

Confiabilidade e validade do *BESTest* e do *MiniBESTest* em hemiparéticos crônicos

Reliability and validity of the BESTest and MiniBESTest in chronic hemiparesis

Carla Bambirra¹, Lívia de Castro Magalhães², Fátima Rodrigues-de-Paula³

RESUMO

Objetivo. Avaliar a confiabilidade e validade das versões brasileiras do BESTest e MiniBESTest em hemiparéticos crônicos. **Método.** As confiabilidades teste-reteste e interexaminadores foram avaliadas pelo coeficiente kappa ponderado. A validade de construto foi analisada através do modelo Rasch em 40 hemiparéticos ($58,8 \pm 12,8$ anos). **Resultados.** As confiabilidades teste-reteste e interexaminadores apresentaram concordância de moderada a quase perfeita para o BESTest ($0,48 \leq Kw \leq 1,0$; $p < 0,05$) e forte a quase perfeita para o MiniBESTest ($0,62 \leq Kw \leq 1,0$; $p < 0,05$). Os itens dos instrumentos dividiram as pessoas em, aproximadamente, três níveis de habilidade. A estabilidade das medidas dos indivíduos e dos itens variou de 0,89 a 0,96 para os instrumentos. No BESTest, o item mais difícil foi “sentar no chão e levantar” e no MiniBESTest “correção com passo lateral à direita”. O item mais fácil para ambos foi “permanecer de olhos abertos, superfície firme”. O nível de dificuldade dos itens dos instrumentos foi apropriado para o nível de habilidade dos indivíduos. Dois itens do BESTest apresentaram padrão errático. No MiniBESTest não houve item errático. **Conclusões.** As versões brasileiras do BESTest e do MiniBESTest são confiáveis e válidas para a avaliação do equilíbrio em hemiparéticos crônicos. Recomenda-se atenção quanto aos itens erráticos e cautela na interpretação do escore total do BESTest.

Unitermos. Equilíbrio Postural, Avaliação da Deficiência, Paresia, Acidente Vascular Cerebral, Psicometria

Citação. Bambirra C, Magalhães LC, Rodrigues-de-Paula F. Confabilidade e validade do BESTest e do MiniBESTest em hemiparéticos crônicos.

Trabalho realizado na Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brasil.

Suporte financeiro. CNPq e Fapemig

ABSTRACT

Objective. To evaluate the reliability and construct validity of the Brazilian versions of the BESTest and MiniBESTest in chronic hemiparesis. **Method.** The test-retest and interrater reliability were assessed using the weighted kappa coefficient. The Rasch model was used for analysis of construct validity in 40 hemiparetic subjects (58.8 ± 12.8 years). **Results.** Moderate to almost perfect agreement was observed for the BESTest for test-retest and interrater ($0.48 \leq Kw \leq 1.0$; $p < 0.05$). Strong to almost perfect agreement was observed for Mini-BESTest for test-retest and interrater reliability ($0.62 \leq Kw \leq 1.0$; $p < 0.05$). The items of the BESTest and MiniBESTest divided people in about three skill levels. The individuals and items stability measurement ranged from 0.89 to 0.96 for the instruments. In BESTest, the most difficult item was “sit on the floor and stand up” and in Mini-BESTest “compensatory stepping correction lateral to the right.” The easier item for both instruments was “stand up on firm surface, eyes open.” The level of difficulty of the items of both instruments was appropriate for the skill level of individuals. Two items of the BESTest were erratic. In MiniBESTest were not observed erratic items. **Conclusion.** The Brazilian versions of the BESTest and MiniBESTest are reliable and have adequate construct validity to evaluate balance in chronic hemiparetic individuals. It is recommended, however, caution in interpreting the total score of the BESTest.

Keywords. Postural Balance, Disability Evaluation, Paresis, stroke, Psychometrics

Citation. Bambirra C, Magalhães LC, Rodrigues-de-Paula F. Reliability and validity of the BESTest and MiniBESTest in chronic hemiparesis.

Endereço para correspondência:

Fátima Rodrigues-de-Paula
Universidade Federal de Minas Gerais
Departamento de Fisioterapia
Avenida Antônio Carlos, 6627.
CEP 31210-901, Belo Horizonte-MG
e-mail: fatimarp@globo.com ; fatimarp@ufmg.br

1. Fisioterapeuta, Centro de Reabilitação Noroeste da Rede do Sistema Único de Saúde, Belo Horizonte-MG, Brasil.

2. Terapeuta Ocupacional, Professora Titular do Departamento de Terapia Ocupacional, UFMG, Belo Horizonte-MG, Brasil.

3. Fisioterapeuta, Doutora, Professora Associada do Departamento de Fisioterapia, UFMG, Belo Horizonte-MG, Brasil.

Original

Recebido em: 07/06/14

Aceito em: 03/02/15

Conflito de interesses: não

INTRODUÇÃO

O equilíbrio é um construto complexo que envolve a contribuição de vários sistemas fisiológicos¹. As desordens do equilíbrio podem ser resultantes do envelhecimento ou de doenças, tais como o acidente vascular cerebral (AVC)². Os déficits de equilíbrio nos indivíduos com hemiparesia decorrente de AVC estão relacionados a alto risco de quedas³. As consequências das quedas após AVC envolvem limitação da atividade, restrição da participação social, aumento da dependência e desenvolvimento do medo de cair^{3,4}.

A avaliação do equilíbrio é importante tanto para o diagnóstico das desordens do equilíbrio quanto para o planejamento de intervenções¹. Instrumentos utilizados para a avaliação do equilíbrio em indivíduos com hemiparesia, como a Escala de Equilíbrio de *Berg* e a Escala de *Fugl-Meyer* - seção equilíbrio, apresentam efeito teto e solo nesta população^{5,6}, o que pode limitar a abrangência dos indivíduos avaliados e a acurácia da avaliação.

O *Balance Evaluation Systems Test (BESTest)* é um instrumento que tem como objetivo ajudar na identificação dos subsistemas do controle postural que podem ser responsáveis pela alteração do equilíbrio funcional para que o tratamento possa ser direcionado aos subsistemas mais envolvidos no *déficit* do equilíbrio¹. O *BESTest* consiste de 36 tarefas organizadas em 27 itens, os quais são agrupados em seis seções: restrições biomecânicas, limites de estabilidade/verticalidade, ajustes posturais antecipatórios, respostas posturais reativas, orientação sensorial e estabilidade na marcha com e sem uma tarefa cognitiva. Cada item é pontuado em uma escala ordinal de quatro pontos: de zero (pior desempenho) a três (melhor desempenho)¹. A validade concorrente foi avaliada pela correlação entre o *BESTest* e a *Activities-specific Balance Confidence* em indivíduos com condições neurológicas diversas e indivíduos sem alterações neurológicas, com coeficiente de correlação de *Spearman* de 0,64¹.

Uma versão reduzida do *BESTest* (*MiniBESTest*) foi desenvolvida com o intuito de facilitar a utilização do instrumento na prática clínica⁷. O *MiniBESTest* requer um tempo de 10-15 minutos para ser administrado e apresentou excelentes propriedades psicométricas em indivíduos com diagnósticos neurológicos diversos⁷ e com doença de Parkinson^{8,9}. Recentemente, o *BESTest* e o *Mi-*

niBESTest foram traduzidos e adaptados transculturalmente para o português-Brasil em um estudo com idosos e indivíduos com doença de Parkinson, apresentando evidência de adequada validade de construto e estabilidade das respostas nestas populações¹⁰.

A confiabilidade de uma medida define sua reproduzibilidade e a validade de construto determina a habilidade do instrumento em medir um conceito abstrato¹¹. Na validação de novos instrumentos, há uma tendência crescente para a utilização do modelo *Rasch*⁷, por este modelo fornecer uma representação mais acurada do processo de medição¹². Dessa forma, este estudo objetivou avaliar a confiabilidade e a validade do *BESTest* e do *MiniBESTest* em indivíduos com hemiparesia, utilizando o modelo *Rasch* para avaliação da validade de construto destes instrumentos.

MÉTODO

Estudo metodológico realizado no Centro de Reabilitação Noroeste na cidade de Belo Horizonte, Minas Gerais, Brasil, durante o período de abril a setembro de 2012. O projeto foi aprovado pelos comitês de Ética da Universidade Federal de Minas Gerais e da Secretaria Municipal de Saúde de Belo Horizonte (parecer nº 0680.0.203.410-11).

Amostra

A amostra foi de conveniência, composta por indivíduos com hemiparesia crônica. O cálculo amostral baseou-se na recomendação de 10 indivíduos para cada categoria de escore do instrumento¹³⁻¹⁵, sendo necessários, dessa forma, 40 indivíduos para obter calibrações estáveis dos itens na avaliação do *BESTest*. Para o *MiniBESTest*, seriam necessários 30 indivíduos hemiparéticos.

Os participantes foram recrutados a partir de convite oral, sendo incluídos no estudo aqueles que atenderam aos seguintes critérios: diagnóstico médico prévio de AVC; tempo de evolução após AVC de, pelo menos, seis meses; idade igual ou superior a 20 anos; de ambos os性os; com hemiparesia caracterizada por escore inferior a 30, em um total de 34 pontos, na escala de avaliação dos membros inferiores de *Fugl-Meyer* (EFM)¹⁶ e capazes de deambular seis metros sem assistência humana, com ou

sem utilização de órtese ou dispositivo de auxílio.

Foram excluídos os indivíduos que apresentaram: afasia de compreensão, definida pela resposta inadequada ao comando: “levante seu braço não comprometido e abra sua mão não comprometida”¹⁷; outras doenças neurológicas; déficit visual não corrigido; alterações músculo-esqueléticas não relacionadas ao quadro de AVC com presença de dor e/ou instabilidade clínica que contraindicasse a realização do protocolo proposto. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

Procedimento

Dados antropométricos, clínicos e demográficos foram coletados por meio da aplicação de questionário previamente elaborado. Os indivíduos foram caracterizados quanto ao nível de atividade física segundo recomendações estabelecidas pelo *Physical Activity Trends/ USA/1990-1998*¹⁸. A EFM foi utilizada para caracterizar o retorno da função motora dos membros inferiores¹⁹.

As versões brasileiras do *BESTest* e do *MiniBESTest* foram aplicadas conforme recomendações prévias¹⁰. Antes do início da coleta foi realizado o treinamento das examinadoras através das instruções padronizadas e do vídeo interativo e educacional *BESTest DVD-ROM*. O *BESTest* e o *MiniBESTest* foram pontuados simultaneamente de acordo com o critério de cada um dos instrumentos. Pequenas adaptações foram necessárias em relação ao posicionamento do membro superior mais afetado de alguns indivíduos, que não apresentavam movimentação ativa neste membro. Nos itens realizados normalmente com os braços cruzados (sentado para de pé, inclinação lateral e verticalidade) estes indivíduos mantinham a mão menos afetada segurando a mais afetada junto ao tronco e em itens realizados com as mãos na cintura, o membro superior plégico foi mantido em sua postura habitual. No item “levantar o braço” o membro superior plégico era posicionado apoiado no halter, de forma a acompanhar o movimento de elevação realizado pelo outro membro.

Análise estatística

Confiabilidade

Para o estudo da confiabilidade participaram um total de 20 hemiparéticos, sendo 10 indivíduos partici-

pantes da análise da confiabilidade teste-reteste e 10 da análise da confiabilidade interexaminadores. A confiabilidade teste-reteste foi realizada por meio da aplicação do *BESTest* e *MiniBESTest* pelo mesmo avaliador, por duas vezes, num intervalo de sete dias. A confiabilidade interexaminadores foi realizada por meio da aplicação do *BESTest* e do *MiniBESTest* por duas examinadoras, as quais alternaram as instruções ao participante e observaram simultaneamente. Não houve nenhum tipo de comunicação entre as observadoras durante a pontuação dos instrumentos, de forma que uma examinadora não tinha conhecimento da pontuação da outra.

Como o *BESTest* é composto por uma escala ordinal com mais de três categorias de resposta foi utilizado o coeficiente *Kappa* ponderado para medida de confiabilidade teste-reteste e interexaminadores, sendo realizada análise para cada item²⁰. O coeficiente *Kappa* ponderado também foi utilizado para as medidas de confiabilidade do *MiniBESTest*. A interpretação dos resultados do *Kappa* ponderado foi baseada nos seguintes pontos de corte: ≤ 0 = concordância pobre; 0,01 a 0,20 = fraca; 0,21 a 0,40 = razoável; 0,41 a 0,60 = moderada; 0,61 a 0,80 = forte; e 0,81 a 1,0 = quase perfeita²⁰. O pacote estatístico *StatsDirect*, versão 2.7.2 foi utilizado para a análise das confiabilidades.

Análise Rasch

As versões brasileiras do *BESTest* e do *MiniBESTest* foram avaliadas com o uso do modelo *Rasch*, por meio do programa computadorizado específico *Winsteps*, versão 3.74.0/2012²¹. O modelo *Rasch* é um modelo probabilístico baseado no princípio de que quanto mais hábil é a pessoa, mais chances ela terá de obter sucesso em qualquer item, e quanto mais difícil o item, menores são as chances de qualquer pessoa obter sucesso¹². No modelo *Rasch*, quando os itens de um instrumento medem o mesmo construto, é possível alinhar a dificuldade dos itens e o nível de habilidade dos indivíduos em um contínuo linear simples, dividido em intervalos iguais (*logits*) que podem ser utilizados para medir a habilidade das pessoas naquela dimensão^{12,22}. No presente estudo, ao se avaliar o equilíbrio através do *BESTest* e do *MiniBESTest*, espera-se que quanto melhor a habilidade de equilíbrio do indivíduo, maior a probabilidade do mesmo em rece-

ber escores altos em todos os itens das escalas. Por outro lado, quanto mais fácil o item, maior a probabilidade de qualquer pessoa receber escore alto no item.

Quando os indivíduos e os itens podem ser calibrados em um mesmo contínuo, o modelo permite determinar se o nível de dificuldade dos itens é apropriado para o nível de habilidade dos indivíduos. Além disso, os itens devem ser distribuídos de forma que cubram toda a variação de habilidade de equilíbrio apresentada pelos indivíduos, sem a presença de efeito teto e solo¹².

O programa *Winsteps* oferece parâmetros como o MnSq (*goodness-of-fit*) e o valor z, que expressam a relação entre o escore esperado e o escore obtido. Se a relação está de acordo com os pressupostos do modelo, o MnSq tem valor igual a 1,0, com variação de $\pm 0,3$ ou $\pm 0,412$. Valor muito alto de MnSq ($>1,3$) acompanhado por valor z $> 2,0$, sinaliza itens erráticos ou que não medem a mesma dimensão. Valores de MnSq muito baixos ($<0,7$) associados a z $< -2,0$ sugerem pouca variabilidade de escores e indicam que o item tem pouco potencial para discriminar diferentes níveis de habilidade¹². Os itens erráticos são aqueles em que, inesperadamente, pessoas com pior habilidade receberam escores altos ou vice-versa. Como o escore errático indica maior problema na definição do item, que pode comprometer o conceito de construto unidimensional, foram assinalados apenas os itens com valores de MnSq altos em seus dois formatos, *Infit* e *Outfit*, que sinalizam flutuações nas pontuações próximas ao nível de habilidade do indivíduo ou do nível de dificuldade do item e a presença de escores extremos, respectivamente^{12,15}. A literatura sugere que quando cerca de mais de 5% do número total de itens não se enquadram no modelo, isso é uma indicação de que os itens da escala não combinam para medir um conceito unidimensional¹⁵. Caso contrário, pode-se inferir que os itens em conjunto colaboram para definir uma dimensão ou construto^{10,15}. Como o *BESTest* tem 36 tarefas, espera-se que não mais de 1,8 ou, arredondando, dois itens não se enquadrem no modelo. Para o *MiniBESTest*, que contém 16 tarefas, espera-se que não mais de 0,8, ou um item, não se enquadre no modelo. O mesmo raciocínio se aplica ao padrão de respostas dos indivíduos avaliados. Outro dado a ser considerado é o resultado de análise factorial (componente principal), incluída no programa *Winsteps*. A variância

explicada pelas medidas $> 50\%$ e eigenvalue da variância do primeiro resíduo < 2 dão suporte a unidimensionalidade da escala²².

O modelo *Rasch* também reporta o erro associado à calibração dos itens e dos indivíduos, que informa a precisão das medidas obtidas. A partir dessa margem de erro obtém-se o “índice de separação” que é utilizado para estimar em quantos níveis de habilidade os itens dividem a amostragem, o que pode ser calculado com uso da fórmula: $H = (4G + 1)/3$, onde H = número de estratos e G = índice de separação¹². Espera-se que um teste divida os participantes em pelo menos três níveis de habilidade¹². Outros índices relacionados são os coeficientes de confiabilidade das medidas das pessoas e da calibração dos itens, os quais fornecem o grau de consistência ou estabilidade das estimativas, com variação de zero a um¹². Os coeficientes maiores que 0,80 são considerados bons e maiores que 0,90 excelentes⁷.

RESULTADOS

A Tabela 1 apresenta as características antropométricas, demográficas e clínicas da amostra. Participaram do estudo 40 indivíduos com hemiparesia crônica. Destes, 18 indivíduos (45%) faziam uso de bengala (canadense ou simples), 35% eram casados, 10% viúvos, 17,5% divorciados ou separados e 37,5% solteiros. Um indivíduo era analfabeto, 18 indivíduos tinham até quatro anos de estudo, 11 estudaram de cinco a oito anos, sete concluíram o ensino médio e três concluíram o ensino superior.

Confiabilidade

Os valores do coeficiente *kappa* ponderado quadrático foram adequados para ambos os instrumentos e estão descritos na Tabela 2. Foi observada concordância de moderada a quase perfeita para o *BESTest* e de forte a quase perfeita para o *MiniBESTest*, tanto para a confiabilidade teste-reteste quanto para a interexaminadores.

Análise *Rasch*

No *BESTest*, o valor do coeficiente de confiabilidade da medida dos indivíduos foi 0,94 e a estimativa da estabilidade de calibração dos itens foi de 0,96. No

Tabela 1. Média, desvio padrão ou distribuição de frequência (%) das características dos participantes do estudo (n = 40).

Variável	Valor
Idade (anos)	58,8±12,8 (26 a 87)
Sexo	
Masculino	21 (52,5%)
Feminino	19 (47,5%)
IMC	25,8±4,3(15,2 a 35,5)
Escore na escala de membros inferiores da EFM	20,7±7,8 (7 a 30)
Nível de atividade física	
Vigoroso	0
Moderado	5 (12,5%)
Insuficiente	3 (7,5%)
Inativo	32 (80%)
Tempo de evolução pós-AVC (meses)	80,6±76,6 (6 a 399)
Tipo de AVC	
Hemorrágico	10 (25%)
Isquêmico	29 (72,5%)
Não informado	1 (2,5%)
Hemiparesia	
Direita	23 (57,5%)
Esquerda	17 (42,5%)

IMC = Índice de massa corporal; EFM = Escala dos membros inferiores de Fugl-Meyer; AVC = Acidente vascular cerebral.

MiniBESTest, o valor do coeficiente de confiabilidade das medidas dos indivíduos foi 0,89 e da calibração dos itens foi 0,94. O índice de separação dos indivíduos foi de 4,10 para o *BESTest* e de 2,91 para o *MiniBESTest*, indicando que os itens dividiram as pessoas em 5,8 e 4,21 níveis de habilidade, respectivamente. O índice de separação dos itens foi 5,16 no *BESTest* e 4,01 no *MiniBESTest*, dividindo os itens do *BESTest* em 7,21 níveis de dificuldade e do *MiniBESTest* em 5,68 níveis de dificuldade.

A calibração dos itens para o *BESTest* e o *MiniBESTest* está apresentada em ordem decrescente de dificuldade nas Tabelas 3 e 4, respectivamente. Os itens do *BESTest* cobrem 4,97 *logits*, a medida média dos indiví-

duos ficou 0,39 *logits* acima da média dos itens e o item mais difícil foi “sentar no chão e levantar”. No *MiniBESTest* os itens cobrem 9,07 *logits*, a medida média dos pacientes ficou apenas 0,17 *logits* abaixo da média dos itens e “correção com passo lateral à direita” foi o item mais difícil. O item mais fácil para ambos os instrumentos foi “permanecer de olhos abertos em superfície firme”. Em geral, as seções do *BESTest* apresentaram combinação de itens, distribuídos ao longo do continuo de equilíbrio, com diferentes níveis de dificuldade.

Dos 36 itens do *BESTest*, dois foram erráticos (não se encaixaram nas expectativas do modelo): “força lateral de quadril e tronco” e “alcance funcional lateral à esquerda”. Somente um indivíduo apresentou padrão errático de resposta no *BESTest*. No *MiniBESTest* não foram observados itens nem indivíduos com padrão errático.

A análise fatorial do *BESTest* indica que, embora a variância explicada pelas medidas seja de 56,5%, o *eigenvalue* do primeiro resíduo é de 3,9, o que indica que o teste tem mais de uma dimensão. Foi feita análise subsequente das subescalas, mas nenhuma delas apresentou qualidades psicométricas adequadas como escala independente. No *MiniBESTest*, a variância explicada pelas medidas foi de 53,5% e o *eigenvalue* de 3,3. A análise subsequente indicou que, embora seja possível criar uma escala unidimensional com apenas 11 itens, isso excluiria itens que tem relevância clínica (“sentado para de pé” e os itens de correção com passo compensatório à frente, para trás, lateral à direita e à esquerda).

A Figura 1 representa os contínuos do nível de dificuldade dos itens e de habilidade dos indivíduos do *BESTest* e do *MiniBESTest*. Não foram observados efeitos teto ou solo em ambos os instrumentos.

Tabela 2. Confiabilidade teste-reteste e interexaminadores dos instrumentos analisados.

Instrumentos	Confiabilidade	
	Teste-reteste	Interexaminadores
BESTest	0,55 ≤ K_w ≤ 1,0*	0,48 ≤ K_w ≤ 1,0*
MiniBESTest	0,62 ≤ K_w ≤ 1,0*	0,71 ≤ K_w ≤ 1,0*

K_w = coeficiente de kappa ponderado quadrático; *p<0,05.

Tabela 3. Calibração dos itens do BESTest.

Item	Seções do teste	Calibração	Erro	Infit		Outfit	
				MnSq	Z	MnSq	Z
5) Sentar no chão e levantar	RB	1,52	0,21	1,46	2,0	1,26	0,9
11.a) De pé em uma perna - E	TR	1,48	0,2	0,76	-1,2	1	0,1
11.b) De pé em uma perna - D	TR	1,48	0,2	1,16	0,8	1,87	2,5
4) Força lateral de quadril e tronco*	RB	1,43	0,2	1,49	2,2	2,95	4,8
18.b) Correção com passo lateral - D	RR	1,23	0,2	1,34	1,6	1,13	0,6
25) Passar sobre obstáculos	EM	1,15	0,2	0,67	-1,8	0,63	-1,6
12) Tocar degrau alternadamente	TR	1,11	0,2	0,39	-4	0,38	-3,2
18.a) Correção com passo lateral - E	RR	1,07	0,2	1,24	1,2	1,08	0,4
2) Alinhamento do centro de massa	RB	0,96	0,2	0,89	-0,5	0,8	-0,8
17) Correção com passo para trás	RR	0,92	0,2	1,27	1,3	1,17	0,8
27)TUG dupla tarefa	EM	0,92	0,2	0,54	-2,7	0,55	-2,2
3) Força e amplitude de tornozelo	RB	0,88	0,2	0,59	-2,3	0,66	-1,6
15) Resposta no lugar para trás	RR	0,84	0,2	1,48	2,1	1,39	1,6
16) Correção com passo compensatório	RR	0,84	0,2	1,37	1,7	1,21	0,9
19.d) Olhos fechados, espuma	OS	0,65	0,2	0,80	-1	0,89	-0,4
10) Ficar na ponta dos pés	TR	0,57	0,2	0,71	-1,5	0,71	-1,3
24) Andar e girar sobre o eixo	EM	0,38	0,2	0,59	-2,3	0,55	-2,3
26) TUG	EM	0,30	0,2	0,39	-3,8	0,37	-3,7
8.b) Alcance funcional lateral - D	LE	0,22	0,2	1,38	1,7	1,34	1,5
21) Marcha em superfície plana	EM	0,18	0,2	0,25	-5,2	0,29	-4,4
7) Alcance funcional para frente	LE	0,14	0,2	0,73	-1,3	0,77	-1
8.a) Alcance funcional lateral - E*	LE	0,14	0,2	1,5	2,1	1,54	2,1
23) Andar com viradas de cabeça	EM	-0,12	0,2	0,78	-1,0	0,82	-0,7
22) Mudança na velocidade da marcha	EM	-0,25	0,21	0,9	-0,4	0,93	-0,2
14) Resposta no lugar para frente	RR	-0,29	0,21	1,13	0,6	1,19	0,8
1) Base de apoio	RB	-0,38	0,21	1,47	1,9	1,67	2,4
6.b) Inclinação lateral - D	LE	-0,57	0,21	0,67	-1,6	0,77	-0,9
6.a) Inclinação lateral - E	LE	-0,60	0,22	1,35	1,4	1,32	1,3
19.c) Olhos abertos, espuma	OS	-0,83	0,22	1,40	1,6	1,20	0,8
20) Inclinação, olhos fechados	OS	-1,18	0,23	1,14	0,6	1,77	2,4
13) Levantar o braço	TR	-1,67	0,25	1,36	1,2	0,96	0
19.b)Olhos fechados superfície firme	OS	-1,75	0,28	0,9	-0,2	1,07	0,3
6.d) Verticalidade sentado - D	LE	-2,25	0,29	1,25	0,8	1,42	1
9) Sentado para de pé	TR	-2,25	0,34	1,25	0,8	0,74	-0,5
6.c)Verticalidade sentado - E	LE	-2,8	0,41	1,21	0,6	1,02	-0,2
19.a) Olhos abertos superfície firme	OS	-3,45	0,53	0,8	-0,3	0,75	-0,1

*Itens erráticos MnSq > 1,3 e Z > 2. D = direita; E = esquerda. TUG = “Timed up and go”; RB = Restrições biomecânicas; TR = Transições-Ajustes posturais antecipatórios; RR = Respostas posturais reativas; EM = Estabilidade na marcha; OS = Orientação sensorial; LE = Limites de estabilidade-Verticalidade.

Tabela 4. Calibração dos itens do *MINIBESTest*.

Item	Calibração	Erro	<i>Infit</i>		<i>Outfit</i>	
			MnSq	Z	MnSq	Z
6.b) Correção com passo lateral - D	1,15	0,29	1,22	1	0,96	0,1
3.a) De pé em uma perna - E	1,06	0,29	0,51	-2,8	0,76	0,98
3.b) De pé em uma perna - D	0,98	0,29	1,13	0,7	1,44	1,1
6.a) Correção com passo lateral - E	0,81	0,29	0,99	0	0,83	-0,3
13) Passar sobre obstáculos	0,81	0,29	0,65	-1,8	0,63	-1
2) Ficar na ponta dos pés	0,73	0,28	1,34	1,6	1,13	0,5
5) Correção com passo para trás	0,73	0,28	1,11	0,6	0,99	0,1
12) Andar e girar sobre o eixo	0,65	0,28	0,73	-1,4	0,69	-0,8
14) TUG dupla tarefa	0,57	0,28	0,63	-2	0,64	-1
4) Correção com passo compensatório	0,49	0,28	1,49	2	1	1
11) Andar com viradas de cabeça	-0,21	0,28	1,12	0,6	1,13	0,6
8) Olhos fechados, espuma	-1,13	0,3	1,01	0,1	0,97	0
10) Mudança na velocidade da marcha	-1,22	0,3	1,15	0,7	1,29	1,1
9) Inclinação, olhos fechados	-2,27	0,35	1,05	0,3	1,26	0,7
1) Sentado para de pé	-3,17	0,42	0,78	-0,6	0,42	-1
7) Olhos abertos superfície firme	-7,92	1,88			Medida mínima	

D = direita; E = esquerda. TUG = "Timed up and go".

DISCUSSÃO

O presente estudo é o primeiro a investigar as propriedades do *BESTest* em indivíduos com hemiparesia decorrente de AVC utilizando o modelo *Rasch*. A versão original do *BESTest* já foi aplicada em indivíduos com condições neurológicas diversas¹, em indivíduos com doença de Parkinson^{8,23}, fibromialgia²⁴ e esclerose múltipla²⁵. Os dados deste estudo forneceram evidência quanto à confiabilidade e a validade de construto do *BESTest* e do *MiniBESTest* em indivíduos com hemiparesia crônica.

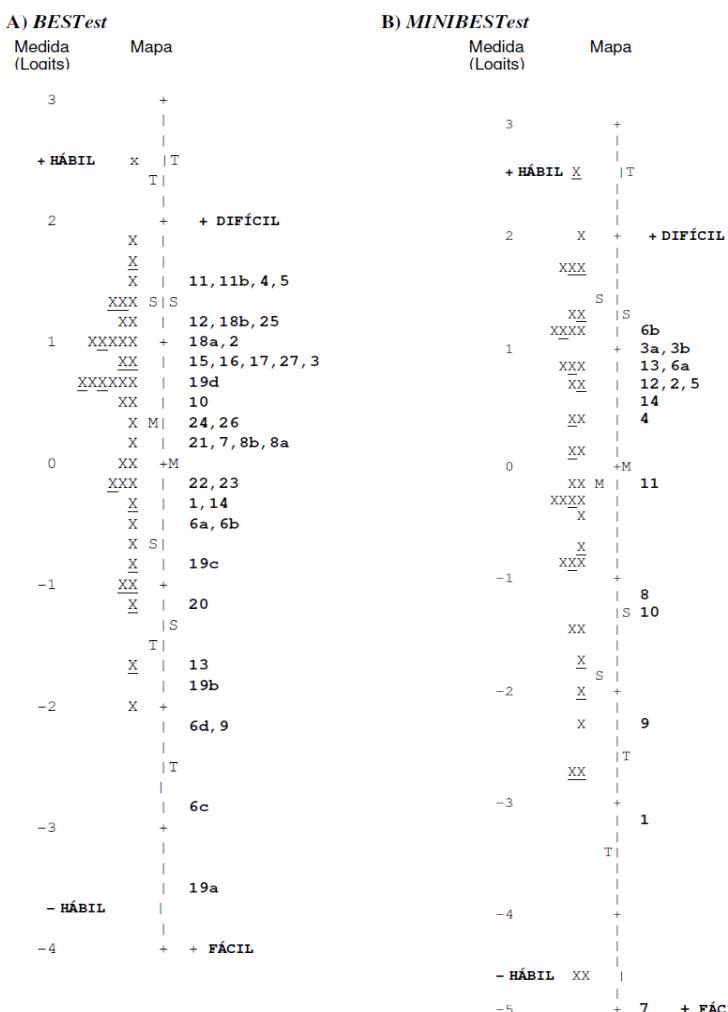
Na análise da confiabilidade, foram observados valores adequados do coeficiente *Kappa* ponderado para todos os itens dos instrumentos. O número de categorias de resposta em um instrumento afeta o coeficiente *Kappa* ponderado, dado que este coeficiente penaliza de forma diferenciada, considerando a magnitude da discordância observada²⁰. Sendo assim, a maior variação observada para a concordância no *BESTest* (moderada a quase perfeita) em relação ao *MiniBESTest* (forte a quase perfeita) pode ser explicada pela presença de uma categoria de

resposta a mais no *BESTest*. Estudos prévios, com outras populações, avaliaram a confiabilidade interexaminadores da versão original do *BESTest*^{1,8,9} e do *MiniBESTest*⁸ por meio do coeficiente de correlação intraclasse (CCI), sendo observados valores altos de CCI para ambos os instrumentos. A confiabilidade teste-reteste foi avaliada para as versões brasileiras dos instrumentos, também por meio do CCI (do escore total), em idosos e indivíduos com doença de Parkinson, com valores considerados quase perfeitos¹⁰.

Os índices de confiabilidade do padrão de respostas da análise *Rasch* para os instrumentos também foram altos, considerados de bons a excelentes, o que significa que as respostas dos indivíduos podem ser reproduzidas em aplicações subsequentes dos testes¹². Os instrumentos também cumpriram a expectativa de dividir os indivíduos em pelo menos três níveis de habilidade e os itens em pelo menos dois níveis de dificuldade¹².

Quanto ao nível de dificuldade dos itens dos instrumentos em relação ao nível de habilidade da amostra,

Figura 1. Mapa representativo da distribuição dos indivíduos e itens no contínuo de equilíbrio, medido pelo *BESTest* (A) e pelo *MiniBESTest* (B).



Cada X representa um indivíduo, sublinhados indicam idade ≥ 60 anos. Ao longo da linha central: M= média; S= desvio padrão; T= 2 desvios padrão. Os números à direita do contínuo representam os itens em cada um dos instrumentos.

percebe-se a presença de indivíduos no topo do contínuo sem itens alinhados (Figura 1). Entretanto, não houve efeito teto e as médias de calibração dos indivíduos ficaram bem próximas da média de dificuldade dos itens, o que sugere que os itens dos testes estão adequados para medir o desempenho da amostra, abrangendo todo o contínuo de habilidade dos indivíduos com hemiparesia crônica. A presença de vários itens do *BESTest* em um mesmo nível de dificuldade pode ser justificada por serem pertencentes a diferentes seções do instrumento. Os itens muito fáceis, observados na parte inferior da Figura 1,

constituem uma vantagem, pois permitem que os testes possam ser utilizados em indivíduos mais debilitados.

Em relação à hierarquia da dificuldade dos itens (Tabela 3), o item mais difícil para os hemiparéticos no *BESTest* foi “sentar no chão e levantar”, item não incluído no *MiniBESTest*. Não foram identificados estudos que fornecessem informações específicas quanto à biomecânica do movimento de sentar no chão e levantar em indivíduos com hemiparesia. Entretanto, sabe-se que esta é uma atividade de grande demanda mecânica. Levantar-se do chão foi a limitação mais frequente observada

em um estudo realizado com indivíduos com desordens musculo-esqueléticas diversas, como osteoartrite e lombalgia²⁶, o que demonstra que esta é uma tarefa difícil não só para indivíduos com hemiparesia. O item mais difícil no *MiniBESTest* (Tabela 4) foi “correção com passo lateral à direita”, seguido pelos itens de apoio unipodal à esquerda e à direita e pelo item “correção com passo lateral à esquerda”. Os itens referentes ao apoio unipodal e de correção com passo lateral também ficaram entre os mais difíceis no *BESTest*. A assimetria funcional presente na hemiparesia representa um desafio para o controle dos passos compensatórios²⁷. Indivíduos com hemiparesia são geralmente relutantes em iniciar passos compensatórios com o membro parético²⁷. Em um estudo realizado com hemiparéticos, quando o membro inferior preferencialmente utilizado para a resposta do passo foi bloqueado foram observadas respostas não apropriadas em 21% das tentativas, demonstrando que os indivíduos com hemiparesia apresentam limitada capacidade de adaptação nas respostas com passo compensatório²⁷. Atividades que envolvem apoio unipodal também são desafiadoras ao equilíbrio devido à diminuição da base de suporte¹ e à frequente incapacidade do membro parético em controlar o equilíbrio na postura de pé²⁸. Inversamente, “permanecer de olhos abertos em superfície firme” foi o item mais fácil tanto no *BESTest* quanto no *MiniBESTest*, sendo uma atividade que envolve menores demandas de controle postural¹. De uma forma geral, a hierarquia e a adequação dos itens derivada da análise Rasch forneceu evidência quanto à validade de construto para ambos os instrumentos.

Comportamento errático foi detectado em dois itens do *BESTest* (item no. 4 e 8a), estando dentro do limite esperado para o instrumento. Ao examinar as pontuações do item “força lateral de quadril e tronco”, observou-se que três indivíduos com pobre habilidade de equilíbrio obtiveram escore correspondente a comprometimento leve neste item. O escore recebido equivale a conseguir abduzir ambos os quadris para levantar o pé do chão durante dez segundos, com apoio nas mãos do examinador, sem conseguir manter o tronco na vertical. Um aspecto que pode ter contribuído para tal resposta inesperada é o fato dos participantes deste estudo terem comprometimento crônico, e, possivelmente, apresentarem estratégias compensatórias bem desenvolvidas. No

item “alcance funcional lateral à esquerda” observou-se resposta inesperada em um indivíduo que, apesar de ter uma alta habilidade de equilíbrio, recebeu pontuação igual a zero neste item. Este indivíduo não apresentava movimentação ativa no membro superior esquerdo, não sendo capaz de realizar o alcance lateral. Neste caso, um fator não relacionado diretamente ao controle postural influenciou a resposta. O padrão errático de resposta sinaliza que se deve dar mais atenção na interpretação do escore de indivíduos com hemiparesia nesse item, sendo importante tentar identificar o que pode ter contribuído para o comportamento inesperado observado^{12,15}. Em relação ao uso do *BESTest* em hemiparéticos, recomenda-se observar se os fatores associados aos comportamentos inesperados estão relacionados ao equilíbrio ou a ausência de movimentação comumente encontrada nesses indivíduos.

A análise fatorial evidenciou que ambos os instrumentos foram multidimensionais. O equilíbrio é um construto amplo que envolve vários sistemas¹. Como o *BESTest* é dividido em seções, pode-se considerar que o instrumento foi desenvolvido com o objetivo de abranger vários aspectos relacionados ao equilíbrio, ou seja, a partir de um modelo multidimensional. Dessa forma, os resultados indicativos da multidimensionalidade confirmam a sua função original. Porém, a divisão dos itens segundo os resultados da análise de componente principal, não seguiu o modelo original, agrupando itens de diferentes seções sem significado teórico evidente. Esperava-se que o *MiniBESTest* fosse unidimensional, como em sua versão original, já que o instrumento foi construído a partir do construto “equilíbrio dinâmico”⁷. Entretanto, o instrumento também foi multidimensional e a análise evidenciou que alguns itens poderiam ser removidos sem grande prejuízo para as qualidades psicométricas. Com exceção de passar de sentado para de pé, todos os itens que poderiam ser removidos são da escala de respostas posturais reativas. Entretanto, a exclusão de itens em uma escala já existente deve ser o último recurso a ser considerado porque os itens podem ser importantes clinicamente²⁹. No presente estudo, se os itens fossem removidos, não seriam avaliadas as respostas posturais reativas, que são importantes para avaliar como o indivíduo reage após uma perturbação do equilíbrio¹. Dessa forma, tanto ques-

tóes psicométricas quanto clínicas devem ser consideradas, não sendo recomendada exclusão de nenhum item. Outro ponto a ser considerado é que, dado o tamanho da amostra, pequenas flutuações no padrão de pacientes individuais podem ter influenciado o resultado global.

O *MiniBESTest*, quando comparado ao *BESTest*, se adequou melhor ao modelo, devido a ausência de itens erráticos. Por outro lado, sendo o *BESTest* um instrumento voltado para diagnóstico dos déficits de equilíbrio, a análise *Rasch* mostrou que seus itens estão menos espaçados e cobrem uma maior porção do contínuo de habilidade. Em um instrumento voltado para triagem como o *MiniBESTest*, a calibração dos itens é mais espaçada, o que implica em menor precisão das medidas.

Uma limitação para o uso do *BESTest* na prática clínica é o tempo despendido para sua aplicação. Estudos anteriores relataram tempo de aplicação variando de 20 a 31 minutos^{1,10} até 45 a 60 minutos⁹. Neste estudo, o tempo gasto foi de 30 a 45 minutos. Uma possibilidade seria a aplicação das seções separadamente¹. Os dados do presente estudo, no entanto, não dão suporte à divisão das seções como proposta no teste. Assim, são necessários novos estudos que avaliem a validade e a utilidade clínica das seções quando utilizadas de forma independente nessa população.

Uma limitação deste estudo foi a amostra restrita a hemiparéticos crônicos. Seriam interessantes novos estudos incluindo indivíduos em fase aguda e subaguda após AVC com diferentes níveis de habilidade. São importantes, também, novos estudos que avaliem outras propriedades psicométricas destes instrumentos em indivíduos com hemiparesia, como responsividade e validade de critério preditiva. Como o *BESTest* apresenta a proposta de direcionar o tratamento na prática clínica¹, são também necessários estudos controlados e randomizados para verificar se o tratamento direcionado às seções do *BESTest* será realmente mais eficaz.

CONCLUSÃO

As versões brasileiras do *BESTest* e do *MiniBESTest* mostraram ser confiáveis e capazes de discriminar diferentes níveis de habilidade entre os indivíduos, sem ser observado efeito teto ou solo.

De uma forma geral, a hierarquia dos itens e adequação ao modelo *Rasch* forneceu evidência quanto à validade de construto dos instrumentos. Recomenda-se, porém, cautela na interpretação do escore total do *BESTest* e atenção em relação aos itens identificados como erráticos. Dado que a multidimensionalidade não impede o uso de um instrumento como ferramenta clinicamente válida³⁰, os dois instrumentos parecem ser úteis para a avaliação do equilíbrio em indivíduos com hemiparesia crônica.

REFERÊNCIAS

1. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BES-Test) to differentiate balance deficits. *Phys Ther* 2009;89:484-98. <http://dx.doi.org/10.2522/ptj.20080071>
2. Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. *Eur J Phys Rehabil Med* 2010;46:239-48.
3. Campbell GB, Matthews JT. An integrative review of factors associated with falls during post-stroke rehabilitation. *J Nurs Scholarsh* 2010;42:395-404. <http://dx.doi.org/10.1111/j.1547-5069.2010.01369.x>
4. Schmid AA, Rittman M. Consequences of poststroke falls: activity limitation, increased dependence, and the development of fear of falling. *Am J Occup Ther* 2009;63:310-6. <http://dx.doi.org/10.5014/ajot.63.3.310>
5. Fryberg GE, Lindmark B, Lanshammar H, Borg J. Correlation between clinical assessment and force plate measurement of postural control after stroke. *J Rehabil Med* 2007;39:448-53. <http://dx.doi.org/10.2340/16501977-0071>
6. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: A systematic review. *Phys Ther* 2008;88:559-66. <http://dx.doi.org/10.2522/ptj.20070205>
7. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BES-Test. *J Rehabil Med* 2010;42:323-31. <http://dx.doi.org/10.2340/16501977-0537>
8. Leddy AL, Crowner BE, Earhart GM. Utility of the Mini-BESTest, BES-Test, and BESTest sections for balance assessments in individuals with Parkinson disease. *J Neurol Phys Ther* 2011;35:90-7. <http://dx.doi.org/10.1097/NPT.0b013e31821a620c>
9. Parminder K, Padgett JV, Jacobs, SLK. Is the BESTest at its best? A suggested briefversion based on interrater reliability, validity, internal consistency, and theoretical construct. *Phys Ther* 2012;92:1197-207. <http://dx.doi.org/10.2522/ptj.20120056>
10. Maia AC, Rodrigues-de-Paula F, Magalhães LC, Teixeira RL. Cross-cultural adaptation and analysis of psychometric properties of the Balance Evaluation Systems Test and MiniBESTest in the elderly and individuals with Parkinson's disease: application of the Rasch model. *Braz J Phys Ther* 2013;17:195-217. <http://dx.doi.org/10.1590/S1413-35552012005000085>
11. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. 2ed. New Jersey: Prentice Hall Health, 2000, 742p.
12. Velozo CA, Forsyth K, Kielhofner G. Objective measurement: the influence of item response theory on research and practice. In: Kielhofner G. Research in Occupational Therapy: Methods of Inquiry for enhancing Practice. Philadelphia: F. A. Davis Company, 2006, p.177-200.

13. Linacre JM. Sample size and item calibration stability. *Rasch Meas Trans* 1994;7:328.
14. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3:85-106.
15. Saliba VA, Faria CD, Laurentino GE, Cassiano JG, Teixeira-Salmela LF. Adaptação transcultural e análise das propriedades psicométricas da versão brasileira do instrumento Motor Activity Log. *Rev Panam Salud Pública* 2011;30:262-71. <http://dx.doi.org/10.1590/S1020-49892011000900011>
16. Neckel N, Pelliccio M, Nichols D, Hidler J. Quantification of functional weakness and abnormal synergy patterns in the lower limb of individuals with chronic stroke. *J Neuroeng Rehabil* 2006;3:1-11. <http://dx.doi.org/10.1186/1743-0003-3-17>
17. Teixeira-Salmela LF, Devaraj R, Olney SJ. Validation of the human activity profile stroke: a comparison of observed, proxy and self-reported scores. *Disabil Rehabil* 2007;29:1518-24. <http://dx.doi.org/10.1080/09638280601055733>
18. Centers for Disease Control and Prevention. Physical activity trends - United States, 1990 -1998. Morbidity and Mortality Weekly Report (MMWR) 2001;50:166-9.
19. Michaelsen SM, Rocha AS, Knabben RJ, Rodrigues LP, Fernandes CGC. Tradução, adaptação e confiabilidade interexaminadores do manual de administração da escala de Fugl-Meyer. *Braz J Phys Ther* 2011;15:80-8.
20. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257-68.
21. Linacre JM, Wright BD (Endereço na Internet). Winsteps Rasch Measurement Version 3.74.0 (Software). Disponível em: <http://www.winsteps.com>
22. Chen HF, Wu CY, Lin KC, Chen HC, Chen CP, Chen CK. Rasch validation of the Streamlined Wolf Motor Function Test in people with chronic stroke and subacute stroke. *Phys Ther* 2012;92:1017-26. <http://dx.doi.org/10.2522/ptj.20110175>
23. Leddy AL, Crowner BE, Earhart GM. Functional gait assessment and Balance Evaluation System Test: Reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson Disease who fall. *Phys Ther* 2011;91:102-12. <http://dx.doi.org/10.2522/ptj.20100113>
24. Jones KD, Horak FB, Winters KS, Morea JM, Bennett RM. Fibromyalgia is associated with impaired balance and falls. *J Clin Rheumatol* 2009;15:16-21. <http://dx.doi.org/10.1097/RHU.0b013e318190f991>
25. Jacobs JV, Kasser SL. Balance impairment in people with multiple sclerosis: Preliminary evidence for the Balance Evaluation Systems Test. *Gait Posture* 2012;36:414-8. <http://dx.doi.org/10.1016/j.gaitpost.2012.03.026>
26. Banerjee A, Jadhav SL, Bhawalkar JS. Limitations of activities in patients with musculoskeletal disorders. *Ann Med Health Sci Res* 2012;2:5-9.
27. Mansfield A, Inness EL, Lakhani B, McIlroy WE. Determinants of limb for preference initiating compensatory stepping poststroke. *Arch Phys Med Rehabil* 2012;93:1179-84. <http://dx.doi.org/10.1016/j.apmr.2012.02.006>
28. Genthon N, Rougier P, Gissot A, Froger J, Pélié J, Pérennou D. Contribution of each lower limb to upright standing in stroke patients. *Stroke* 2008;39:1793-9. <http://dx.doi.org/10.1161/STROKEAHA.107.497701>
29. Nilsson AL, Tennant A. Past and present issues in Rasch analysis: the functional independence measure (FIM) revisited. *J Rehabil Med* 2011;43: 884-91. <http://dx.doi.org/10.2340/16501977-0871>
30. Lovaglio PG, Monzani EM. Validation aspects of the health of the nation outcome scales. *Inter J Mental Health Sys* 2011;5:1-11. <http://dx.doi.org/10.1186/1752-4458-5-20>