Available online at www.sciencedirect.com

Resuscitation Plus

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

Clinical paper

Predicting transfers to intensive care in children using CEWT and other early warning systems

Kevin McCaffery [a], *Kyle A. Carey* [b], *Victoria Campbell* [a], *Shaune Gifford* [a], *Kate Smith* [a], *Dana Edelson* [b], *Matthew M. Churpek* [c,d], *Anoop Mayampurath* [c,d,*]

Abstract

Background and Objective: The Children's Early Warning Tool (CEWT), developed in Australia, is widely used in many countries to monitor the risk of deterioration in hospitalized children. Our objective was to compare CEWT prediction performance against a version of the Bedside Pediatric Early Warning Score (Bedside PEWS), Between the Flags (BTF), and the pediatric Calculated Assessment of Risk and Triage (pCART).

Methods: We conducted a retrospective observational study of all patient admissions to the Comer Children's Hospital at the University of Chicago between 2009–2019. We compared performance for predicting the primary outcome of a direct ward-to-intensive care unit (ICU) transfer within the next 12 h using the area under the receiver operating characteristic curve (AUC). Alert rates at various score thresholds were also compared.

Results: Of 50,815 ward admissions, 1,874 (3.7%) experienced the primary outcome. Among patients in Cohort 1 (years 2009–2017, on which the machine learning-based pCART was trained), CEWT performed slightly worse than Bedside PEWS but better than BTF (CEWT AUC 0.74 vs. Bedside PEWS 0.76, P < 0.001; vs. BTF 0.66, P < 0.001), while pCART performed best for patients in Cohort 2 (years 2018–2019, pCART AUC 0.84 vs. CEWT AUC 0.79, P < 0.001; vs. BTF AUC 0.67, P < 0.001; vs. Bedside PEWS 0.80, P < 0.001). Sensitivity, specificity, and positive predictive values varied across all four tools at the examined thresholds for alerts.

Conclusion: CEWT has good discrimination for predicting which patients will likely be transferred to the ICU, while pCART performed the best.

Keywords: Risk management, Critical care, Electronic health records, Pediatrics

Introduction

Hospitalized children who experience unrecognized or poorly managed deterioration are at high risk for mortality and long-term morbidity.[1–3] Studies have demonstrated that these children often show signs of impending deterioration hours in advance.[2,4,5] Therefore, accurate and early recognition of deterioration may enable intervention that could improve outcomes in children.[5]

Early warning (EW) systems are often used for early recognition of deterioration in children. Initial algorithms, such as Monaghan's Pediatric Early Warning Score (PEWS, later referred to as the Brighton PEWS) or the Cardiff & Vale score, have evolved into current standards such as Bedside PEWS.[6–9] Other early warning systems, such as the Australian Between the Flags, are single track-and-trigger in that a threshold breach of a single vital sign will trigger an escalation in patient care.[10] The Children's Early Warning Tool (CEWT) is used in over 100 public and private hospitals in Australia and other sites around the world to ensure timely recognition and escalation of care for deteriorating children. The CEWT is a combination system that integrates an aggregated weighted score with a track-and-trigger system. Designed through human factors research,[11,12] it facilitates longitudinal tracking of scores and provides tiered escalation guidance (see **Supplementary Fig. 1**).

CEWT was initially implemented in the state of Queensland, Australia in 2010, with revisions over time based on recommendations from incidents and clinician users. It has a proven track record with regard to patient outcomes, as the Queensland Health Department requires reporting and expert panel review of serious adverse event incidents such as death or permanent disability. [13,14] A recent report indicates that there were 36 serious adverse event incidents across the state between 2012 and 2017 that were related to deterioration. Fourteen were associated with the incorrect use of CEWT (not used, no score generated, wrong chart used, incorrect calculation, not

* Corresponding author at: 265H WARF Office Building, 610 Walnut Street, Madison, WI 53726.
E-mail address: mayampurath@wisc.edu (A. Mayampurath).

https://doi.org/10.1016/j.resplu.2023.100540
Received 21 August 2023; Received in revised form 15 November 2023; Accepted 13 December 2023

escalated), with no indication that the set thresholds were inadequate.[13] However, though the CEWT scoring algorithm that is the foundation of the tool was developed from local, manually collected observed data, it has not been tested on a large dataset in an external setting.

The main objective of this study was to compare the predictive accuracy of the CEWT scoring system with that of other commonly used pediatric early warning scores and with pCART, a machine learning model we previously developed,[15] in identifying children at risk for being transferred to the ICU within 12 h. Additionally, we compare the performance of these tools in predicting critical deterioration, defined as an ICU transfer followed by initiation of mechanical ventilation, vasopressor administration, or death within 12 h of transfer.[16] Finally, to better understand the clinical utility of these tools, we compare the efficiency of alerts for derangement for each scoring system in terms of escalation burden, timeliness, and sensitivity measures at different thresholds.

Materials and methods

Setting and study population

We conducted an observational cohort study of all consecutive pediatric (age < 18 years) medical ward admissions to the University of Chicago Medicine (UCM) Comer Children's Hospital from April 1, 2009, to December 31, 2019. Birth encounters for newborn patients were excluded. UCM is a tertiary care center with approximately 5,000 pediatric admissions per year, with rapid response systems implemented since 2008. The Brighton pediatric early warning system was implemented in 2013 and was utilized for decisions regarding transfer to the ICU. All data elements were collected from the electronic health record (EHR [EPIC; Verona, WI]) or the hospital administration database. The study was approved by the local Institutional Review Board (IRB# 18–0645).

CEWT scoring system

Each score component within the CEWT scoring system is set based on age-based reference values, and a total cumulative CEWT score is calculated across all assessments. The tool, designed with human factors principles,[11,12] provides an observation chart, age-based reference ranges, color-coding of score components that promote easy identification of abnormal physiological measures, and a guided list of customizable interventions based on the range of the total score (see **Supplementary Fig. 1**). For example, CEWT scores of 4–5 (mild severity or early derangement) require escalation to a ward attending within 30 min, while scores of 6–7 (moderate derangement) require escalation to a medical registrar or ward attending within 15 min and initiation of a call to the emergency rescue team otherwise. CEWT scores of $\geq$8 (severe derangement) requires automatic notification to emergency rescue teams within the pediatric hospital. Thus, CEWT allows for the measurement of the severity of illness and guidance and provides guidance for management or intervention based on risk.

Outcomes

Our primary outcome of interest was a direct ward-to-ICU transfer within 12 h. The choice of 12 h was based on several prior studies.[9,15,17] Transfers from the ward to other hospital locations, such as a surgical operating room (OR), and then to the ICU did not constitute our primary outcome. Our secondary outcome of interest was critical deterioration, defined as the primary outcome followed by initiation of mechanical ventilation, administration of vasopressors, or mortality within 12 h of ICU admission.

Early warning tools

We compared four early warning tools: CEWT, BTF, Bedside PEWS, and pCART. We used modified versions of CEWT, Bedside PEWS, and BTF that included respiratory rate, oxygen saturation, temperature, heart rate, supplemental oxygen requirement, blood pressure, and neurological assessment using the alert-verbal-pain-unresponsive (AVPU) scale. Capillary refill time, pain scores, and noted indications of respiratory distress were not included as these could not be reliably extracted from the electronic health record. **Supplementary Table 1** lists scoring components for every EW tool used in this study. Data regarding the distribution and percentage of admissions with missing values across all variables in our cohort are provided in **Supplementary Table 2**.

Early warning thresholds

We set the following types of EW thresholds a priori based on how these scores are used clinically (see Table 1). Pediatric EW systems vary in not only their afferent arms (detection tools for deterioration) but in their efferent arms (recommended escalation in observation frequency and expertise required to review the patient). Mild derangement alerts (BTF: Yellow threshold breach, Bedside PEWS score: 3–4, CEWT score: 4–5, pCART 85–91) recommend escalation to a ward doctor (this may be a senior experienced nurse in some hospitals depending on resources and local procedure). Moderate derangement alerts (CEWT only) recommend escalation to a senior ward doctor (CEWT score: 6–7). Finally, severe derangement alerts (BTF: Red/3 x Yellow/severe threshold breach, Bedside PEWS score of at least 5, CEWT score of at least 8 or a purple/severe threshold breach, and pCART score of at least 92) recommend escalation to emergency response personnel, usually involving critical care expertise. Therefore, BTF, Bedside PEWS, and pCART have a 2-tiered medical expertise escalation response, while CEWT has a 3-tiered recommended medical response. Thresholds of CEWT were based on operational criteria and scoring recommendations used in Queensland Health. Thresholds for Bedside PEWS were based on prior published literature after modifying for the unavailable elements.[8,9,18,19] Thresholds for pCART were set based on current operational criteria in place at the University of Chicago, which were established independently of this analysis to balance sensitivity with available response resources.

Statistical analysis

Patient characteristics for the cohort were presented as counts, percentages, or median with inter-quartile range (IQR) based on the characteristic. We divided our cohort into two parts: Cohort 1 (years 2009–2017) and Cohort 2 (years 2018–2019) based on the admission date. The longitudinal split corresponded to derivation and validation datasets for pCART.[15] Missing data were handled by first carrying forward the last recorded observation followed by median imputation. Medians were calculated from Cohort 1 and used to impute missing data for Cohort 1 and 2. All observations within 12 h leading up to the outcome were marked as 1 for having the outcome, while others were marked as 0. We compared the area under the receiver operating characteristic curve (AUC) for CEWT, BTF,

Table 1 – Alert/escalation thresholds used by each early warning tool in this study.

	Early/mild derangement alert (escalation to medical officer)	Higher severity alert (escalation to advanced medical officer)	Severe derangement alert (escalation to Medical Emergency team)
BTF	Yellow (or "Not applicable"?)		Red or 3 yellows (or "Not applicable"?)
Bedside PEWS	3–4		>=5
CEWT	4–5	6–7	>=8 or any single parameter in the purple ("E") zone
pCART	85–91		>=92

and Bedside PEWS for Cohort 1, and CEWT, BTF, Bedside PEWS, and pCART for Cohort 2. This ensures that pCART performance was not assessed during the period when it was derived. DeLong's method was used to assess statistical significance for comparing AUC measures. We also compared the time to event (the difference between the time of first reaching a threshold and the actual outcome), sensitivity, and specificity between the tools. Finally, we compared the percentage of positive alerts against sensitivity for all tools at the mild and severe derangement alerts. All analyses were performed using Stata version 15.1 (Stata Corps; College Station, Texas) and R version 6.2 (R Project for Statistical Computing, Vienna, Austria), and a two-sided $P < 0.05$ was utilized to indicate statistical significance.

Results

Study population

Our cohort consisted of 50,815 patients (Cohort 1, years 2009–2017: 39,805 patients, Cohort 2, years 2018–2019: 11,010 patients). Of these, 1,874 (3.7%) experienced a ward-to-ICU transfer event at least once during their admissions. Patients in our study cohort were 55% male, 60% African American, and the median age was 5 years (see Table 2). Less than 1% suffered in-hospital mortality, and the top prior pediatric complex chronic condition was dependence upon medical technology.[20,21] A comparison of characteristics and outcomes among patients who went to the ICU and those who stayed on the ward is given in **Supplementary Table 3**.

Performance of early warning tools

Performance metrics of all early warning tools for predicting ICU transfer events and critical deterioration events are shown in Table 3. In Cohort 1, the CEWT model demonstrated slightly lower performance to Bedside PEWS (AUC 0.74 vs. 0.76, $P < 0.001$) but outperformed BTF (AUC 0.74 vs. 0.66, $P < 0.001$) for predicting ICU transfers within 12 h. **Supplementary Fig. 2** depicts the AUC performance of BTF, Bedside PEWS, and CEWT across all years in Cohort 1 for predicting ICU transfer events. While BTF performance stayed consistent across all years, CEWT and Bedside PEWS were marginally higher in the post-2013 period when Brighton PEWS was implemented in our center. In Cohort 2, the CEWT model demonstrated improved performance to BTF (AUC 0.79 vs. 0.67, $P < 0.001$) and similar performance to Bedside PEWS (AUC 0.79 vs. 0.80, $P < 0.001$) but lower performance compared to pCART (AUC 0.79 vs. 0.84, $P < 0.001$) at discriminating patients who were transferred to the ICU within 12 h from those who stayed on the ward. However, CEWT and PEWS were closer to pCART in terms of performance for predicting critical deterioration events (Table 3).

Sensitivity and specificity metrics at clinically used thresholds for all four early warning scores for the primary outcome within Cohort 2 are given in Table 4. The CEWT has the most sensitive (84%) initial notification alert, with a score of 1 (requiring the ward senior nurse to be aware), but a lower ward doctor/emergency response team alert sensitivity compared to the other tools. The most comparable thresholds in terms of specificity across the four tools were a BTF Red, Bedside PEWS severe at $\geq$5, CEWT early at $\geq$4, and pCART severe at $\geq$92, with pCART demonstrating the highest sensitivity. At a similar sensitivity of 40% (CEWT early at $\geq$ 4, BTF Red), CEWT had a higher specificity than BTF (94% vs. 84%). In addition, the positive predictive values for BTF, Bedside PEWS, CEWT, and pCART were 2%, 3%, 8%, and 5% at early derangement thresholds and 3%, 5%, 12%, and 7% at severe derangement thresholds respectively. The positive predictive value for CEWT at the higher severity thresholds (CEWT was 13%). Table 4 also depicts the median time to ICU transfer event from when thresholds are reached. At the early derangement thresholds, the median times to event for all tools were largely similar (10–16 h). At the higher severity threshold, CEWT had a median time to event of 6 h. At the severe derangement threshold, CEWT had a considerably lower median time to event than other alerts at higher thresholds (4 h vs. 16 h [BTF], vs. 11 h [Bedside PEWS], vs. 11 h [pCART]).

The EW tools also varied in the rate of threshold breaches. CEWT, which triggers pop-alerts in the digital system, had a similar number of alert triggers to BTF, triggering an alert in 45% and 43% of observations, respectively. However, CEWT had a significantly lower ward doctor escalation frequency (6% of included observations) compared to BTF (43%), Bedside PEWS (36%) and pCART (17%), and emergency response team escalation frequency (0.7% of included observations), compared to BTF (17%), Bedside PEWS (12%) and pCART (9%). Figs. 1 and 2 compare sensitivity against the alerting efficiency for all four EW tools by depicting the rates of alerts, i.e., the proportion of scores above thresholds, for the early and severe derangement thresholds.

Discussion

In this retrospective single-center study aimed at external assessment of CEWT, we demonstrate that CEWT has slightly lower performance than Bedside PEWS but outperformed BTF in discriminating patients likely to require ICU transfer and critical deterioration within 12 h compared to patients who remain on the ward. The pCART

Table 2 – Characteristics of patients in study cohort.

Demographic	Value	All encounters (n = 50815)
Age, years	Median (IQR)	5 (1, 12)
Sex	Female	22,907 (45%)
	Male	27,907 (55%)
	Unknown	1 (0%)
Race	White	13,949 (27%)
	Black/African American	30,403 (60%)
	Asian/Middle Eastern	924 (2%)
	American Indian or Alaska Native	112 (0%)
	Native Hawaiian or other Pacific Islander	56 (0%)
	More than one Race	2981 (6%)
	Unknown	2390 (5%)
Ethnicity	Hispanic	6107 (12%)
	Not Hispanic	42,862 (84%)
	Unknown	1846 (4%)
Encounter LOS (hours)	Median (IQR)	48 (26, 95)
ICU ever	N (%)	10,403 (20%)
Died ever	N (%)	63 (0%)
ICU outcome ever	N (%)	1874 (4%)
Critical event outcome ever	N (%)	270 (1%)
Most common prior pediatric chronic condition	Technology Dependence Conditions	5550 (11%)
	Neuromuscular Conditions	4678 (9%)
	Gastrointestinal Condition	4084 (8%)
	Cardiovascular Disease	3116 (6%)
	Metabolic Conditions	2241 (4%)

Table 3 – Distributions and AUCs for included early warning tool algorithms for Cohort 1 (years 2009–2017) and Cohort 2 (years 2018–2019) for primary and secondary outcomes. pCART AUCs are only shown for Cohort 2, as it was trained using data from Cohort 1.

Score	Cohort	ICU transfer AUC	Critical event AUC
BTF	Cohort 1	0.66 (0.66, 0.67)	0.72 (0.71, 0.73)
	Cohort 2	0.67 (0.66, 0.68)	0.72 (0.69, 0.74)
Bedside PEWS	Cohort 1	0.76 (0.76, 0.77)	0.81 (0.80, 0.82)
	Cohort 2	0.80 (0.80, 0.81)	0.84 (0.82, 0.86)
CEWT	Cohort 1	0.74 (0.74, 0.74)	0.79 (0.78, 0.80)
	Cohort 2	0.79 (0.78, 0.79)	0.82 (0.80, 0.84)
pCART	Cohort 2	0.84 (0.83, 0.84)	0.86 (0.84, 0.87)

model had the highest AUC, however, superiority cannot be quantified due to the exclusion of key scoring elements from the other tools and because pCART was derived using data from the same hospital. With fewer alerts and the load sharing of multiple severity thresholds, the resource implications make CEWT an attractive tool for the early identification and escalation of deterioration.

Prior research on pediatric EW tools for hospitalized children has focused on establishing predictive performance through analysis of model accuracy.[8,15,17,22–26] Additionally, the validation of these tools in other settings, such as the emergency department, has also been explored.[2,27,28] Other focus areas include adherence and impact on patient outcomes.[18,29,30] However, little research has been conducted on understanding the efficiency and timeliness of alerts made by these tools that indicate the likelihood of deterioration. For example, it is crucial that a balance be maintained between the number of alerts that an early warning tool issues and the true positive rate of these alerts. Too few alerts run the risk of unrecognized deterioration. However, too many alerts may lower the overall impact of the tool as it would be a significant contributor to alarm fatigue.[31] The timeliness of the alerts is also an essential factor wherein too early an alert may lead to unnecessary interventions, while an alert issued too late will not prevent adverse outcomes.[32] Our study fills this critical gap in knowledge by analyzing and comparing four pediatric early warning tools as an effective means of alerting care personnel about the likelihood of being transferred to the ICU within 12 h.

We note that three of the four tools are based on age and vital signs, whereas pCART also incorporates laboratory results as additional features. In our previous study, we observed a significant increase in AUC when information from labs was included in comparison to a vital sign-only model.[15] Further, we also observed a significant increase in AUC when extending from standard regression-based methods of deriving prediction models to more complex

Table 4 – Characteristics of different thresholds of the studied scoring systems using whether an ICU transfer event occurred within 12 h.

Tool	Score	Encounters (%)	Scores above threshold (%)	Median (IQR) hours to outcome	Sensitivity	Specificity
Any derangement threshold						
CEWT	>=1	9803 (89%)	151,193 (45%)	16 (6, 48)	84%	56%
Early derangement threshold						
BTF	Yellow/Moderate	9885 (90%)	146,903 (43%)	16 (7, 49)	69%	57%
Bedside PEWS	>=3	7986 (73%)	120,912 (36%)	15 (6, 48)	82%	65%
CEWT	>=4	3195 (29%)	21,120 (6%)	10 (3, 37)	40%	94%
pCART	>=85	4903 (45%)	56,059 (17%)	13 (5, 39)	64%	84%
Moderate severity threshold						
CEWT	>=6	1138 (10%)	5663 (2%)	6 (2, 27)	18%	99%
Severe derangement threshold						
BTF	Red/3 yellow/Severe	6923 (63%)	56,630 (17%)	16 (4, 53)	43%	84%
Bedside PEWS	>=5	4242 (39%)	40,024 (12%)	11 (4, 37)	53%	89%
CEWT	>=8	440 (4%)	2445 (0.7%)	4 (1, 26)	7%	99%
pCART	>=92	3188 (29%)	31,906 (9%)	11 (4, 32)	53%	91%

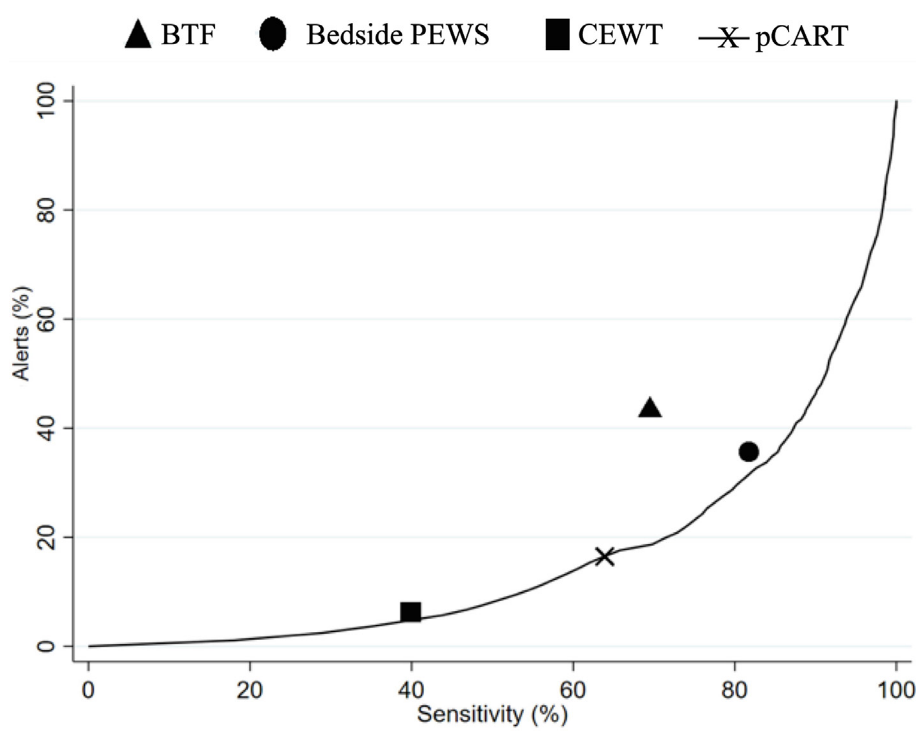

Fig. 1 – Sensitivity vs. Alerts for four early warning tools at thresholds corresponding to early derangement.

machine learning frameworks such as using gradient-boosted machines. Additionally, pCART was trained on the patients from the same hospital. These reasons, i.e., incorporating laboratory results as predictors, use of machine learning techniques, and internal derivation, may explain pCART's higher performance for discriminating patients at risk of being transferred to the ICU within 12 h from those likely to stay in the ward. However, for the outcome of critical deterioration events, which is less influenced by the ICU admission criteria variability, CEWT and Bedside PEWS were closer to pCART, and may be more comparable to pCART if EW score contributions of capillary refill and respiratory distress had been included for CEWT and Bedside PEWS in the analysis.

We compared the rate of threshold breaches, which trigger both alerts (automatic notifications, intrusive pop-ups, or passive screen item color changes depending on facility capability and preference) and the requirement for escalation to a ward doctor or emergency response team (which can disrupt workflows, cause conflict or impact patient care). We note that at thresholds corresponding to any derangement (CEWT score at or above 1, which triggers pop-up alerts in the digital system), the alert rate was similar to BTF.

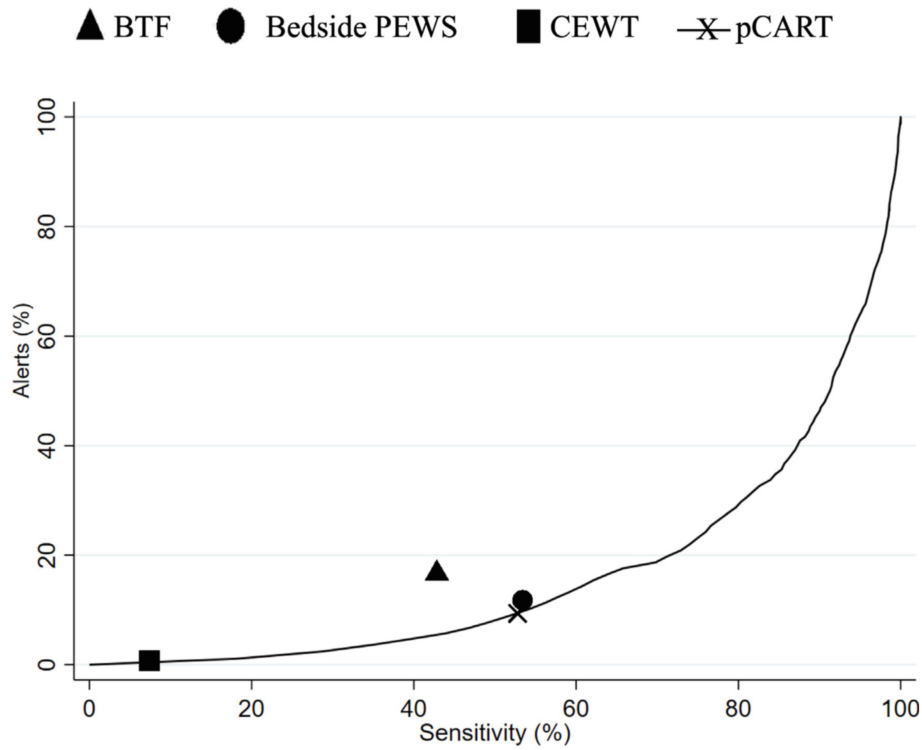

Fig. 2 – Sensitivity vs. Alerts for four early warning tools at thresholds corresponding to severe derangement.

However, at early derangements (requiring escalation to a doctor), CEWT issued the fewest alerts compared to the other tools. The median times to event at this threshold were similar, with CEWT recognizing events at 10 h in comparison to over 13 h for other tools. Notably, the availability of a moderate severity threshold within the CEWT system that requires escalation to a more senior ward officer ensures that appropriate resources are diverted to respond to the deteriorating patient. A three-level alert system such as CEWT is likely to provide additional flexibility in designing more efficient interventions based on patient condition and available resources compared to other two-level alert systems. They are also more likely to catch and arrest the decline prior to an emergency-level severe derangement alert. At these severe levels, which require emergency response team escalation, CEWT results in less than 10% of the alerts generated by the other tools. With a greater positive predictive value, there is a significant loss in sensitivity, but this must be considered in the context of the limitations described below (e.g., underestimation of sensitivity) and the observed success of the long-standing and widely implemented CEWT system.[14] Further, the addition of a higher severity alert in addition to the early and severe derangement alert that are common in offers clinicians and care providers flexibility in designing responses and interventions.

This study has a number of limitations. Most importantly, as for all algorithm validation studies, these recognition tools are only part of any EW System, which is made up of both recognition and response. The response component (efferent limb) varies between regions and institutions in escalation requirements, expected levels of expertise, and nursing vs doctor response processes, and the actions carried out depend on these as well as the clinical acumen of the responders. As a result, a direct relationship between the accuracy and escalation burden identified here and patient outcomes cannot be assumed. The recommended escalation may or may not occur, escalation may be encouraged in the absence of any abnormal vital signs if a clinician is concerned, and the success of the actions taken will vary due to patient and clinical factors. True full EW system outcome studies are required but are expensive and complex. There are other limitations. We only utilize available elements for the BTF, CEWT, and Bedside PEWS scoring schemes. As a result, scoring criteria regarding capillary refill time and indications of pain or respiratory distress, which are key markers of deterioration in children were not able to be assessed. While this facilitates the comparison of only objective elements within early warning tools, potentially decreasing variability,[33] our comparative study is limited in not using the original score definitions derived and validated with both subjective and objective elements. Another limitation of our study is that it is a retrospective analysis at a single institution. However, this is the first external performance assessment study of the CEWT scoring system. A further limitation is that transfer to ICU was used as a proxy for the occurrence of deterioration. The actual process of deterioration in a patient's health is a continuum, and the decision to transfer to ICU is likely to be multi-factorial, involving clinical acumen and resource availability. It was also not possible to determine the reason for ICU transfer and the impact of the deterioration on mortality due to the limited retrospective nature of our dataset. Prospective studies that capture the reason for ICU transfer through documentation of details surrounding the deterioration event will aid in understanding the differences between the tools with more nuance.[34] The secondary outcome of predicting critical deterioration requiring organ support better reflects significant deterioration, and as mentioned, the accuracy difference was less marked for this outcome. Future studies are needed to evaluate the impact of CEWT implementation on patient outcomes, such as improvements in mortality or length of stay, as well as assessment of the impact on hospital staff in terms of decreased alert burden and user satisfaction.

Conclusion

We analyzed the performance of the CEWT model in predicting ward-to-ICU transfer events among hospitalized children. CEWT was superior to BTF but slightly inferior to Bedside PEWS. pCART had a higher AUC, but superiority can't be confirmed due to the exclusion of key CEWT/Bedside PEWS scoring components and because pCART was internally derived. At the clinically utilized thresholds, CEWT had far fewer escalation threshold breaches than the other tools which have significant resource advantages. Further evaluation studies are needed to examine the impact of CEWT on pediatric outcomes and hospital operations.

CRediT authorship contribution statement

Kevin McCaffery: Conceptualization, Methodology, Formal Analysis, Investigation, Writing - Review & Editing, Funding Acquisition, Supervision. **Kyle Carey:** Data Curation, Software, Formal Analysis, Investigation, Writing - Review & Editing. **Victoria Campbell:** Formal Analysis, Writing - Review & Editing. **Shaune Gifford:** Methodology, Project Administration, Writing - Review & Editing. **Kate Smith:** Methodology, Project Administration, Writing - Review & Editing. **Dana Edelson:** Methodology, Formal Analysis, Writing - Review & Editing. **Matthew Churpek:** Methodology, Formal Analysis, Writing - Review & Editing. **Anoop Mayampurath:** Conceptualization, Methodology, Software, Formal Analysis, Investigation, Writing - Original Draft, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Drs. Churpek and Edelson have a patent (#11,410,777) for risk stratification algorithms for hospitalized patients. Dr. Edelson has received research support and honoraria from Philips Healthcare (Andover, MA). Dr. Edelson has ownership interest in AgileMD (San Francisco, CA), which licenses eCART, a patient risk analytic. Dr. Mayampurath is supported by a career development award from the National Heart, Lung, and Blood Institute (K01HL148390). Dr. Churpek is supported by a research grant from NHLBI (R01 HL157262).].

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resplu.2023.100540.

Author details

[a]*Queensland Health Patient Safety Centre, Brisbane, Queensland, Australia* [b]*Department of Medicine, University of Chicago, Chicago IL, United States* [c]*Department of Medicine, University of Wisconsin-Madison, Madison, WI, United States* [d]*Department of Biostatistics & Medical Informatics, University of Wisconsin-Madison, Madison, WI, United States*

REFERENCES

1. Tress EE, Kochanek PM, Saladino RA, Manole MD. Cardiac arrest in children. J Emerg Trauma Shock 2010;3:267–72. https://doi.org/10.4103/0974-2700.66528.
2. Lambert V, Matthews A, MacDonell R, Fitzsimons J. Paediatric early warning systems for detecting and responding to clinical deterioration in children: a systematic review. BMJ Open 2017;7. https://doi.org/10.1136/bmjopen-2016-014497.
3. Berg RA, Nadkarni VM, Clark AE, et al. Incidence and outcomes of cardiopulmonary resuscitation in pediatric intensive care units. Crit Care Med. 2016;44:798–808. https://doi.org/10.1097/CCM.0000000000001484.
4. Agulnik A, Gossett J, Carrillo AK, Kang G, Morrison RR. Abnormal vital signs predict critical deterioration in hospitalized pediatric hematology-oncology and post-hematopoietic cell transplant patients. Front Oncol. 2020;10:354. https://doi.org/10.3389/fonc.2020.00354.
5. Bonafide CP, Localio A, Roberts KE, Nadkarni VM, Weirich CM, Keren R. Impact of rapid response system implementation on critical deterioration events in children. JAMA Pediatr 2014;168:25–33. https://doi.org/10.1001/jamapediatrics.2013.3266.
6. Monaghan A. Detecting and managing deterioration in children. Paediatr Nurs 2005;17:32–5. https://doi.org/10.7748/paed2005.02.17.1.32.c964.
7. Edwards ED, Powell CVE, Mason BW, Oliver A. Prospective cohort study to test the predictability of the Cardiff and Vale paediatric early warning system. Arch Dis Child 2009;94:602. https://doi.org/10.1136/adc.2008.142026.
8. Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the Bedside Paediatric Early Warning System score. Crit Care 2009;13:R135. https://doi.org/10.1186/cc7998.
9. 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. https://doi.org/10.1186/cc10337.
10. Pain C, Green M, Duff C, et al. Between the flags: implementing a safety-net system at scale to recognise and manage deteriorating patients in the New South Wales Public Health System. Int J Qual Health Care. 2017;29:130–6. https://doi.org/10.1093/intqhc/mzw132.
11. Horswill MS, Preece MHW, Hill A, et al. Human factors research regarding observation charts: research project overview. Aust Comm Saf Qual Health Care 2010:1–20.
12. Hill A, Horswill MS, Preece MHW, Watson MO. The development of the adult deterioration detection system (ADDS) chart. Published online May 1, 2010:1-26.
13. Coulthard MG, Osborne JM, McCaffery K, McAuley SA, McEniery JA. Multi-incident analysis of reviews of serious adverse clinical events in children with serious bacterial infection and/or sepsis in Queensland, Australia between 2012 and 2017. J Paediatr Child Health. 2022;58:497–503. https://doi.org/10.1111/jpc.15759.
14. Clinical Incident Subcommittee, Queensland Paediatric Quality Council. *Multi-Incident Analysis of SAC1 Clinical Incident Reports Involving Children 2015-2019.*; 2023. https://www.childrens.health.qld.gov.au/wp-content/uploads/2015-2019-Multi-Incident-Analysis-of-SAC1-reports.pdf.
15. Mayampurath A, Sanchez-Pinto LN, Hegermiller E, et al. Development and external validation of a machine learning model for prediction of potential transfer to the ICU. Pediatr Crit Care Med 2022;23:514–23. https://doi.org/10.1097/PCC.0000000000002965.
16. Bonafide CP, Roberts KE, Priestley MA, et al. Development of a pragmatic measure for evaluating and optimizing rapid response systems. Pediatrics 2012;129. https://doi.org/10.1542/peds.2011-2784 e874.

17. Mayampurath A, Jani P, Dai Y, Gibbons R, Edelson D, Churpek MM. A Vital sign-based model to predict clinical deterioration in hospitalized children. Pediatr Crit Care Med 2020;21:820–6. https://doi.org/10.1097/PCC.0000000000002414.
18. 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. https://doi.org/10.1001/jama.2018.0948.
19. Romaine ST, Sefton G, Lim E, et al. 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. BMJ Open 2021;11. https://doi.org/10.1136/bmjopen-2020-044091 e044091.
20. Feudtner C, Feinstein JA, Zhong W, Hall M, Dai D. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. BMC Pediatr 2014;14:199. https://doi.org/10.1186/1471-2431-14-199.
21. Feinstein JA, Russell S, DeWitt PE, Feudtner C, Dai D, Bennett TD. R Package for pediatric complex chronic condition classification. JAMA Pediatr 2018;172:596–8. https://doi.org/10.1001/jamapediatrics.2018.0256.
22. Wellner B, Grand J, Canzone E, et al. Predicting unplanned transfers to the intensive care unit: a machine learning approach leveraging diverse clinical elements. JMIR Med Inf 2017;5:e45.
23. Zhai H, Brady P, Li Q, et al. Developing and evaluating a machine learning based algorithm to predict the need of pediatric intensive care unit transfer for newly hospitalized children. Resuscitation 2014;85:1065–71. https://doi.org/10.1016/j.resuscitation.2014.04.009.
24. Rubin J, Potes C, Xu-Wilson M, et al. An ensemble boosting model for predicting transfer to the pediatric intensive care unit. Int J Med Inf. 2018;112:15–20. https://doi.org/10.1016/j.ijmedinf.2018.01.001.
25. Gorham T, Rust S, Bambach S, et al. The deterioration risk index: predicting pediatric inpatient deterioration with machine learning and the electronic health record. Pediatrics 2020;146:229–30. https://doi.org/10.1542/peds.146.1_MeetingAbstract.229.
26. Lin EJD, Hefner JL, Zeng X, et al. A deep learning model for pediatric patient risk stratification. Am J Manag Care 2019;25:e310–5.
27. Seiger N, Maconochie I, Oostenbrink R, Moll HA. Validity of different pediatric early warning scores in the emergency department. Pediatrics 2013;132:e841–50. https://doi.org/10.1542/peds.2012-3594.
28. Brown SR, Martinez Garcia D, Agulnik A. Scoping Review of Pediatric Early Warning Systems (PEWS) in resource-limited and humanitarian settings. Front Pediatr 2019;6. Available from: https://www.frontiersin.org/article/10.3389/fped.2018.00410.
29. Gawronski O, Ferro F, Cecchetti C, et al. Adherence to the bedside paediatric early warning system (BedsidePEWS) in a pediatric tertiary care hospital. BMC Health Serv Res 2021;21:852. https://doi.org/10.1186/s12913-021-06809-2.
30. Kowalski RL, Lee L, Spaeder MC, Moorman JR, Keim-Malpass J. Accuracy and Monitoring of Pediatric Early Warning Score (PEWS) scores prior to emergent pediatric Intensive Care Unit (ICU) transfer: retrospective analysis. JMIR Pediatr Parent 2021;4. https://doi.org/10.2196/25991 e25991.
31. Poole S, Shah N. Addressing vital sign alarm fatigue using personalized alarm thresholds. Pac Symp Biocomput Pac Symp Biocomput. 2018;23:472–83.
32. Winter MC, Kubis S, Bonafide CP. Beyond reporting early warning score sensitivity: the temporal relationship and clinical relevance of “True Positive” alerts that precede critical deterioration. J Hosp Med 2019;14. https://doi.org/10.12788/jhm.3066.
33. Chapman SM, Wray J, Oulton K, Pagel C, Ray S, Peters MJ. ‘The Score Matters’: wide variations in predictive performance of 18 paediatric track and trigger systems. Arch Dis Child. 2017;102:487–95. https://doi.org/10.1136/archdischild-2016-311088.
34. Tucker KM, Brewer TL, Baker RB, Demeritt B, Vossmeyer MT. Prospective evaluation of a pediatric inpatient early warning scoring system. J Spec Pediatr Nurs 2009;14:79–85. https://doi.org/10.1111/j.1744-6155.2008.00178.x.