

Validação brasileira dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília do *Patient-Reported Outcomes Measurement Information System (PROMIS)*

Brazilian validation of the item banks on Sleep
Disturbance and Wake Disturbance in the
*Patient-Reported Outcomes Measurement
Information System (PROMIS)*

Validación brasileña de los bancos de ítems
Disturbio del Sueño y Disturbio de la Vigilia
en el *Patient-Reported Outcomes Measurement
Information System (PROMIS)*

Zilma Maria Severino Silva e Costa ¹
Rogério de Melo Costa Pinto ¹
Tânia Maria da Silva Mendonça ¹
Carlos Henrique Martins da Silva ¹

doi: 10.1590/0102-311X00228519

Resumo

O objetivo deste estudo foi validar a versão brasileira dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília do Patient-Reported Outcomes Measurement Information System (PROMIS). A validação desses bancos de itens foi feita por meio da Teoria Clássica dos Testes, Teoria de Resposta ao Item, calibração dos itens conforme o modelo de resposta gradual proposto por Samejima e análise do Funcionamento Diferencial do Item (DIF), sendo a amostra constituída por 627 participantes com mais de 18 anos. Em relação à qualidade dos dados, observou-se que não houve dados perdidos, e cinco itens do banco Distúrbio do Sono e quatro itens do banco Distúrbio da Vigília tiveram efeito piso, e nenhum dos bancos apresentou efeito teto. Os dois bancos mostraram excelente confiabilidade com uma escala homogênea e boa estabilidade temporal. Com a Análise Fatorial Confirmatória, através dos índices de qualidade de ajuste absoluto e incremental, verificou-se um bom ajuste para os dois bancos, o que revela que os instrumentos são válidos e suficientemente unidimensionais. A calibração evidenciou que as propriedades psicométricas dos itens foram satisfatórias com boa cobertura do construto sono-vigília. No tocante ao DIF, tanto o banco Distúrbio do Sono quanto o banco Distúrbio da Vigília mostraram ser homogêneos para os diferentes grupos testados. Pode-se concluir que a versão para a língua portuguesa falada no Brasil dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília do PROMIS demonstrou ser uma medida confiável, precisa e válida, com análises psicométricas robustas.

Estudos de Validação; Transtornos do Sono-Vigília; Qualidade de Vida

Correspondência

Z. M. S. Silva e Costa
Universidade Federal de Uberlândia,
Al. Mariana 364, Uberlândia, MG 38411-262, Brasil.
zilmaseverino@hotmail.com

¹ Universidade Federal de Uberlândia, Uberlândia, Brasil.



Introdução

Sono e vigília são dois estados neurocomportamentais de fundamental importância para os seres humanos, regulados por processos homeostáticos e circadianos ¹. O sono é um estado complexo, ativo e muito importante para o desenvolvimento normal do cérebro, da memória e do aprendizado, pois é no período de sono noturno que são liberados hormônios imprescindíveis para o desenvolvimento, maturação e manutenção da saúde, refletindo diretamente na qualidade de vida do indivíduo. Já a vigília é um estado comportamental ativo de interação com o meio ambiente ¹ e está diretamente correlacionada com o sono. A ausência total ou parcial de sono noturno de qualidade pode ocasionar distúrbios do sono e da vigília, os quais estão fortemente relacionados à saúde e podem levar a fatores de risco, como obesidade, hipertensão, vulnerabilidade ao resfriado comum, depressão e mortalidade por todas as causas ¹. O distúrbio do sono mais frequente na população é a insônia, que atinge de 30% a 40% dos adultos ²; e em relação ao distúrbio da vigília, a sonolência diurna excessiva é a mais comum e atinge de 5% a 15% da população adulta ³.

O sono pode ser avaliado por diversas ferramentas, como a eletroencefalografia, mapeamento cerebral e polissonografia, que são os exames existentes com melhor acurácia para caracterizar os padrões de distúrbios do sono e distúrbios da vigília na amostra clínica e epidemiológica ¹, mas que são pouco acessíveis devido ao seu alto custo. Sendo assim, os instrumentos de autorrelato são as ferramentas mais práticas e as mais usadas na identificação da função sono-vigília por possuírem menor custo com a qualidade necessária. Os instrumentos de autorrelato disponíveis são o *Pittsburgh Sleep Quality Index* (PSQI) ⁴ e o *Epworth Sleepiness Scale* (ESS) ⁵, que foram traduzidos para a língua portuguesa, validados para a cultura brasileira ⁶ e são amplamente utilizados. Porém, mesmo com o seu uso generalizado, essas medidas encontram desafios no âmbito metodológico relacionados às propriedades psicométricas ^{1,7,8} que podem interferir no resultado final da avaliação, pois favorecem o aparecimento de vies nas medidas.

Em 2004, Instituto Nacional de Saúde dos Estados Unidos (NIH) lançou um programa de pesquisa chamado *Patient-Reported Outcomes Measurement Information System* (PROMIS) ⁹, cujo objetivo era criar itens e desenvolver questionários de autorrelato para medir os principais domínios de saúde e, assim, fornecer aos médicos e pesquisadores acesso a medidas eficientes, precisas e válidas que corroboram as avaliações de saúde para aperfeiçoar os atendimentos, diagnósticos clínicos e tratamentos das diversas doenças crônicas. Alguns instrumentos já estão disponíveis em várias línguas para serem utilizados no formato impresso e também por meio eletrônico, que são os testes adaptativos computadorizados (CAT – *computer adaptive test*), os quais fornecem relatórios instantâneos da saúde do paciente. Dentre esses bancos de itens estão o Distúrbio do Sono e o Distúrbio da Vigília do PROMIS, que foram construídos para aprimorar a mensuração da função sono-vigília. Esses bancos foram desenvolvidos para calibrar itens individuais de alta precisão, por meio de um processo sistemático de revisões de literatura, consenso de especialistas, métodos de pesquisa qualitativa, métodos da Teoria Clássica dos Testes (TCT) e da Teoria de Resposta ao Item (TRI) ¹, a qual representa o principal diferencial desses bancos em relação aos instrumentos de autorrelato já existentes. A TRI concentra-se nas propriedades particulares de cada item e sua utilização na construção dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília apresenta diversas vantagens, tais como a probabilidade do indivíduo acertar cada um dos itens de um instrumento ocorre em função do seu traço latente e independe dos itens; os parâmetros de um item não dependem dos outros itens do instrumento nem da amostra de sujeitos utilizada; e são utilizados itens com dificuldade tal que esteja ao redor do tamanho da habilidade de um sujeito para avaliar a mesma, de modo que esta avaliação não seja influenciada pelo teste ¹⁰.

Os bancos de itens Distúrbio do Sono e Distúrbio da Vigília já foram traduzidos e adaptados para a língua portuguesa falada no Brasil, mas para serem utilizados é necessário fazer a sua validação. Portanto, o objetivo deste trabalho é validar a versão brasileira dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília do PROMIS.

Metodologia

Desenho do estudo e participantes

Estudo metodológico e transversal de validação realizado após aprovação pelo Comitê de Ética em Pesquisa da Universidade Federal de Uberlândia (CEP/UFU), sob o registro CEP 443/10.

De acordo com a literatura, para a determinação do tamanho da amostra para estudos de validação, recomenda-se uma proporção de 1:5 a 1:10 de cada item a ser validado por participante^{11,12,13}. Dessa forma, a amostra foi constituída por 627 participantes maiores de 18 anos, alfabetizados e que estavam em acompanhamento clínico de rotina em diversas especialidades no complexo hospitalar da UFU. O participante foi informado que deveria responder a todos os itens dos questionários, caso não, seria excluído da pesquisa.

Os dados foram colhidos por uma só pesquisadora e o início do preenchimento dos instrumentos foi precedido pela leitura e assinatura do Termo de Consentimento Livre e Esclarecido (TCLE), e também foi aplicado um questionário para mapear as características sociodemográficas da amostra.

Instrumentos

Os participantes responderam à versão traduzida e adaptada culturalmente para a língua portuguesa falada no Brasil do banco de itens Distúrbio do Sono e do banco de itens Distúrbio da Vigília do PROMIS¹⁴. O banco de itens Distúrbio do Sono é composto por 27 itens, sendo 7 reversos, e o Distúrbio da Vigília por 16 itens, com 3 reversos, tendo os dois período recordatório de sete dias. Ambos estão na escala Likert, que apresenta variação de intensidade de um a cinco, e para os itens reversos a pontuação é invertida. No banco de itens Distúrbio do Sono há três níveis de respostas, sendo eles, de 1 (nem um pouco) a 5 (muitíssimo), de 1 (nunca) a 5 (sempre) e de 1 (muito fraca) a 5 (muito boa). E no banco de itens Distúrbio da Vigília são dois níveis de respostas, de 1 (nem um pouco) a 5 (muitíssimo) e de 1 (nunca) a 5 (sempre). Os escores do banco Distúrbio do Sono variam de 27 a 135 pontos e do banco Distúrbio da Vigília de 16 a 80 pontos, de modo que a qualidade do sono aumenta com a diminuição da pontuação¹.

Análise estatística

• Teoria Clássica dos Testes (TCT)

a) Qualidade dos dados

A qualidade dos dados foi examinada por meio da ocorrência de dados perdidos e da presença de efeito piso e teto, através do software SPSS, versão 20.0 (<https://www.ibm.com/>). Os dados perdidos não devem ultrapassar em 20% o número total de participantes, e para efeito piso e teto é necessário não haver mais de 50% dos respondentes em níveis extremos dos itens^{15,16,17}.

b) Confiabilidade

A confiabilidade foi verificada por meio da consistência interna do item e do teste-reteste, ambos calculados com o software SPSS, versão 20.0. A consistência interna do item foi avaliada pelo coeficiente alfa (α) de Cronbach, sendo que valores acima de 0,70 são considerados bons e valores próximos a 1,00 são considerados ideais¹⁸. A confiabilidade pelo teste-reteste mede a correlação entre as aplicações em dois pontos distintos do tempo, a fim de obter a mesma resposta. Neste estudo o intervalo foi de 20 dias, aplicado em 10% da amostra, e sua medida foi feita através do coeficiente de correlação intra-classe (CCI). Os valores do coeficiente podem ser classificados como < 0,40, pobre reprodutibilidade; entre 0,40 e 0,75, boa reprodutibilidade; e maior que 0,75, excelente reprodutibilidade¹¹.

c) Validade

A validade testada foi a validade de constructo, que é a capacidade que o instrumento possui de estabelecer e testar hipóteses ou modelos sobre o seu comportamento, e é considerada a forma mais fundamental de validade dos instrumentos¹⁹. A validade de construto foi avaliada pela Análise Fatorial Confirmatória (AFC) e foi utilizado o software LISREL, versão 8.80 (<http://www.ssicentral.com/lisrel/>).

- **Teoria de Resposta ao Item (TRI)**

Os pressupostos básicos da TRI são unidimensionalidade, independência local e monotonicidade, e deveriam ser atendidos para dar continuidade ao processo de validação dos instrumentos. Além disso, é importante ressaltar que a AFC é utilizada tanto para verificar a validade do instrumento na TCT como a unidimensionalidade na TRI, sendo a conexão entre as duas teorias.

a) Unidimensionalidade

A unidimensionalidade foi verificada através da AFC e foi feita por meio do software LISREL versão 8.80, para verificar o ajuste ao modelo de cada um dos bancos por intermédio de dois índices de ajuste absoluto e dois índices de ajuste incremental. Os índices de ajuste absoluto utilizados foram a *Root Mean Square Error of Approximation* (RMSEA), sendo que valores < 0,08 indicam ajuste adequado e valores < 0,06 indicam bom ajuste; e a *Standardized Root Mean Square Residual* (SRMR), em que valores < 0,08 indicam bom ajuste. Já os índices de ajuste incremental foram o *Comparative Fit Index* (CFI), com valores desejáveis > 0,90; e o *Non-Normed Fit Index* (NNFI) ou *Tucker-Lewis Index* (TLI), ideal para valores > 0,90¹³. Para estimar tais índices utilizou-se o método *Diagonally Weighted Least Squares* (DWLS) e, atendidos todos os requisitos, pode-se concluir que os instrumentos são suficientemente unidimensionais.

b) Independência local

No estudo foi feita a análise dos resíduos da matriz de correlação da AFC através do coeficiente de correlação de Spearman (ρ), com o software SPSS, versão 20.0. Segundo Buysse et al.¹, correlações acima de 0,30 são geralmente usadas para identificar dependência local. Mas, a natureza dos bancos de itens Distúrbio do Sono e Distúrbio da Vigília produz mais pares de itens localmente dependentes se comparados a outros tipos de bancos de itens. Dessa forma, utilizou-se um valor menos restritivo, de $\rho > 0,50$, para verificar a presença de dependência local e a necessidade de exclusão do item¹.

c) Monotonicidade

Para verificar a monotonicidade foi utilizado o software Multilog, versão 7.0 (<https://www.environmental-expert.com/software/multilog-controller-software-394690>), e executado o Modelo de Resposta Gradual (MRG) proposto por Samejima²⁰, que considera um valor positivo para o parâmetro a (discriminação), o que resulta em uma monotonicidade crescente. A análise da curva característica do item (CCI) gerada permitiu verificar se os itens foram bem ajustados e, conseqüentemente, se este pressuposto foi devidamente atendido²¹.

Calibração

Os bancos de itens Distúrbio do Sono e Distúrbio da Vigília foram calibrados individualmente conforme o MRG proposto por Samejima²⁰, usando-se o software Multilog, versão 7.0, que estima o grau de discriminação, parâmetro a , e o grau de dificuldade, parâmetro b , dos itens com base nas respostas dos participantes. O parâmetro a representa a capacidade do item em diferenciar os respondentes de diferentes níveis de traço latente, determinando a qualidade do item. Parâmetro a com valor entre 0,01 e 0,34 é classificado como muito baixo; de 0,35 a 0,64, baixo; de 0,65 a 1,34, moderado; de 1,35

a 1,69, alto; e, acima de 1,70, muito alto ²¹. O parâmetro *b* está localizado ao longo do continuum da escala de aptidão da variável latente sono-vigília. Para que os itens sejam idealmente perfeitos devem ter abrangência suficiente de modo que contenham os graus de dificuldade do construto medido ²².

Funcionamento Diferencial do Item (DIF)

Para um mesmo item, os participantes com mesmo traço latente devem possuir a mesma resposta. Porém, há casos em que dois subgrupos possuem respostas diferentes, mesmo tendo traços latentes iguais. Isso ocorre pois o item funciona de forma diferente para os subgrupos, como por exemplo, homem e mulher, saudável e portador de doença crônica, entre outros ²³. Há dois tipos de DIF, uniforme ou consistente e não uniforme ou inconsistente. O DIF uniforme ocorre quando o item se difere entre os dois grupos no parâmetro *b*, e no DIF não uniforme a diferença está no parâmetro *a* ¹⁰. Neste estudo, foi analisada a presença de DIF para os bancos Distúrbio do Sono e Distúrbio da Vigília nos grupos homem *versus* mulher, saudável *versus* portador de doença crônica e língua portuguesa *versus* língua inglesa. O critério aplicado para a detecção de DIF foi o pseudo $R^2 \geq 0,02$ ^{24,25}, e seu impacto foi medido pela análise do pseudo R^2 de Nagelkerke, sendo $< 0,13$ insignificante; entre 0,13 e 0,26, moderado; e $> 0,26$, grande ²³. O software utilizado foi o R (<http://www.r-project.org>), com o pacote *lordif* e na versão 3.6.1 ²⁶.

Resultados

A amostra foi constituída por 627 participantes e os dados sociodemográficos estão na Tabela 1.

Qualidade dos dados

Todos os itens foram respondidos por todos os participantes, portanto, não houve dados perdidos. O efeito piso foi observado em cinco itens do banco Distúrbio do Sono, *Sleep 108*, *Sleep 67*, *Sleep 70*, *Sleep 72* e *Sleep 93*, como também em quatro itens do banco Distúrbio da Vigília, *Sleep 06*, *Sleep 07*, *Sleep 10* e *Sleep 30*. Já o efeito teto não foi observado em nenhum dos dois bancos de itens.

Tabela 1

Dados sociodemográficos da amostra.

Características	Valor
Idade média [anos] (DP)	42,41 (14,75)
Sexo (%)	
Feminino	52,79
Estado civil (%)	
Casado	50,40
Raça (%)	
Branca	55,30
Escolaridade (%)	
Ensino Médio completo (12 anos de estudos)	46,25
Ensino Superior completo (17 anos de estudos)	23,13
Renda familiar (%)	
Até dois salários mínimos	53,40
Saúde (%)	
Boa	53,40
Doença crônica autorrelatada (%)	
Presença	57,90

DP: desvio padrão.

Confiabilidade

O coeficiente alfa de Cronbach foi $\alpha = 0,96$ e o CCI = 0,96 (IC95%: 0,95;0,96).

Validade (TCT) e Unidimensionalidade (TRI)

Para o banco Distúrbio do Sono os índices foram RMSEA = 0,099, SRMR = 0,066, CFI = 0,970 e NNFI = 0,970, e para o banco Distúrbio da Vigília os índices foram RMSEA = 0,110, SRMR = 0,081, CFI = 0,940 e NNFI = 0,930.

Independência local

No banco Distúrbio do Sono 21 pares de itens apresentaram correlação $\rho > 0,30$, e no banco Distúrbio da Vigília foram cinco. Nos dois bancos não houve nenhum par de itens com $\rho > 0,50$.

Monotonicidade

As curvas características de cada item, do banco Distúrbio do Sono e do banco Distúrbio da Vigília, estão nas Figuras 1 e 2, respectivamente. A curva total de informação do teste do banco Distúrbio do Sono evidenciou que a escala de aptidão da amostra está entre os valores -1,5 e 3, e para o banco Distúrbio da Vigília a escala possui valores entre -1 e 3.

Figura 1

Curvas características de cada item do banco Distúrbio do Sono.

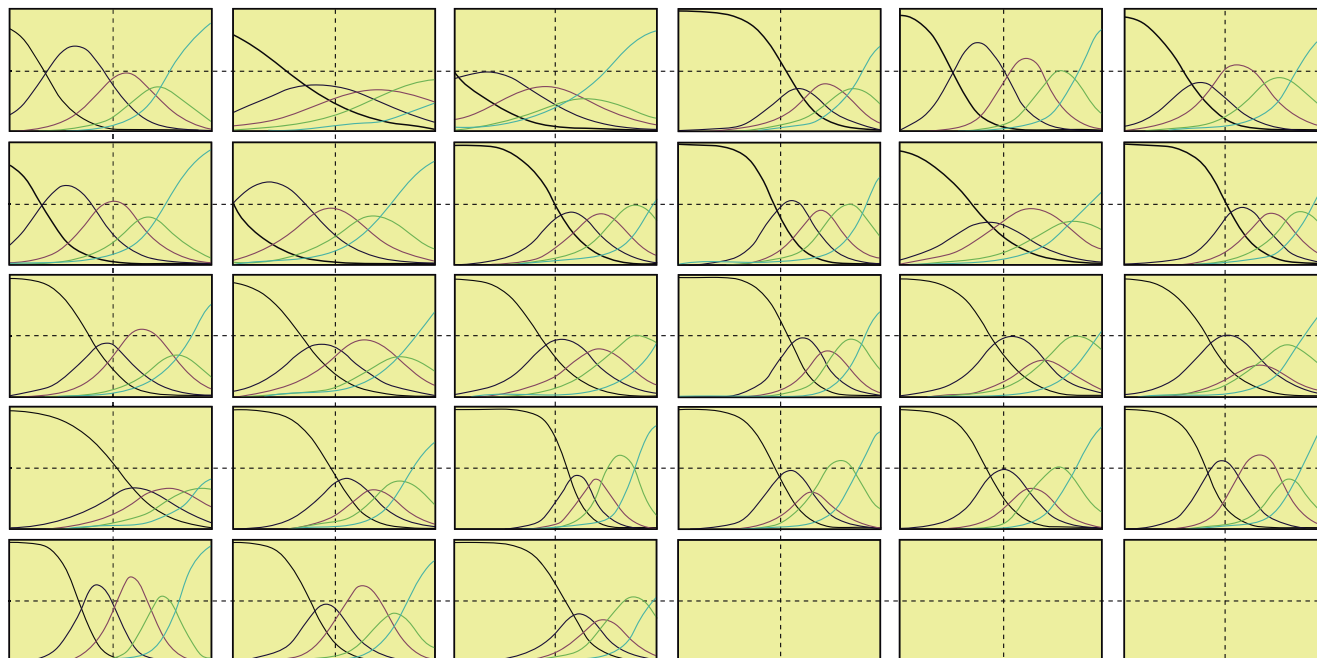
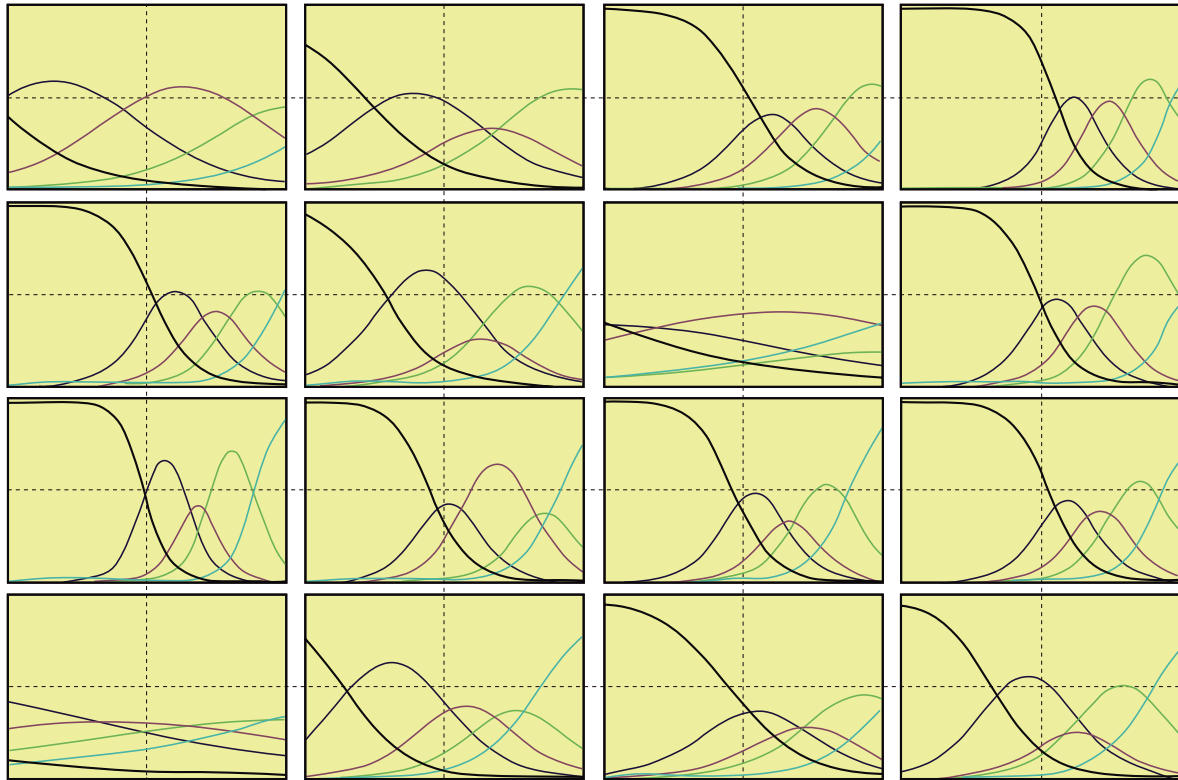


Figura 2

Curvas características de cada item do banco Distúrbio da Vigília.



Calibração

Os valores das estimativas do parâmetro a do banco de itens Distúrbio do Sono variaram de 0,83 (*Sleep 106*) a 3,19 (*Sleep 72*) e os valores das estimativas do parâmetro b variaram de -3,22 (*Sleep 107*) a 4,46 (*Sleep 106*), conforme a Tabela 2. Já os valores das estimativas do parâmetro a do banco de itens Distúrbio da Vigília variaram de 0,32 (*Sleep 119*) a 3,52 (*Sleep 27*) e os valores das estimativas do parâmetro b variaram de -10,38 (*Sleep 119*) a 4,61 (*Sleep 19*), de acordo com a Tabela 3. A média das estimativas do parâmetro b do banco de itens Distúrbio do Sono variou de -0,67 (*Sleep 107*) a 1,69 (*Sleep 70*) na população brasileira e de -0,02 (*Sleep 115*) a 1,91 (*Sleep 70*) na população americana (Tabela 2). Para o banco de itens Distúrbio da Vigília, a média variou de -1,75 (*Sleep 119*) a 1,73 (*Sleep 07*) na população brasileira e de -0,10 (*Sleep 120*) a 1,57 (*Sleep 33*) na população americana (Tabela 3).

Funcionamento Diferencial do Item (DIF)

No banco Distúrbio do Sono, em relação ao grupo homem *versus* mulher e saudável *versus* portador de doença crônica, nenhum item apresentou DIF. Em relação ao grupo de língua portuguesa *versus* língua inglesa, seis itens apresentaram DIF, *Sleep 109*, *Sleep 115*, *Sleep 116*, *Sleep 42*, *Sleep 44* e *Sleep 45*. Apenas o *Sleep 116* apresentou DIF não uniforme e todos tiveram impacto insignificante. Já no banco Distúrbio da Vigília, para o grupo homem *versus* mulher nenhum item apresentou DIF; para o saudável *versus* portador de doença crônica, apenas o *Sleep 04* apresentou DIF uniforme, mas com impac-

Tabela 2Estimativas dos parâmetros *a* e *b* do banco Distúrbio do Sono do Brasil e dos Estados Unidos.

Itens	<i>a</i>		<i>b1</i>		<i>b2</i>		<i>b3</i>		<i>b4</i>		Média <i>b</i>	
	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos
<i>Sleep 105</i>	1,85	2,45	-2,08	1,20	-0,17	-0,15	0,99	0,72	1,80	1,59	0,14	0,24
<i>Sleep 106</i>	0,83	1,51	-1,50	-0,65	0,45	0,44	2,19	1,59	4,46	2,61	1,40	1,00
<i>Sleep 107</i>	1,04	1,57	-3,22	-1,57	-1,15	-0,35	0,33	0,66	1,36	1,92	-0,67	0,17
<i>Sleep 108</i>	1,90	2,30	0,16	-0,29	0,93	0,69	1,82	1,45	2,57	2,33	1,37	1,05
<i>Sleep 109</i>	2,29	3,39	-1,58	-0,29	0,08	0,00	1,29	1,08	2,24	1,90	0,51	0,67
<i>Sleep 110</i>	1,75	2,17	-1,33	-1,56	-0,35	-0,16	1,06	0,77	2,17	1,81	0,39	0,22
<i>Sleep 115</i>	1,86	2,77	-2,20	-1,25	-0,57	-0,34	0,67	0,43	1,53	1,09	-0,14	-0,02
<i>Sleep 116</i>	1,41	2,58	-3,19	-1,35	0,83	-0,34	0,56	0,49	1,72	1,28	-0,44	0,02
<i>Sleep 125</i>	1,91	1,91	-0,10	-0,14	0,86	0,49	1,78	1,32	2,90	2,07	1,36	0,94
<i>Sleep 20</i>	2,34	2,80	-0,19	-0,56	0,82	0,33	1,64	0,98	2,58	1,74	1,21	0,62
<i>Sleep 42</i>	1,31	2,09	-1,00	-1,10	0,09	0,05	1,59	0,94	2,68	1,84	0,84	0,43
<i>Sleep 44</i>	2,13	2,51	0,01	-0,46	0,97	0,31	1,80	0,98	2,68	1,72	1,37	0,64
<i>Sleep 45</i>	1,91	2,18	-0,67	0,03	0,28	0,85	1,63	1,55	2,37	2,35	0,90	1,20
<i>Sleep 50</i>	1,49	1,19	-1,01	-0,98	0,20	0,33	1,58	1,76	2,47	3,30	0,81	1,10
<i>Sleep 65</i>	1,65	1,64	-0,49	0,21	0,74	1,13	1,74	2,02	3,12	2,96	1,28	1,58
<i>Sleep 67</i>	2,46	2,37	0,22	0,28	1,08	1,02	1,73	1,62	2,58	2,37	1,40	1,32
<i>Sleep 68</i>	1,72	1,77	-0,36	0,22	0,88	1,21	1,58	1,95	2,88	2,73	1,25	1,53
<i>Sleep 69</i>	1,72	1,75	-0,61	-0,57	0,68	0,41	1,29	1,04	2,38	1,81	0,94	0,67
<i>Sleep 70</i>	1,36	1,40	0,15	0,67	1,19	1,56	2,21	2,26	3,20	3,13	1,69	1,91
<i>Sleep 71</i>	1,86	1,52	-0,10	-0,19	0,85	0,95	1,56	1,76	2,44	2,72	1,19	1,31
<i>Sleep 72</i>	3,19	2,47	0,29	0,03	0,88	0,66	1,45	1,37	2,39	2,30	1,25	1,09
<i>Sleep 78</i>	2,38	1,99	-0,15	-0,02	0,72	0,89	1,24	1,61	2,36	2,22	1,04	1,18
<i>Sleep 86</i>	2,15	1,85	-0,49	-0,58	0,51	0,66	1,16	1,37	2,21	2,30	0,85	0,24
<i>Sleep 87</i>	2,48	2,19	-0,63	-0,90	0,39	0,10	1,56	1,00	2,27	1,78	0,90	1,00
<i>Sleep 90</i>	3,10	3,66	-0,99	-0,61	0,02	0,16	1,18	0,96	1,97	1,62	0,55	0,17
<i>Sleep 92</i>	2,34	2,17	-0,72	-0,55	0,19	0,36	1,47	1,31	2,21	2,24	0,79	1,05
<i>Sleep 93</i>	2,12	1,97	0,25	0,22	1,04	1,03	1,74	1,65	2,85	2,47	1,47	0,67

to insignificante; e quanto à língua portuguesa *versus* língua inglesa, dois itens apresentaram DIF, *Sleep 04* e *Sleep 19*, ambos não uniformes com impacto insignificante.

Discussão

O estudo foi composto por uma amostragem de conveniência, o que resultou em uma proporcionalidade entre os sexos, sendo 52,79% da população composta por mulheres. A idade média dos participantes foi de 42,41 anos, sendo 50,4% casados e 55,3% da raça branca. A renda familiar predominante foi de até dois salários mínimos com 53,4% dos respondentes, e 53,4% também disseram apresentar uma boa saúde, apesar de 57,9% relatarem ter pelo menos um tipo de doença crônica não transmissível. Em relação aos anos de estudos, 46,25% têm o ensino médio completo e 23,13% têm o ensino superior completo, portanto, 69,38% apresentam bom nível de formação escolar, com capacidade de interpretar bem os itens.

A qualidade dos dados foi adequada em relação aos dados perdidos e também ao efeito piso e teto. Todos os participantes responderam corretamente aos questionários e não houve dados perdidos, pois a pesquisadora orientou quanto ao preenchimento total dos questionários e feita a verificação

Tabela 3Estimativas dos parâmetros *a* e *b* do banco Distúrbio da Vigília do Brasil e dos Estados Unidos.

Itens	<i>a</i>		<i>b1</i>		<i>b2</i>		<i>b3</i>		<i>b4</i>		Média <i>b</i>	
	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos	Brasil	Estados Unidos
<i>Sleep 04</i>	0,92	1,83	-3,53	-1,68	-0,57	-0,11	2,19	1,17	4,32	2,19	0,60	0,39
<i>Sleep 06</i>	1,09	2,24	-1,77	-1,29	0,41	0,27	1,65	1,07	3,95	2,11	1,06	0,54
<i>Sleep 07</i>	1,81	2,20	0,13	-0,14	1,09	0,93	2,12	1,73	3,57	2,55	1,73	1,27
<i>Sleep 10</i>	2,77	3,45	0,32	0,10	1,13	0,97	1,89	1,65	2,91	2,38	1,56	1,28
<i>Sleep 11</i>	2,24	3,40	0,12	-0,09	1,14	0,88	1,91	1,58	2,95	2,28	1,53	1,16
<i>Sleep 18</i>	1,62	2,67	-1,35	-1,54	0,49	0,18	1,13	0,94	2,64	1,90	0,73	0,37
<i>Sleep 19</i>	0,41	1,43	-4,65	-0,44	-1,33	0,69	2,79	1,88	4,61	3,18	0,36	1,33
<i>Sleep 25</i>	2,39	3,76	-0,09	-0,09	0,78	0,84	1,57	1,53	3,09	2,25	1,34	1,13
<i>Sleep 27</i>	3,52	4,82	-0,06	0,10	0,86	1,02	1,37	1,61	2,40	2,22	1,14	1,24
<i>Sleep 29</i>	2,29	3,66	-0,32	-0,05	0,47	0,74	1,84	1,65	2,53	2,47	1,13	1,20
<i>Sleep 30</i>	2,33	2,92	-0,20	-0,03	0,72	0,89	1,32	1,56	2,34	2,33	1,05	1,19
<i>Sleep 33</i>	2,38	2,60	0,16	0,36	0,97	1,26	1,66	1,99	2,70	2,68	1,37	1,57
<i>Sleep 119</i>	0,32	1,67	-10,38	-1,58	-2,91	-0,39	1,05	0,52	5,23	1,39	-1,75	-0,02
<i>Sleep 120</i>	1,44	1,87	-2,24	-1,51	-0,14	-0,48	1,03	0,39	2,11	1,19	0,19	-0,10
<i>Sleep 123</i>	1,25	1,18	-0,30	-0,15	0,93	1,04	1,84	2,02	3,43	2,99	1,48	1,48
<i>Sleep 124</i>	1,63	1,72	-1,09	-1,27	0,48	0,12	1,10	0,80	2,49	1,66	0,75	0,33

ao término de cada aplicação. Isso assinala o cuidado dos pesquisadores como também a dedicação dos respondentes que participaram da pesquisa. Os itens do banco Distúrbio do Sono como Distúrbio da Vigília não apresentaram efeito teto e o efeito piso foi verificado em 18,51% do banco Distúrbio do Sono e 25% do banco Distúrbio da Vigília, o que representa um baixo impacto e sugere que os níveis de resposta dos instrumentos em geral estão bem distribuídos²⁷. Os itens do banco Distúrbio do Sono que apresentaram efeito piso foram o *Sleep 108*: O meu sono foi agitado; *Sleep 67*: Fiquei preocupado(a) por não conseguir adormecer; *Sleep 70*: Senti tristeza ao deitar; *Sleep 72*: Fiz um grande esforço para adormecer; e *Sleep 93*: Tive medo de não conseguir voltar a dormir depois de ter acordado. Os itens do Distúrbio da Vigília foram o *Sleep 06*: Tive sono durante o dia; *Sleep 07*: Tive dificuldade em ficar acordado(a) durante o dia; *Sleep 10*: Tive dificuldade em fazer coisas porque estava com sono; e *Sleep 30*: Senti que estava irritado(a) porque dormi mal. O valor máximo detectado para o efeito piso no banco Distúrbio do Sono foi no item *Sleep 72*, com 58,4%, e no banco Distúrbio da Vigília no item *Sleep 07*, com 60,6%. Entretanto, há na literatura frequências de até 70% para os efeitos piso e teto, pois suas consequências devem ser consideradas a partir do tipo de estudo e das características da amostra em questão²⁸.

Pode-se observar que nesses itens as pessoas responderam no nível 1 (nem um pouco) nos dois bancos de itens, o que significa que os participantes não apresentavam distúrbios nem do sono nem da vigília. Portanto, o efeito piso está diretamente relacionado à boa qualidade do sono, pois esta é melhor com a diminuição da pontuação. Dessa forma, os instrumentos cumprem o seu objetivo de discriminar os respondentes que tiveram escores mais altos como sendo os que apresentam diagnóstico de distúrbios relacionados ao sono-vigília.

Ao avaliar a confiabilidade da consistência interna do item verificou-se uma excelente confiabilidade, isto significa que os instrumentos são homogêneos e os itens medem o mesmo construto sono-vigília¹⁸. A confiabilidade teste-reteste ou reprodutibilidade da versão brasileira foi considerada excelente, de acordo com o valor de referência, e indica que os instrumentos apresentam boa estabilidade temporal.

A validade da TCT e a unidimensionalidade da TRI foram testadas por meio dos índices de qualidade de ajuste absoluto e incremental da AFC, os quais foram bons e indicam um bom ajuste para os

dois bancos. No banco Distúrbio do Sono (RMSEA = 0,099 e SRMR = 0,066) e no banco Distúrbio da Vigília (RMSEA = 0,110 e SRMR = 0,081) verificou-se que houve um bom ajuste, pois a amostra foi inferior a 1.000 participantes e os instrumentos com pouco número de itens¹³. Quando se compara com a validação dos instrumentos originais, banco Distúrbio do Sono (RMSEA = 0,140 e SRMR = 0,086) e banco Distúrbio da Vigília (RMSEA = 0,157 e SRMR = 0,820), pode-se concluir que houve um bom ajuste dos instrumentos para a população brasileira nos índices de ajuste absoluto. Em relação aos índices de ajuste incremental, banco Distúrbio do Sono (CFI = 0,970 e NNFI = 0,970) e banco Distúrbio da Vigília (CFI = 0,940 e NNFI = 0,930), tais índices estão bem próximos de 1, o que indica qualidade de ajuste muito boa para a população brasileira¹³. Na validação dos instrumentos originais, os índices do banco Distúrbio do Sono (CFI = 0,843 e NNFI = 0,957) e do banco Distúrbio da Vigília (CFI = 0,812 e NNFI = 0,955) evidenciam que os resultados encontrados para a amostra brasileira são adequados. Dessa maneira, pode-se concluir que os instrumentos são válidos, isto é, realmente medem o construto sono-vigília que se propuseram a medir, e também são suficientemente unidimensionais, ou seja, há apenas um traço latente dominante responsável pelo desempenho do teste e este fator dominante que foi medido.

Através da análise dos resíduos da matriz de correlação da AFC, por meio do coeficiente de correlação de Spearman, apurou-se que 19 pares de itens do banco Distúrbio do Sono e cinco pares do Distúrbio da Vigília apresentaram $\rho > 0,30$, o que são menos de 3% e 2% dos pares de itens, respectivamente. Mas segundo Buysse et al.¹, esse valor deve ser menos restritivo para os instrumentos Distúrbio do Sono e Distúrbio da Vigília e, por isto, considera-se $\rho > 0,50$. Dessa forma, tanto o banco Distúrbio do Sono como o banco Distúrbio da Vigília não registraram pares de itens localmente dependentes, o que mostra que os instrumentos apresentam um bom ajuste e cumprem com mais um dos pressupostos da TRI.

Pela análise da curva característica de cada item observou-se que o pressuposto da monotonicidade foi devidamente atendido tanto para o banco Distúrbio do Sono quanto para o banco Distúrbio da Vigília, pois foi constatado que todos os itens apresentaram estimativas do parâmetro *a* (discriminação) positivas, em ambos os bancos. Três itens do banco Distúrbio do Sono, *Sleep 106*, *Sleep 107* e *Sleep 42*, apresentaram valores de *a* considerados moderados, com boa discriminação; quatro itens, *Sleep 116*, *Sleep 50*, *Sleep 65* e *Sleep 72*, valores muito bons, com alto poder de discriminação; e vinte itens, valores excelentes, com muito alto poder de discriminação. No banco Distúrbio da Vigília, o *Sleep 119* apresentou discriminação muito baixa; o *Sleep 19* registrou baixa discriminação; três itens, *Sleep 04*, *Sleep 06* e *Sleep 123*, apresentaram valores considerados moderados, com boa discriminação; três itens, *Sleep 18*, *Sleep 120* e *Sleep 124*, valores muito bons, com alto poder de discriminação; e oito itens, valores excelentes, com muito alto poder de discriminação. Como todos os itens registraram estimativas do parâmetro *a* positivas, houve uma convergência e formaram-se as curvas, apenas dois itens do banco Distúrbio da Vigília, *Sleep 19* e *Sleep 119*, registraram baixa e muito baixa discriminação, respectivamente, e apresentaram curvas com baixa inclinação. Pôde-se verificar também que a curva total de informação do teste, que representa o somatório das informações de todos os itens, teve os valores na escala de aptidão da variável latente sono-vigília entre -1,5 e 3 e -1 e 3, para o banco Distúrbio do Sono e para o banco Distúrbio da Vigília, respectivamente. Ambas as curvas estão posicionadas mais para a direita, o que evidencia que os itens funcionam melhor para as pessoas com boa qualidade de sono-vigília do que para aquelas que apresentam distúrbios do sono e da vigília.

Por meio da calibração pôde-se verificar que as propriedades psicométricas dos itens foram satisfatórias, como também extrair informações sobre os que mais discriminaram e os que registraram estimativas do parâmetro de dificuldade (*b*) maior. O parâmetro *b* posiciona os itens ao longo do construto e corresponde ao valor do traço latente para o qual a probabilidade de responder a uma determinada categoria de resposta do item é de 50%. Neste estudo, pode-se afirmar que os itens realizaram uma boa cobertura do continuum do traço latente sono-vigília, uma vez que os valores de *b* variaram de -3,22 a 4,46 para o banco de itens Distúrbio do Sono e de -10,38 a 4,61 para o banco Distúrbio da Vigília.

Também por meio da calibração dos itens foi possível fazer um paralelo entre as estimativas de parâmetros geradas pela população brasileira e a americana. Dos 27 itens do banco Distúrbio do Sono, 21 estão na mesma categoria de discriminação tanto na amostra brasileira quanto na americana, e no banco Distúrbio da Vigília, dos 16 itens, nove estão na mesma categoria. Por exemplo, o item *Sleep*

125: Senti que estava péssimo(a) ao acordar, e o *Sleep* 65: Senti que estava fisicamente tenso(a) ao deitar, obtiveram o mesmo valor nas duas amostras. Assim, verifica-se que os itens apresentaram o mesmo padrão de comportamento nas duas populações, e isto reflete o rigor do processo de tradução e adaptação transcultural para a língua portuguesa falada no Brasil. Pode-se perceber, então, que houve um equilíbrio entre os instrumentos brasileiro e americano, o que confirma a calibração dos itens dos dois bancos na população brasileira (Tabelas 2 e 3).

Os bancos Distúrbio do Sono e Distúrbio da Vigília apresentaram homogeneidade para os diferentes grupos testados, homem *versus* mulher, saudável *versus* portador de doença crônica e língua portuguesa *versus* língua inglesa. Pois, apesar da presença de DIF em alguns itens, ele mostrou-se insignificante, já que os valores do pseudo R^2 de Nagelkerke foram inferiores a 0,13. Os métodos para a detecção e mensuração do impacto do Funcionamento Diferencial do Item são vários e, portanto, não há um consenso na literatura, já que cada método produz um resultado ²⁵.

Pode-se concluir que a versão para a língua portuguesa falada no Brasil do banco de itens Distúrbio do Sono e do banco de itens Distúrbio da Vigília apresenta adequadas propriedades psicométricas avaliadas tanto pela TCT quanto pela TRI, e mostrou ser uma medida confiável, precisa e válida, com análises psicométricas robustas. Portanto, poderá ser utilizada em pesquisas e estudos clínicos, como também em diagnósticos clínicos e tratamentos de doenças crônicas.

Colaboradores

Z. M. S. Silva e Costa foi responsável pelo trabalho de campo, análise do material, levantamento bibliográfico e redação do texto do artigo. R. M. C. Pinto foi responsável por importantes contribuições para a discussão, revisão do conteúdo intelectual e referências bibliográficas do artigo. T. M. S. Mendonça e C. H. M. Silva contribuíram na revisão do conteúdo e referências bibliográficas do artigo. Todos os autores aprovaram a versão final do artigo.

Informações adicionais

ORCID: Zilma Maria Severino Silva e Costa (0000-0001-6585-1365); Rogério de Melo Costa Pinto (0000-0002-3397-5803); Tânia Maria da Silva Mendonça (0000-0001-8392-672X); Carlos Henrique Martins da Silva (0000-0001-8634-6878).

Agradecimentos

À Fundação de Amparo a Pesquisa do Estado de Minas Gerais e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo apoio financeiro disponibilizado. À Zhamayra de Azambuja pela contribuição na execução dos softwares.

Referências

1. Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep* 2010; 33:781-92.
2. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev* 2002; 6:97-111.
3. Ohayon M.M. From wakefulness to excessive sleepiness: what we know and still need to know. *Sleep Med Rev* 2008; 12:129-41.
4. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989; 28:193-213.
5. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991; 14:540-5.
6. Bertolazi AN. Tradução, adaptação cultural e validação de dois instrumentos de avaliação do sono: escala de sonolência de Epworth e índice de qualidade de sono de Pittsburgh. <https://www.lume.ufrgs.br/bitstream/handle/10183/14041/000653543.pdf> (acessado em 10/Jul/2016).
7. Miletin MS, Hanly PJ. Measurement properties of the Epworth sleepiness scale. *Sleep Med* 2003; 4:195-9.
8. Cole JC, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh sleep quality index in older adults. *Sleep* 2006; 29:112-6.
9. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res* 2007; 16:133-41.

10. Pasquali L. TRI – Teoria de resposta ao item: teoria, procedimentos e aplicações. Brasília: Laboratório de Avaliação e Instrumentação Psicológica, Universidade de Brasília; 2007.
11. Nunnally JC, Bernstein IH. Psychometric theory. 3rd Ed. New York: McGraw-Hill; 1994.
12. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods* 1999; 4:84-99.
13. Hair Jr. JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Análise multivariada de dados. 6a Ed. Porto Alegre: Bookman; 2009.
14. Costa ZMSS, Pinto RMC, Mendonça TMS, Silva CHM. Tradução e adaptação cultural para a língua portuguesa dos domínios distúrbios do sono e distúrbios da vigília do *Patient-Reported Outcomes Measurement Information System* (PROMIS). *Cad Saúde Pública* 2014; 30:1391-401.
15. McHorney CA, Ware Jr. JE, Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994; 32:40-66.
16. Cramer JA. Principles of health-related quality of life: assessment in clinical trials. *Epilepsia* 2002; 43:1084-95.
17. Nagl M, Gramm L, Heyduck K, Glattacker M, Farin E. Development and psychometric evaluation of a German version of the PROMIS item banks for satisfaction with participation. *Eval Health Prof* 2015;38:160-80.
18. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16:297-334.
19. Pasquali L. *Psicometria*. *Rev Esc Enferm USP* 2009; 43:992-9.
20. Samejima F. Estimation of latent ability using a response pattern of graded scores. <https://www.psychometricsociety.org/sites/default/files/pdf/MN17.pdf> (acessado em 25/Jan/2017).
21. Baker FB. The basics of item response theory. <http://echo.edres.org:8080/irt/baker/final.pdf> (acessado em 21/Mai/2019).
22. Bortolotti SLV, Tezza R, Andrade DF, Bornia AC, Sousa AF. Relevance and advantages of using the item response theory. *Quality & Quantity* 2013; 47:2341-60.
23. Ladwig R. Detecção de funcionamento diferencial do item através da regressão logística e da teoria da resposta ao item: uma interface gráfica. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2012.
24. Paz SH, Spritzer KL, Morales LS, Hays RD. Age-related differential item functioning for the patient-reported outcomes information system (PROMIS) physical functioning items. *Prim Health Care* 2013; 3:131.
25. Wong AWK, Lai JS, Correia H, Cella D. Evaluating psychometric properties of the Spanish-version of the pediatric functional assessment of chronic illness therapy-perceived cognitive function (pedsFACIT-PCF). *Qual Life Res* 2015; 24:2289-95.
26. Choi SW, Gibbons LE, Crane PK. Lordif: an R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *J Stat Softw* 2011; 39:1-30.
27. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health* 2015; 18:25-34.
28. Hyland ME. A brief guide to the selection of quality of life instrument. *Health Qual Life Outcomes* 2003; 1:24.

Abstract

This study aimed to validate the Brazilian version of the item banks for Sleep Disturbance and Wake Disturbance from the Patient-Reported Outcomes Measurement Information System (PROMIS). Validation of these item banks used classical test theory, item response theory, calibrations of items according to the gradual response model proposed by Samejima, and analysis of Differential Item Functioning (DIF), with a sample consisting of 627 participants over 18 years of age. As for data quality, there were no missing data, five items from the Sleep Disturbance bank and four items from the WE bank showed a floor effect, and none of the banks showed a ceiling effect. The two banks displayed excellent reliability with homogeneous scale and good temporal stability. With confirmatory factor analysis, using absolute and incremental goodness-of-fit indices, both banks displayed good fit, showing that the instruments are valid and sufficiently unidimensional. The calibration showed that the items' psychometric properties were satisfactory with good coverage of the sleep-wake construct. As for DIF, both the Sleep Disturbance and Wake Disturbance banks proved to be homogeneous for the different test groups. In conclusion, the version in Brazilian Portuguese for the item banks on Sleep Disturbance and Wake Disturbance in the PROMIS proved to be a reliable, precise, and valid instrument with robust psychometric analyses.

Validation Study; Sleep Wake Disorders; Quality of Life

Resumen

El objetivo de este estudio fue validar la versión brasileña de los bancos de ítems Disturbio del Sueño y Disturbio de la Vigilia del Patient-Reported Outcomes Measurement Information System (PROMIS). La validación de esos bancos de ítems se realizó mediante teoría clásica de los tests, teoría de respuesta al ítem, calibración de los ítems, según el modelo de respuesta gradual propuesto por Samejima, así como el análisis del Funcionamiento Diferencial del Ítem (DIF), estando la muestra constituida por 627 participantes con más de 18 años. En relación con la calidad de los datos, se observó que no hubo datos perdidos y cinco ítems del banco Disturbio del Sueño y cuatro ítems del banco Disturbio de la Vigilia tuvieron efecto suelo, además ninguno de los bancos presentó un efecto techo. Los dos bancos mostraron excelente confiabilidad con una escala homogénea y buena estabilidad temporal. Con el análisis factorial confirmatorio, usándose los índices de calidad de ajuste absoluto e incremental, se verificó un buen ajuste para los dos bancos, lo que revela que los instrumentos son válidos y suficientemente unidimensionales. La calibración evidenció que las propiedades psicométricas de los ítems fueron satisfactorias con una buena cobertura del constructo sueño-vigilia. En lo que se refiere al DIF, tanto el banco Disturbio del Sueño como el Disturbio de la Vigilia mostraron ser homogéneos para los diferentes grupos testados. Se puede concluir que la versión para la lengua portuguesa hablada en Brasil de los bancos de ítems Disturbio del Sueño y Disturbio de la Vigilia del PROMIS demostró ser una medida confiable, precisa y válida, con análisis psicométricos robustos.

Estudio de Validación; Trastornos del Sueño-Vigilia; Calidad de Vida

Recebido em 28/Nov/2019
Versão final reapresentada em 28/Mar/2020
Aprovado em 02/Abr/2020