Original Article

TEXTO & CONTEXTO ENFERMAGEM
TEXT & CONTEXT NURSING TEXTO & CONTEXTO ENFERMERÍA

TRANSLATION, ADAPTATION AND VALIDATION OF THE *FULL OUTLINE OF UNRESPONSIVENESS* SCALE INTO BRAZILIAN PORTUGUESE

Fabiana Cristina Pires Bernardinelli[1]
Gustavo Correa de Amorin[2]
Vanderlei José Haas[1]
Cássia Regina Vancini Campanharo[3]
Maria Helena Barbosa[1]
Suzel Regina Ribeiro Chavaglia[1]

[1]Universidade Federal do Triângulo Mineiro, Programa de Pós-graduação em Atenção à Saúde. Uberaba, Minas Gerais, Brasil.
[2]Universidade Federal do Triângulo Mineiro, Programa de Residência Enfermagem em Urgência/Trauma. Uberaba, Minas Gerais, Brasil.
[3]Universidade Federal de São Paulo, Escola Paulista de Enfermagem, Programa de Pós-graduação em Enfermagem. São Paulo, São Paulo, Brasil.

ABSTRACT

Objective: to translate, culturally adapt and validate the *Full Outline of UnResponsiveness* scale into Brazilian Portuguese.
Method: a methodological study carried out at the Clinical Hospital of *Universidade Federal do Triângulo Mineiro*, Uberaba, Brazil, through the following stages: translation, synthesis, evaluation by the experts' committee, back-translation, consensus, semantic evaluation and pre-test. A sample of 188 adult patients was reached. Data collection took place between August and December 2020. Concurrent criterion validity was analyzed by comparing the *Full Outline of UnResponsiveness* scale with the Glasgow Coma Scale by means of *Spearman*'s and *Pearson*'s correlation coefficients; and predictive validity analysis was performed with Cox Regression, Sensitivity and Specificity and Area Under the *Receiver Operating Characteristic* Curve. The *Cronbach*'s alpha, weighted Kappa and Intraclass Correlation coefficients were also adopted for interobserver reliability.
Results: *Spearman*'s test for the motor and eye response items, respectively, resulted in 0.81 and 0.96, and *Pearson*'s test for the total score was 0.97. A relative risk of 0.80, 95.5% specificity, 51.6% sensitivity and accuracy of 0.80 (95% CI: 0.688–0,905, p<0.001) were obtained. *Cronbach*'s alpha was 0.94, weighted Kappa varied from 0.89 to 1.0, and ICC resulted in 0.99.
Conclusion: the *Full Outline of UnResponsiveness* scale (Brazilian version), maintained four domains and the 20 items from the original scale, making it appropriate for use in Brazil and contributing to the assessment of the level of consciousness and prognosis of adult patients in severe conditions.

DESCRIPTORS: Validation study. Psychometry. Nursing. Level of consciousness. Adult.

HOW CITED: Bernardinelli FCP, Amorim GC, Haas VJ, Campanharo CRV, Barbosa MH, Chavaglia SRR. Translation, adaptation and validation of the *Full Outline of UnResponsiveness* scale into Brazilian portuguese. Texto Contexto Enferm [Internet]. 2022 [cited YEAR MONTH DAY]; 31:e20210427. Available from: https://doi.org/10.1590/1980-265X-TCE-2021-0427en

TRADUÇÃO, ADAPTAÇÃO E VALIDAÇÃO DA ESCALA *FULL OUTLINE OF UNRESPONSIVENESS* PARA O PORTUGUÊS DO BRASIL

RESUMO

Objetivo: traduzir, adaptar culturalmente e validar a escala *Full Outline of UnResponsiveness* para o português do Brasil.
Método: estudo metodológico realizado no Hospital de Clínicas da Universidade Federal do Triângulo Mineiro, Uberaba, Brasil, por meio das etapas: tradução, síntese, avaliação pelo comitê de especialistas, retrotradução, consenso, avaliação semântica e pré-teste. Alcançou-se uma amostra de 188 pacientes adultos. A coleta de dados ocorreu entre agosto e dezembro de 2020. Analisou-se a validade de critério concorrente comparando a escala *Full Outline of UnResponsiveness* com a Escala de Coma de Glasgow por meio dos coeficientes de correlação de *Spearman* e *Pearson*, e a validade preditiva com a Regressão de Cox, Sensibilidade e Especificidade e Área Sob a Curva *Receiver Operating Characteristic*. Adotaram-se, também, o alfa de *Cronbach* e os coeficientes Kappa ponderado e de Correlação Intraclasse para a confiabilidade interobservador.
Resultados: o teste de *Spearman* para os itens resposta motora e ocular, respectivamente, resultou-se em 0,81 e 0,96, e o de *Pearson* para o escore total em 0,97. Obteve-se um risco relativo de 0,80, especificidade de 95,5%, sensibilidade de 51,6% e acurácia de 0,80 (IC95%: 0,688–0,905, p<0,001). O alfa de *Cronbach* foi de 0,94, o *Kappa* ponderado variou entre 0,89 e 1,0 e o ICC resultou em 0,99.
Conclusão: a escala *Full Outline of UnResponsiveness* - versão brasileira, manteve quatro domínios e os 20 itens da escala original, tornando-se apropriada para utilização no Brasil e contribuindo para a avaliação do nível de consciência e prognóstico de pacientes adultos em condição grave.

DESCRITORES: Estudo de validação. Psicometria. Enfermagem. Estado de consciência. Adulto.

TRADUCCIÓN, ADAPTACIÓN Y VALIDACIÓN DE LA ESCALA *FULL OUTLINE OF UNRESPONSIVENESS* AL PORTUGUÉS DE BRASIL

RESUMEN

Objetivo: traducir, adaptar culturalmente y validar la escala *Full Outline of UnResponsiveness* al portugués de Brasil.
Método: estudio metodológico realizado en el Hospital de Clínicas de la *Universidade Federal do Triângulo Mineiro*, Uberaba, Brasil, por medio de las siguientes etapas: traducción, síntesis, evaluación a cargo del comité de especialistas, retrotraducción, consenso, evaluación semántica y prueba previa. Se llegó a una muestra de 188 pacientes adultos. La recolección de datos tuvo lugar entre agosto y diciembre de 2020. Se analizó la validez de criterio concurrente comparando la escala *Full Outline of UnResponsiveness* con la Escala de Coma de Glasgow por medio de los coeficientes de correlación de *Spearman* y *Pearson*, y el análisis de la validez predictiva se efectuó con la Regresión de Cox, Sensibilidad y Especificidad y Área por debajo de la Curva *Receiver Operating Characteristic*. También se adoptaron el alfa de *Cronbach* y los coeficientes Kappa ponderado y de Correlación Intraclase para determinar la confiabilidad interobservador.
Resultados: en la prueba de *Spearman* para los ítems de respuesta motora y respuesta ocular, respectivamente, se obtuvieron valores de 0,81 y 0,96, y el coeficiente de *Pearson* para la puntuación total fue de 0,97. Se obtuvo un riesgo relativo de 0,80, especificidad del 95,5%, sensibilidad del 51,6% y precisión de 0,80 (IC 95%: 0,688–0,905, p<0,001). El alfa de *Cronbach* fue de 0,94, el índice *Kappa* ponderado varió entre 0,89 y 1,0 y el resultado del ICC fue 0,99.
Conclusión: la escala *Full Outline of UnResponsiveness* (Versión brasileña), mantuvo cuatro dominios y los 20 ítems de la escala original, lo que la hace apropiada para ser utilizada en Brasil y contribuye a la evaluación del nivel de consciencia y del pronóstico de pacientes adultos en condiciones de gravedad.

DESCRIPTORES: Estudio de validación. Psicometría. Enfermería. Nivel de consciencia. Adulto.

INTRODUCTION

The growing need to obtain good quality health care has encouraged researchers around the world to develop instruments capable of assessing level of consciousness in critically-ill patients, in order to accurately identify and monitor their clinical conditions and support the professionals' clinical judgment on solid scientific evidence[1].

A patient is considered to be in a serious condition when there is imminent risk of death or physiological deterioration of organs and systems, due to trauma or other diseases, and also due to the presence of an unstable hemodynamic state, possibility of circulatory shock, or hemodynamic compensation through increasing and high doses of vasoactive drugs or other forms of cardiovascular support[2].

Assessment of the level of consciousness in critically-ill patients is based on the interpretation of the changes in their brain function, generally identified by the Glasgow Coma Scale (GCS), published in 1974 and updated in 2018, to monitor evolution of the level of consciousness and assist in obtaining the patients' clinical prognoses[3–4].

Regarding the GCS, it is necessary to emphasize that, even though it is commonly adopted in the clinical practice in a global context, it is an instrument not yet validated for Brazilian Portuguese, which presents diagnostic limitations such as the impossibility of verbal assessment in patients on Invasive Mechanical Ventilation (IMV) and the absence of brainstem reflex tests, factors that can compromise accuracy of the neurological assessment[5].

With the intention of filling the existing gaps in the evaluation of consciousness obtained by the GCS, in 2005, in the city of Rochester (Minnesota, United States of America), a scale called Full Outline of UnResponsiveness (FOUR) was developed and validated in several clinical settings[6–8], consisting of four domains: eye response, motor response, brainstem reflexes and breathing, each one with five items, classified from zero to four points[6].

The following stand out among the advantages presented by the FOUR scale: simple handling and interpretation of the results, in addition to easy memorization of its items, added to the possibility of evaluating important clinical variables such as brainstem reflex and the patient's breathing pattern, even in the presence of endotracheal tubes, and exceeding in amplitude and quality of the consciousness level assessment when compared to other existing scales for the same purpose, supporting its adoption[9]. However, it is considered that, even in the face of its accuracy, scientific studies that address effectiveness of this scale to assess consciousness in critically-ill patients are internationally incipient and non-existent in the national context, as the scale has not been validated and translated into Brazilian Portuguese[9].

Given the importance of obtaining validated instruments for Brazilian Portuguese that provide an accurate assessment of the level of consciousness in adult patients in severe clinical conditions, this study aimed at translating, culturally adapting and validating the Full Outline of UnResponsiveness scale into Brazilian Portuguese.

METHOD

This is a methodological study[10] on the translation, cultural adaptation and validation of the FOUR scale to the Brazilian context, which was grounded on a method developed by international authors[11].

The FOUR scale consists of four domains characterized by eye response, motor response, brainstem reflexes and breathing, each one with five answer options varying from zero to four points. The values obtained in the answers can range between zero and 16 points, and a score of 16 represents the patient's highest level of consciousness[6]. It should be noted that, in addition to this scale, data collection was carried out with the aid of the GCS, made up of 15 items distributed into three domains: eye response, verbal response and motor response[4].

The data were collected from August to December 2020 in the Adult Emergency Service (*Pronto Socorro Adulto*, PSA) of the public Clinical Hospital belonging to *Universidade Federal do Triângulo Mineiro* (HC-UFTM), Uberaba, Brazil, which has 302 beds, 32 of them for emergency care for adult patients.

The target population consisted of critically-ill patients[2] admitted to the PSA according to the following inclusion criteria: patients aged at least 18 years old, either hemodynamically unstable or compensated with the use of vasoactive drugs or other forms of cardiovascular support. Patients using sedative medications were excluded, given the impossibility of applying the GCS in this condition.

A non-probabilistic and sequential sample was adopted during the data collection period due to the impossibility of randomly selecting the sample by means of a draw. For sample size selection, the *Power Analysis and Sample Size* tool (version 13) was used, and the Intraclass Correlation Coefficient (ICC) was considered between the expected adherence scores (ICC=0.9) and between the level of consciousness scores, assuming a minimum value of ICC=0.75 for an *a priori* power of 90% and obtaining a minimum sample size of 36 patients for interobserver reliability. The following was considered for concurrent and predictive criterion validity: 24.6% incidence of death in critically-ill patients, precision of 4.5%, and a 95% confidence interval for a finite population of 400 hospitalizations a year, reaching a minimum sample of 188 individuals. A significance level of α=0.05 was also considered.

The translation, cultural adaptation and validation process took place after authorization from the main author of the scale, following these steps: 1) translation of the scale into Brazilian Portuguese; 2) synthesis - obtaining the first consensus version for Portuguese; 3) evaluation by the experts' committee; 4) back-translation into the original language; 5) subsequent consensus reached by the translators of the Portuguese versions when compared to the original; 6) performing the semantic evaluation of the items; and 7) pre-test[11]. It is noted that the patients' participation occurred in the pre-test and in the validation process by means of the test of the psychometric properties, as shown in Figure 1.

Initially, the FOUR scale was translated from its Original Version (OV) into the target language, Brazilian Portuguese, with the support of two specialists in the English language, giving rise to Version in Portuguese, Translation 1 (VPT1) and to Version in Portuguese, Translation 2 (VPT2) of the scale proposed. The translated versions were compared, obtaining a consensus of the scale in Brazilian Portuguese, called Version in Portuguese, Consensus 1 (VPC1), which was forwarded for evaluation by an experts' committee, comprised by 15 professional nurses and/or physicians with more than five years of experience in emergency services, located through the Lattes Platform according to the pre-established framework[12].

The scale was evaluated, *a priori*, in view of the clarity and precision of its criteria[13]. In this way, the semantic, idiomatic, experimental and conceptual equivalences of the scale were evaluated, compiling the experts' suggestions to generate Version in Portuguese, Consensus 2 (VPC2).

The back-translation was performed based on VPC2 and forwarded to two translators, born in the USA and living in Brazil (where they work in the area of English language teaching), who were blinded to the objectives of the current study and to the Original Version (OV) of the scale, performing the back-translations individually, and having as outcomes Version in English, Translator 1 (VET1) and Version in English, Translator 2 (VET2) of the scale.

After the translations were completed, the researchers responsible for the current study met with both translators to present the purpose of the research, the original version of the scale and its objective for the health area. The two versions of the scale (VET1 and VET2) were compared, evaluated and, finally, the Final Version in English (FVE) was prepared.

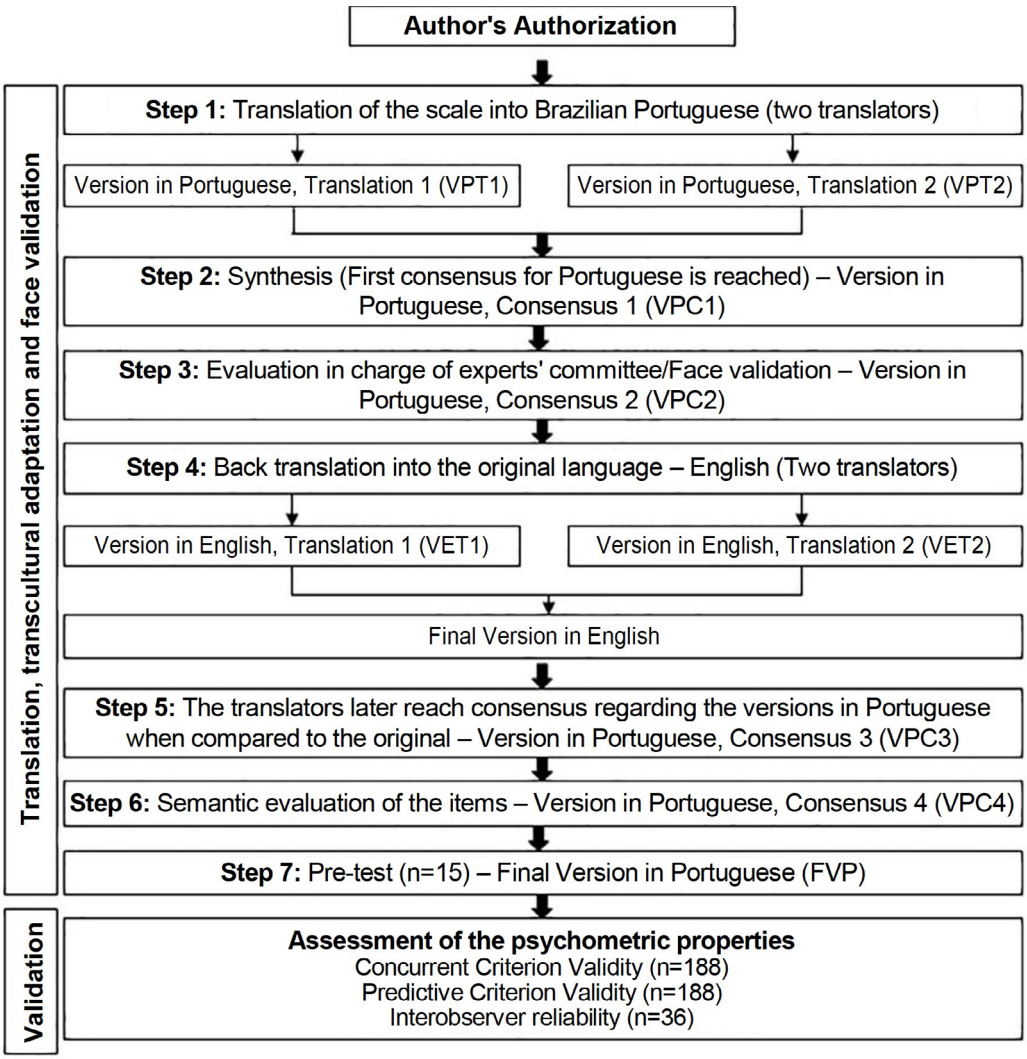


Figure 1 - Flowchart of the methodological path for the translation, cross-cultural adaptation and validation of the Full Outline of UnResponsiveness scale. Uberaba, MG, Brazil, 2021.

Subsequently, a copy of the original instrument (OV) was forwarded to each translator to compare it with the FVE, and from that point, each part of the scale was read and compared with the FVE, analyzing its equivalence and obtaining the new version in Portuguese, called Version in Portuguese, Consensus 3 (VPC3). Semantic evaluation was carried out through the translators' discussion, selecting the phrases expressed in the best ways, which comprised Version in Portuguese, Consensus 4 (VPC4). It is noted that, although the FVE was emailed to the main author in order to obtain his agreement, there was no response from him.

The data collection team consisted of two nurses, experts in urgency and emergency, which led us to select these professionals, as they were able to apply the instrument in the sector of choice, characterized by the PSA. They were theoretically and practically trained to use the instrument to characterize the subjects and the scales.

A pre-test was carried out with a sample of 15 participants referring to the target population, meeting the inclusion and exclusion criteria proposed to allow creating the scale's Final Version in Portuguese (FVP). For the evaluation of the metric properties, the FOUR scale (Brazilian version) was applied in order to verify predictive and concurrent criterion validity and interobserver reliability.

In the predictive criterion validity process, the association of the FOUR scale (Brazilian version) was tested with the death outcome, in which each patient was followed-up for 30 days, in order to

verify if the scale can predict death in the face of low scores. To verify the correlation degree between the FOUR scale (Brazilian version) and the GCS, the concurrent criterion validity procedure was performed, in which the scores of both scales were compared in order to obtain similar final scores. In turn, interobserver reliability was investigated by two nurses, applying the FOUR scale (Brazilian version) in critically-ill patients admitted to the PSA, as it has been validated in other contexts for this population[6–8], independently, and within a maximum interval of 10 minutes.

The findings were analyzed using the *MedCalc* statistical software for weighted Kappa and the *Statistical Package for the Social Sciences* (SPSS) statistical software, version 20.0 for Windows. At a first moment, face validation was performed, adopting the Content Validity Index (CVI) for analysis by item and for the entire construct, considering values above 0.80 or 80% as acceptable[13].

Subsequently, a univariate analysis of the findings was performed, including absolute and relative frequency distributions for categorical variables, and central tendency (mean, median) and variability (range of variation and standard deviation) measures for quantitative variables.

In the concurrent criterion validity analysis, *Spearman*'s Correlation Coefficient was used for the individual items and *Pearson*'s Correlation Coefficient for the total scores of the scale. For predictive validity, Cox Regression was used, sensitivity and specificity were analyzed by cross-tabulation and accuracy through the analysis of the Area Under the *Receiver Operating Characteristic* (ROC).

In the interobserver reliability analysis, specifically regarding the internal consistency of the scale items, *Cronbach*'s alpha coefficient was chosen, used to measure the correlation degree between items with values that varied between zero and one. The weighted *Kappa* coefficient for the individual items and the Intraclass Correlation Coefficient (ICC) for the reliability of the total scores of both observers were also used.

This research was approved by the Research Ethics Committee of HC-UFTM under opinion number 3,998,265 and adjusted in accordance with the provisions set forth in CONEP Resolution 466/12. It is noted that consent was obtained from the participants that agreed to participate in the research and who were in due physical and mental conditions to consent and sign the Free and Informed Consent Form (FICF). In the case of those who were unable to consent and sign the informed consent, authorization was asked to their guardians.

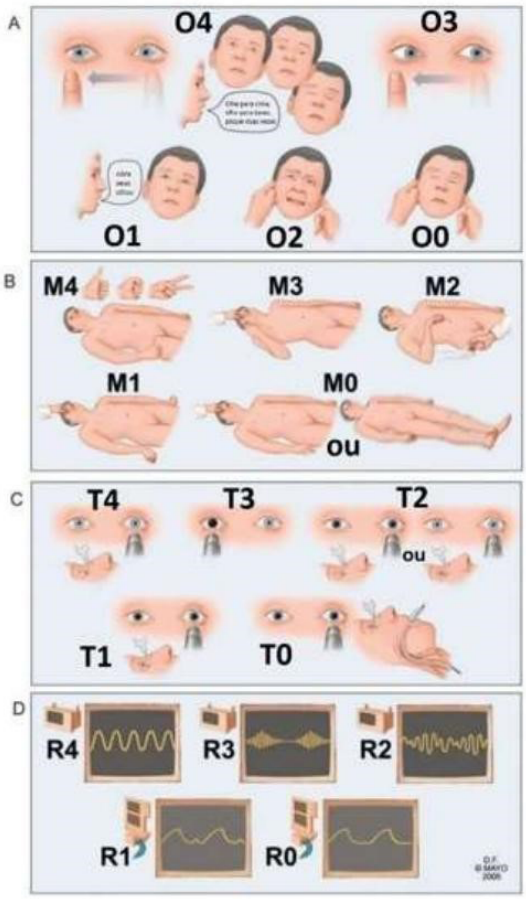
RESULTS

Through the synthesis process, VPC1 of the scale proposed was obtained and submitted to 15 (100%) experts. Of this total, 10 (66.70%) were female and 5 (33.30%) were male, with a mean age of 41.67 years old (SD=8.90; range of 28-62 years old). The majority, represented by 14 (93.30%) specialists, consisted of nurses with a mean time since graduation 19.20 years (SD=8.79; range of 7-41 years) and a mean of 14.93 years of experience in emergency services (SD=7.59; range of 5-34 years).

Based on the findings obtained in the cultural, semantic, conceptual and idiomatic equivalence and face validity assessments, the CVI per item was calculated, which varied from 93% to 100%, to later calculate the overall CVI of the scale, which resulted in 95%. The main changes were related to exchanging some words for their synonyms that best suited the Brazilian context, mainly changes in the domains configured by eye response and breathing pattern.

Subsequently, the pre-test was carried out with 15 (100%) patients: 4 (26.70%) of them female and 11 (73.30%) male, with a mean age of 66.7 years old (SD=15.6 and range of 38-90 years old). This phase did not result in changes in the scale, and the FVP was called *Escala Full Outline of UnResponsiveness - Versão para o Português Brasileiro* or *Escala FOUR - Versão Brasileira*, which is presented in Figure 2 below.

Escala *Full Outline of UnResponsiveness* – versão para o português brasileiro



Resposta ocular
O4 = Pálpebras abertas, olhos acompanham ou piscam ao comando verbal
O3 = Pálpebras abertas, mas não acompanha com o olhar
O2 = Pálpebras fechadas, mas abre ao ouvir voz alta
O1 = Pálpebras fechadas, mas abre em resposta ao estimulo doloroso
O0 = Pálpebras permanecem fechadas ao estímulo doloroso

Resposta motora
M4 = Levanta o polegar ou o punho ou sinal da paz
M3 = Localiza a dor
M2 = Flexão em resposta ao estímulo doloroso
M1 = Extensão em resposta ao estímulo doloroso
M0 = Sem resposta ao estímulo doloroso ou mioclonias generalizadas

Reflexos tronco encefálicos
T4 = Reflexos pupilares e corneanos presentes
T3 = Uma pupila dilatada e fixa
T2 = Reflexos pupilares ou corneanos ausentes
T1 = Reflexos pupilares e corneanos ausentes
T0 = Reflexos pupilares, corneanos e de tosse ausentes

Padrão respiratório
R4 = Não intubado, padrão respiratório regular
R3 = Não intubado, padrão respiratório de *Cheyne-Stokes*
R2 = Não intubado, padrão respiração irregular
R1 = Respira acima da frequência do ventilador
R0 = Apneia ou respira na frequência do ventilador

Instruções: (A) Resposta ocular (O): Uma pontuação de O4 indica pelo menos três excursões voluntárias. Se as pálpebras estiverem fechadas, o examinador deve abri-las e realizar o exame mediante movimentação com um dedo ou objeto. A avaliação com a abertura de uma pálpebra será suficiente em casos de edema palpebral ou trauma facial. Uma pontuação de O3 indica a ausência de movimentação voluntária com os olhos abertos. Uma pontuação de O2 indica pálpebras se abrindo mediante voz alta. Uma pontuação de O1 indica pálpebras abertas com estímulo de dor. Uma pontuação de O0 indica ausência de abertura palpebral mesmo com dor. (B) Resposta motora (M): Uma pontuação de M4 indica que o paciente evidenciou pelo menos uma das três posições da mão (polegar para cima, punho ou sinal de paz) com ambas as mãos. Uma pontuação de M3 (localização) indica que o paciente tocou a mão do examinador após um estímulo doloroso comprimindo a articulação temporomandibular ou nervo supraorbital. Uma pontuação de M2 indica qualquer movimento de flexão dos membros superiores. Uma pontuação de M1 indica resposta extensora à dor. Uma pontuação de M0 indica que não há resposta motora à dor. (C) Reflexos do tronco encefálico (T): Examine os reflexos das pupilas e das córneas. De preferência, os reflexos das córneas são testados incutindo entre duas e três gotas de solução salina estéril na córnea a uma distância de 10 a 15 cm (isso minimiza o trauma da córnea resultante de repetidas análises). Também podem ser utilizados cotonetes esterilizados. O reflexo da tosse para a aspiração traqueal é testado somente quando ambos os reflexos estão ausentes. Uma pontuação de T4 indica a presença de reflexos das pupilas e das córneas. Uma pontuação de T3 indica uma pupila dilatada e fixa. Uma pontuação de T2 indica que os reflexos das pupilas ou os reflexos das córneas estão ausentes. Uma pontuação de T1 indica que tanto os reflexos das pupilas como os das córneas estão ausentes. Uma pontuação de T0 indica ausência dos reflexos das pupilas e das córneas, assim como do reflexo da tosse (usando aspiração traqueal). (D) Padrão respiratório (R): determine o padrão de respiração espontânea em um paciente sem intubação e classifique simplesmente como regular (R4) ou irregular (R2), ou ainda respiração de *Cheyne-Stokes* (R3). Em pacientes submetidos à ventilação mecânica, avalie o formato da onda de pressão do padrão de respiração espontânea ou o acionamento do ventilador pelo paciente (R1). O monitor do ventilador que exibe padrões respiratórios pode ser usado para identificar os ciclos respiratórios gerados pelo paciente no ventilador. Não se devem fazer quaisquer ajustes no ventilador enquanto o paciente estiver sendo classificado, porém a classificação é feita preferencialmente com $PaCO_2$ dentro dos limites normais. Um teste padrão de apneia pode ser necessário quando o paciente respira na frequência do ventilador (R0).

Figure 2 - Full Outline of UnResponsiveness Scale (Brazilian Version). Uberaba, MG, Brazil, 2021.

The FOUR scale (Brazilian version) was applied to 188 (100%) patients, of which 112 (59.6%) were male and 76 (40.4%) were female, with a mean age of 63.26 years old (SD=14.77, range of 19-97 years old). The majority, represented by 80 (42.60%) patients, received a clinical diagnosis of Acute Myocardial Infarction (AMI), followed by stroke with 19 (10.1%); 152 (80.00%) were hemodynamically compensated patients and in use of other forms of cardiovascular support (130, 69.10%) as presented in Table 1.

Table 1 - Distribution of the study participants' clinical variables referring to the assessment of the scale's psychometric properties. Uberaba, MG, Brazil, 2021. (n=188)

Variable	N	%
Type of problem according to the diagnosis		
Cardiovascular disorders	118	62,76%
Neurological disorders	36	19,14%
Pulmonary or respiratory disorders	9	4,78%
Sepsis and infection	8	4,25%
Hepatic disorders	5	2,65%
Fractures	3	1,59%
Neoplasms	3	1,59%
Abdominal disorders	2	1,59%
Renal problems	2	1,06%
Hematological disorders	1	0,53%
External problems or violence	1	0,53%
Hemodynamic state		
Hemodynamically unstable	36	19,10%
Hemodynamically compensated	152	80,90%
Drugs		
Use of vasoactive drugs		
Yes	111	59,00%
No	77	41,00%
Use of other forms of cardiovascular support		
Yes	130	69,10%
No	58	30,90%

To verify concurrent criterion validity, the FOUR scale scores were compared to the GCS score, noticing that the higher the score, the better the patient's response and the lower the chances of evolving to death, and that the lower the score, the worse the response.

Spearman's Correlation Coefficient identified a strong correlation in the eye and motor responses[14], as presented in Table 2.

Table 2 - Presentation of *Spearman*'s correlation (r_s) for the eye response and motor response individual items. Uberaba, MG, Brazil, 2021. (n=188)

Itenm	r_s	p
Eye response	0,96	<0,001
Motor response	0,81	<0,001

The brainstem reflex and breathing pattern items were not calculated due to the absence of similar criteria in the GCS. (r)=0.97 (p<0.001) was obtained for the total scores, which also characterized a strong correlation[14].

As for predictive criterion validity, the association of the FOUR scale (Brazilian version) with the death outcome was tested, encompassing 31 (16.5%) patients. The relative risk of predicting death was 0.78 (p<0.001), indicating that, for each additional point on the scale, the risk of death is reduced by nearly 20%. The relative risk considered for the current study was less than 1, representing a protective factor as the scale score increases.

For sensitivity and specificity, cutoff points pre-defined by a validation study of the FOUR scale (Chinese version)[8] were used, which considered the score of 13 points to predict death. Given this evidence, 95.5% specificity was obtained for scores equal to or greater than 13 points, as well as 51.6% sensitivity for scores up to 12 points in the current study.

In terms of accuracy, values considered high (AUC > 0.8) were evidenced, indicating the instrument's good precision[15]. The area under the ROC curve between the scale and the death outcome was 0.80 (95% CI: 0.688–0.905, p<0.001), that is, in 80% of the times when the FOUR scale (Brazilian version) was used, it is possible to discriminate between true positives and true negatives, and it will present false results 20% of the times.

In internal consistency (n=188), the *Cronbach*'s alpha value for the four items that comprise the FOUR scale (Brazilian version) was 0.94, configuring very high consistency between the items[16]. The interobserver reliability analysis (n=36) was performed at two moments: initially, the individual items and the significance level were calculated for each item of the 4 domains in the instrument. Afterwards, the final score of the FOUR scale (Brazilian version) obtained by an observer were considered.

For reliability of the items, the eye response, motor response and brainstem reflexes were analyzed, obtaining a value of k=1.0; in turn, the breathing pattern resulted in k=0.89, classified as almost perfect reliability[17]. The instrument's overall agreement scores resulted in ICC=0.99, showing excellent reliability[18].

DISCUSSION

Assessment of the level of consciousness is part of the care for critically-ill patients and requires the adoption of scales capable of supporting adequate care and therapy for each case[19]. In this context, the current study makes science in Health and Nursing unique by translating, adapting and validating the Full Outline of UnResponsiveness scale into Brazilian Portuguese, capable of identifying and evaluating level of consciousness in adult patients[6] more comprehensively than the existing scales, mainly because it considers parameters that ensure accuracy in the assessment, such as breathing pattern and brainstem reflex, as well as it allows analyzing the verbal pattern of patients in use of endotracheal devices, criteria not addressed by the instruments commonly used for this purpose.

A rigorous methodological path was followed to proceed with the translation, adaptation and validation of the FOUR scale into Brazilian Portuguese;[10–11] this procedure has been adopted in other health realities and scenarios due to its reliability for clinical practice.

The results identified by the face validation of the current research corroborate a Brazilian study carried out in Rio Grande do Norte, which proposed the validation of a Nursing care protocol aimed at septic patients, obtaining an overall CVI of the construct of almost perfect agreement, a result that helps to consider its suitability for the clinical practice[20].

Another study, carried out at *Universidade Federal de Santa Catarina*, considered a CVI above 90% in the process of adapting and validating the *Patient Measure of Safety* Questionnaire into Brazilian Portuguese[21].

As it is considered a clinimetric scale, characterized as a construct with easy interpretation of its clinical phenomena containing objective, direct and clear variables and little sensitive to changes, the FOUR scale did not require profound changes[22].

To assess the psychometric properties of a scale, it is necessary to apply it to a target population, characterized in the current study by adult, critically-ill and hospitalized patients, mainly due to AMI and stroke. This context is corroborated by a validation study of the FOUR scale for the Spanish version, which mostly included stroke victims[23], and by an Australian survey, which compared the GCS to the FOUR scale, based on a sample in which critically-ill patients who were victims of AMI and stroke prevailed[24].

A strong correlation was identified regarding concurrent criterion validity for items of the FOUR scale (Brazilian version), a condition similar to other validation studies of the FOUR scale that also presented a strong correlation when comparing the GCS and FOUR scales[7–8,23,25].

The strong correlation evidenced in the total scores of the FOUR scale (Brazilian version) was also found in a study carried out in an ICU of a hospital in Arizona, United States, which compared the FOUR scale to the GCS and showed satisfactory psychometric properties for its adoption in the clinical practice[26].

From the perspective of predictive criterion validity, the FOUR scale (Brazilian version) was associated with the death outcome, identifying that among the 188 (100%) patients, 31 (16.5%) evolved to death and presented a higher risk for this outcome the lower the scale score. A study developed with 359 (100%) patients in Uganda, Africa, was similar to this context, and compared the predictive power of the GCS with the FOUR scale, evidencing death in 144 (40.1%) patients with a risk 2.64 times higher the lower the score obtained in the FOUR scale[27].

Assessing the accuracy, sensitivity and specificity of a scale is useful to determine its performance when applied to the target population[28]. The Brazilian Portuguese version of the FOUR scale presented good accuracy, a condition that is similar to a validation study of the FOUR scale for the Chinese language, which obtained an AUC of 0.834 (95% CI: 0.740–0.928), considered good[8].

As the population addressed for validation of the FOUR scale for Brazilian Portuguese differed from the population profile identified in other studies[7–8], when assessing the sensitivity and specificity of the scale proposed, the decision was to adopt the cutoff point defined by a Chinese study carried out with 120 critically-ill patients with clinical diagnoses similar to those of the current study, which established a score of 13 (79% sensitivity; 72% specificity) to predict death[8]. When establishing a comparison between this and the baseline study, it was detected that 22% of the patients evolved to death[8] and, in view of the above, it is possible to evidence a smaller number of patients who obtained a score below 13 points.

Also regarding the cutoff point established, a review about the historical context of the FOUR scale is similar to the current research by highlighting that the score to determine the risk of in-hospital mortality is calculated based on a sum of 12 points, which is compared to score 8 of the GCS. It is noted that the main author of the scale discourages the use of the sum of scores due to the fact that a one-point decrease in any component has significant clinical relevance[29].

Determining reliability is essential to verify the ability of a scale to reproduce a result consistently in time and space by means of different observers, indicating the quality of the construct; and one of the ways to measure this criterion is through internal consistency, as it indicates how homogeneous the items of a scale are and how they are correlated with each other[13].

The high internal consistency presented by the Brazilian version of the FOUR scale was also identified in other studies that proposed to translate, adapt and validate the aforementioned instrument[25,30], as well as the assessment of the items' interobserver reliability, characterized as perfect

agreement and corroborated by a validation survey of the FOUR scale in an Australian emergency department;[24] and the overall reliability of the ICC, which was similar to a validation study of the FOUR scale for the Italian language[25].

Other validation studies of the FOUR scale in different languages presented excellent reliability of the scale when comparing the score generated by observers 1 and 2[8,23].

The findings presented in the current study are relevant to the Health and Nursing areas because they provide a useful scale, easy to remember and simple to apply for the national scenario, which evidences important information about the level of consciousness and prognosis of adult patients in serious conditions and in a hospital environment.

The main limitation of this study was the fact that the data were collected in a single teaching institution. In view of this, it is suggested to carry out multicenter research studies on the theme in order to facilitate generalization of the findings.

CONCLUSION

This study makes available the FOUR scale (Brazilian version) for its use in Brazil, maintaining the 20 items and four domains of the original scale. The analyses adopted to verify predictive criterion validity indicated that the scale is capable of predicting undesirable outcomes such as death. There was a strong correlation between the GCS and the FOUR scale (Brazilian version). Reliability presented satisfactory indices, proving to be a valid, reliable and useful scale to assess level of consciousness in adult patients.

It is suggested that further research studies be carried out on the FOUR scale (Brazilian version), in order to foster scientific deepening of this theme and to determine a cutoff point for the scale, capable of more precisely indicating the necessary course of action in the face of changes in the level of awareness of adult patients.

REFERENCES

1. Halvachizadeh S, Baradaran L, Cinelli P, Pfeifer R, Sprengel K, Pape H-C. How to detect a polytrauma patient at risk of complications: a validation and database analysis of four published scales. PLoS One [Internet]. 2020 [cited 2021 Oct 9];15(1):e0228082. Available from: https://doi.org/10.1371/journal.pone.0228082
2. Gaspar A, Azevedo P, Roncon-Albuquerque R Jr. Non-invasive hemodynamic evaluation by Doppler echocardiography. Rev Bras Ter Intensiva [Internet]. 2018 [cited 2021 Oct 9]; 30(3):385-93. Available from: https://doi.org/10.5935/0103-507X.20180055
3. Oliveira DMP, Pereira CU, Freitas ZMP. Neurological evaluation about nursing knowledge of the patient with traumatic brain injury. Rev Enferm UFPE On Line [Internet]. 2016 [cited 2021 Oct 9];10(5):4249-54. Available from: https://doi.org/10.5205/reuol.9284-81146-1-SM.1005sup201611
4. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet [Internet]. 1974 [cited 2021 Oct 9];2(7872):81-4. Available from: https://doi.org/10.1016/s0140-6736(74)91639-0
5. Brennan PM, Murray GD, Teasdale GM. Simplifying the use of prognostic information in traumatic brain injury. Part 1: The GCS-Pupils score: an extended index of clinical severity. J Neurosurg [Internet]. 2018 [cited 2021 Oct 8]; 128(6):1612-20. Available from: https://doi.org/10.3171/2017.12.JNS172780
6. Wijdicks EFM, Bamlet WR, Maramattom BV, Manno EM, Mcclelland RL. Validation of a new coma scale: the FOUR score. Ann Neurol [Internet]. 2005 [cited 2021 Oct 9];58(4):585-93. Available from: https://doi.org/10.1002/ana.20611

7. Momenyan S, Mousavi SM, Dadkhahtehrani T, Sarvi F, Heidarifar R, Kabiri F, et al. Predictive validity, and inter-rater reliability of the Persian version of full outline of unresponsiveness among unconscious patients with traumatic brain injury in an intensive care unit. Neurocrit Care [Internet]. 2017 [cited 2021 Oct 9];27(2):229-36. Available from: https://doi.org/10.1007/s12028-016-0324-0

8. Peng J, Deng Y, Chen F, Zhang X, Wang X, Zhou Y, et al. Validation of the Chinese version of the FOUR score in the assessment of neurosurgical patients with different level of consciousness. BMC Neurol [Internet]. 2015 [cited 2021 Oct 10];15:254. Available from: https://doi.org/10.1186/s12883-015-0508-9

9. Suresh V, Yaddanapudi LN, Podder S. Full Outline of UnResponsiveness score versus Glasgow Coma Scale in critically ill patients with altered sensorium: a comparison of inter-observer variability and outcomes. Indian J Anaesth [Internet]. 2019 [cited 2021 Oct 9];63(8):640-7. Available from: https://doi.org/10.4103/ija.IJA_377_19

10. Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment. Health Policy [Internet]. 1991 [cited 2022 Feb 7];19(1):33-44. Available from: https://doi.org/10.1016/0168-8510(91)90072-6

11. Ferrer M, Alonso J, Prieto L, Plaza V, Monsó E, Marrades R, et al. Validity and reability of the St Geroge's Respiratory Questionaire after adaptation to a different language and culture: the spanish example. Eur Respir J [Internet]. 1996 [cited 2021 Oct 9];9(6):1160-6. Available from: https://doi.org/10.1183/09031936.96.09061160

12. Fehring RJ. The Fehring model. In: Carrol-Jonhnson RM, Paquete M, editors. Classification of nursing diagnoses: proceedings of the Tenth Conference. Philadelphia, PA(US): J.B. Limppincott; 1994. p. 55-62.

13. Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Epidemiol Serv Saúde [Internet]. 2017 [cited 2021 Oct 9];26(3):649-59. Available from: https://doi.org/10.5123/s1679-49742017000300022

14. Cohen, J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ(US): Erlbaum; 1988. 579 p.

15. Krohling LL, Paula KMP, Behlau M. ROC curve of the Pediatric Voice Related Quality-of-Life Survey (P-VRQOL). Codas [Internet]. 2016 [cited 2021 Oct 9];28(3):311-3. Available from: https://doi.org/10.1590/2317-1782/20162015103

16. Cronbach LJ, Schönemann P, Mckie D. Alpha coefficients for stratified parallel tests. Educ Psychol Meas [Internet]. 1965 [cited 2021 Oct 9];25(2):291-312. Available from: https://doi.org/10.1177/001316446502500201

17. Landis JR, Koch GG. The measurement of observer agrément for categorical data. Biometrics. 1977;33(1):159-74.

18. Fleiss JL. The design and analysis of clinical experiments. New York, NY(US): Wiley; 1999. 448 p.

19. Barros WCTS, Dal Sasso GTMS, Alvarez AG, Ramos SF, Martins SR. App to evaluate the level of consciousness in adults: technological production in nursing. Cogitare Enferm [Internet]. 2019 [cited 2021 Oct 3];24:e60338. Available from: https://doi.org/10.5380/ce.v24i0.60338

20. Pedrosa KKA, Oliveira AS, Machado RC. Validation of a care protocol for the septic patient in the Intensive Care Unit. Rev Bras Enferm [Internet]. 2018 [cited 2021 Oct 9];71(3):1106-14. Available from: https://doi.org/10.1590/0034-7167-2017-0312

21. Mello JF, Barbosa SFF. National Early Warning Score 2: transcultural adaptation to Brazilian Portuguese. Texto Contexto Enferm [Internet]. 2021 [cited 2021 Oct 9];30:e20180322. Available from: https://doi.org/10.1590/1980-265X-TCE-2018-0322

22. Pais-Ribeiro JL. Medida na avaliação psicológica. Psic, Saúde Doenças [Internet]. 2013 [cited 2021 Oct 9];14(1):245-63. Available from: http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1645-00862013000100016&lng=pt&nrm=iso

23. Idrovo L, Fuentes B, Medina J, Gabaldón L, Ruiz-Ares G, Abenza MJ, et al. Validation of the FOUR Score (Spanish Version) in acute stroke: an interobserver variability study. Eur Neurol [Internet]. 2010 [cited 2021 Oct 9];63(6):364-9. Available from: https://doi.org/10.1159/000292498

24. Kevric J, Jelinek GA, Knott J, Weiland TJ. Validation of the Full Outline of Unresponsiveness (FOUR) Scale for conscious state in the emergency department: comparison against the Glasgow Coma Scale. Emerg Med J [Internet]. 2011 [cited 2021 Oct 5];28(6):486-90. Available from: https://doi.org/10.1136/emj.2009.085845

25. Marcati E, Ricci S, Casalena A, Toni D, Carolei A, Sacco S. Validation of the Italian version of a new coma scale: the FOUR score. Intern Emerg Med [Internet]. 2012 [cited 2021 Oct 9];7(2):145-52. Available from: https://doi.org/10.1007/s11739-011-0583-x

26. Adcock AK, Kosiorek H, Parikh P, Chauncey A, Wu Q, Demaerschalk BM. Reliability of robotic telemedicine for Assessing Critically Ill Patients with the Full Outline of UnResponsiveness Score and Glasgow Coma Scale. Telemed J E Health [Internet]. 2017 [cited 2021 Oct 9];23(7):555-60. Available from: https://doi.org/10.1089/tmj.2016.0225

27. Abdallah A, Demaerschalk BM, Kimweri D, Aden AA, Butterfiel R, Assimwe SB, et al. A comparison of the Full Outline of Unresponsiveness (FOUR) and Glasgow Coma Scale (GCS) scores in predicting mortality among patients with reduced level of consciousness in Uganda. Neurocrit Care [Internet]. 2020 [cited 2021 Oct 7];32(3):734-41. Available from: https://doi.org/10.1007/s12028-019-00806-4

28. Patino CM, Ferreira JC. Entendendo os testes diagnósticos: parte 2. J Bras Pneumol [Internet]. 2017 [cited 2021 Oct 9];43(6):408. Available from: https://doi.org/10.1590/S1806-37562017000000424

29. Wijdicks EFM. Clinical scales for comatose patients: the Glasgow Coma Scale in historical context and the new FOUR Score. Rev Neurol Dis. 2006;3(3):109-17.

30. Wolf CA, Wijdicks EFM, Bamlet WR, Mcclelland RL. Further validation of the FOUR score coma scale by intensive care nurses. Mayo Clin Proc [Internet]. 2007 [cited 2021 Oct 29];82(4):435-8. Available from: https://doi.org/10.4065/82.4.435

NOTES

ORIGIN OF THE ARTICLE

Extracted from the dissertation -Translation, Cultural Adaptation and Validation of *the Full Outline of UnResponsiveness* Scale into Brazilian Portuguese, presented to the Graduate Program in Health Care of *Universidade Federal do Triângulo Mineiro*. in 2021

CONTRIBUTION OF AUTHORITY

Study design: Bernardinelli FCP, Chavaglia SRR.
Data collection: Bernardinelli FCP, Amorim GC.
Data analysis and interpretation: Bernardinelli FCP, Haas VJ, Chavaglia SRR.
Discussion of the results: Bernardinelli FCP, Chavaglia SRR.
Writing and/or critical review of the content: Bernardinelli FCP, Chavaglia SRR, Haas VJ, Campanharo CRV, Barbosa MH, Amorim GC.
Review and final approval of the final version: Bernardinelli FCP, Chavaglia SRR, Haas VJ, Campanharo CRV, Barbosa MH, Amorim GC.

FUNDING INFORMATION

This paper was carried out with the support of *Coordenação de Aperfeiçoamento de Nível Superior* - Brazil (CAPES) - Funding Code 001.

APPROVAL OF ETHICS COMMITTEE IN RESEARCH

Approved by the Research Ethics Committee of the Clinical Hospital of *Universidade Federal do Triângulo Mineiro*, opinion No.3,998,265 and Certificate of Presentation for Ethical Appreciation No. 27172719.0.0000.8667.

CONFLICT OF INTEREST

There is no conflict of interest.

HISTORICAL

Received: November 25, 2021.
Approved: March 10, 2022.

CORRESPONDING AUTHOR

Suzel Regina Ribeiro Chavaglia
suzel.ribeiro@yahoo.com.br