Pediatric Alert Score (EPA) performance in clinical deterioration

Desempenho do Escore Pediátrico de Alerta (EPA) de deterioração clínica

Desempeño del Sistema de Alerta Precoz Infantil (SAPI) de deterioración clínica

Thaiane de Lima Oliveira[1] https://orcid.org/0000-0001-9151-8283
Juliana Freitas Oliveira Miranda[1] https://orcid.org/0000-0001-7659-3103
Alan Peter Monaghan[2] https://orcid.org/0000-0001-7437-8768
Renata Costa Silva[3] https://orcid.org/0000-0003-4905-2672
Ana Kelly de Araujo Santana[3] https://orcid.org/0000-0001-5755-3376
Marina Vieira Silva[3] https://orcid.org/0000-0002-5848-0620
José de Bessa Junior[1] https://orcid.org/0000-0003-4833-4889
Ana Paloma Martins Rocha Ribeiro[1] https://orcid.org/0000-0002-4596-4300

How to cite:
Oliveira TL, Miranda JF, Monaghan AP, Silva RC, Santana AK, Silva MV, et al. Pediatric Alert Score (EPA) performance in clinical deterioration. Acta Paul Enferm. 2023;36:eAPE00872.

DOI
http://dx.doi.org/10.37689/acta-ape/2023AO008722

Keywords
Clinical deterioration; Early warning score; Child hospitalized; Pediatric nursing; Validation study

Descritores
Deterioração clínica; Escore de alerta precoce; Criança hospitalizada; Enfermagem pediátrica; Estudo de validação

Descriptores
Deterioro clínico; Puntuación de alerta temprana; Niño hospitalizado; Enfermería pediátrica; Estudio de validación

Submitted
May 16, 2022

Accepted
September 27, 2022

Corresponding author
Juliana Freitas Oliveira Miranda
E-mail: julidefreitas@hotmail.com/jofmiranda@uefs.br

Associate Editor (Peer review process):
Kelly Pereira Coca
(https://orcid.org/0000-0002-3604-852X)
Escola Paulista de Enfermagem, Universidade Federal de São Paulo, São Paulo, SP, Brazil

Abstract

Objective: To assess the Pediatric Alert Score (EPA) accuracy, usefulness, reproducibility and applicability in identifying clinical deterioration in hospitalized children and adolescents.

Methods: This is a prospective diagnostic test study, carried out between October/2018 and October/2019, to measure EPA diagnostic accuracy in a sample of 240 children, and its reproducibility and applicability in a sample of 60 children. Data were processed and analyzed on MedCalc and VassarStats.net.

Results: At cut-off point ≥ 3, the score had a sensitivity of 73.6%, specificity of 95.7%, positive predictive value of 83%, negative predictive value of 92.7, area under the ROC curve of 93.6%, estimated prevalence of 19.6%, positive probability ratio of 17.1, positive post-test probability of 77.8%, simple Kappa of 0.946.

Conclusion: The study provides evidence on EPA high accuracy, usefulness and reproducibility in identifying clinical deterioration in a Brazilian pediatric hospital setting, and considered the instrument applicable in the context of the research.

Resumo

Objetivo: Avaliar a acurácia, utilidade, reprodutibilidade e aplicabilidade do Escore Pediátrico de Alerta (EPA) na identificação da deterioração clínica em crianças e adolescentes hospitalizados.

Métodos: Estudo de teste diagnóstico, prospectivo, realizado entre outubro/2018 a outubro/2019, para medir a acurácia diagnóstica do EPA em uma amostra de 240 crianças, e sua reprodutibilidade e aplicabilidade em uma amostra de 60 crianças. Os dados foram processados e analisados no *MedCalc* e *VassarStats.net*.

Resultados: No ponto de corte ≥ 3, o escore apresentou sensibilidade de 73,6%, especificidade de 95,7%, valor preditivo positivo de 83%, valor preditivo negativo de 92,7, área sob a curva ROC de 93,6%, prevalência estimada pelo teste de 19,6%, razão de probabilidade positiva 17,1, probabilidade pós-teste positivo de 77,8%, *kappa simples* de 0,946.

Conclusão: O estudo fornece evidências sobre a elevada acurácia, utilidade e reprodutibilidade do EPA na identificação da deterioração clínica em um cenário hospitalar pediátrico brasileiro, e considerou o instrumento aplicável no contexto da pesquisa.

Resumen

Objetivo: Evaluar la precisión, utilidad, reproducibilidad y aplicabilidad del Sistema de Alerta Precoz Infantil (SAPI) en la identificación del deterioro clínico en niños y adolescentes hospitalizados.

[1]Universidade Estadual de Feira de Santana, Feira de Santana, BA, Brazil.
[2]University of Brighton, Brighton, United Kingdom.
[3]Hospital Estadual da Criança, Feira de Santana, BA, Brazil.
Conflicts of interest: nothing to declare.

Métodos: Estudio de prueba diagnóstica, prospectiva, realizada entre octubre de 2018 y octubre de 2019, para medir la precisión diagnóstica del SAPI en una muestra de 240 niños y su reproducibilidad y aplicabilidad en una muestra de 60 niños. Los datos fueron procesados y analizados en *MedCalc* y *VassarStats.net*.

Resultados: En el punto de corte ≥ 3, el puntaje presentó una sensibilidad del 73,6 %, especificidad del 95,7 %, valor predictivo positivo del 83 %, valor predictivo negativo de 92,7, área bajo la curva ROC del 93,6 %, prevalencia estimada por la prueba del 19,6 %, razón de probabilidad positiva 17,1, probabilidad posprueba positiva del 77,8 %, *kappa simple* de 0,946.

Conclusión: El estudio presenta evidencias sobre la elevada precisión, utilidad y reproducibilidad del SAPI en la identificación del deterioro clínico en un escenario hospitalario pediátrico brasileño, por lo que el instrumento se consideró aplicable en el contexto de la investigación.

Introduction

Early recognition of clinical worsening in hospitalized children is paramount for proper and timely management of deterioration.[1] In this perspective, the Pediatric Early Warning Score (PEWS) are instruments developed to systematize the clinical assessment and assist the health team in the early identification of deteriorating children, supporting decision-making and the implementation of care aimed at reducing the probability of unfavorable outcomes.[2]

PEWS are tools of simple application, based on clinical parameters and vital signs that are easy to measure, without the aid of complex and expensive equipment, and are currently considered afferent branches of rapid response systems to pediatric clinical deterioration.[3,4]

The first PEWS, published in 2005, was the Brighton Paediatric Early Warning Score (BPEWS). It is a scoring system composed of three assessment components: neurological, cardiovascular and respiratory, in addition to the use of nebulization and the presence of persistent vomiting in the postoperative period. The score can vary between 0 and 13 points, and from 3 points, the higher the score, the greater the risk of deterioration. [1,5,6]

In Brazil, BPEWS went through the process of translation, adaptation and validity for the Brazilian context (BPEWS-Br), and presented encouraging results, becoming the first pediatric early warning score validated in a Brazilian setting with the aim of assisting in the early detection of clinical worsening in children in the hospital setting.[1]

After BPEWS-Br validity, its authors, based on an extensive discussion, raised the need for adjustments to the instrument, in order to include new indicators for assessment of pediatric clinical deterioration described in the literature and better serve the study context. Although the BPEWS-Br performed satisfactorily, it did not contain indicators considered to be important predictors of clinical worsening in children, such as certain criteria that can help in screening patients with suspected sepsis. The purpose was to have a valid tool that was more adequate to the reality of the service, the team, the patients and the context of Brazilian public health.[1]

A new instrument, called the Pediatric Alert Score (EPA), was developed from the BPEWS-Br. Initially, it was validated from the standpoint of content, presenting a General Content Validity Index of 0.97, considered easy to use, with good structure and presentation, in addition to contemplating relevant indicators of clinical relevance and applicable to the study site.[7] However, its performance in recognizing clinical deterioration needs to be measured in order to produce evidence to strengthen and sustain its use in the national scenario.

It is estimated that, since the creation and validity of the first PEWS, more than 30 scores are in use, each with different parameters assessed, formats, as well as degree of validity.[8] Evidence on the validity and clinical utility of pediatric alert systems is still limited, given the large number of scores in use. Therefore, research to obtain robust, valid and clinically significant results to assess PEWS in different scenarios is necessary. Understanding and defining the properties that lead to high performance can support a more evidence-based approach to the development of future PEWS.[9] Given the variability of these instruments, the decision to adopt a PEWS needs to be accompanied by evidence on its validity in care contexts.

This study aimed to assess EPA accuracy, usefulness, reproducibility and applicability in identi-

fying clinical deterioration in hospitalized children and adolescents.

Methods

This is a quantitative, prospective, diagnostic test study to verify EPA performance in identifying clinical deterioration in children and adolescents in a hospital context. The study was conducted based on the Standards for Reporting of Diagnostic Accuracy Studies (STARD) recommendations.[10]

The research site was a large maternal-infant hospital, a reference for the care of children and adolescents under 16 years of age, with 240 beds, located in the municipality of Feira de Santana, the second largest city in the state of Bahia, Brazil.

The study's reference population consists of children and adolescents aged 1 month to 15 years, admitted to the hospital's clinical-surgical and emergency units. The following were excluded from the sample: newborns, patients aged ≥ 16 years who remained under follow-up at the hospital due to the diagnosis; patients with medical discharge prescribed in medical records; patients with a diagnosis of heart disease described in the medical records; undergoing cancer treatment; in isolation and using invasive mechanical ventilation. It was decided to exclude newborns, patients with heart disease and cancer because there are already validated proposals in the literature for scores for this population.[11-13] Patients with cancer were also excluded because they were undergoing chemotherapy and with more restricted handling due to low immunity. Patients with indication for isolation were excluded due to the risk of cross-infection during data collection.

A sample of 240 patients was calculated to verify the EPA diagnostic accuracy and usefulness. The sample calculation was made using the formula: $N=1.96^2[0.17(1-0.17)/(0.05^2)$, adding another 10% to the value, considering the losses. The proportion of clinical deterioration adopted for the sample calculation was 17%, found in a previous study.[1] The Confidence Interval (CI) spectrum was 0.10. The CI semi-amplitude as an acceptable error was 0.05. The value of Z was 1.96.[14] To verify the EPA reproducibility among observers and applicability, 60 children and adolescents were assessed by convenience sampling.

Three measurement and data collection instruments were used: the reference standard or gold standard, the EPA and the instrument for collecting sociodemographic and clinical data for sample characterization. The reference standard for determining the deterioration of children and adolescents participating in the study was the set of clinical criteria that make up the Primary Clinical Assessment of the Critically Ill Child recommended by the American Heart Association (AHA) and the American Academy of Pediatrics.[15]

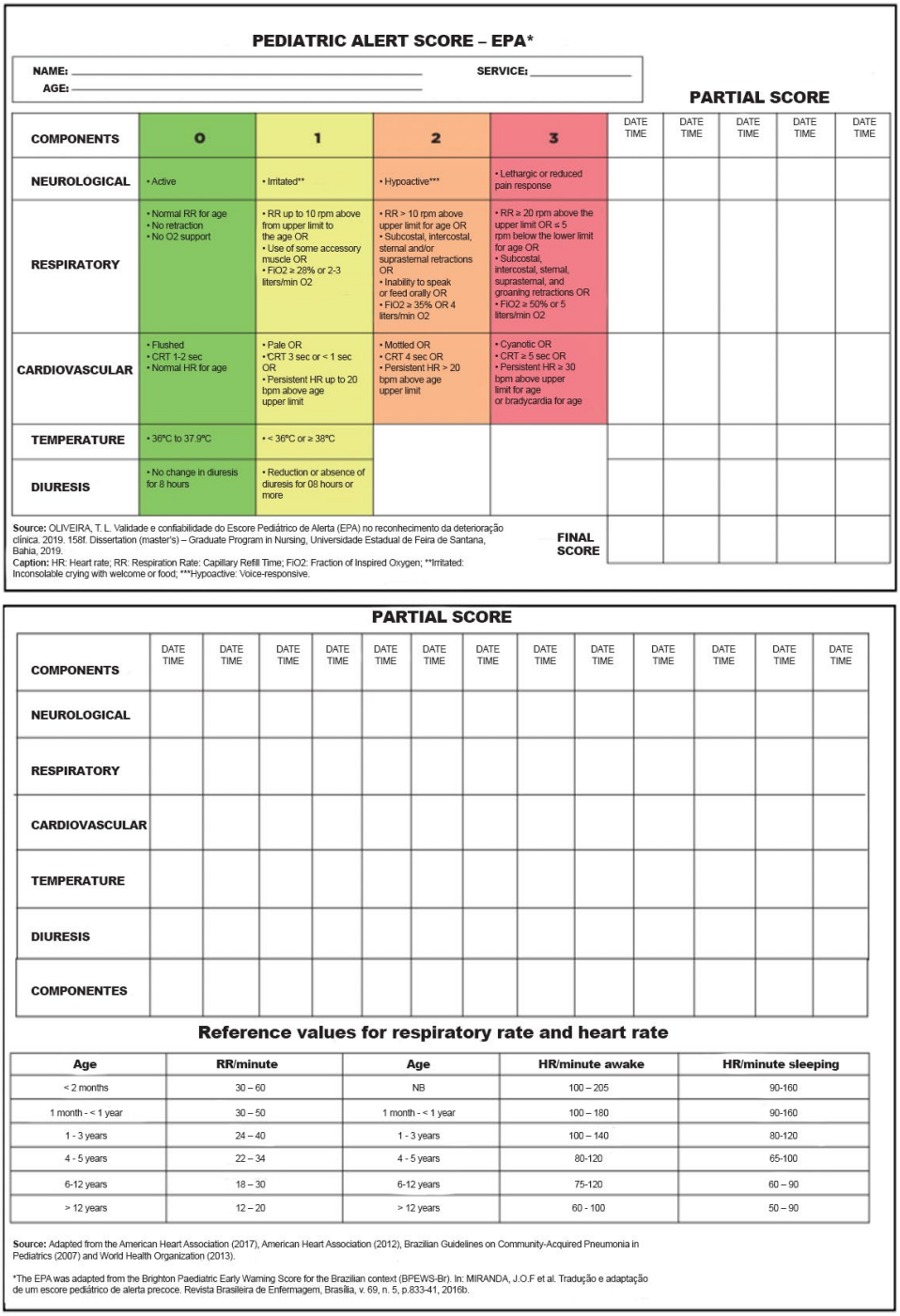
Among the AHA criteria, for neurological assessment, the Pediatric AVPU Response Scale (Alert, Voice, Pain, Unresponsive) was used in order to optimize the assessment time. Blood pressure exclusion was chosen, since hypotension in children is characterized as a late sign of cardiovascular decompensation.[5] Furthermore, the lack of a consistent routine for blood pressure measurement in children in the wards could be a complicating factor for data collection.

From a broad discussion among researchers about the adopted reference standard,[15] iit was defined that the patient with 3 or more altered indicators in the reference standard would be classified as "with signs of deterioration".

The EPA (Figure 1) consists of clinical criteria for neurological, respiratory (respiratory rate, breathing pattern, oxygen support) and cardiovascular assessment (skin color, capillary refill time, heart rate, temperature and urine output), whose score can range from 0 to 11 points.[7]

Assessments of clinical criteria for applying the reference standard and EPA were standardized from an operational manual, in order to calibrate observers and minimize measurement bias. Additionally, a pilot test was applied to 20 patients to resolve possible doubts at the time of application, as well as to adapt data collection to the inpatient dynamics unit, to the multidisciplinary team's work and, mainly, to children and their companions.

Data collection occurred from October/2018 to October/2019. The 240 children and adolescents



PEDIATRIC ALERT SCORE – EPA*

NAME: _____ SERVICE: _____
AGE: _____

PARTIAL SCORE

COMPONENTS	0	1	2	3	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME
NEUROLOGICAL	• Active	• Irritated**	• Hypoactive***	• Lethargic or reduced pain response					
RESPIRATORY	• Normal RR for age • No retraction • No O2 support	• RR up to 10 rpm above from upper limit to the age OR • Use of some accessory muscle OR • FiO2 ≥ 28% or 2-3 liters/min O2	• RR > 10 rpm above upper limit for age OR • Subcostal, intercostal, sternal and/or suprasternal retractions OR • Inability to speak or feed orally OR • FiO2 ≥ 35% OR 4 liters/min O2	• RR ≥ 20 rpm above the upper limit OR ≤ 5 rpm below the lower limit for age OR • Subcostal, intercostal, sternal, suprasternal, and groaning retractions OR • FiO2 ≥ 50% or 5 liters/min O2					
CARDIOVASCULAR	• Flushed • CRT 1-2 sec • Normal HR for age	• Pale OR • CRT 3 sec or < 1 sec OR • Persistent HR up to 20 bpm above age upper limit	• Mottled OR • CRT 4 sec OR • Persistent HR > 20 bpm above age upper limit	• Cyanotic OR • CRT ≥ 5 sec OR • Persistent HR ≥ 30 bpm above upper limit for age or bradycardia for age					
TEMPERATURE	• 36ºC to 37.9ºC	• < 36ºC or ≥ 38ºC							
DIURESIS	• No change in diuresis for 8 hours	• Reduction or absence of diuresis for 08 hours or more							
				FINAL SCORE					

Source: OLIVEIRA, T. L. Validade e confiabilidade do Escore Pediátrico de Alerta (EPA) no reconhecimento da deterioração clínica. 2019. 158f. Dissertation (master's) – Graduate Program in Nursing, Universidade Estadual de Feira de Santana, Bahia, 2019.
Caption: HR: Heart rate; RR: Respiration Rate: Capillary Refill Time; FiO2: Fraction of Inspired Oxygen; **Irritated: Inconsolable crying with welcome or food; ***Hypoactive: Voice-responsive.

PARTIAL SCORE

COMPONENTS	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME	DATE TIME
NEUROLOGICAL													
RESPIRATORY													
CARDIOVASCULAR													
TEMPERATURE													
DIURESIS													
COMPONENTES													

Reference values for respiratory rate and heart rate

Age	RR/minute	Age	HR/minute awake	HR/minute sleeping
< 2 months	30 – 60	NB	100 – 205	90-160
1 month - < 1 year	30 – 50	1 month - < 1 year	100 – 180	90-160
1 - 3 years	24 – 40	1 - 3 years	100 – 140	80-120
4 - 5 years	22 – 34	4 - 5 years	80-120	65-100
6-12 years	18 – 30	6-12 years	75-120	60 – 90
> 12 years	12 – 20	> 12 years	60 - 100	50 – 90

Source: Adapted from the American Heart Association (2017), American Heart Association (2012), Brazilian Guidelines on Community-Acquired Pneumonia in Pediatrics (2007) and World Health Organization (2013).

*The EPA was adapted from the Brighton Paediatric Early Warning Score for the Brazilian context (BPEWS-Br). In: MIRANDA, J.O.F et al. Tradução e adaptação de um escore pediátrico de alerta precoce. Revista Brasileira de Enfermagem, Brasília, v. 69, n. 5, p.833-41, 2016b.

Figure 1. Pediatric Alert Score – EPA (front and back)

who participated in verification of EPA diagnostic accuracy and usefulness were selected by drawing lots from the list of patients admitted to the units. The EPA and reference standard application was blindly performed by a nurse with experience in pediatrics and by a pediatrician, respectively, respecting an interval of 5 to 10 minutes between measurements. The signs of clinical deterioration identified in patients were formally communicated to the on-call team and the relevant conducts were performed.

Data collection to verify reproducibility was performed in 60 children and adolescents by two nurses with experience in pediatrics who applied

the EPA with an interval of 5 minutes between measurements, independently and blindly. Also, nurses timed the time of patient assessment and application of the score to assess its applicability in the study context.

The data obtained were entered into two databases built in Excel 16.27 to compare the information and identify possible typing errors. Once this was done, they were processed electronically in MedCalc® Statistical Software version 20.007 in order to estimate the EPA accuracy indicators.

Absolute and relative frequencies were calculated for qualitative variables. For continuous quantitative variables, means, medians and measures of dispersion (standard deviation and interquartile range) were calculated.

The indicators adopted to measure the EPA accuracy and usefulness were: sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), Receiver Operating Characteristic Curve (ROC curve), positive likelihood ratio (PLR), negative likelihood ratio (NLR), positive post-test probability (+PTP) and negative post-test probability (-PTP). In addition to these, the prevalence determined by the reference standard and the estimated prevalence were calculated.

To measure reproducibility, the Simple Kappa and Weighted Kappa coefficients of agreement were calculated using the VassarStats.net program. For the calculation of Simple Kappa, child classification was dichotomized into "with signs of deterioration" and "without signs of deterioration", considering the cut-off point ≥ 3 for the presence of clinical deterioration. As for the Weighted Kappa, the natural hierarchy (relative to the severity of the event) between the score categories was maintained, which ranged from 0 to 8 points in the studied sample.

To assess the Kappa index results, the following reference criteria were adopted to interpret the degree of agreement: < 0.00 (poor), 0.00–0.20 (mild), 0.21–0.40 (fair), 0.41–0.60 (moderate), 0.61–0.80 (substantial), 0.81–1.00 (almost perfect).[16]

The EPA applicability was measured by the average time, in minutes, spent by the nurses to apply the score.

The study was approved by the Research Ethics Committee, CAAE (*Certificado de Apresentação para Apreciação Ética* - Certificate of Presentation for Ethical Consideration) 79484117.2.0000.0053, and complied with all ethical precepts for research involving human beings.

Results

The sample of 240 children and adolescents ranged in age between 1 month and 15 years, with a median of 4 years (interquartile range: 7 years). Most children had a clinical diagnosis as the cause of hospitalization (93.0%), with a prevalence of infections and problems related to the respiratory system. More than half of the sample had no comorbidities (68.0%) and a history of previous hospitalization (59.0%).

Table 1 describes the EPA accuracy and usefulness indicators for the cut-off points found. Scores ≥ 2 and ≥ 3 were considered the best cut-off points. EPA ≥ 2 had a sensitivity of 86.8%, specificity of 87.7% and PPV of 66.7%. For EPA ≥ 3, sensitivity and specificity were 73.6% and 95.7%, respectively, and PPV was 83%.

The estimated prevalence (EP) of patients with EPA-triggered clinical deterioration ≥ 3 was 19.6%, closer to the prevalence determined by the reference standard (22.1%) when compared to EPA EP ≥ 2 (28.7%).

Regarding EPA usefulness, Odds Ratios were used to calculate the +PTP and -PTP. +PTP of clinical deterioration increased the pre-test probability (prevalence estimated by the reference standard) from 22.1% to 59% in the presence of a score ≥ 2 and to 82.7% in the presence of a score ≥ 3. PTP reduced the pretest probability from 22.1% to 3% if EPA < 2, and to 6% if EPA < 3.

Thus, in the sample assessed by the study, a score ≥ 3 can be adopted as the best cut-off point, understanding that this value added better accuracy and utility to the test.

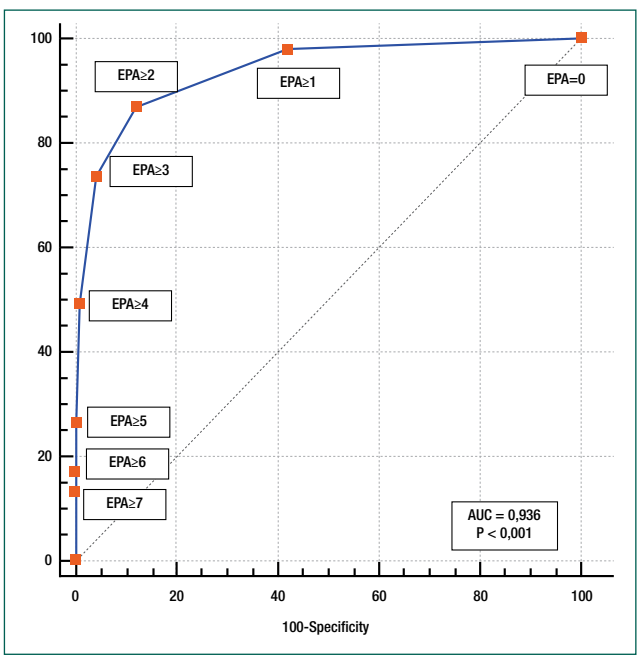
The instrument's overall performance in discriminating the presence and absence of clinical deterioration was assessed from the area under the

Table 1. EPA accuracy and usefulness indicators

EPA	Sp	95% CI	E	95% CI	PPV	95% CI	NPV	95% CI	PLR	95% CI	NLR	95% CI	+PTP	-PTP	EP
0	100	93.3 - 100	0.0	0.0 - 2.0	22.1	22.1 - 22.1	-	-	1	1.0 - 1.0	-	-	21.8	-	100
≥ 1	98.1	89.9 - 100	57.7	50.3 - 64.9	39.7	35.7 - 43.9	99.1	93.9 - 99.9	2.3	2.0 - 2.8	0.03	0.0 - 0.2	39.1	0.8	54.6
≥ 2	86.8	74.7 - 94.5	87.7	82.1 - 92.0	66.7	57.4 - 74.8	95.9	92.1 - 97.9	7.06	4.7 - 10.5	0.15	0.0 - 0.3	66.3	4	28.7
≥ 3	73.6	59.7 - 84.7	95.7	91.7 - 98.1	83.0	70.8 - 90.7	92.7	89.1 - 95.3	17.2	8.6 - 34.5	0.28	0.2 - 0.4	82.7	6	19.6
≥ 4	49.1	35.1 - 63.2	98.9	96.2 - 99.9	92.9	76.1 - 98.1	87.3	84.0 - 89.9	45.9	11.2 - 187	0.51	0.4 - 0.7	92.7	12.4	11.7
≥ 5	26.4	15.3 - 40.3	99.5	97.1 – 100	93.3	65.3 - 99.0	82.7	80.2 - 84.9	49.4	6.6 -367.1	0.74	0.6 - 0.9	93.2	17.1	6.2
≥ 6	17.0	8.1 - 29.8	100	98.0 – 100	100	98.0 - 100	81	79.0 - 82.8	-	-	0.83	0.7 - 0.9	-	18.8	3.7
≥ 7	13.2	6.54 - 24.8	100	98.0 – 100	100	98.0 - 100	-	-	-	-	0.87	0.8 - 0.9	-	19.5	2.9

S: Sensitivity; Sp: Specificity; CI: Confidence Interval; PPV: positive predictive value; PPV: negative predictive value; PLR: positive probability likelihood; NLR: negative probability likelihood; EP: estimated prevalence; +PTP: positive post-test probability; -PTP: negative post-test probability.

ROC curve between the EPA and the reference standard (Figure 2). The value found was 0.936 (95% CI: 0.89 - 0.96), indicating that, in 93.6% of cases, EPA will be able to correctly discriminate children in clinical deterioration, and in only 6.4% the instrument will give false results. From the ROC curve, it is also possible to state that the best cut-off point of EPA was 3, since it is closer to the upper left corner of the graph. Point 3 provided 4.3% of false positive results, while point 2 provided 12.3%.



**Receiver Operating Characteristic Curve*

Figure 2. ROC curve* between EPA and reference standard

To assess EPA application reproducibility among nurses, the simple Kappa coefficient was calculated, from the dichotomization of the score at point 3 [0.946 (CI: 0.842 - 1.00)], and the weighted Kappa coefficient [0.834 (CI: 1.00)] 0.731 - 0.9162)], maintaining the EPA cut-off points measured in the sample. To verify EPA applicability in the study scenario, the average time spent for patient assessment and application of the instrument by the two nurses was measured, being 4.2 minutes (SD: ±1.0) and 4.7 minutes (SD: 1.0) ±1.8), respectively, with a minimum time of two and a maximum of eleven minutes. Children with EPA ≥ 3 required a longer average time for assessment by one of the nurses (5.0 minutes) when compared to children with EPA ≤ 2 (4.6 minutes), however, for the second nurse there was no difference in this regard (4.2 minutes). Thus, it may be that the time for application of EPA is influenced by the degree of patients' clinical deterioration. Still on EPA applicability, nurses participating in the study indicated that the use of practical clinical criteria and that they do not require more sophisticated equipment are positive characteristics for the application of the score in the context. However, they brought as a limitation the need for support from the medical team for their effective implementation.

Discussion

A basic principle for adopting an alert score is its ability to accurately and reliably identify signs of clinical worsening in a patient. Studies to verify the performance of PEWS, generally used as indicators of validity the measures of sensitivity, specificity, predictive values and/or area under the ROC curve.[1,6,17–24] In this study, for EPA validity, in addition to these indicators, Odds Ratios and positive and negative post-test probabilities were used.

The best EPA cut-off point found in this study, considered capable of alerting possible clinical deterioration, was score 3, weighing the balance between sensitivity, specificity and an excellent area under the ROC curve (AUC).[25-27] The AUC found in the EPA (93.6%) was similar to that of the BPEWS-Br (91.9%) and higher than that found in other studies, whose values ranged from 69% to 89%.[1,6,17-19,21-24]

There is no standard of acceptability between sensitivity and specificity, but it is suggested that both have a balance between the highest possible values.[25-27] The sensitivity and specificity of EPA ≥ 3 were 73.6% and 95.7%. Studies that validated BPEWS in other settings, adopting different cut-off points ranging from 1 to 4, found similar results for sensitivity and specificity, respectively: 79.8% and 65.2% for a cut-off point ≥ 1; 58% and 75%, 72% and 88%, 68.4% and 81.6% for a cut-off point ≥ 2; 62% and 89% for a cut-off point ≥ 2.5; 90.2% and 74.4%, 32% and 93% and 73.9% and 95.5% for a cut-off point ≥ 3; and 14.8% and 97.7% for a cut-off point ≥ 4.[1,6,17-19,21-24]

To estimate pediatric clinical deterioration prediction, positive (PPV) and negative predictive values (NPV) were calculated in order to measure the probability of patients worsening or not in the face of positivity or negativity of the score.[26] Considering the cut-off point ≥ 3, EPA presented a PPV of 83%, higher than that found in some studies that validated adapted versions of BPEWS: 36% and 15% at cut-off point ≥ 2, 77.3% and 5.8% at cut-off point ≥ 3 and 58% at cut-off point ≥ 4.[1,6,17,23,24] The EPA NPV ≤ 3 was 92.7% and four studies found values close to or higher for the assessed PEWS: 94.7% and 99.8% at cut-off point ≥ 3; 96% and 98% at cut-off point ≥ 2.[1,6,17,23]

In order to assess EPA usefulness, the probability ratio (PR) was calculated, whose values at the cut-off point ≥ 3 were 17.2 (PLR) and 0.28 (NLR). Good diagnostic tests must present a PLR > 10, indicating that its positive result contributes significantly to the diagnosis in question, proving to be a useful instrument. For NLR values, the closer to zero, the lower the probability of disease in the face of a negative test.[27]

Among the studies that validated the BPEWS, four were identified, which worked with RP. For PLR, only the BPEWS-Br presented a value similar to EPA, of 16.6, and the other studies presented values < 10, with a variation of cut-off points between ≥1 and ≥ 3: 6.2, 2, 3 and 4.72. For NLR, most studies found values close to those found in EPA: 0.27, 0.32, 0.3 and one study found a higher value of 0.73.[1,17,21,22]

From the LR, one can find the post-test probability (PTP). This indicator also endorses the usefulness of a diagnostic test, as it informs how much the test result will increase (+PTP) or decrease (-PTP) the prevalence of the disease (pre-test probability) in that scenario.[27] Considering the pre-test probability of clinical deterioration of 22.1% found by the reference standard, EPA +PTP was 82.7% and -PTP was 6% at cut-off point ≥ 3. These values are similar to those found in the only study that assessed the BPEWS usefulness in Brazil, finding a +PTP of 80% and a -PTP of 6% for the same cut-off point.[1]

The EP of clinical deterioration by EPA was also calculated in this study, reflecting the ratio between the number of people with clinical deterioration according to the EPA and the population at risk. The prevalence of deterioration determined by the reference standard was 22.1%, whereas the EP by EPA ≥ 3 was 19.6%. In the BPEWS-Br validity study, considering a score ≥ 3, the EP by EPA was 16.2% and that determined by the reference standard was 17%.[1] The study that validated the first modified version of the BPEWS adopted transfer to the ICU as a reference standard and found an EP of 24.2% for a score ≥ 3.[6]

EPA reproducibility was measured by the Kappa coefficient. The EPA showed a simple Kappa of 0.946 and a weighted Kappa of 0.824. Therefore, the instrument showed an almost perfect degree of agreement, both for simple Kappa and for weighted Kappa.[16]

Some studies have verified the BPEWS reproducibility in the validity process. Two of them found simple Kappa values considered substantial: 0.75 and 0.74; and two others found almost perfect values: 0.85 and 0.91.[17,22,28,29] Only one

study that validated the BPEWS also calculated the weighted Kappa, finding a value of 0.80.[29] Another study that sought to assess inter-examiner reliability among nurses using PEWS systems in central Denmark found that nurses assigned the same aggregate score to two PEWS models in 76% of cases. The intraclass correlation coefficients for the aggregate score in the two PEWS models were 0.98 and 0.95.[30]

A PEWS must be objective, easy and quick to apply so as not to generate work overload or extra work for the team.[5,29] In this study, the EPA applicability was related to the time taken for its application, and the average total time taken for application (4.2 and 4.8 minutes) was slightly higher than that found for the BPEWS-Br (4.1 and 3.5 minutes).[29] This can probably be associated with the fact that the EPA aggregates more clinical indicators than the BPEWS.

The time taken to apply the original BPEWS was described as 30 seconds. This time was reduced as the evaluator became familiar with the instrument.[5] To apply the EPA in the study, even if patients were monitored, heart and respiratory rates were counted, in addition to measuring the axillary temperature for one minute, according to the equipment's instruction manual. Therefore, it is believed that, when incorporated into nurses' assessment routine and there is familiarity with the score, the time spent for its application can be gradually reduced.

A characteristic of EPA, as well as BPEWS, which can facilitate and speed up its application is the non-inclusion of blood pressure (BP) as a clinical criterion, since its measurement in children requires time, appropriate equipment, standardization of size of cuffs and a consistent routine for gauging. These factors could hamper the EPA implementation in many Brazilian realities.

Regarding the EPA limitations, in this research, the score was not validated for use in newborns, children with heart disease and oncological diseases, and its application for these populations is not recommended, which creates a gap for the possibility of new studies in this perspective. For these populations, there are scores validated in international settings that need to be tested in Brazilian hospitals, such as The Neonatal Trigger Score for neonates, the Cardiac Children's Hospital Early Warning Score for patients with heart disease and The Children's Hospital Early Warning Score for patients with cancer.[11-13]

Medical team support was pointed out by the nurses as a limitation for the effective EPA implementation. The usefulness of PEWS is reported in terms of identifying sick children in clinical deterioration, timely intervention by the multidisciplinary team, effective communication and trust in child care. However, its use is still limited due to the variation of scores and configurations.[31] This scenario raises the need for standardization according to each context.

Authors suggest as research priorities on the PEWS studies that seek to: determine its predictive characteristics in different clinical settings, populations and organizational structures; assess the effect on mortality and morbidity; investigate barriers and facilitating factors for its implementation; consider developing a national PEWS to standardize in practice; and understand the role of human factors in its implementation and explore the role of technology in identifying and scaling children at risk of deterioration.[10] Thus, there are still many gaps to be filled in terms of scientific knowledge and production of evidence on PEWS, as well as challenges to be overcome, especially in the national context.

Regarding the use of these scores by Brazilian services, a cross-sectional, retrospective study assessed nurses' adherence to filling out the PEWS in an Emergency Service based on an analysis of 1,219 medical records and found 75.2% of compliance, which points to the need for actions for professional education, with the purpose of improving indicators related to the application of PEWS and guaranteeing safe care.[32] Authors claim that it is possible to implement a PEWS in environments with limited resources. However, this is a time and energy consuming process, and requires an active and involved team for successful implementation.[33]

Pediatric alert scores have been highlighted for their performance in the early recognition of clinical worsening, as they provide quantifiable evidence

of deterioration. However, it is important to note that the PEWS will not identify all children at risk of clinical worsening, either because of the speed or the mechanism involved in the deterioration, staff training is essential to recognize common patterns of deterioration and not rely on PEWS as the only screening mechanism. However, they can be a valuable adjunct to clinical decision making, particularly for less experienced staff, as they are often accompanied by an escalation algorithm that indicates the action to be taken on each score.[10]

Given the above, PEWS can not only improve the power of early recognition of clinical instability, but also increase the team's situational awareness and standardize communication, so that all teams, with different levels of training, understand the severity of the situation. When properly validated, and combined with care algorithms, they can trigger timely actions and help manage clinical deterioration.

Conclusion

The present study provides evidence on EPA high accuracy, usefulness and reproducibility in identifying clinical deterioration in a Brazilian pediatric hospital setting, and considered the instrument applicable in the context of the research. It sought to follow all the methodological rigor of a diagnostic test study, but it is limited because it is single-center research and has not been validated, so far, for use in newborns, children with heart disease and with oncological diseases. Based on the evidence presented, the EPA was considered a valid, useful, reliable and applicable instrument for the early recognition of clinical deterioration in children and adolescents, being able to give the nurse greater autonomy, and, if linked to a care algorithm, guide support actions and improve decision-making at critical moments with the multidisciplinary team. It is noteworthy that the improvement of care for patients in clinical deterioration transcends the application of an alert score. It is necessary to think of a multifaceted system, triggered by early recognition and that includes the implementation of appropriate care, trained personnel, adequate resources and audits to improve treatment and increase patient safety. The next challenge is the standardization of a single EPA on the national scene, establishing a common language among professionals to recognize and promptly respond to deterioration in children and adolescents. Moreover, studies are needed on the impact of implementing these instruments in care settings.

Acknowledgments

To the Brazilian National Council for Scientific and Technological Development (CNPq - *Conselho Nacional de Desenvolvimento Científico e Tecnológico*), Ministry of Science, Technology, Innovations and Communications (MCTIC) and the Ministry of Health (MoH), called MCTIC/CNPq 28/2018 – Project Financing "*Reconhecimento da deterioração clínica pediátrica no contexto hospitalar da saúde da criança no município de Feira de Santana – Bahia - Brasil*" from the *Universidade Estadual de Feira de Santana* (UEFS), Process 405101/2018-0. To the UEFS. To the *Hospital Estadual da Criança* and all its staff.

Collaborations

Oliveira TL, Miranda JFO, Monaghan AP, Silva RC, Santana AKA, Silva MV, Bessa Junior J and Ribeiro APMR declare that they contributed to study design, data analysis and interpretation, article writing, review, relevant interpretation of the intellectual content and approval of the final version to be published.

References

1. Miranda JO, Camargo CL, Nascimento CL, Portela DS, Monaghan A. Accuracy of a pediatric early warning score in the recognition of clinical deterioration. Rev Lat Am Enfermagem. 2017;25:e2912.
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(3):e014497. Review.

3. Jensen CS, Aagaard H, Olesen HV, Kirkegaard H. A multicentre, randomised intervention study of the Paediatric Early Warning Score: study protocol for a randomised controlled trial. Trials. 2017;18(1):267.
4. Gawronski O, Parshuram C, Cecchetti C, Tiozzo E, Degli Atti ML, Dall'Oglio I, et al. Qualitative study exploring factors influencing escalation of care of deteriorating children in a children's hospital. BMJ Paediatr Open. 2018;2(1):1–8.
5. Monaghan A. Detecting and managing deterioration in children. Paediatr Nurs. 2005;17(1):32-5. Review.
6. 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(2):79-85.
7. Oliveira TL, Miranda JO, Sobrinho CL, Mendoza RF, Monaghan AP, Alencar AA. Desenvolvimento e validação de conteúdo do escore pediátrico de alerta. Rev Soc Bras Enferm Ped. 2021;21(2):91-101.
8. Roland D, Oliver A, Edwards ED, Mason BW, Powell CV. 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(1):26-9.
9. Chapman SM, Maconochie IK. Early warning scores in paediatrics: an overview. Arch Dis Child. 2019;104(4):395-9. Review.
10. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799.
11. Simeone S, Platone N, Serra N, Assanta N, Guillari A, Rea T, et al. Cardiac Children's Hospital Early Warning Score: Italian Validation. J Pediatr Nurs. 2020;51:e21-6.
12. Mortensen N, Augustsson JH, Ulriksen J, Hinna UT, Schmölzer GM, Solevåg AL. Early warning- and track and trigger systems for newborn infants: a review. J Child Health Care. 2017;21(1):112-20. Review.
13. Agulnik A, Forbes PW, Stenquist N, Rodriguez-Galindo C, Kleinman M. Validation of a Pediatric Early Warning Score in Hospitalized Pediatric Oncology and Hematopoietic Stem Cell Transplant Patients. Pediatr Crit Care Med. 2016;17(4):e146-53.
14. Andrade AL, Zicker F. Métodos de investigação epidemiológica em doenças transmissíveis. Brasília (DF): Organização Pan-Americana da Saúde;Fundação Nacional da Saúde; 1997. 182 p.
15. American Heart Association. Suporte avançado de vida em pediatria manual do profissional. Suporte avançado de vida em pediatria manual do profissional. EUA: Ourora; 2017.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
17. 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(11):1249–56.
18. Zhai H, Brady P, Li Q, Lingren T, Ni Y, Wheeler DS, 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(8):1065–71.
19. Skaletzky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. Clin Pediatr (Phila). 2012;51(5):431–5.
20. Akre M, Finkelstein M, Erickson M, Liu M, Vanderbilt L, Billman G. Sensitivity of the pediatric early warning score to identify patient deterioration. Pediatrics. 2010;125(4):e763-9.
21. Seiger N, Maconochie I, Oostenbrink R, Moll HA. Validity of different pediatric early warning scores in the emergency department. Pediatrics. 2013;132(4):841–50.
22. Lillitos PJ, Hadley G, Maconochie I. Can paediatric early warning scores (PEWS) be used to guide the need for hospital admission and predict significant illness in children presenting to the emergency department? An assessment of PEWS diagnostic accuracy using sensitivity and specificity. Emerg Med J. 2016;33(5):329-37,
23. Lampin ME, Duhamel A, Behal H, Recher M, Leclerc F, Leteurtre S. Use of paediatric early warning scores in intermediate care units. Arch Dis Child. 2020;105(2):173-9.
24. Bradman K, Borland M, Pascoe E. Predicting patient disposition in a paediatric emergency department. J Paediatr Child Health. 2014;50(10):E39–44.
25. Ferreira JC, Patino CM. Understanding diagnostic tests. Part 1. J Bras Pneumol 2017;43(5):330.
26. Patino CM, Ferreira JC. Understanding diagnostic tests. Part 2. J Bras Pneumol. 2017;43(6):408.
27. Ferreira JC, Patino CM. Understanding diagnostic tests. Part 3. J Bras Pneumol. 2018;44(1):4.
28. Bell D, Mac A, Ochoa Y, Gordon M, Gregurich MA, Taylor T. 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(6):e2-9.
29. Miranda JO, Camargo CL, Nascimento Sobrinho CL, Oliveira TL, Matos AC, Portela DS. Reproducibility and Applicability of a Pediatric Score of Clinical Deterioration Warning. Rev Min Enferm. 2019;23:1-6.
30. Jensen CS, Aagaard H, Olesen HV, Kirkegaard H. Inter-rater reliability of two paediatric early warning score tools. Eur J Emerg Med. 2019;26(1):34-40.
31. Duraisamy R, Balasubramanian B, Palanisamy S. Utility of pediatric early warning scoring system in predicting clinical deterioration in children: a review. J Basic Clin Applied Health Science. 2019;2(3):115–8.
32. Ribeiro SP, Vieira LB, Strada JK, Freitas CA, Almeida VM, Wegner W. Aplication of transference note and paediatric early warning score in the Pediatric Emergency Department of Hospital de Clínicas de Porto Alegre. Sci Med. 2021;31(1):e39407.
33. van der Fluit KS, Boom MC, Brandão MB, Lopes GD, Barreto PG, Leite DC, et al. How to implement a PEWS in a resource-limited setting: a quantitative analysis of the bedside-PEWS implementation in a hospital in northeast Brazil. Trop Med Int Health. 2021;26(10):1240-7.