REPRODUCIBILITY AND APPLICABILITY OF A PEDIATRIC SCORE OF CLINICAL DETERIORATION WARNING

ABSTRACT

Object: to measure the reproducibility and applicability of the Brighton Pediatric Early Warning Score to the Brazilian context (BPEWS-Br) in order to detect clinical deterioration. Method: a study to test the performance of BPEWS-Br regarding its reproducibility and applicability. Two trained nurses randomly assigned a score to 50 children from zero to 10 years old with a three to five minute interval between evaluations. To verify the applicability, nurses timed the score assignment. Data were processed in SPSS and VassarStats.net. Reproducibility was measured by simple Kappa and weighted Kappa scores. The mean was calculated regarding the time of the score. Results: simple Kappa was 0.85 and weighted Kappa was 0.80. The average time required for the nurses to evaluate and use BPEWS-Br was 4.14 and 3.48 minutes. Conclusion: BPEWS-Br proved to be reliable and feasible to recognize warning signs of clinical deterioration in the children studied.

Keywords: Reproducibility of Results; Feasibility Studies; Child, Hospitalized; Pediatric Nursing.

RESUMO

Objetivo: medir a reprodutibilidade e aplicabilidade do Brighton Paediatric Early Warning Score para o contexto brasileiro (BPEWS-Br) no reconhecimento da deterioração clínica. Método: estudo para testar o desempenho do BPEWS-Br quanto à sua reprodutibilidade e aplicabilidade. Duas enfermeiras treinadas aplicaram o escore em 50 crianças de zero a 10 anos de idade e intervalo de três a cinco minutos entre as avaliações. Para verificar a aplicabilidade, as enfermeiras registraram o tempo de aplicação do escore. Os dados foram processados no SPSS e VassarStats.net. A reprodutibilidade foi medida pelos índices Kappa simples e ponderado. O tempo médio para avaliação e aplicação do BPEWS-Br pelas enfermeiras foi de 4.14 e 3.48 minutos. Conclusão: o BPEWS-Br mostrou-se confiável e viável para reconhecer sinais de alerta de deterioração clínica nas crianças estudadas.

Palavras-chave: Reproducibilidade dos Testes; Estudos de Viabilidade; Criança Hospitalizada; Enfermagem Pediátrica.

INTRODUCTION

The Pediatric Early Warning Scores (PEWS) are tools developed to assist in the early detection of deterioration of the clinical conditions of hospitalized children that are under observation of the health team, in order to provide immediate assistance. Since 2005, many PEWS have been published in the international literature, and among them stands out the Brighton Pediatric Early Warning Score (BPEWS).

The BPEWS or Monaghan PEWS, as it is also known, is basically based on the assessment of neurological, cardiovascular, and respiratory signs of the child. It ranges from zero to 13 points and its minimum score for the risk of clinical deterioration is three points. The clinical indicators that compose the instrument are: the child’s spontaneous neurological response or neurological response to stimuli, skin color, capillary refill time (CRT), heart rate (HR), respiratory rate (RR), use of accessory muscles, need for oxygen support or nebulization, and post-surgical vomiting.

This instrument was described as valid and reliable to identify signs of clinical deterioration in hospitalized children in non-Brazilian contexts. It is a simple, fast-paced score that evaluates clinical criteria in the child. These are characteristics that should be considered for its use in the Brazilian hospital environment, especially in the public sector, whose human, material, and technological resources deficit is a reality.

The BPEWS was translated, adapted, and validated to the Brazilian context in 2016, being necessary to verify its reproducibility and applicability in the detection of clinical deterioration in children. Reproducibility or reliability is defined as the ability of a test to show consistent results, performed independently, under the same conditions, being one of the properties required for adopting a measurement tool in the health care.

In addition to validity and reliability, another important feature of PEWS is the time taken for its use, considering that it should not generate extra work for the nursing team, especially in units where there is work overload.

The article aimed to measure the reproducibility and applicability of the Brighton Pediatric Early Warning Score to the Brazilian context (BPEWS-Br) in order to detect clinical deterioration.

METHOD

This is a study to measure the performance of a diagnostic test conducted to test the inter-rater reproducibility and applicability (time of use) of the translated and adapted version of the Brighton Pediatric Early Warning Score to the Brazilian context (BPEWS-Br).

Fifty children with ages ranging from zero to 10 years old participated in the study and were randomly selected; they were hospitalized in a pediatric referral hospital located in the municipality of Feira de Santana, a city with approximately 600 thousand inhabitants in the countryside of Bahia, Brazil. A sample of 50 children was chosen in order to verify the reliability of BPEWS-Br, and other studies adopted similar samples.

Inclusion criteria were: children aged zero to 10 years old, hospitalized in clinic-surgical wards and emergency observation/stabilization units, regardless of length of hospital stay. Exclusion criteria were: age ≥ 11 years old, medical discharge prescribed in medical records, children with heart disease, hospitalized at the Oncology unit and/or in isolation.

Children with heart disease were excluded because already exists in the international literature a proposed scale for this population. Children of the Oncology unit and in isolation were excluded because they were under treatment, with restricted manipulation due to low immunity and risk of cross infection during data collection.

**Table 1 - Translated, adapted, and validated version of the Brighton Pediatric Early Warning Score to the Brazilian context**

<table>
<thead>
<tr>
<th>Components</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Partial Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurological</td>
<td>Active</td>
<td>Sleepy / hypoactive</td>
<td>Irritable</td>
<td>Lethargic/confused or reduced response to pain</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Pink or CRT 1-2 sec</td>
<td>Pale or CRT 3 sec or HF above the upper limit for the age</td>
<td>Mottled or CRT 4 sec or HF ≥ 20 bpm above the upper limit for the age</td>
<td>Grey/cyanotic or CRT ≥ 5 sec or HR ≥ 30 bpm above the upper limit for the age or bradycardia for the age</td>
<td></td>
</tr>
<tr>
<td>Respiratory</td>
<td>Normal HR for the age, no retraction</td>
<td>HR above the upper limit for the age, using accessory muscles or FiO₂ ≥ 30% or 4 liters/min of O₂</td>
<td>HR ≥ 20 bpm above the upper limit for the age, subcostal, intercostal, and furcula retraction or FiO₂ ≥ 40% or 6 liters/min of O₂</td>
<td>HR ≤ 5 bpm below the upper limit for the age, subcostal, intercostal, furcula, and sternum retraction and moaning or FiO₂ ≥ 50% or 8 liters/min of O₂</td>
<td></td>
</tr>
</tbody>
</table>

Add 2 extra points if there was nebulization for up to 15 minutes or persistent vomiting after the surgery

**FINAL SCORE**

*Fonte: Miranda JOF et al.*
To test the inter-rater reliability of BPEWS-Br, two nurses, specialists in Pediatrics, were properly trained. They answered a pre-test on clinical evaluation of critically ill children and the use of pediatric early warning scores to detect clinical deterioration, achieving 50%. Then they underwent a theoretical and practical training with the same themes organized in five meetings. After the training, the nurses answered the post-test and achieved 90%.

As a strategy to calibrate the measurements in the training phase, the nurses, together with the responsible researcher, read and discussed the operational manual created to systematize the evaluation criteria for children and the use of BPEWS-Br. The score was then assigned by the two nurses in a pilot test with 10 children and the remaining doubts were answered.

After the training phase, data were collected on 50 children. Two instruments were used: BPEWS-Br and the instrument for collection of identification, socio-demographic and clinical data. The BPEWS-Br variables have a discrete categorical, ordinal, and interval nature, collected from the clinical examination of the child by the nurses. The identification, socio-demographic, and clinical variables of the children were categorical and nominal and were collected from medical record data and data provided by the parents or guardians.

The BPEWS-Br was randomly assigned with an interval of three to five minutes between the nurses’ evaluations. The period spent in the BPEWS-Br was timed by the nurses in order to evaluate the applicability of the score.

Based on the study that verified the accuracy/validity of BPEWS-Br, the score ≥ 3 was defined to determine signs of clinical deterioration. In these cases, the on-call nurse was advised to evaluate and proceed according to the routine of the service.

It is important to note that inter-rater reliability depends on accurate operational definitions of variables measured and observers trained to use the instrument. The agreement between independent observers when applying the scoring criteria is the most important factor. For evaluation of the Kappa index, reference criteria proposed for interpretation of the agreement level was adopted: < 0.00 (poor), 0.00-0.20 (slight), 0.21-0.40 (fair), 0.41-0.60 (moderate), 0.61-0.80 (substantial), 0.81-1.00 (almost perfect).

**RESULTS**

**Characterization of the sample**

The characterization of the socio-demographic and the clinical profile of the 50 children evaluated in the study is described in Table 2. Regarding socio-demographic data, the majority were < 6 years old (66%), brown or black (82%), as declared by the accompanying persons, with income <1 minimum wage, which reflects a socially vulnerable population, a common characteristic of the population assisted in the Brazilian public services. Regarding the clinical profile, 84% were hospitalized for clinical reasons, 52% were hospitalized for 15 days or more, and 44% had a history of previous hospitalization.

Table 2 - Distribution of the socio-demographic and clinical characteristics of the children evaluated. Feira de Santana, Bahia, Brazil, 2015

<table>
<thead>
<tr>
<th>Socio-demographic and clinical characteristics (n=50)</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 to 10</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>3 to 5</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>1 to 2</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>&lt; 1</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Black/brown</td>
<td>41</td>
<td>82</td>
</tr>
<tr>
<td><strong>Income (minimum wages)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 to 4</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>1 to less than 2</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Less than 1</td>
<td>41</td>
<td>82</td>
</tr>
</tbody>
</table>

**DATA are presented in tables.**

Following the ethical questions of research with children, the parents/guardians were submitted to the application of the free and informed consent and, children> 6 years old, to the Consent to Treatment. The research was approved by the Comitê de Ética da Escola de Enfermagem of Nursing of the Universidade Federal da Bahia, Brazil, registered in the Comissão Nacional de Ética em Pesquisa under opinion number 964.177 and the Certificate of Presentation for Ethical Consideration (CAAE – “Certificado de Apresentação para Consideração Ética”, in Portuguese language) 40030314.7.0000.5531.
According to Table 3, the coefficients of agreement of the final scores among the nurses’ evaluations were 0.85 (simple Kappa) and 0.80 (weighted Kappa). The interpretation of the simple Kappa coefficient showed an almost perfect agreement (0.81-1.00). On the other hand, the weighted Kappa showed substantial agreement (0.61-0.80).14

As for partial neurological, cardiovascular, and respiratory scores, the coefficients were 1.00, 0.54, and 0.93 (weighted Kappa), respectively. Thus, the agreement measures calculated by the weighted Kappa were perfect (1.00) regarding the neurological score, moderate (0.41-0.60) regarding the cardiovascular score, and almost perfect (0.81-0.99) regarding the respiratory score.

The classification of the evaluated children and the average time of use of BPEWS-Br according to the final scores calculated by the nurses are described in Table 4. Nurses 1 and 2 identified, respectively, that 3 (6%) and 4 (8%) of the 50 children evaluated were showing warning signs.

The average time during which the nurses used the BPEWS-Br was 4.14 minutes and 3.48 minutes, with a minimum time of two minutes and a maximum of six. The nurses needed more time to evaluate and use the BPEWS-Br in children with a final score ≥ 3 (5 and 4.25 minutes) in comparison with children with a score ≤ 2 (4.09 and 3.41 minutes), thus the time seemed to vary according to the final score.

### DISCUSSION

**Reliability of BPEWS-Br**

Some studies that measured the interobserver reliability of adapted/modified BPEWS versions found good agreement indicators and considered the score to be reliable. The first study that evaluated the use of an adapted version of BPEWS for detection of clinical deterioration in hospitalized children measured its reliability between two nurses. They assigned the score independently in 55 patients with a few minutes interval. The intraclass correlation coefficient was 0.92, evidencing a high interobserver reliability.15

Another study examined the psychometric properties of the Texas Children’s Hospital Pediatric Advanced Warning Score (TCH PAWS) as an indicator of clinical deterioration in 150 infants and children. The TCH PAWS is an instrument modified from an adapted version of BPEWS. For the calculation of interobserver reliability, two nurses evaluated the patients and assigned the score at the same time. The agreement measure calculated by the intraclass correlation coefficient was 0.92, evidencing a high interobserver reliability.16
Reproducibility and applicability of a pediatric score of clinical deterioration warning

By modifying BPEWS for use in children with heart disease—aiming at validating the Cardiac Children’s Hospital Early Warning Score (C-CHEWS) tool in order to identify clinical deterioration in hospitalized pediatric patients with heart disease, the study measured the interobserver reliability in a sample of 37 children. An agreement was found among the scores in 67% of the times, with Kappa of 0.50, which is considered moderate. However, when the score was categorized and classified as ≥ 3 (the first cut-off point in the C-CHEWS algorithm that triggers a response) and ≤ 2, the scores matched 100% of the times, with Kappa of 1.00 (perfect).13

With the purpose of exploring whether BPEWS assigned to children in the emergency could predict the need for ICU hospitalization or clinical deterioration in hospitalized patients, a study also found its inter-rater reliability, finding a ICC of 0.91, considered excellent.17

In this study, a perfect agreement was found for the neurological score and almost perfect for the respiratory score. However, for the cardiovascular score, the agreement was moderate. These data may be related to measurement of capillary refill time (CRT), since this was the clinical sign that varied the most in the evaluation of the nurses.

The CRT, measured in seconds, consists in the time it takes for a distal capillary bed to recover its color after sufficient pressure is applied to cause bleaching. It can be measured by different techniques, and its result is susceptible to certain factors such as age, ambient, skin, and body temperature, lighting, as well as duration, quantity, and place of application of pressure.18 The study that investigated the graduation of CRT by a group of nine nursing assistants and 37 nurses revealed that the value of Kappa for normality was 0.56 and the intraclass correlation coefficient was 0.62. In view of these results, the authors concluded that CRT should be used with caution in clinical practice.19

Based on the data from this study and the final score agreement, BPEWS-Br proved to be a reliable instrument for detection of warning signs of clinical deterioration among the children studied.

**Time interval of BPEWS-Br**

The time required to assign the BPEWS was described by its author in the original study of the score. The time it took to calculate the score was 30 seconds from a standard set of observations, which time was reduced as the nurse became familiar with the scoring system.2

It should be noted that in this study the vital data necessary for assigning the score (respiratory and cardiac frequencies) were measured at the time of the evaluation, even if some child was monitored, which required more time to evaluate and use the instrument. Thus, if the use of BPEWS-Br was linked to control of vital signs, a routine already established in the hospitalization units, the average time spent evaluating and assigning the score could have been reduced.

Early warning scoring systems, because of their objectivity and ease of use (about 15-20 seconds), helped to increase confidence in nurses to detect children with risk of deterioration.16 In addition, the use of early warning instruments promotes fast and efficient communication among nurses, physicians, and the health team.20

No studies were found that reported the total average time spent evaluating children and assigning the score. However, this is a prerequisite that needs to be considered. The BPEWS was chosen by some researchers, who considered it an instrument easily adaptable to the assistant nursing workflow,4 and can be assigned quickly and accurately by nurses with work overload in emergency units.17

The authors suggested some standards for selection and use of a PEWS: validity; ease of use; practicality; generalization for any kind of child care; incorporation of other observations and assessment scales. In addition, a PEWS must be properly taught to nurses prior to its implementation so that it can strengthen professional relationships and communication.21

This study provides evidence on the reproducibility and applicability of BPEWS-Br in a Brazilian context. Few studies using BPEWS presented detailed description of the process followed by its reliability measure, and few researches described the time required for this score application.

The training for assignment of alert scores should be guided by well-designed operational protocols. The measurement of the clinical criteria of the score needs to be well standardized and thoroughly known by the professionals who will assign it, so that their values are similar and do not generate very distant interpretations. Both the well-executed training and the involvement of the health team in the use of a PEWS are important prerequisites for the success of its implementation and possible impact on the service.

**CONCLUSION**

The adoption of a PEWS involves important criteria that not only validity and reproducibility, such as ease of use, practicality, and training. This study intended to follow these criteria.

Based on the BPEWS-Br clinical parameters, the accuracy already described in another research, and its reproducibility indicators and average time of application found in this study, it can be considered a valid, reliable, and feasible instrument to measure clinical deterioration in hospitalized children, which help to stimulate its use in pediatric hospital services throughout the country.
Reproducibility and applicability of a pediatric score of clinical deterioration warning

This is the first research on the reproducibility of BPEWS-Br in the Brazilian scenario, which show the need for future studies to strengthen the evidence so that this score can be used.

ACKNOWLEDGMENTS

To the children and their accompanying persons that participated in the study. To the Children’s State Hospital and its team, for the hospitality during the research.

REFERENCES


This is an open-access article distributed under the terms of the Creative Commons Attribution License.