Original Article
doi: https://doi.org/10.1590/1983-1447.2022.20210329.en
☑ Knowledge translation and advances in health and nursing practices

Cross-cultural adaptation and validation of the Hamilton Early Warning Score for Brazil

Adaptação transcultural e validação da Hamilton Early Warning Score para o Brasil

Adaptación transcultural y validación del Hamilton Early Warning Score para Brasil

Luana Vilela e Vilaça[a]
Fabiana Cristina Pires Bernardinelli[a]
Allana dos Reis Correa[b]
Rosali Isabel Barduchi Ohl[c]
Elizabeth Barichello[a]
Suzel Regina Ribeiro Chavaglia[a]

How to cite this article:
Vilaça LV, Bernardinelli FCP, Correa AR, Ohl RIB, Barichello E, Chavaglia SRR. Cross-cultural adaptation and validation of the Hamilton Early Warning Score for Brazil. Rev Gaúcha Enferm. 2022;43(spe):e20210329. doi: https://doi.org/10.1590/1983-1447.2022.20210329.en

ABSTRACT

Objective: Cross-culturally adapt and validate, for Portuguese, the Hamilton Early Warning Score to detect clinical deterioration in emergency services.

Method: Methodological study comprising the stages of translation, synthesis, back translation, expert committee (n=13), pre-test, submission, and analysis of the measurement properties in a sample of 188 patients. The Canadian Acute Scale Triage was compared with the Hamilton Early Warning Score. The Weighted Kappa Coefficient, Intraclass and Pearson Correlation Coefficient, Binary Logistic Regression and the Area Under the Receiver Operating Characteristic Curve were used for data analysis.

Results: The Hamilton Early Warning Score showed excellent reliability, α=0.924 (p<0.001). The construct validity identified a strong and negative correlation r=-0.75 and the predictive one presented an odds ratio of 1.63, 95% CI (1.358-1.918) (p<0.001).

Conclusion: The Hamilton Early Warning Score in Portuguese is valid and reliable to recognize patients in a condition of clinical deterioration in emergency services.

Keywords: Validation study. Early warning. Clinical deterioration. Emergency medical services. Cross-cultural comparison. Patient safety.

RESUMO

Objetivo: Adaptar transculturalmente e validar, para a língua portuguesa, a *Hamilton Early Warning Score* para detectar a deterioração clínica em serviços de emergência.

Método: Estudo metodológico compreendendo as etapas de tradução, síntese, retrotradução, comitê de especialistas (n=13), pré-teste, envio e análise das propriedades de medidas em uma amostra composta por 188 pacientes. Comparou-se a *Canadian Acute Scale Triage* com a *Hamilton Early Warning Score*. Foram utilizados o Coeficiente *Kappa* Ponderado, Coeficiente de Correlação Intraclasse e de *Pearson*, Regressão Logística Binária e a Área Sob a Curva *Receiver Operating Characteristic* para a análise dos dados.

Resultados: A *Hamilton Early Warning Score* apresentou confiabilidade excelente, ou seja, α=0,924 (p<0,001). A validade de construto identificou correlação forte e negativa r=-0,75 e a preditiva apresentou um *odds ratio* de 1,63, IC 95% (1,358-1,918) (p<0,001).

Conclusão: A *Hamilton Early Warning Score* em português é válida e confiável para reconhecer pacientes em condição de deterioração clínica em serviços de emergência.

Palavras-chave: Estudo de validação. Alerta rápido. Deterioração clínica. Serviços médicos de emergência. Comparação transcultural. Segurança do paciente.

RESUMEN

Objetivo: Adaptar y validar transculturalmente, para portugués, el *Hamilton Early Warning Score* para detectar el deterioro clínico en los servicios de emergencia.

Método: Estudio metodológico que comprende las etapas de traducción, síntesis, retrotraducción, comité de expertos (n=13), pretest, envío y análisis de las propiedades de medición en una muestra de 188 pacientes. Se comparó la *Canadian Acute Scale Triage* con la *Hamilton Early Warning Score*. Para el análisis de datos se utilizaron el Coeficiente Kappa Ponderado, el Coeficiente de Correlación Intraclase y de Pearson, la Regresión Logística Binaria y el Área Bajo la Curva Característica de Operación del Receptor.

Resultados: El *Hamilton Early Warning Score* mostró excelente confiabilidad, α=0,924 (p<0,001). La validez de constructo identificó una correlación fuerte y negativa r=-0,75 y la predictiva presentó un *odds ratio* de 1,63, IC 95% (1.358-1.918) (p<0,001).

Conclusión: El *Hamilton Early Warning Score* en portugués es válido y confiable para reconocer pacientes en deterioro clínico en servicios de emergencia.

Palabras clave: Estudio de validación. Alerta temprana. Deterioro clínico. Servicios médicos de urgencia. Comparación transcultural. Seguridad del paciente.

[a] Universidade Federal do Triângulo Mineiro (UFTM). Uberaba, Minas Gerais, Brasil.
[b] Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, Minas Gerais, Brasil.
[c] Universidade Federal de São Paulo (UNIFESP). São Paulo, São Paulo, Brasil.

INTRODUCTION

Patients in severe or potentially severe condition are often referred to Emergency Medical Services (EMS). In these services, especially in the emergency room, admitted patients are vulnerable to the worsening of the clinical condition and may present a state of clinical or physiological deterioration characterized by instability of heart rate, respiratory rate, blood pressure, temperature and level of consciousness[1].

The condition of clinical deterioration is sometimes not recognized by the health team, especially in patients who are not continuously monitored, progressing to critical adverse events, such as Cardiopulmonary Arrest (CPA) and death [2]. CPA in a hospital environment rarely occurs suddenly, in which the instability of signs and symptoms can be identified up to 24 hours before this condition[3].

The early identification of the instability of signs and symptoms through surveillance and adequate monitoring are pointed out as the first conduct to consider to avoid the occurrence of CRA[4–5]. In view of this situation, it is recommended the use of Rapid Response Systems (RRS) in order to promote patient safety[6–8].

The RRS consist of two parameters, namely: the afferent pathway that corresponds to the instruments that will identify the clinical deterioration; and the efferent pathway, which is made up of professionals from the rapid response team, activated to carry out interventions that stabilize the clinical worsening of the patient[5,8]. Regarding the afferent pathway, at the end of the 20th century, early warning scales emerged, which are instruments used in a hospital environment to identify patients at risk of progressing to clinical deterioration. By the use of these scales/instruments, when indicating the first signs of instability of vital signs or symptoms compatible with worsening, the health team is able to take timely conduct to stabilize the clinical condition and, thus, contribute to the patient safety[1,6].

These instruments are referred to in the literature as early warning systems or screening and activation systems, known by the English terms Early Warning Scores (EWS) and Trick and Track Systems (TTS) respectively[5–6,8].

Among the early warning instruments, the Hamilton Early Warning Score (HEWS), built and validated in 2015 by a multiprofessional team in Canada, with the purpose of predicting the clinical deterioration of patients admitted to the emergency room, stands out for presenting reliability, accuracy, sensitivity, and specificity in predicting in-hospital mortality. In addition, it stands out for its applicability and easy handling, since it comprises easily accessible vital signs that are measured during the shift by the nursing team and by the systematization of the conducts that are indicated according to the final score[9–11].

Since then, HEWS has been applied to critically ill patients with sepsis in intensive care units, emergency medical services and wards in the early identification of clinical deterioration and sepsis, which directly contributes to patient safety[9–11].

This scale comprises seven physiological parameters, each one with a score between 0 and 3, in which 0 is the normal range and 3 is the value of the greatest abnormality[9]. The total risk score ranges from 0 to 21 points and is defined by the sum of the scores for each physiological parameter, in which the higher the score, the greater the risk of clinical deterioration[9]. The HEWS authors defined conducts according to the score: for a score of 3 points, the nursing technician must alert the supervising nurse and increase the monitoring frequency; for a score of 4, a first-year medical resident must be called in to assess the patient; a score of 5 points was described as the threshold for clinical deterioration and indicates activation of the last year resident or the rapid response team[9,11].

The HEWS is a relevant scale for clinical practice and can be used to improve resource management, quality of care and, mainly, patient safety, given its potential to predict the physiological deterioration of patients and prevent undesirable and irreversible outcomes that cause damage to health[9–11].

Despite the relevance of its use in clinical practice, a systematic review, with the aim of evaluating the ability of EWS to predict clinical deterioration in hospitalized patients, found out that there is no scientific production that evidences the adaptation and validation of the HEWS in Brazil. In order to assess patients at risk of clinical deterioration admitted to EMS[12], a fact that led to the development of this research, the following guiding question was identified: "the Brazilian version of the HEWS is valid and reliable to recognize patients in clinical deterioration?". In view of the above, the objective of this study was to cross-culturally adapt and validate the Hamilton Early Warning Score into Portuguese to detect clinical deterioration in emergency services.

METHOD

This is a methodological study of cross-cultural adaptation and validation of the HEWS into Brazilian Portuguese in emergency services.

Data were collected from August to December 2020, in an Adult Emergency Room (AER) of a public hospital located in the city of Uberaba, Minas Gerais, Brazil. The hospital serves 27 municipalities that make up the Triângulo Sul macro-region, being the only public hospital that offers tertiary care

of high complexity. In view of the last report that we had access to, the AER of the present hospital has 32 beds, an average of 39.51 patient/day hospitalization, with an average stay of 2.59 days.

The target population consisted of adult and elderly patients admitted to the AER. Inpatients aged 18 years or older were included. Pregnant women with peculiar physiological and clinical deterioration characteristics, patients using advanced airways and patients using sedation were excluded, since the HEWS scale is not applied to such group of patients due to the particularities of the methods for measuring the neurological status and oxygen supply.

Regarding the sample, it was configured as non-probabilistic and sequential during the data collection period due to the impossibility of drawing the sample. For the composition of the sample size, it was used the Power Analysis and Sample Size tool – version 13, and an Intraclass Correlation Coefficient (ICC) was considered between the expected adherence scores of ICC=0.9 and between the scores of the neurological status, assuming a minimum value of ICC=0.75 for an a priori power of 90% and obtaining a minimum sample size of 37 patients for interobserver reliability. For construct criterion validity and predictive criterion validity, it was considered an incidence of deaths of 24.6%, a precision of 4.5% and a confidence interval of 95% for a finite population of 400 hospitalizations per year, reaching a minimum sample of 188 individuals. A significance level corresponding to p = 0.05 was also considered.

The process of cross-cultural adaptation took place after authorization by the instrument's main author, Fox Robichaud, and was based on the methodological path proposed by international literature[13] and consisted of six stages: (1) initial translation; (2) synthesis of translations; (3) back translation; (4) expert committee; (5) pre-test; and (6) forwarding to the original person responsible for the instrument. The participation of the patients occurred in the pre-test and in the validation process through the psychometric properties test as shown in Figure 1.

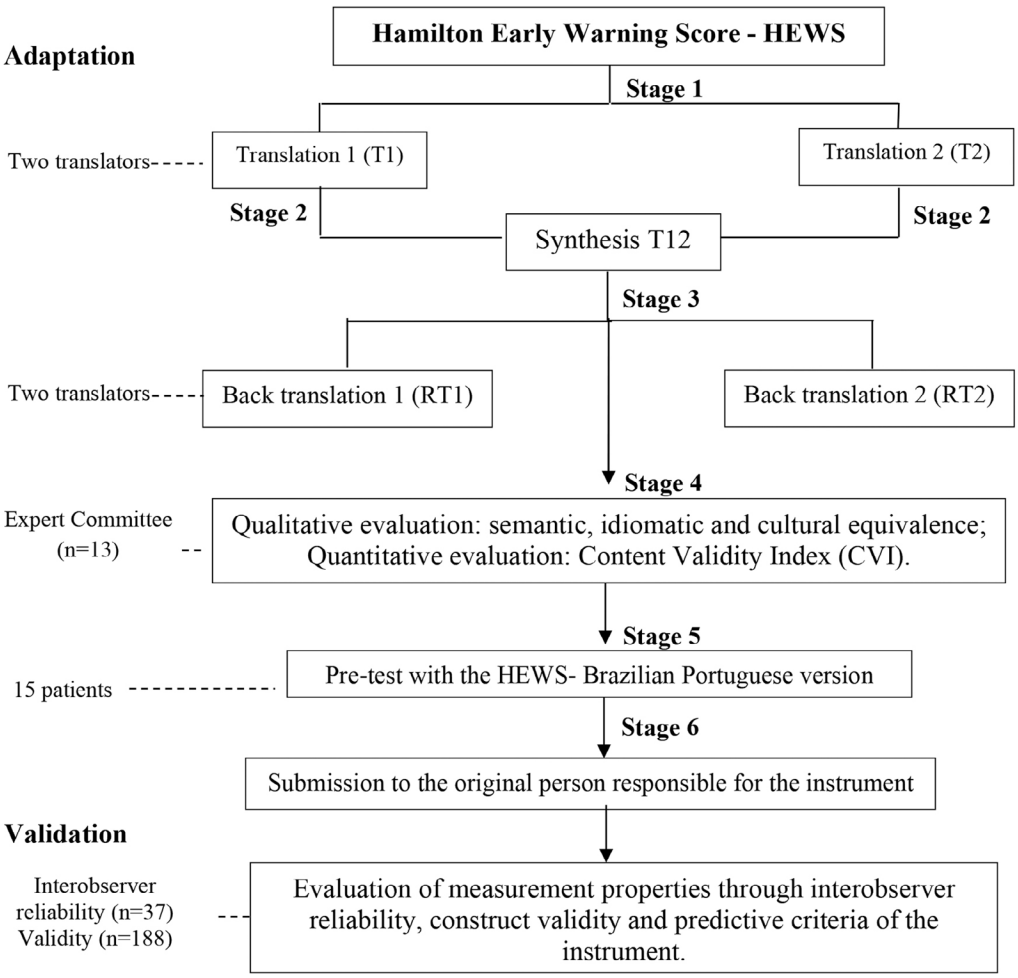

Figure 1 – Stages of cross-cultural adaptation and validation. Uberaba, Minas Gerais, Brazil, 2020
Source: Research data, 2020.

In stage 1, the scale was translated by two blinded professionals as for the objectives of the scale, independently and from the original language into Brazilian Portuguese. It is noteworthy that one of the professionals had no experience in the health area, was Brazilian, fluent in English, graduated in languages and with experience in translation; the other, in turn, had experience in health area, with a doctorate in medicine, with knowledge on the topic addressed, with experience in an English-speaking country and experienced in cultural adaptation processes. The two professionals generated the T1 and T2 versions.

In this sequence, in stage 2, versions T1 and T2 were submitted to the synthesis to reconcile discrepancies carried out by the researchers of this study, resulting in version T12. In stage 3, the synthesis back translation process was performed (version T12) through two different translators from stage 1, independently, from Portuguese to English and both were blinded as to the objectives of the scale. It is evident that the first translator is Canadian with experience in Brazil, graduated in languages, fluent in English and Portuguese and with experience in translation. In turn, the second had English as his mother language and was fluent in Portuguese, graduated in the health area and experienced in translation. Two versions were generated, namely: RT1 and RT2.

In stage 4, the evaluation by the expert committee was performed, to which T1, T2, version T12, RT1 and RT2 were submitted for consideration by the committee via an online platform to verify their conceptual, semantic, idiomatic, and cultural equivalence. A convenience sample was adopted to compose the expert committee, which was formed by 13 nurses, physicians and professionals of languages, masters and/or doctors and with expertise in urgency and emergency, located through the *Lattes* Platform and selected those who obtained the minimum value of 5 points according to the pre-established framework [14–15].

Using an electronic form, the clarity and accuracy of each domain of the instrument were evaluated using a Likert-type scale, with scores between 1 and 4, where 1 = non-equivalent item; 2 = item needs major revision to assess equivalence; 3 = equivalent item, needs minor changes; and 4 = absolutely equivalent item. Items that receive a score of 1 or 2 should be reviewed.

To quantify the degree of agreement between experts, the Content Validity Index (CVI) was used, which assesses the percentage in agreement with the aspects of the instrument and its items through the sum of the answers "3" and "4" in each item on the scale and divided by the total number of responses/experts. For this study, it was considered a CVI greater than 0.80 or 80%[16]. In this sequence, still, in stage 4, the items that obtained a CVI of less than 80% went to a second round of qualitative evaluation, in order to reconcile residual discrepancies through a meeting via an online platform with three experts[14] who were part of the expert committee, producing the final version to be submitted to the pre-test.

In stage 5, the pre-test was conducted, in which the HEWS – Brazilian version was applied by two nurses experts in urgency and emergency, since they were in direct contact with critically ill patients and were able to work with this public. It is noteworthy that the pre-test was applied to a convenience sample composed of 15 patients[17] hospitalized in the Emergency Room of this study. The application of the instrument, in a minimum sample of individuals (10 to 40 people) that includes aspects similar to the target audience, is advised to assess the understanding of the instrument and application time, in addition to verifying the understanding of the items and terms[13].

Finally, in stage 6, the scale was sent via email to the author responsible for the scale, however, after three attempts, no response was received and the process of investigating its measurement properties continued. It is worth mentioning that non-compliance with this stage was indifferent to the reliability of the adaptation process since the changes preserved the original terms of the scale. Given the above, the interobserver reliability, construct validity and predictive criteria of the scale were analyzed.

A priori, interobserver reliability was performed, investigated by two previously trained nurses, experts in urgency and emergency, who applied the HEWS scale – Brazilian version independently and without interaction, at different times, however, in a maximum interval of 10 minutes.

In assessing the validity of the construct, through the hypothesis test, in order to measure the phenomenon of clinical severity and risk of death, the HEWS – Brazilian version and the Canadian Triage Acuity Scale (CTAS)[10] were applied in 188 patients, in order to compare them due to their similarities, since both classify the patient's grade of severity and instability; the first through vital signs, the second through signs and symptoms and previous clinical history to establish treatment priorities[10].

The CTAS was created in 1997 with the purpose of classifying patients into levels of severity, establishing treatment priorities to organize emergency medical services and adapting the waiting time for medical care according to the patient's clinical condition. Its classification is made according to five levels of severity, as follows: level 1 resuscitation, which refers to life-threatening situations or imminent signs of risk of deterioration of the clinical condition; level 2, emergency, conditions that are potentially life-threatening or require rapid intervention; level 3, urgency, conditions that can progress

to a serious problem; level 4, semi-urgency, in which there is potential for complications or complications related to the patient's age; and, finally, level 5, non-urgent, which refers to non-urgent acute conditions or chronic problems without signs of deterioration[10].

Physiological parameters, signs and symptoms of the patient were collected at hospital admission for CTAS at the same time with the calculation of the HEWS – Brazilian version score. In summary, for the validity of the predictive criterion, it was investigated the outcome CRA, death and referral to the Intensive Care Unit (ICU) after the 24-hour period of calculating the HEWS – Brazilian version score.

Reliability analysis was performed with the support of MedCalc statistical software, considering a Weighted Kappa Coefficient for ordinal variables and an Intraclass Correlation Coefficient (ICC) for the total acceptable risk score, that is, the one that obtains a value greater than 0.75[18], with a significance level of p=0.05.

The data analysis of construct validity evaluated Pearson's Correlation to set the degree of relationship between the two instruments. Data analysis of Predictive Criteria Validity comprised binary logistic regression, sensitivity, and specificity by cross-tabulation for the cut-off points of HEWS – Brazilian version 3, 6 and 9 and accuracy through analysis of the Area Under the Curve (AUC) Receiver Operating Characteristic (ROC) through the statistical software Statistical Package for the Social Sciences (SPSS) version 21.

This research was approved by the Research Ethics Committee of the *Hospital de Clínicas* of the *Universidade Federal do Triângulo Mineiro* (HC-UFTM) under opinion number 3,903,656 and complied with the ethical principles of resolution 466/2012.

RESULTS

In the process of translation, synthesis, back translation and evaluation by the expert committee, the terms that generated divergences and presented CVI less than 0.80 were "room air", which ranged between environment air, location air and ambient air and "unresponsive" ranging from unresponsive, unconscious, and no response. Such divergences were discussed in a consensual meeting between the researchers and three experts, in which the terms ambient air and unresponsive were chosen due to the understanding that they best adapt to the Brazilian context.

The generated versions were sent for content validation by the expert committee composed of 13 experts, the majority, 9 (69.2%) had doctorate, 10 (76.9%) worked in teaching and researching and 11 (84, 6%) had more than two years of experience in the urgency and emergency area.

The quantitative evaluation obtained a considerable CVI of 0.89 and there was an indication of clarification of the mnemonic exposed in the AVPU footer (patient alert, responds to verbal stimulus, responds to pain stimulus or unresponsive to stimulus). However, three terms had a CVI lower than 0.80; these were addressed in a consensus meeting with experts with 100% agreement. Subsequently, the instrument was submitted to the pre-test, which was applied by two nurses to a population of 15 patients. It should be noted that the researchers did not point out difficulties or suggestions regarding the terms of the scale, which resulted in the final version of the HEWS – Brazilian version or Hamilton Early Alert Scale presented in table 1, below.

The adapted version was submitted to the analysis of the measurement properties through application of the scale in 188 patients, of which the majority were male (62.2%) and elderly with a median of 61.5 years. Most, 140 people (74.45%), had at least one comorbidity. From these, 102 patients (54.25%) reported cardiovascular diseases and 32 (17.0%) endocrine diseases. Regarding the clinical characterization, several hospitalization diagnoses were raised, the most recurrent being those related to cardiovascular diseases, that is, 63 (32.9%), followed by gastrointestinal 33 (17.6%) and external causes or violence 21 (11.2%).

A priori, from the 188 patients, 37 participated in the interobserver reliability analysis applied by two researchers. After performing the analysis, the values obtained expressed an excellent correlation of 0.924 (0.858 – 0.9603) with a significance level of p<0.001 for the total score and high to very high agreement from 0.83 to 1.0 (p<0.001) for the individual items of the scale. The heart rate, systolic blood pressure and oxygen supply domains had a total agreement of 1.0. The items with the lowest value were respiratory rate (0.831) and oxygen saturation (0.834).

In this sequence, the construct validity had a total of 179 patients due to the loss of 9 patients. It is noteworthy that the loss of less than 5% of the sample does not affect the statistical power of the study, since the construct validity requires a considerably smaller sample than the predictive validity for the analysis. In the analyses, a correlation of *r* -0.75 was observed between the HEWS and the CTAS, obtaining a statistically significant result (p<0.001) and a strong and negative correlation between the two instruments.

Finally, the Predictive Criterion Validity was performed, which had a sample of 188 patients and an incidence of 28 cases (14.9%) of critical events (CPA, death, and ICU referral). From these, 10 (5.3%) died, 2 (1.1%) had CPA with return of spontaneous circulation and 16 (8.8%) were referred to the ICU. For all the statistical analysis, the critical events mentioned above were considered as the outcome in a combined way.

Table 1 – HEWS scale adapted to Brazilian Portuguese. Uberaba, Minas Gerais, Brazil, 2020

Hamilton Early Warning Scale							
	3	**2**	**1**	**0**	**1**	**2**	**3**
Heart rate		<40	41-50	51-100	101-110	111-130	>130
Systolic blood pressure	<70	71-90		91-170		171-200	>200
Respiratory rate	<8	8-13		14-20		21-30	>30
Temperature	<35º	35.1º-36.4º	36.5º-38º	38.1º-39º	>39.1º		
Oxygen saturation	<85		85-91	>92			
Oxygen supply				Ambient air	≤5l/min or ≤50% per mask		>5l/min or >50% per mask
Neurological status		Positive CAM		Alert	Verbal	Pain	Unresponsive

Neurological status according to the Confusion Assessment Tool (CAM) scale and assessment of the patient's response to stimuli according to the AVPU mnemonic (patient alert, responds to verbal stimulus, responds to painful stimulus or unresponsive to stimulus).

Source: Research data, 2020.

In the Binary Logistic Regression, it was obtained the value of Odds Ratio (OR) 1.63; 95% CI, (1.358-1.918); $p < 0.001$, showing that the chances of occurrence of the investigated adverse events increase by 60% for each additional point on the HEWS scale score.

Regarding the classification of risk of clinical deterioration identified in this population through the application of the HEWS, the majority 97 (51.6%) were evaluated as low risk (0 to 2 points), followed by 50 (26.6%) intermediate risk (3 to 5 points), 23 (12.2%) high risk (6 to 8 points) and 18 (9.6%) very high risk (above 9 points).

The results achieved for sensitivity and specificity according to the cut-off point of the HEWS risk classification showed that an ideal cut-off value is situated between a score of 3 to 6, as shown in Table 2.

The analysis of accuracy through the area under the ROC curve obtained values considered high (ASC > 0.8), indicating that the instrument's accuracy is satisfactory, as shown in table 3.

Table 2 – Distribution of sensitivity and specificity indicators of the HEWS instrument – Brazilian version through cross-tabulation for different cutoff points. Uberaba, Minas Gerais, Brazil, 2020

Validity indicators	Hamilton Early Warning Scale cutoff points		
	≥3 (intermediate risk)	**≥6 (high risk)**	**≥9 (very high risk)**
Sensitivity	96.4%	71.4%	46.4%
Specificity	60.0%	86.9%	96.9%
False positives	40.0%	13.1	03.1%
False negatives	03.6%	28.6	53.6%

Source: Research data, 2020.

Table 3 – Distribution of the overall accuracy of the HEWS – Brazilian version regarding the investigated outcomes. Uberaba, Minas Gerais, Brazil, 2020

Combined outcome	Overall Accuracy
Cardiopulmonary arrest, death and transfer to the Intensive Care Unit	**0.89** (0.831 – 0.953, 95% CI; p<0.001)
Individual outcomes	**Accuracy**
Cardiopulmonary arrest	**0.98** (0.970 – 1.0; 95% CI; p<0.001)
Death	**0.98** (0.975 – 1.0; 95% CI; p<0.001)
Transfer to the Intensive Care Unit	**0.76** (0.676 – 0.854; 95% CI; p<0.001)

Source: Research data, 2020.

DISCUSSION

When thinking about the dynamics involved in the care for critically ill patients, there is a need for instruments and scales that identify the risk of clinical deterioration and help the management of available resources according to the degree of instability of the patient, with the intention of preventing the occurrence of adverse events and provide patient safety[1–4].

It is recommended that tools and work strategies to be used based on scientific evidence to support health care and management[19]. However, in Brazilian practice, the use of clinical instruments can still be observed, in particular, the Early Warning Scores (EWS), translated literally and without prior validation for the target population in which it will be applied[1,20].

In this sense, the process of cross-cultural adaptation and validation of the Canadian HEWS scale was given for the context of urgency and emergency with regard to EMS. Authors argue that, based on this process, the instrument can be applied to the proposed group, providing a lower degree of cultural bias for the Brazilian context, which allows comparing the characteristics of individuals, in order to provide greater equity in the process of applying the scale[1,13].

In view of the above, the present study is unprecedented and relevant by providing a national early warning scale to assist in the recognition of patients in clinical deterioration, risk management and determination of conducts, such as continuous monitoring, bed complexity, adequate interval for nursing observations and the need for medical review, improving the care provided, with a view to resource management, quality of care and patient safety.

It is noteworthy that few articles are related to the process of cross-cultural adaptation of EWS. In addition to the HEWS, the only early warning scale that was submitted to a methodological process of adaptation for Brazil was the National Early Warning Score 2 (NEWS 2) for use in adults[19] and the Brighton Pediatric Early Warning Score (BPEWS) for pediatric use[21]. When comparing the content validation results of the HEWS – Brazilian version and NEWS 2 scales, both scales obtained a good validity index, greater than 0.8, which demonstrates that these scales are clear and precise[20].

For the HEWS – Brazilian version, the terms that needed adaptation were consistent with those adapted in the Brazilian version of NEWS 2, such as: the description of the AVPU mnemonic (alert, verbal, pain and unresponsive) and the expression "ambient air"[20].

The assessment of the neurological status through the AVPU is not common in the nursing routine in Brazil, however, in order not to interfere with the original version, it was decided not to replace it, but to describe the items in full in order to make it clearer for those who will use the scale, as well as in the adaptation process of NEWS 2[20].

Regarding the pilot test, a study that adapted the BPEWS for pediatric use corroborates this research by not evidencing possible scale adjustments after the pilot test; this fact indicates the efficiency and reliability of the content validation process by the expert committee[21].

Still, about the pilot test, the research in which the NEWS 2 was adapted, as well as the present study, showed satisfactory results that demonstrate that these scales are easy to apply, access and handle, in addition to their items being accessible, that is, easy to understand, which contributes to the work of the care team, especially the nursing team, as well as expanding the academic scope for research that proposes to adapt and validate this scale in other contexts[20].

As for interobserver reliability, the scale obtained "high" to "very high" agreement (0.831 and 1.0) for individual items and excellent reproducibility of the final score. The result corroborates a Canadian study, in which the HEWS obtained a considerably satisfactory interobserver reliability, with a Kappa 0.89[10].

Regarding the vital parameters, the items that showed the greatest inconsistency were respiratory rate (0.831) and oxygen saturation (0.834). For NEWS 2, the item with the lowest agreement was related to the "oxygen supply"[20]. These results corroborate reports that identified negligence in the measurement and recording of respiratory rate by the care team, which causes possible health implications and causes direct interference to patient safety[22].

Another factor to be considered is that oxygen saturation, as well as respiratory rate, are among the main indicators of clinical deterioration, as they are the first data to present instability[2]. Such information shows the need for reflection on the relevance and correct measurement and recording of such vital data[5].

Regarding the assessment of construct validity, the present research identified a strong and negative correlation (r=-0.75) between HEWS – Brazilian version and CTAS. The application of EWS in screening combined with recurrent use during hospitalization is indicated by several authors, since the risk classification, by itself, is not enough to distinguish between stable and unstable patients during the observation period[3,10].

This investigation identified OR 1.63, 95% CI (1.358-1.918), p<0.001 in the binary logistic regression analysis to predict clinical deterioration. This indicates that the chances of occurrence of the investigated adverse events increase by 60% for each additional point in the HEWS scale score – Brazilian version.

These findings are consistent with another study that applied an EWS in an emergency department and pointed out that 40% of CRA cases presented a one point increase in the risk classification, one hour before the critical event[7]. Evidence shows that 60% to 80% of patients who develop clinical deterioration show changes in vital signs at least four hours before episodes such as CRA[4].

Regarding the final score of the scale, there is always an exchange between sensitivity and specificity according to the value obtained, and an ideal cutoff point with good accuracy is sought for the investigated result[5]. One of the validation studies of HEWS in Canada, evaluated the cutoff points that mark the limit for change in risk classification 3, 6 and 9, obtaining sensitivity of 54%, 19% and 9% and specificity of 63%, 89%, 98%, respectively[10].

This study followed the same pattern in the literature, as while the cutoff point increased, sensitivity decreased, and specificity increased. For scores 3, 6 and 9, more significant values were obtained than the aforementioned studies of 96.4%, 71.4% and 46.4% for sensitivity and 60%, 86.9% and 96.9% for specificity respectively.

For an EWS, the specificity demonstrates patients who were not considered to be at risk for clinical deterioration. Sensitivity, in turn, indicates the number of unstable patients that were correctly identified by the instrument. Regarding the values found in this study, the lower the HEWS score (three points), the greater the sensitivity (96.4%) and the higher the percentage of false positives (40%). On the other hand, the higher the score (nine points), the greater the specificity (96.9%) and the higher the number of false negatives (53%).

In this sense, one of the consequences of establishing the cutoff point for an EWS is the unnecessary workload it can generate (false positives) and the number of deteriorating patients, which may not be identified (false negatives). Authors argue that the more specific the EWS, the fewer activations can be triggered; in view of this question, more specific scores would be the most useful[5,9].

Although sensitivity and specificity values were presented, a cutoff point was not determined for use in this population, by understanding that this is a determination to be made by the work team according to the profile of the patients treated and the institution's resources, which corroborates the observation made in a study with the Modified Early Warning Scores (MEWS), conducted in Ribeirão Preto, São Paulo[1].

Investigations conducted with the HEWS portray this variation of the ideal score. The HEWS score 5 was considered the best cutoff point and had a sensitivity of 75.9% and specificity of 67.6% in a retrospective Canadian study[9]. On the other hand, in a study carried out in China, the ideal value for the HEWS cutoff point was 8, with a sensitivity of 80% and specificity of 89.69%[19], which provides patient safety by offering a sensitive and specific scale in detecting clinical deterioration.

Regarding the accuracy through the area under the ROC curve, this study obtained a good accuracy of 0.89 (0.831 – 0.953, 95% CI; p<0.001), almost reaching an excellent value

of 0.9, which corroborates a study carried out in Spain that found a good accuracy of 0.89 (0.81 – 0.96, 95% CI; p<0.001) [23]. Also, a recent research conducted in China with HEWS obtained a similar value of 0.821 (95% CI: 0.748-0.895)[24]. In Canada, an investigation demonstrated a good discriminative ability to predict the occurrence of a critical event among septic patients with a value of 0.82 (95% CI: 0.75-0.90)[10]. In another study carried out in Canada, the area under the ROC curve was 0.76 (95% CI: 0.75–0.77), for HEWS in the overall population and among patients with suspected infection it was 0.79 (95% CI: 0.76-0.81)[9].

A new aspect to be considered is the coronavirus 2019 pandemic (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARSCoV-2), which has substantially increased the number of hospitalizations requiring critical care and, consequently, the lack of availability of beds in ICUs.

Thus, early recognition of patients requiring ICU admission is a critical step in the treatment of patients with COVID-19, which can be signaled by the application of EWS[25]. A comparative study with five EWS indicated the HEWS as one of the scales with the best predictive value in this group of patients[24].

Studies carried out in Brazil, with the objective of investigating the potential of HEWS and MEWS, showed significant results when detecting improvements in patient safety as a result of the early recognition of physiological deterioration and the reduction of unfavorable clinical outcomes, often characterized by death[1,24,26].

The HEWS is shown to be a differentiated scale because it includes a variation in the score according to the volume of oxygen offered, indicating the patient who is in greater respiratory instability[24]. Another differential was to consider the presence of mental confusion or delirium, since patients with acute illness may manifest changes in mental status as a result of hypoxia, hypotension, sepsis or metabolic disorders[10,27].

It is worth noting that the complexity of comorbidities and drug use interfere with the patient's physiological response, thus, this interference may imply the reliability of the HEWS – Brazilian version. In view of the exposed, the nursing team needs to associate their knowledge and experiences with these tools in order to identify values that reflect normality, or not, according to the patient's individuality[5,27].

CONCLUSION

The HEWS was adapted to Brazilian Portuguese, validated and proved to be a reliable instrument to identify clinical deterioration in patients admitted to Emergency Services, specifically, emergency rooms.

A priori, stands out as a limitation of the present study is that data collection was carried out in only one EMS, which makes it difficult to generalize the findings. In addition, although the investigation has correlated the occurrence of critical events with a higher risk score identified by the HEWS scale – Brazilian version, as it is a methodological validation study, it was not evaluated its application within the systematized flowchart of conducts by the health care team, in order to investigate its impact on reducing critical events.

This study contributes to health and nursing by presenting a potential for knowledge translation obtained to clinical practice, insofar as it provides a cross-culturally adapted and validated scale, capable of identifying the clinical deterioration of patients and subsidizing the best practices in this perspective by contributing to the systematized work process, which consists of activities that enable nurses to detect, solve and/or forward patient demands, minimizing the risks arising from the care provided and, mainly, contributing to patient safety.

It is recommended that methodologically well-designed studies be conducted, with the purpose of apply the HEWS scale to other sectors and populations in a hospital environment, as well as to measure the influence of the parameters used and the interference of the patient profile according to the use of medication, age and comorbidities presented. It is also suggested to carry out studies on the application of the scale and its flowchart of conducts in care practice to evaluate its performance in reducing CPA and death.

REFERENCES

1. Montenegro SMSL, Miranda CH. Evaluation of the performance of the modified early warning score in a Brazilian public hospital. Rev Bras Enferm. 2019;72(6):1428-34. doi: https://doi.org/10.1590/0034-7167-2017-0537.
2. Considine J, Street M, Hutchinson AM, Mohebbi M, Rawson H, Dunning T, et al. Vital sign abnormalities as predictors of clinical deterioration in subacute care patients: A prospective case-time-control study. Int J Nurs Stud. 2020;108:103612. doi: https://doi.org/10.1016/j.ijnurstu.2020.103612.
3. Srivilaithon W, Amnuaypattanapon K, Limjindaporn C, Imsuwan I, Daorattanachai K, Dasanadeba I, et al. Predictors of in-hospital cardiac arrest within 24 h after emergency department triage: a case–control study in urban Thailand. Emerg Med Australas. 2019;31:843-50. doi: https://doi.org/10.1111/1742-6723.13267.
4. Andersen LW, Holmberg MJ, Berg KM, Donnino MW, Granfeldt A. In-hospital cardiac arrest – a review. JAMA. 2019;321(12):1200-10. doi: https://doi.org/10.1001/jama.2019.1696.
5. Grant S, Crimmons S. Limitations of track and trigger systems and the National Early Warning Score. part 2: sensitivity versus specificity. Br J Nurs. 2018;27(12):705-10. doi: https://doi.org/10.12968/bjon.2018.27.12.705.
6. Spencer W, Smith J, Date P, Tonnerre E, Taylor DMC. Determination of the best early warning scores to predict clinical outcomes of patients in the emergency department. Emerg Med J. 2019;36:716-721. doi: https://doi.org/10.1136/emermed-2019-208622.

7. Green C, Chaudhri S, Premchand M, Peet C, Sadiq A, Strain S, et al. Clinical appropriateness of the use of early warning scores in medical wards. Clin Med. 2020;20(Suppl 2):s82-s83. doi: https://doi.org/10.7861/clinmed.20-2-s82.
8. Haegdorens F, Van Bogaert P, Roelant E, De Meester K, Misselyn M, Wouters K, et al. The introduction of a rapid response system in acute hospitals: a pragmatic stepped wedge cluster randomised controlled trial. Resuscitation. 2018;129:127-34. doi: https://doi.org/10.1016/j.resuscitation.2018.04.018.
9. Fernando SM, Fox-Robichaud AE, Rochwerg B, Cardinal P, Seely AJE, Perry JJ, et al Prognostic accuracy of the Hamilton Early Warning Score (HEWS) and the National Early Warning Score 2 (NEWS2) among hospitalized patients assessed by a rapid response team. Crit Care. 2019;23:60. doi: https://doi.org/10.1186/s13054-019-2355-3.
10. Skitch S, Tam B, Xu M, McInnis L, Vu A, Fox-Robichaud A. Examining the utility of the Hamilton Early Warning Scores (HEWS) at triage: Retrospective pilot study in a Canadian emergency department. CJEM. 2018;20(2):266-74. doi: https://doi.org/10.1017/cem.2017.21.
11. Tam B, Xu M, Kwong M, Wardell C, Kwong A, Fox-Robichaud A. The admission Hamilton Early Warning Score (HEWS) predicts the risk of critical event during hospitalization. Can Journ Gen Int Med. 2016;11(4):24-7. doi: https://doi.org/10.22374/cjgim.v11i4.190.
12. Brekke IJ, Puntervoll LH, Pedersen PB, Kellett J, Brabrand M. The value of vital sign trends in predicting and monitoring clinical deterioration: a systematic review. PLoS One. 2019;14(1):e0210875. doi: https://doi.org/10.1371/journal.pone.0210875.
13. Beaton D, Bombardier C, Guillemin F, Ferraz MB. Recommendations for the Cross-Cultural Adaptation of the DASH & QuickDASH Outcome Measures [Internet]. Toronto: Institute for Work & Health; 2007 [cited 2021 Jun 11]. Available from: https://dash.iwh.on.ca/sites/dash/files/downloads/cross_cultural_adaptation_2007.pdf.
14. Alexandre NMC, Coluci MZO. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. Ciênc Saúde Coletiva. 2011;16(7):3061-8. doi: https://doi.org/10.1590/S1413-81232011000800006.
15. Fehring RJ. The Fehring model. In: Carrol-Johnson RM, Paquete M. Classification of nursing diagnoses: proceeding of the tenth conference. Philadelphia: Lippincott Williams & Wilkins; 1994. p. 55-62.
16. Polit DF, Beck CT. The contente validity index: are you sure you know what's being reported? critique and recommendations. Res Nurs Health. 2006;29(5):489-97. doi: https://doi.org/10.1002/nur.20147.
17. Malhotra NK. Pesquisa de marketing: uma orientação aplicada. Porto Alegre: Bookman; 2008.
18. Fayers PM, Machin D. Quality of life: the assesment, analysis and interpretation. Chichester, England: John Wiley & Sons, 2007.
19. Oliveira F, Kuznier TP, Souza CC, Chianca TCM. Theoretical and methodological aspects for the cultural adaptation and validation of instruments in nursing. Texto Contexto Enferm. 2018;27(2):e4900016. doi: http://doi.org/10.1590/0104-070720180004900016.
20. Oliveira, APA; Urbanetto, JS; Caregnato, RTA. National Early Warning Score 2: ranscultural adaptation to Brazilian Portuguese. Rev Gaúcha Enferm. 2020;41:e20190424. doi: https://doi.org/10.1590/1983-1447.2020.20190424.
21. Miranda JOF, Camargo CL, Sobrinho CLN, Portela DS, Monaghan A, Freitas KS, et al. Translation and adaptation of a pediatric early warning score. Rev Bras Enferm. 2016;69(5):833-41. doi: http://doi.org/10.1590/0034-7167-2015-0096.
22. Al-Kalaldeh M, Suleiman K, Abu-Shahroor L, Al-Mawajdah H. The impact of introducing the Modified Early Warning Score 'MEWS' on emergency nurses' perceived role and self-efficacy: a quasi-experimental study. Int Emerg Nurs. 2019;45:25-30. doi: https://doi.org/10.1016/j.ienj.2019.03.005.
23. Martín-Rodríguez F, Castro-Villamor MÁ, Del Pozo Vegas C, Martín-Conty JL, Mayo-Iscar A, Delgado Benito JF, et al. Analysis of the early warning score to detect critical or high-risk patients in the prehospital setting. Intern Emer Med. 2019;14(4):581-9. doi: http://doi.org/10.1007/s11739-019-02026-2.
24. Hu H, Yao N, Qiu Y. Predictive value of 5 Early Warning Scores for critical COVID-19 patients. Disaster Med Public Health Prep. 2022;16(1):232-9. doi: https://doi.org/10.1017/dmp.2020.324.
25. Swiss Society of Intensive Care Medicine. Recommendations for the admission of patients with COVID-19 to intensive care and intermediate care units (ICUs and IMCUs). Swiss Med Wkly. 2020;150:w20227. doi: https://doi.org/10.4414/smw.2020.20227.
26. Nascimento JSG, Macedo GO, Borges GB. Predictive power of the early warning scale for clinical deterioration of critically ill patients. Rev Enferm UFSM. 2020;10(e5):1-14. doi: https://doi.org/10.5902/2179769238300.
27. Mohammed MA, Faisal M, Richardson D, Scally A, Howes R, Beatson K, et al. The inclusion of delirium in version 2 of the National Early Warning Score will substantially increase the alerts for escalating levels of care: findings from a retrospective database study of emergency medical admissions in two hospitals. Clin Med. 2019;19(2):104-8. doi: https://doi.org/10.7861%2Fclinmedicine.19-2-104.

Funding/Acknowledgment:
Coordination for the Improvement of Higher Education Personnel (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* – CAPES).

Authorship contribution:
Project administration: Suzel Regina Ribeiro Chavaglia.
Formal analysis: Luana Vilela e Vilaça.
Conceptualization: Luana Vilela e Vilaça, Suzel Regina Ribeiro Chavaglia.
Data curation: Luana Vilela e Vilaça, Fabiana Cristina Pires Bernardinelli.
Writing-original draft: Luana Vilela e Vilaça, Suzel Regina Ribeiro Chavaglia, Fabiana Cristina Pires Bernardinelli, Allana dos Reis Correa, Isabel Barduchi Ohl e Elizabeth Barichello.
Writing-review & editing: Luana Vilela e Vilaça, Suzel Regina Ribeiro Chavaglia, Fabiana Cristina Pires Bernardinelli, Allana dos Reis Correa, Rosali Isabel Barduchi Ohl e Elizabeth Barichello.
Investigation: Luana Vilela e Vilaça, Fabiana Cristina Pires Bernardinelli.
Methodology: Luana Vilela e Vilaça, Suzel Regina Ribeiro Chavaglia.
Supervision: Suzel Regina Ribeiro Chavaglia.
Visualization: Luana Vilela e Vilaça, Fabiana Cristina Pires Bernardinelli, Allana dos Reis Correa, Rosali Isabel Barduchi Ohl, Elizabeth Barichello, Suzel Regina Ribeiro Chavaglia.

The authors declare that there is no conflict of interest.

Corresponding author:
Suzel Regina Ribeiro Chavaglia
E-mail: suzel.ribeiro@yahoo.com.br

Associate editor:
João Lucas Campos de Oliveira

Received: 02.02.2022
Approved: 05.16.2022

Editor-in-chief:
Maria da Graça Oliveira Crossetti