in hospitalized pediatric patients: the EPOCH randomized clinical trial. *JAMA* 2018;319:1002–12.

37 Jensen CS, Olesen HV, Aagaard H, *et al*. Comparison of two pediatric early warning systems: a randomized trial. *J Pediatr Nurs* 2019;44:e58–65.

38 Thomas-Jones E, Lloyd A, Roland D, *et al*. A prospective, mixed-methods, before and after study to identify the evidence base for the core components of an effective Paediatric Early Warning System and the development of an implementation package containing those core recommendations for use in the UK: Paediatric early warning system - utilisation and mortality avoidance- the PUMA study protocol. *BMC Pediatr* 2018;18:244.

39 RCPCH. High dependency care for children - time to move on (2014). https://www.rcpch.ac.uk/resources/high-dependency-care-children-time-move (accessed July 11, 2019).