







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

Fabiana Cristina Pires Bernardinelli¹ 
Gustavo Correa de Amorim² 
Vanderlei José Haas¹ 
Cássia Regina Vancini Campanharo³ 
Maria Helena Barbosa¹ 
Suzel Regina Ribeiro Chavaglia¹ 

¹Universidade Federal do Triângulo Mineiro, Programa de Pós-graduação em Atenção à Saúde. Uberaba, Minas Gerais, Brasil.

²Universidade Federal do Triângulo Mineiro, Programa de Residência Enfermagem em Urgência/Trauma. Uberaba, Minas Gerais, Brasil.

³Universidade Federal de São Paulo, Escola Paulista de Enfermagem, Programa de Pós-graduação em Enfermagem. São Paulo, São Paulo, 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.

COMO CITAR: Bernardinelli FCP, Amorim GC, Haas VJ, Campanharo CRV, Barbosa MH, Chavaglia SRR. Tradução, adaptação e validação da escala *Full Outline of UnResponsiveness* para o português do Brasil. *Texto Contexto Enferm [Internet]*. 2022 [acesso MÊS ANO DIA]; 31:e20210427. Disponível em: <https://doi.org/10.1590/1980-265X-TCE-2021-0427pt>

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

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.

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.

INTRODUÇÃO

A crescente necessidade de obter-se uma assistência à saúde de qualidade vem instigando pesquisadores, em todo o mundo, a desenvolver instrumentos capazes de avaliar o nível de consciência de pacientes graves, a fim de identificar e monitorar suas condições clínicas com precisão e embasar o julgamento clínico dos profissionais em sólidas evidências científicas¹.

Considera-se um paciente em estado grave quando há risco iminente de perder a vida ou deterioração fisiológica de órgãos e sistemas, decorrente de trauma ou outras doenças, ainda, pela presença de um estado hemodinâmico instável, possibilidade de choque circulatório, ou compensação hemodinâmica por meio de doses crescentes e altas de drogas vasoativas ou outras formas de suporte cardiovascular².

A avaliação do nível de consciência de pacientes graves se dá pela interpretação das alterações da sua função cerebral, geralmente, identificada pela Escala de Coma de Glasgow (ECG), publicada em 1974 e atualizada no ano de 2018, para acompanhar a evolução do nível de consciência e auxiliar na obtenção de prognóstico clínico dos pacientes³⁻⁴.

Sobre a ECG, é necessário ressaltar que, mesmo sendo comumente adotada na prática clínica em um contexto global, trata-se de um instrumento ainda não validado para o português do Brasil, que apresenta limitações diagnósticas, como a impossibilidade da avaliação verbal em pacientes sob uso de Ventilação Mecânica Invasiva (VMI) e a ausência de testes de reflexos de tronco encefálico, fatores que podem comprometer a precisão da avaliação neurológica⁵.

Na intencionalidade de suprir as lacunas existentes na avaliação da consciência obtida pela ECG, desenvolveu-se em 2005, na cidade de Rochester - Minnesota, situada nos Estados Unidos da América, uma escala denominada *Full Outline of UnResponsiveness* (FOUR), validada em vários contextos clínicos⁶⁻⁸ e composta de quatro domínios: resposta ocular, resposta motora, reflexos de tronco cerebral e respiração, cada um constituído por cinco itens, classificados de zero a quatro pontos⁶.

Dentre as vantagens apresentadas pela escala FOUR, destacam-se o simples manejo e interpretação dos resultados, além de fácil memorização de seus itens, somada à possibilidade de avaliar variáveis clínicas importantes, como o reflexo tronco encefálico e o padrão respiratório do paciente, mesmo na presença de cânulas endotraqueais, e sobrepor em amplitude e qualidade da avaliação do nível de consciência quando comparada a outras escalas existentes para o mesmo fim, sustentando a sua adoção⁹. Considera-se, no entanto, que mesmo em face de sua acurácia, estudos científicos que tratam da efetividade dessa escala para avaliar a consciência de pacientes graves no âmbito internacional são incipientes e inexistentes no contexto nacional, visto que esta não foi validada e traduzida para o português do Brasil⁹.

Diante da importância de obterem-se instrumentos validados para o português do Brasil que propiciem acurácia à avaliação do nível de consciência de pacientes adultos em estados clínicos de gravidade, este estudo teve como objetivo traduzir, adaptar culturalmente e validar a escala *Full Outline of UnResponsiveness* para o português do Brasil.

MÉTODO

Trata-se de um estudo metodológico¹⁰ de tradução, adaptação cultural e validação da escala FOUR para o contexto brasileiro, que foi fundamentada no método desenvolvido por autores internacionais¹¹.

A escala FOUR é composta de quatro domínios caracterizados pela resposta ocular, resposta motora, reflexos de tronco cerebral e respiração e constituídos por cinco opções de resposta cada um, que variam de zero a quatro pontos. Os valores obtidos nas respostas podem flutuar entre zero a 16 pontos, e o escore 16 representa o mais alto nível de consciência do paciente⁶. Cabe ressaltar que, além dessa escala, utilizou-se na coleta de dados a ECG, composta de 15 itens distribuídos em três domínios: resposta ocular, resposta verbal e resposta motora⁴.

Os dados foram coletados no período de agosto a dezembro de 2020, em um Pronto Socorro Adulto (PSA) de um hospital público denominado Hospital de Clínicas da Universidade Federal do Triângulo Mineiro (HC-UFTM), Uberaba, Brasil, que abrange 302 leitos, destes, 32 para atendimentos emergenciais de pacientes adultos.

A população-alvo foi composta de pacientes graves² admitidos no PSA de acordo com os seguintes critérios de inclusão: pacientes com idade igual ou superior a 18 anos, hemodinamicamente instáveis ou compensados hemodinamicamente com uso de drogas vasoativas ou outras formas de suporte cardiovascular. Excluíram-se pacientes em uso de medicações sedativas, visto a impossibilidade de aplicação da ECG nesta condição.

Adotou-se uma amostra do tipo não probabilística e sequencial durante o período de coleta de dados pela impossibilidade de sortear a amostra. Para a seleção do tamanho amostral, utilizou-se a ferramenta *Power Analysis and Sample Size* - versão 13, e considerou-se um Coeficiente de Correlação Intraclasse (CCI) entre os escores de adesão esperados de CCI=0,9 e entre os escores de nível de consciência, admitindo-se um valor mínimo de CCI=0,75 para um poder apriorístico de 90% e obtendo-se um tamanho amostral mínimo de 36 pacientes para confiabilidade interobservador. Para a validade de critério concorrente e preditiva, considerou-se uma incidência de óbitos de pacientes graves de 24,6%, uma precisão de 4,5% e um intervalo de confiança de 95% para uma população finita de 400 internações ao ano, chegando-se a uma amostra mínima de 188 indivíduos. Considerou-se, ainda, um nível de significância $\alpha = 0,05$.

O processo de tradução, adaptação cultural e validação ocorreu após autorização do autor principal da escala, cumprindo-se os seguintes passos: 1) tradução da escala para a língua portuguesa brasileira; 2) síntese - obtenção do primeiro consenso para o português; 3) avaliação pelo comitê de especialistas; 4) retrotradução para a língua original; 5) obtenção posterior pelos tradutores de um consenso das versões em português em comparação com a original; 6) realização da avaliação semântica dos itens e 7) pré-teste¹¹. Ressalta-se que a participação dos pacientes ocorreu no pré-teste e no processo de validação por meio do teste das propriedades psicométricas conforme apresentado na Figura 1.

Inicialmente, realizou-se a tradução da escala FOUR em sua Versão Original (VO) para o idioma-alvo, o português do Brasil com o apoio de dois especialistas na língua inglesa, originando a Versão Português Tradução 1 (VPT1) e Versão Português Tradução 2 (VPT2) da escala proposta. As versões traduzidas foram comparadas, obtendo-se um consenso da escala em português do Brasil, nomeada Versão Português Consenso 1 (VPC1), a qual foi encaminhada para avaliação por um comitê de especialistas, composto de 15 profissionais enfermeiros e/ou médicos com tempo de experiência em pronto-socorro superior a cinco anos, localizados por meio da Plataforma *Lattes*, de acordo com o referencial pré-estabelecido¹².

A escala foi avaliada, *a priori*, em face da clareza e precisão de seus critérios¹³. Desta forma, avaliou-se a equivalência semântica, idiomática, experimental e conceitual da escala, compilando-se as sugestões dos especialistas para gerar a Versão Português Consenso 2 (VPC2).

A retrotradução foi realizada com base na VPC2 e encaminhada a dois tradutores, nascidos nos EUA e residentes no Brasil (local em que desempenham atividade de ensino da língua inglesa), cegados para os objetivos do presente estudo e para a Versão Original (VO) da escala, realizando-se as retrotraduções de forma individual, e tendo como desfechos a Versão Inglês Tradutor 1 (VIT1) e Versão Inglês Tradutor 2 (VIT2) da escala.

Com o término das traduções, as pesquisadoras responsáveis pelo presente estudo se reuniram com os dois tradutores para apresentar a intencionalidade da pesquisa, a versão original da escala e a sua finalidade para a área da saúde. As duas versões da escala (VI1 e VI2) foram comparadas, avaliadas e, por fim, elaborou-se a Versão Inglês Final (VIF).

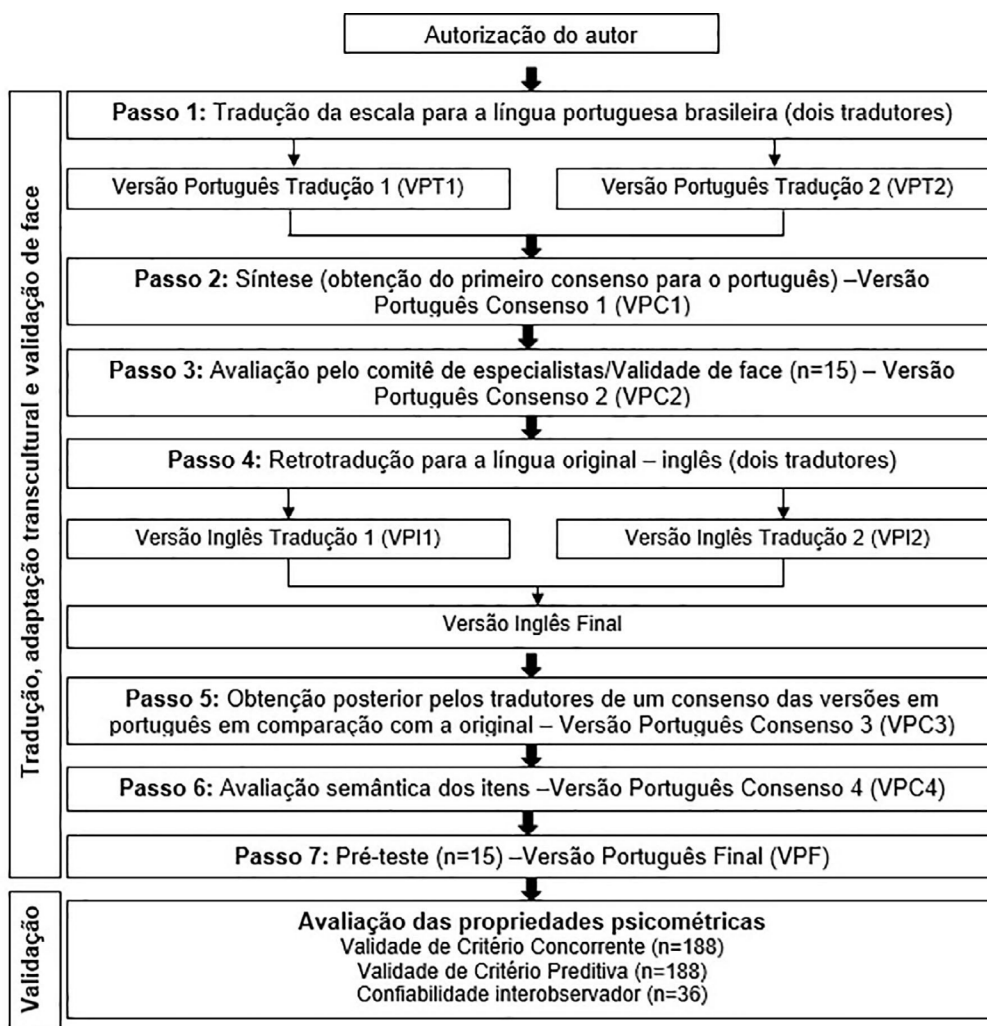


Figura 1 - Fluxograma da trajetória metodológica de tradução, adaptação transcultural e validação da escala *Full Outline of UnResponsiveness*. Uberaba, MG, Brasil, 2021.

Na sequência, encaminhou-se a cada um dos tradutores a cópia do instrumento original (VO) para a comparação com a VIF, e a partir de então, cada parte da escala foi lida e comparada com a VIF, analisando-se a sua equivalência, e obtendo-se a nova versão em português, denominada Versão Português Consenso 3 (VPC3). A avaliação semântica foi realizada por meio da discussão dos tradutores, selecionando-se as frases de melhor expressão, que compuseram a Versão Português Consenso 4 (VPC4). Salienta-se que, apesar de a VIF ter sido enviada por *e-mail* ao autor principal, a fim de obter a sua concordância, não houve resposta do autor.

A equipe de coleta de dados foi composta de dois enfermeiros, especialistas em urgência e emergência, o que nos levou a optar por estes profissionais, visto que estes estavam aptos para aplicar o instrumento no setor de escolha, caracterizado pelo PSA. Eles foram capacitados de forma teórica e prática para utilizar o instrumento de caracterização dos sujeitos e as escalas.

Realizou-se um pré-teste com uma amostra de 15 participantes referentes à população-alvo, cumprindo-se os critérios de inclusão e exclusão propostos para viabilizar a Versão Português Final (VPF) da escala. Para a avaliação das propriedades métricas, aplicou-se a escala FOUR - versão brasileira a fim de verificar a validade de critério preditiva, concorrente e a confiabilidade interobservador.

Na validade de critério preditiva testou-se a associação da escala FOUR – versão brasileira, com o desfecho óbito, no qual cada paciente foi acompanhado por 30 dias, a fim de verificar se a escala é capaz de prever o óbito diante de escores baixos. Para verificar o grau de correlação entre a escala

FOUR – versão brasileira e a ECG, realizou-se a validade de critério concorrente, em que foram comparados os escores de ambas as escalas com o intuito de obter escores finais semelhantes. Já, a confiabilidade interobservador foi investigada por dois enfermeiros, aplicando-se a escala FOUR - versão brasileira em pacientes graves admitidos no PSA, visto que ela foi validada em outros contextos para esta população⁶⁻⁸, de maneira independente, em um intervalo máximo de 10 minutos.

Analisaram-se os achados, utilizando-se o *MedCalc Statistical Software* para o Kappa ponderado e o *software* estatístico *Statistical Package for the Social Sciences* (SPSS) versão 20.0 Windows. Realizou-se, primeiramente, a validação de face adotando-se o Índice de Validade de Conteúdo (IVC) para análise por item e total do constructo, considerando-se aceitável um valor superior a 0,80 ou 80%¹³.

Em seguida, realizou-se a análise univariada dos achados, incluindo-se as distribuições de frequência absolutas e relativas para variáveis categóricas, e as medidas de tendência central (média, mediana) e de variabilidade (amplitudes de variação e desvio-padrão) para variáveis quantitativas.

Na análise da validade de critério concorrente, utilizou-se o Coeficiente de Correlação de *Spearman* para os itens individuais e o Coeficiente de Correlação de *Pearson* para escores totais da escala. Para a preditiva, utilizou-se a Regressão de Cox, sensibilidade e especificidade por tabulação cruzada e acurácia por meio da análise da Área Sob a Curva (ASC) *Receiver Operating Characteristic* (ROC).

Na análise da confiabilidade interobservador, especificamente, quanto à consistência interna dos itens da escala, optou-se pelo coeficiente alfa de *Cronbach*, utilizado para medir o grau de correlação entre itens com valores que variaram entre zero e um. Utilizou-se, ainda, o coeficiente *Kappa* ponderado para os itens individuais e o Coeficiente de Correlação Intraclasse (CCI) para a confiabilidade dos escores totais dos dois observadores.

Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa do HC-UFTM sob o parecer número 3.998.265 e ajustada nas determinações da Resolução CONEP 466/12. Salienta-se que o consentimento foi obtido dos participantes que aceitaram participar da pesquisa e que possuíam condições físicas e mentais para consentir e assinar o Termo de Consentimento Livre e Esclarecido (TCLE). Àqueles que estavam sem condições de consentir e assinar o TCLE, solicitou-se a autorização aos seus responsáveis.

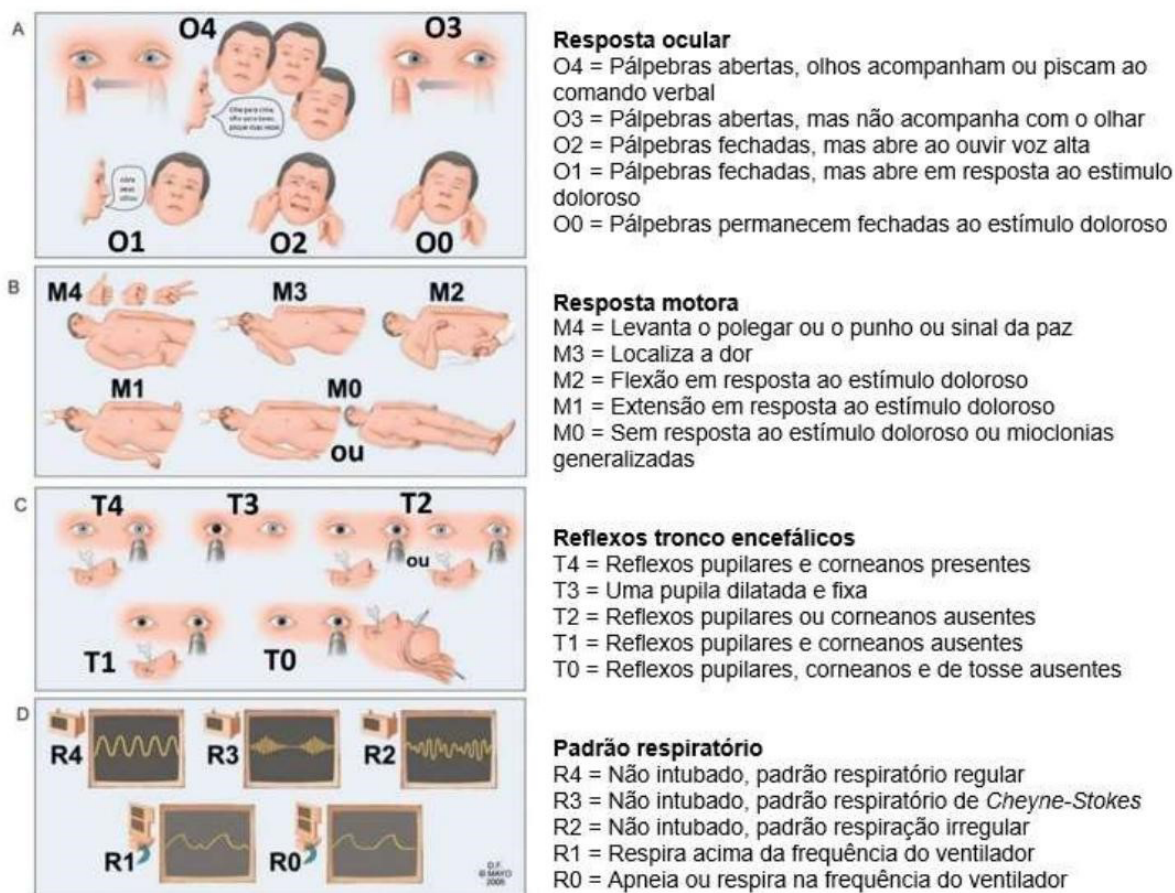
RESULTADOS

Por meio do processo de síntese obteve-se a VPC1 da escala proposta, submetida a 15 (100%) especialistas. Desse total, 10 (66,70%) eram do sexo feminino e 5 (33,30%) do masculino, com idade média de 41,67 anos (DP= 8,90; variação de 28-62 anos). A maioria, representada por 14 (93,30%) especialistas, foi constituída por enfermeiros com tempo de formação médio de 19,20 anos (DP=8,79; variação de 7-41 anos) e tempo de experiência médio em pronto-socorro de 14,93 anos (DP=7,59; variação de 5-34 anos).

Baseando-se nos achados obtidos nas avaliações das equivalências cultural, semântica, conceitual e idiomática e validade de face, calculou-se o IVC por item, que variou de 93% a 100% e, após o IVC total da escala, que resultou em 95%. As principais modificações foram relacionadas com a troca de palavras por seus sinônimos que melhor se adaptaram ao contexto brasileiro, destacando-se, principalmente, alterações nos domínios configurados pela resposta ocular e o padrão respiratório.

Posteriormente, realizou-se o pré-teste com 15(100%) pacientes, sendo 4 (26,70%) do sexo feminino e 11 (73,30%) do masculino com idade média de 66,7 anos (DP= 15,6 e variação de 38-90 anos). Esta fase não resultou em alterações na escala, e a VPF foi nomeada como Escala *Full Outline of UnResponsiveness* - versão para o português brasileiro ou escala FOUR - versão brasileira apresentada na Figura 2, a seguir.

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



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 inculando 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 PaCO₂ dentro dos limites normais. Um teste padrão de apneia pode ser necessário quando o paciente respira na frequência do ventilador (R0).

Figura 2 - Escala *Full Outline of UnResponsiveness* - Versão Brasileira. Uberaba, MG, Brasil, 2021.

A escala FOUR - versão brasileira foi aplicada em 188(100%) pacientes, destes, 112(59,6%) eram do sexo masculino, 76(40,4%), feminino, com idade média de 63,26 anos (DP= 14,77 e variação de 19 a 97 anos). A maioria, representada por 80(42,60%) pacientes, recebeu diagnóstico clínico de Infarto Agudo do Miocárdio (IAM), seguido do Acidente Vascular Cerebral (AVC) 19(10,1%), estavam compensados hemodinamicamente 152(80,00%) pacientes e em uso de outras formas de suporte cardiovascular 130(69,10%) conforme é apresentado na Tabela 1.

Tabela 1 - Distribuição das variáveis clínicas dos participantes do estudo referente à avaliação das propriedades psicométricas da escala. Uberaba, MG, Brasil, 2021. (n=188)

Variável	N	%
Tipo de agravo de acordo com o diagnóstico		
Desordens cardiovasculares	118	62,76%
Desordens neurológicas	36	19,14%
Desordens pulmonares ou respiratórias	9	4,78%
Sepse e infecção	8	4,25%
Desordens hepáticas	5	2,65%
Fraturas	3	1,59%
Neoplasias	3	1,59%
Desordens abdominais	2	1,59%
Agravos renais	2	1,06%
Desordens hematológicas	1	0,53%
Agravos externos ou violência	1	0,53%
Estado hemodinâmico		
Hemodinamicamente instável	36	19,10%
Compensado hemodinamicamente	152	80,90%
Drogas		
Uso de drogas vasoativas		
Sim	111	59,00%
Não	77	41,00%
Uso de outras formas de suporte cardiovascular		
Sim	130	69,10%
Não	58	30,90%

Para a verificação da validade de critério concorrente, compararam-se os escores da escala FOUR com o escore da ECG, observando-se que quanto maior o escore melhor era a resposta do paciente e menores as chances de ir a óbito, e quanto menor o escore pior era a resposta.

O Coeficiente de Correlação de *Spearman* identificou na resposta ocular e na resposta motora forte correlação¹⁴ conforme apresentado na Tabela 2.

Tabela 2 - Apresentação da correlação de *Spearman* (r_s) para o item individual resposta ocular e resposta motora. Uberaba, MG, Brasil, 2021. (n=188)

Itens	r_s	p
Resposta Ocular	0,96	<0,001
Resposta Motora	0,81	<0,001

Os itens reflexos do tronco encefálico e padrão respiratório não foram calculados devido à ausência de critérios semelhantes na ECG. Para os escores totais obteve-se $(r) = 0,97$ ($p < 0,001$) que também caracterizou uma forte correlação¹⁴.

Quanto à validade de critério preditiva, testou-se a associação da escala FOUR - versão brasileira com o desfecho óbito, abrangendo-se 31 (16,5%) pacientes. O risco relativo de se prever o óbito foi de 0,78 ($p < 0,001$), indicando que para cada ponto adicional da escala, diminui-se o risco de óbito em cerca de 20%. O risco relativo considerado para a presente pesquisa foi menor que 1, representando um fator de proteção à medida que a pontuação da escala aumenta.

Para a sensibilidade e especificidade, utilizaram-se pontos de corte pré-definidos por um estudo de validação da escala FOUR - versão chinesa⁸ que considerou o escore de 13 pontos para prever o óbito. Diante dessa evidência, obteve-se uma especificidade de 95,5% para pontuações iguais ou maiores que 13 pontos e uma sensibilidade de 51,6% para escores gerados até 12 pontos no presente estudo.

Na acurácia, evidenciaram-se valores considerados altos ($ASC > 0,8$) que indicaram a boa precisão do instrumento¹⁵. A área sob a curva ROC entre a escala e o desfecho óbito foi de 0,80 (IC 95%: 0,688 - 0,905, $p < 0,001$), ou seja, em 80% das vezes em que for utilizada a escala FOUR – versão brasileira, conseguir-se-ão discriminar os verdadeiros positivos e os verdadeiros negativos, e em 20% das vezes ela apresentará falsos resultados.

Na consistência interna ($n=188$), o valor do alfa de *Cronbach* para os quatro itens que compõem a escala FOUR - versão brasileira foi de 0,94, configurando como muito alta consistência entre os itens¹⁶. A análise da confiabilidade interobservador ($n=36$) foi realizada em dois momentos, inicialmente, calcularam-se os itens individuais e o nível de significância para cada item dos 4 domínios que o instrumento possui. Depois, consideraram-se os escores de pontuação final da escala FOUR – versão brasileira obtida por observador.

Para a confiabilidade dos itens, analisaram-se a resposta ocular, resposta motora e reflexos do tronco encefálico, obtendo-se valor de $k=1,0$; já o padrão respiratório resultou em um $k=0,89$, classificado como quase perfeita confiabilidade¹⁷. Os escores de concordância geral do instrumento resultaram em $ICC=0,99$, demonstrando uma excelente confiabilidade¹⁸.

DISCUSSÃO

A avaliação do nível de consciência faz parte do atendimento ao paciente grave e requer a adoção de escalas capazes de subsidiar uma assistência e terapêutica adequadas para cada caso¹⁹. Neste contexto, o presente estudo confere ineditismo à ciência em saúde e à enfermagem por traduzir, adaptar e validar para o português do Brasil, a escala *Full Outline of UnResponsiveness*, capaz de identificar e avaliar o nível de consciência de pacientes adultos⁶ de maneira mais abrangente do que as escalas existentes, principalmente, por considerar parâmetros que garantam acurácia na avaliação como o padrão respiratório e o reflexo do tronco encefálico e permitir a análise do padrão verbal de pacientes em uso de dispositivos endotraqueais, critérios não abordados pelos instrumentos comumente utilizados para este fim.

Cumpriu-se um percurso metodológico rigoroso para proceder à tradução, adaptação e validação da escala FOUR para o português do Brasil¹⁰⁻¹¹, procedimento esse que vem sendo adotado em outras realidades e cenários em saúde devido à sua confiabilidade para a prática clínica.

Os resultados identificados pela validação de face da presente pesquisa corroboram um estudo brasileiro, realizado no Rio Grande do Norte, que propôs a validação de um protocolo assistencial de enfermagem voltado à pacientes sépticos, obtendo-se um IVC total do constructo de quase perfeita concordância, resultado que auxilia a consideração de sua adequabilidade à prática clínica²⁰.

Outra pesquisa, realizada na Universidade Federal de Santa Catarina, considerou um IVC acima de 90% no processo de adaptação e validação do Questionário *Patient Measure of Safety* para o português do Brasil²¹.

Por ser considerada uma escala clinimétrica, caracterizada como um constructo de fácil interpretação dos seus fenômenos clínicos com variáveis objetivas, diretas e claras e pouco sensíveis a mudanças, a escala FOUR dispensou profundas alterações²².

Para avaliar as propriedades psicométricas de uma escala, é necessário aplicá-la a uma população-alvo, caracterizada no presente estudo por pacientes adultos, graves e internados, principalmente por IAM e AVC. Corroboram este contexto um estudo de validação da escala FOUR para a versão espanhola, que abrangeu na maioria, pacientes vítimas de AVC²³ e uma pesquisa australiana, que comparou a ECG com a escala FOUR, baseando-se em uma amostra em que prevaleceram pacientes graves, vítimas de IAM e de AVC²⁴.

Identificou-se uma forte correlação quanto à validade de critério concorrente para itens da escala FOUR – versão brasileira, condição semelhante a outros estudos de validação da escala FOUR que também apresentaram uma forte correlação ao comparar a ECG e a escala FOUR^{7-8,23,25}.

A forte correlação evidenciada nos escores totais da escala FOUR – versão brasileira, também foi encontrada num estudo realizado em uma UTI de um hospital do Arizona, Estados Unidos, que comparou a escala FOUR com a ECG e demonstrou propriedades psicométricas satisfatórias para sua adoção na prática clínica²⁶.

Na perspectiva da validade de critério preditiva, associou-se a escala FOUR – versão brasileira ao desfecho óbito, identificando-se que dos 188(100%) pacientes, 31(16,5%) foram a óbito e apresentaram um risco maior para este desfecho quanto menor a pontuação da escala. Assemelhou-se a este contexto um estudo desenvolvido em Uganda, África, com 359(100%) pacientes, que comparou o poder preditivo da ECG com a escala FOUR e evidenciou o óbito em 144(40,1%) pacientes com um risco 2,64 vezes maior quanto mais baixo o escore obtido na escala FOUR²⁷.

Avaliar a acurácia, sensibilidade e especificidade de uma escala é útil para determinar o seu desempenho ao ser aplicada na população-alvo²⁸. A versão para o português do Brasil da escala FOUR apresentou uma boa acurácia, condição que se assemelha a um estudo de validação da escala FOUR para a língua chinesa que obteve uma AUC de 0,834 (IC 95%: 0,740-0,928), considerada boa⁸.

Pelo fato de a população abordada para a validação da escala FOUR para o português do Brasil ter divergido do perfil populacional identificado em outras pesquisas⁷⁻⁸, adotou-se na avaliação da sensibilidade e especificidade da escala proposta, o ponto de corte definido por um estudo chinês, realizado com 120 pacientes graves com diagnósticos clínicos semelhantes aos do presente estudo que estabeleceu o escore 13 (sensibilidade, 79%; especificidade, 72%) para prever o óbito⁸. Ao estabelecer uma comparação entre este e o estudo de base, detectou-se que 22% dos pacientes foram a óbito⁸, e ante o exposto é possível evidenciar um número menor de pacientes que obteve um escore inferior a 13 pontos.

Ainda quanto ao ponto de corte estabelecido, uma revisão acerca do contexto histórico da escala FOUR se assemelha à presente pesquisa ao destacar que a pontuação para determinar o risco de mortalidade intra-hospitalar é realizada com base no somatório de 12 pontos, o que se compara à pontuação 8 da ECG. Salienta-se que o autor principal da escala desencoraja o uso da soma das pontuações em decorrência da diminuição de um ponto em qualquer componente, o qual possui uma significativa relevância clínica²⁹.

Determinar a confiabilidade é essencial para verificar a capacidade de uma escala para reproduzir um resultado de forma consistente no tempo e espaço por meio de observadores diferentes indicando a qualidade do constructo, e uma das formas de mensurar esse critério é por meio da consistência interna, visto que indica o quanto os itens de uma escala são homogêneos e estão correlacionados entre si¹³.

A alta consistência interna apresentada pela versão brasileira da escala FOUR, também foi identificada em outros estudos que se propuseram a traduzir, adaptar e validar a referida escala^{25,30} assim como a avaliação da confiabilidade interobservador dos itens, caracterizada como uma concordância perfeita, corroborada por uma pesquisa de validação da escala FOUR em um departamento de emergência na Austrália²⁴, e da confiabilidade total do ICC que se assemelhou a um estudo de validação da escala FOUR para a língua italiana²⁵.

Outros estudos de validação da escala FOUR em diferentes idiomas apresentaram uma excelente confiabilidade da escala ao comparar o escore gerado pelo observador 1 e observador 2^{8,23}.

Os achados apresentados no presente estudo são relevantes para a área da saúde e enfermagem por fornecer para o cenário nacional uma escala útil, de fácil memorização e simples aplicação, que evidencia informações importantes acerca do nível de consciência e prognóstico de pacientes adultos em condição grave e em ambiente hospitalar.

A limitação deste estudo configurou-se pelo fato de os dados terem sido coletados em uma única instituição de ensino. Diante disso, sugere-se a realização de pesquisas multicêntricas sobre a temática a fim de propiciar a generalização dos achados.

CONCLUSÃO

Este estudo disponibiliza para uso no Brasil a escala FOUR – versão brasileira, mantendo 20 itens e quatro domínios da escala original. As análises adotadas para verificar a validade de critério preditiva indicaram que a escala é capaz de prever desfechos indesejáveis como o óbito. Houve uma correlação forte entre a ECG a escala FOUR – versão brasileira. A confiabilidade apresentou índices satisfatórios, demonstrando-se uma escala válida, confiável e útil para avaliar o nível de consciência de pacientes adultos.

Sugere-se o desenvolvimento de outras pesquisas sobre a escala FOUR – versão brasileira para fomentar o aprofundamento científico desta temática e determinar um ponto de corte para a escala, capaz de indicar, com mais precisão, as condutas necessárias em face das alterações do nível de consciência de pacientes adultos.

REFERÊNCIAS

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 [acesso 2021 Out 9];15(1):e0228082. Disponível em: <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 [acesso 2021 Out 9]; 30(3):385-93. Disponível em: <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 [acesso 2021 Out 9];10(5):4249-54. Disponível em: <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 [acesso 2021 Out 9];2(7872):81-4. Disponível em: [https://doi.org/10.1016/s0140-6736\(74\)91639-0](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 [acesso 2021 Out 8]; 128(6):1612-20. Disponível em: <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 [acesso 2021 Out 9];58(4):585-93. Disponível em: <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 [acesso 2021 Out 9];27(2):229-36. Disponível em: <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 [acesso 2021 Out 10];15:254. Disponível em: <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 [acesso 2021 Out 9];63(8):640-7. Disponível em: 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 [acesso 2022 Fev 7];19(1):33-44. Disponível em: [https://doi.org/10.1016/0168-8510\(91\)90072-6](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 George's Respiratory Questionnaire after adaptation to a different language and culture: the spanish example. *Eur Respir J* [Internet]. 1996 [acesso 2021 Out 9];9(6):1160-6. Disponível em: <https://doi.org/10.1183/09031936.96.09061160>
12. Fehring RJ. The Fehring model. In: Carrol-Johnson RM, Paquete M, editors. *Classification of nursing diagnoses: proceedings of the Tenth Conference*. Philadelphia, PA(US): J.B. Lippincott; 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 [acesso 2021 Out 9];26(3):649-59. Disponível em: <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 [acesso 2021 Out 9];28(3):311-3. Disponível em: <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 [acesso 2021 Out 9];25(2):291-312. Disponível em: <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 [acesso 2021 Out 3];24:e60338. Disponível em: <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 [acesso 2021 Out 9];71(3):1106-14. Disponível em: <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 [acesso 2021 Out 9];30:e20180322. Disponível em: <https://doi.org/10.1590/1980-265X-TCE-2018-0322>
22. Pais-Ribeiro JL. Medida na avaliação psicológica. *Psic, Saúde Doenças* [Internet]. 2013 [acesso 2021 Out 9];14(1):245-63. Disponível em: 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 [acesso 2021 Out 9];63(6):364-9. Disponível em: <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 [acesso 2021 Out 5];28(6):486-90. Disponível em: <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 [acesso 2021 Out 9];7(2):145-52. Disponível em: <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 [acesso 2021 Out 9];23(7):555-60. Disponível em: <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 [acesso 2021 Out 7];32(3):734-41. Disponível em: <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 [acesso 2021 Out 9];43(6):408. Disponível em: <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 [acesso 2021 Out 29];82(4):435-8. Disponível em: <https://doi.org/10.4065/82.4.435>

NOTAS

ORIGEM DO ARTIGO

Extraído da dissertação - Tradução, Adaptação Cultural e Validação para o Português do Brasil da Escala *Full Outline Of Unresponsiveness*, apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Atenção à Saúde, da Universidade Federal do Triângulo Mineiro, em 2021.

CONTRIBUIÇÃO DE AUTORIA

Concepção do estudo: Bernardinelli FCP, Chavaglia SRR.

Coleta de dados: Bernardinelli FCP, Amorim GC.

Análise e interpretação dos dados: Bernardinelli FCP, Haas VJ, Chavaglia SRR.

Discussão dos resultados: Bernardinelli FCP, Chavaglia SRR.

Redação e/ou revisão crítica do conteúdo: Bernardinelli FCP, Chavaglia SRR, Haas VJ, Campanharo CRV, Barbosa MH, Amorim GC.

Revisão e aprovação final da versão final: Bernardinelli FCP, Chavaglia SRR, Haas VJ, Campanharo CRV, Barbosa MH, Amorim GC.

FINANCIAMENTO

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

APROVAÇÃO DE COMITÊ DE ÉTICA EM PESQUISA

Aprovado no Comitê de Ética em Pesquisa do Hospital de Clínicas da Universidade Federal do Triângulo Mineiro, parecer n. 3.998.265, Certificado de Apresentação para Apreciação Ética n. 27172719.0.0000.8667.

CONFLITO DE INTERESSES

Não há conflito de interesses.

EDITORES

Editores Associados: Clemente Neves de Sousa, Monica Motta Lino.

Editor-chefe: Roberta Costa.

HISTÓRICO

Recebido: 25 de novembro de 2021.

Aprovado: 10 de março de 2022.

AUTOR CORRESPONDENTE

Suzel Regina Ribeiro Chavaglia

suzel.ribeiro@yahoo.com.br