

## Psychometric Properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in Community-Dwelling Individuals With Chronic Stroke

Charlotte S.L. Tsang, Lin-Rong Liao, Raymond C.K. Chung, Marco Y.C. Pang

C.S.L. Tsang, MSc, Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong.

L-R. Liao, MPT, Department of Rehabilitation Sciences, Hong Kong Polytechnic University, and Department of Physiotherapy, Guangdong Provincial Work Injury Rehabilitation Hospital, Guangzhou, China.

R.C.K. Chung, PhD, Department of Rehabilitation Sciences, Hong Kong Polytechnic University.

M.Y.C. Pang, PhD, Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong. Address all correspondence to Dr Pang at: marco.pang@polyu.edu.hk.

[Tsang CSL, Liao L-R, Chung RCK, Pang MYC. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. *Phys Ther*. 2013;93:1102-1115.]

© 2013 American Physical Therapy Association

Published Ahead of Print:  
April 4, 2013

Accepted: April 1, 2013  
Submitted: November 13, 2012

**Background.** The Mini-Balance Evaluation Systems Test (Mini-BESTest) is a new balance assessment, but its psychometric properties have not been specifically tested in individuals with stroke.

**Objectives.** The purpose of this study was to examine the reliability and validity of the Mini-BESTest and its accuracy in categorizing people with stroke based on fall history.

**Design.** An observational measurement study with a test-retest design was conducted.

**Methods.** One hundred six people with chronic stroke were recruited. Intrarater reliability was evaluated by repeating the Mini-BESTest within 10 days by the same rater. The Mini-BESTest was administered by 2 independent raters to establish interrater reliability. Validity was assessed by correlating Mini-BESTest scores with scores of other balance measures (Berg Balance Scale, one-leg-standing, Functional Reach Test, and Timed “Up & Go” Test) in the stroke group and by comparing Mini-BESTest scores between the stroke group and 48 control participants, and between fallers ( $\geq 1$  falls in the previous 12 months,  $n=25$ ) and nonfallers ( $n=81$ ) in the stroke group.

**Results.** The Mini-BESTest had excellent internal consistency (Cronbach  $\alpha=.89-.94$ ), intrarater reliability (intraclass correlation coefficient [3,1]=.97), and interrater reliability (intraclass correlation coefficient [2,1]=.96). The minimal detectable change at 95% confidence interval was 3.0 points. The Mini-BESTest was strongly correlated with other balance measures. Significant differences in Mini-BESTest total scores were found between the stroke and control groups and between fallers and nonfallers in the stroke group. In terms of floor and ceiling effects, the Mini-BESTest was significantly less skewed than other balance measures, except for one-leg-standing on the nonparetic side. The Berg Balance Scale showed significantly better ability to identify fallers (positive likelihood ratio=2.6) than the Mini-BESTest (positive likelihood ratio=1.8).

**Limitations.** The results are generalizable only to people with mild to moderate chronic stroke.

**Conclusions.** The Mini-BESTest is a reliable and valid tool for evaluating balance in people with chronic stroke.



Post a Rapid Response to  
this article at:  
[ptjournal.apta.org](http://ptjournal.apta.org)

Stroke is a major cause of disability and global disease burden.<sup>1</sup> Dysfunction in balance control is one of the most common physical impairments observed after stroke.<sup>2,3</sup> Compromised balance ability has been associated with reduced ambulatory function,<sup>4</sup> poorer performance in activities of daily living (ADL),<sup>5</sup> and restricted societal participation.<sup>6</sup> Impaired balance also is a significant predictor of falls<sup>7</sup> and long-term institutionalization.<sup>8</sup>

Much effort has been directed toward enhancing balance function in people with stroke.<sup>9-11</sup> Balance control is complex and involves various aspects such as ability to maintain a body position, postural responses to external perturbations, anticipatory postural adjustments, and sensory integration.<sup>12</sup> To obtain a clearer understanding of balance dysfunctions after a stroke and to better assess the effect of intervention programs, a standardized assessment of balance function is essential. Many clinical tools are available to assess balance in individuals with stroke.<sup>13,14</sup> Some of the most commonly used balance assessment tools in stroke rehabilitation are the Berg Balance Scale (BBS),<sup>15</sup> Functional Reach Test (FRT),<sup>16</sup> Timed "Up & Go" Test (TUG),<sup>17</sup> and one-leg standing (OLS).<sup>18,19</sup> However, they are not without their limitations. For example, important aspects of dynamic balance control that reflect balance challenges during ADL are missing in the BBS.<sup>20</sup> Leroux et al<sup>21</sup> found that among ambulatory patients with chronic stroke, improvement in postural stability observed after exercise intervention was poorly correlated with change in the BBS score. On the other hand, OLS, FRT, and TUG, being single-task assessments, are unable to provide information on which postural control subsystem is dysfunctional and have a limited role in directing treatment.<sup>13</sup> Significant floor or ceiling effects also have been

identified in the BBS, OLS, and FRT.<sup>22-24</sup> Furthermore, the BBS<sup>25,26</sup> and TUG<sup>27</sup> have been criticized for their limited ability to predict falls in people with stroke. Certain balance assessment tools that are specifically designed for people with stroke also have similar limitations. For example, the balance subscale of the Fugl-Meyer test<sup>28</sup> has been shown to have significant floor effects.<sup>22</sup>

The Balance Evaluation Systems Test (BESTest) is a relatively new multi-task balance assessment developed to identify specific postural control problems (ie, biomechanical constraints, stability limits, postural responses, anticipatory postural adjustments, sensory orientation, dynamic balance during gait, and cognitive effects).<sup>20,29</sup> However, this 36-item assessment takes 30 to 35 minutes to complete and may not be feasible in real clinical settings, where time constraint is often a major concern. A shorter version of the test, the 14-item Mini-BESTest, has recently been developed.<sup>20</sup> It takes only 10 minutes to complete, and good intrarater and interrater reliability have been reported in a sample of people with mixed conditions.<sup>30</sup> Recent studies further showed that the Mini-BESTest has good interrater and intrarater reliability and concurrent validity<sup>31,32</sup> and is useful in predicting falls<sup>33,34</sup> in patients with Parkinson disease (PD). However, the psychometric properties of the Mini-BESTest have not been specifically evaluated in the stroke population. Additionally, no study has evaluated the ability of the Mini-BESTest in distinguishing fallers from nonfallers among individuals with stroke. The current study was undertaken to (1) examine the reliability and validity of the Mini-BESTest and (2) compare the Mini-BESTest with 4 other balance measures based on the floor and ceiling effects and on sensitivity and specificity for distinguishing

between individuals with and without a history of falls in a group of community-dwelling people with chronic stroke.

## Method

### Study Overview

This was an observational measurement study. Floor and ceiling effects, reliability (internal consistency, intrarater and interrater), and validity (concurrent, convergent, discriminant, known-groups) of the Mini-BESTest were assessed in a sample of people with stroke. To establish known-groups validity, a control group was included to enable us to assess the differences in Mini-BESTest scores between the stroke group and control group. The ability of the Mini-BESTest to distinguish between people with stroke with and without a history of falls also was examined and compared with that of 4 other balance measures (ie, BBS, TUG, OLS, and FRT). All of the raters involved in the study were physical therapists who had more than 10 years of relevant experience and were well trained to administer all of the balance assessment tools used in this study.

### Participants and Sample Size Calculations

Participants were recruited during the period June 2009 and December 2010. Individuals with stroke were recruited from a local rehabilitation center and community self-help groups on a volunteer basis (ie, convenience sampling). Each participant was interviewed during the first assessment session. Ability to understand verbal instructions was one of the inclusion criteria. An individual was considered to have fulfilled this criterion if he or she managed to carry out a normal conversation with the assessor. Other inclusion criteria for the stroke group were: a diagnosis of stroke for more than 6 months, community-dwelling, and aged 18 years or older. The exclusion criteria

were: pain during performance of daily activities, neurological conditions in addition to stroke, other conditions that affect balance (eg, Ménière disease), and any other serious illnesses that precluded participation. Control individuals were recruited from the community for comparison. The eligibility criteria were the same as those used in the stroke group, except that the control participants did not have a history of stroke. All participants provided written informed consent before enrollment in the study. All procedures were conducted in accordance with the Declaration of Helsinki.

All sample size calculations were done prior to enrollment of participants and were based on an alpha level of .05 (2-tailed) and a power of 0.8 (NCSS and PASS 2005, NCSS LLS Co, Kaysville, Utah). For reliability analysis, a coefficient of .75 or greater was generally considered to be acceptable.<sup>35</sup> Leddy et al<sup>32</sup> found that the Mini-BESTest had excellent intrarater and interrater reliability in people with PD, with intraclass correlation coefficient (ICC) values of .92 and .91, respectively. A similar reliability coefficient was expected in this study. Thus, the acceptable reliability and expected reliability was set at ICC=.75 and ICC=.90, respectively.<sup>32</sup> For establishing interrater reliability between 2 raters, a sample of 26 patients with stroke was required. As establishing intrarater reliability required 2 assessment sessions, a 10% attrition rate was estimated, yielding a minimum sample of 30 participants.

A study by King et al<sup>31</sup> showed a strong correlation between the Mini-BESTest and the BBS in patients with PD ( $r=.79$ ; large effect size). Therefore, for analysis of concurrent and convergent validity, a large effect size was expected when the Mini-BESTest was correlated with other balance and related measures in indi-

viduals with stroke. Using the conventional value of a large effect size ( $r=.5$ ) in the sample size calculation,<sup>35</sup> the minimum number of participants required for the analysis of concurrent validity would be 26.

The Mini-BESTest scores obtained from the stroke group were compared with those from the control group to establish known-groups validity. Horak et al<sup>29</sup> compared the BESTest total score between patients with different balance problems ( $\bar{X}=74.5$ ,  $SD=9.0$ ) and controls without disabilities ( $\bar{X}=90.6$ ,  $SD=4.8$ ), and the effect size was large (Cohen  $d=1.8$ ). We expected the Mini-BESTest to also have good ability to discriminate between the 2 groups. Using the conventional value of a large effect size (Cohen  $d=0.8$ ) for calculation,<sup>35</sup> a minimum of 26 participants per group would be required for this analysis.

We also were interested in determining whether the Mini-BESTest scores and other balance tests could differentiate people with stroke with and without a history of falls. Receiver operating characteristic (ROC) curve plots were used for this analysis.<sup>35</sup> An area under the curve (AUC) value of 0.7 to 0.8 was generally considered to be acceptable.<sup>36</sup> Duncan et al<sup>34</sup> showed that the Mini-BESTest had good ability to identify fallers among patients with PD, with an AUC value of 0.86. The acceptable and expected AUC values thus were set at 0.7 and 0.9, respectively.<sup>36</sup> Previous studies in community-dwelling individuals with stroke demonstrated a fall rate of 23% to 73%.<sup>7,37-39</sup> Assuming that the proportion of fallers was 30% in our stroke group, a minimum of 60 individuals with stroke (fallers:  $n=18$ ; nonfallers:  $n=42$ ) would be required for ROC curve plots. In summary, a minimum of 60 and 26 individuals would be recruited from the stroke and control groups, respectively.

## Procedure

**Stroke group.** In the initial assessment (session 1), relevant demographic data (eg, age, medical history) and fall history were obtained from interviewing the participants. To calculate body mass index (BMI, in  $\text{kg}/\text{m}^2$ ), height (in meters) and weight (in kilograms) were measured with a stadiometer (Health O Meter, Alsip, Illinois). Each participant was evaluated with the Mini-BESTest, 4 additional balance assessments (BBS, FRT, OLS, and TUG) and other measures (Chedoke-McMaster Stroke Assessment, Modified Ashworth Scale [MAS], Activities-specific Balance Confidence [ABC] Scale, Abbreviated Mental Test [AMT], Geriatric Depression Scale-short form [GDS], and Oxfordshire Community Stroke Project Classification). Either rater 1 or rater 2 conducted the assessments in session 1.

The first 30 participants assessed by rater 2 in session 1 also were evaluated with the Mini-BESTest a second time by another independent rater (rater 3) in the same session. Whether rater 2 or rater 3 administered the Mini-BESTest first was determined randomly by drawing lots. Intermittent rest periods were given throughout the session. The typical duration of session 1 was 2.5 hours, including the rest periods. Interrater reliability of the Mini-BESTest was determined by comparing the scores given by raters 2 and 3 in session 1.

The 30 participants with stroke who were evaluated for interrater reliability also participated in the intrarater reliability experiments. A second assessment session (session 2) was held within 10 days after session 1. The participants did not receive any physical therapy intervention during the period between sessions 1 and 2. In session 2, each of the 30 participants was evaluated with the Mini-BESTest once by rater 2. Session 2

was typically 20 minutes in duration. Intrarater reliability was established by comparing the Mini-BESTest scores given by rater 2 in sessions 1 and 2.

**Control group.** The participants in the control group underwent one assessment session conducted by rater 1. Demographic data (eg, age, medical history), height, and weight were obtained using the same methods as in the stroke group described above. The Mini-BESTest was administered once. Comparing the Mini-BESTest scores of the control group with those of the stroke group would be useful in determining the known-groups validity. No other measures were administered to the control group.

## Measures

**Fall history.** Information on fall history was obtained through interview of participants. Those who had experienced one or more falls in the previous 12 months were considered to have a positive fall history.

**Mini-BESTest.** The Mini-BESTest is a 14-item performance-based measure of balance disorders. The tasks involved varied in difficulty and covered different balance subsystems, including responses to external perturbations, anticipatory postural adjustments, stability in gait, and sensory orientation. Each task was rated from an ordinal scale of 0 to 2. Items 3 (stand on one leg) and 6 (compensatory stepping correction in lateral direction) assessed both sides, and only the side with a lower score was used for calculating the total score.<sup>20</sup> When reporting the item scores, however, the results of both the paretic and nonparetic sides were shown for these 2 items. The total score ranged from 0 to 28, with higher scores denoting better balance ability.

**Other balance measures.** The BBS is a 14-item assessment of functional balance. Each task was rated from 0 to 4, yielding a possible maximum total score of 56. Higher scores are indicative of better balance.<sup>15</sup> The BBS has shown good interrater and intrarater reliability (ICC>.90) and concurrent validity (correlation with Postural Assessment Scale for Stroke Patients:  $r=.92-.95$ ) in individuals with stroke.<sup>15,22,40</sup>

The FRT measures balance by assessing the limit of stability.<sup>16</sup> The maximum distance (in centimeters) an individual could reach forward beyond arm's length on a fixed base of support was measured. Its interrater reliability (ICC=.99) and validity (correlation with the BBS:  $r=.619$ ) in people with stroke are well established.<sup>40</sup> A score of 0 cm was given for participants who were unable to maintain the standing position without external support.

The OLS test measures the time (in seconds) an individual can stand on one leg (either side).<sup>18</sup> Participants were asked to stand on one leg with eyes open and hands placed on the hips. Using a stopwatch, timing commenced when the foot left the ground and stopped when the same foot touched the ground, when the individual's hand swung away from the hips, or when OLS was maintained for a period of 1 minute. One-leg standing was tested on both sides in the current study. One-leg standing has shown good intrarater reliability (nonparetic side: ICC=.88, paretic side: ICC=.92) and significant correlation with the BBS ( $r=.65$ ) in people with stroke.<sup>18</sup> A score of 0 second was given for participants who were unable to maintain the standing position without external support.

The TUG measures the time (in seconds) an individual required to get

up from an armed chair, walk 3 m with normal walking pace, turn around, walk back, and sit down again.<sup>17</sup> Use of a walking aid was allowed if necessary. The TUG has shown good test-retest reliability (ICC=.96) and concurrent validity (correlation with Community Balance and Mobility Scale:  $\rho=-.75$ ) in individuals with stroke.<sup>41,42</sup>

**Measures of other related functions.** The Impairment Inventory of the Chedoke-McMaster Stroke Assessment was used to assess the motor recovery of arm, hand, leg, and foot in the stroke group.<sup>43</sup> Each of the 4 body parts was rated on a 7-point scale, with a higher score indicating better motor recovery. Good intrarater (ICC=.98) and interrater reliability (ICC=.97) have been reported in people with stroke.<sup>43</sup>

The MAS, a 6-point ordinal scale, was used for assessing muscle tone around the ankle joint of the affected leg (0=no increase in muscle tone, 4=part rigid in flexion and extension).<sup>44</sup> The intrarater and interrater reliability of the MAS in people with stroke are well established ( $\kappa>.8$ ).<sup>44</sup>

The ABC Scale was used for measuring balance confidence.<sup>45</sup> Participants were asked to rate their confidence in their balance associated with performing 16 listed daily tasks from 0% (absolutely no confidence) to 100% (fully confident). The average score of the 16 items was calculated. The ABC Scale has shown high test-retest reliability (ICC=.87) and concurrent validity (correlation with the BBS:  $\rho=.36$  and with gait speed:  $\rho=.48$ ) among individuals with chronic stroke.<sup>46,47</sup>

**Other measures.** The Oxfordshire Community Stroke Project Classification was used to identify the clinical stroke subtypes.<sup>48</sup> The intrarater agreement and interrater

agreement for the classification was moderate to good, with kappa values of .48 to .83 and .54 to .64, respectively.<sup>49,50</sup>

The AMT was used to assess cognitive function.<sup>51</sup> The AMT has shown good internal consistency (Cronbach  $\alpha$  = .81), interrater reliability (ICC = .99), and concurrent validity (correlation with Mini-Mental State Examination:  $r$  = .86) among older adults.<sup>52</sup> It also is able to differentiate between individuals with and without cognitive impairments ( $P$   $\leq$  .001).<sup>52</sup>

The 15-item GDS was used to indicate the severity of depressive symptoms (0–4 = no depression, 5–10 = mild depression, and  $\geq$  11 = severe depression).<sup>53,54</sup> The GDS has shown good test-retest reliability (ICC = .75) in people with stroke.<sup>54</sup>

### Data Analysis

All statistical analyses were performed using SPSS 18.0 software (SPSS Inc, Chicago, Illinois), unless otherwise indicated. The significance level was set a priori at  $\leq$  .05.

**Floor and ceiling effects.** The skewness ( $\gamma_1$ ) of the distribution of scores was first assessed for each balance measure. Positive skewness reflects a floor effect and negative skewness indicates a ceiling effect for the Mini-BESTest, BBS, OLS, and FRT, whereas the opposite is true for the TUG.<sup>31</sup> R Statistical Software with Bootstrapping methods (version 2.15.2, Bell Laboratories, Murray Hill, New Jersey) was used to compare the degree of skewness in distribution of scores between the Mini-BESTest and other balance measures.<sup>31</sup> To further explore the floor and ceiling effects, the proportion of participants with the lowest and highest possible scores was examined.<sup>23</sup> Floor or ceiling effects greater than 20% were considered to be significant.<sup>23</sup>

**Reliability.** Using the data obtained from the stroke group, the internal consistency of the Mini-BESTest was assessed by Cronbach alpha. Intraclass correlation coefficients were used to determine the intrarater (ICC [3,1]) and interrater (ICC [2,1]) reliability of the Mini-BESTest total score. An ICC  $>$  .75 is indicative of good reliability, and an ICC of .5 to .75 is indicative of moderate reliability.<sup>55</sup> The kappa statistic was used to examine the intrarater and interrater reliability of each individual test item (kappa: .81 = almost perfect agreement, .61–.8 = substantial agreement, .41–.6 = adequate agreement, .21–.4 = fair agreement, and 0–.2 = slight agreement).<sup>55</sup> Using the intrarater reliability results, the minimal detectable change at the 95% confidence interval (MDC<sub>95</sub>) was computed using the following formula<sup>35</sup>:

$$MDC_{95} = 1.96 \times SEM \times \sqrt{2}$$

The standard error of measurement (SEM) value of the Mini-BESTest total score was derived from the following formula<sup>35</sup>:

$$SEM = S_x \sqrt{(1 - r_{xx})},$$

where  $S_x$  is the standard deviation of the Mini-BESTest total score and  $r_{xx}$  is the reliability coefficient.

**Validity.** For the stroke group data, the Spearman rho was used to examine the degree of association of the Mini-BESTest total scores (measured in the first session) with the following: (1) other established balance measures (ie, BBS, FRT, TUG, and OLS) (ie, concurrent validity), (2) instruments measuring attributes that supposedly are related to balance function (ie, Chedoke-McMaster Stroke Assessment leg and foot impairment score and ABC Scale) (ie, convergent validity), and (3) measures that assess unrelated

characteristics (ie, GDS and AMT) (ie, discriminant validity).

In addition to assessing convergence and discrimination, another way to examine the construct validity of the Mini-BESTest was to evaluate the known-groups validity. A test with good known-groups validity should be able to distinguish individuals with good balance ability from those with poor balance ability. Comparisons of Mini-BESTest total and item scores were made between the stroke and control groups, and between participants with and without a history of falls in the stroke group, using the Mann-Whitney  $U$  test, as the total scores were not normally distributed (checked by Kolmogorov-Smirnov test) and the item scores were ordinal in nature. In Mann-Whitney  $U$  test, the between-group comparison was based on rank ordering of the raw scores.<sup>35</sup> Considering the data of the 2 groups together, the scores were ranked from the smallest to largest. For example, the lowest score was assigned the rank of 1, and the next smallest value was assigned the rank of 2. When 2 or more scores were tied, they were each given the same rank, which was the average of the ranks they occupied. For example, if there were 3 scores with the smallest value, they occupied ranks 1, 2, and 3. Thus, they were each given the rank of 2 (the average of 1 + 2 + 3).<sup>35</sup> The rank scores of each group then were summed and divided by the number of participants in the group to yield the mean rank score. A higher mean rank reflected an overall better balance ability as a group.

To further compare the Mini-BESTest with other balance measures in differentiating between people with stroke with and without a history of falls, ROC curves were constructed. The AUC derived from the Mini-BESTest data then was compared with that of other balance measures,

using the chi-square test for comparing the areas under 2 or more correlated ROC curves (SigmaPlot version 12.3, Systat Software Inc, San Jose, California).<sup>56</sup> For each ROC curve, the score that yielded the largest Youden index (sensitivity + [1 – specificity]) was chosen as the cutoff score. The positive and negative likelihood ratios (LR+ and LR–) and their 95% confidence intervals (95% CI) were computed using an online CI calculator.<sup>57</sup> As 4 participants were unable to ambulate without manual assistance and thus did not complete the TUG, their data were not included for the comparison of skewness and AUC between the Mini-BESTest and the TUG.

## Results

A total of 106 individuals with stroke (73 men, 33 women) and 48 controls (28 men, 20 women) participated in the study. The participant characteristics are shown in Table 1. Seventy participants (66.0%) in the stroke group did not require any walking aid for ambulation. Twenty-five individuals (23.6%) in the stroke group had a history of falls, 7 (6.6%) of whom were recurrent fallers (ie, 2 or more falls during the previous 12 months).

Four participants required physical assistance to ambulate and thus were unable to complete the TUG. Three individuals were unable to maintain the standing position without external support and were given a score of 0 for the OLS and FRT. There were no significant differences in any of the demographic variables (eg, age, proportion of men and women, BMI) between the stroke and control groups.

### Score Distribution and Ceiling and Floor Effects

The score distribution of the Mini-BESTest within the stroke group is shown in Figure 1A, and those of the BBS, FRT, TUG, and OLS are shown

**Table 1.**  
Characteristics of Participants<sup>a</sup>

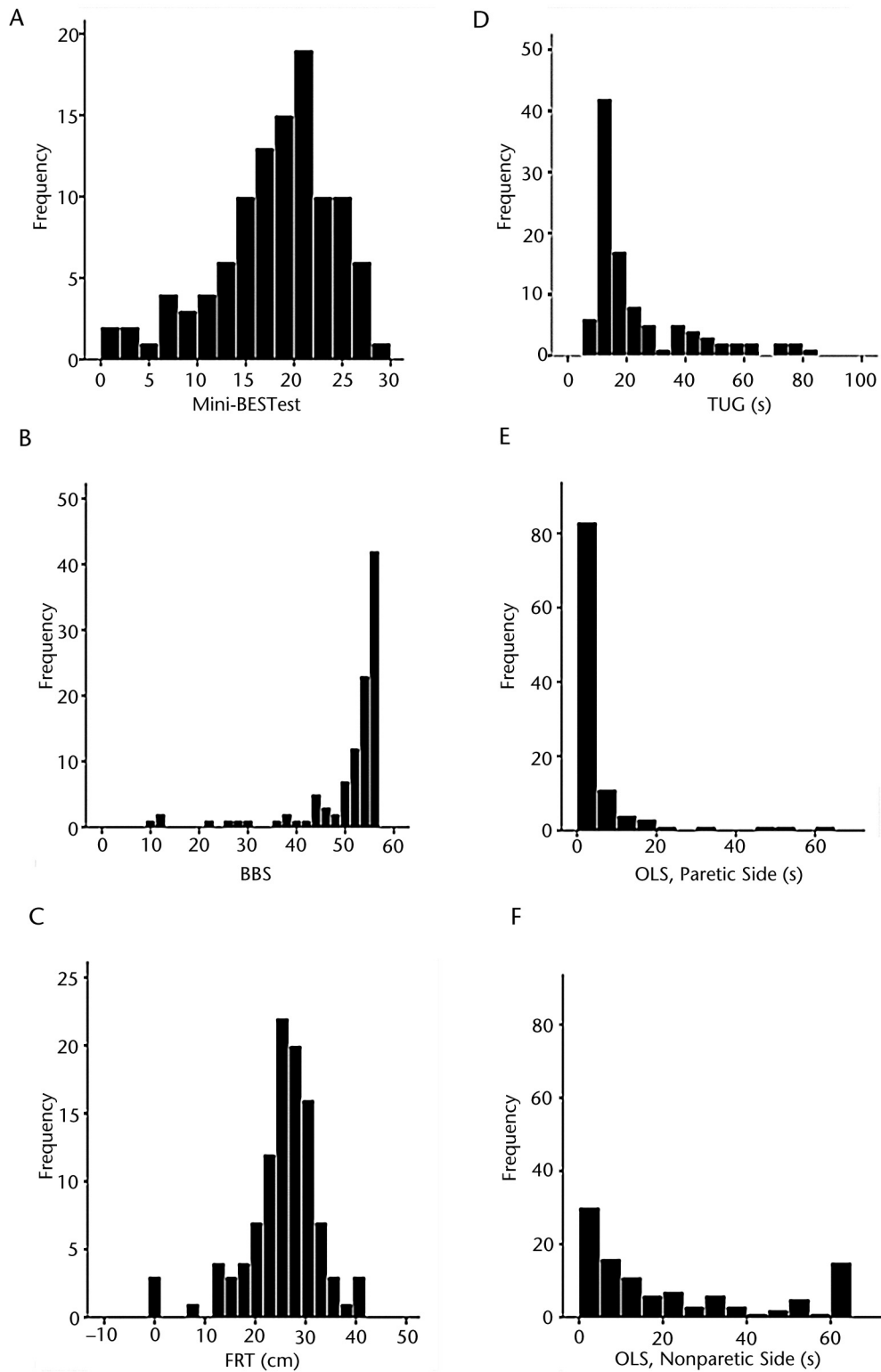
Descriptor	Stroke Group (n=106)	Control Group (n=48)	P
Demographics			
Age, y	57.1 (11.0)	60.2 (9.3)	.09
Sex (male/female), n	73/33	28/20	.20
Body mass index, kg/m <sup>2</sup>	24.9 (3.8)	23.9 (3.1)	.11
Poststroke duration, y, median (IQR)	2.9 (1.2–5.5)		
Hemiplegic side (left/right), n	46/60		
Chedoke McMaster Stroke Assessment, median (IQR)			
Leg (1–7)	4.0 (4.0–5.0)		
Foot (1–7)	3.0 (2.8–4.0)		
Arm (1–7)	3.0 (2.8–5.0)		
Hand (1–7)	3.0 (2.0–5.0)		
Type of stroke			
TACI/PACI/PCI/LCI/hemorrhage/unknown, n	0/15/9/32/46/4		
Modified Ashworth Scale (0–4), median (IQR)	1.5 (1.0–2.0)		
Walking aid for indoor walking			
None/cane/quadrupod/wheelchair/others, n	70/11/14/4/7	0/0/0/0/0	
Geriatric Depression Scale (0–15), median (IQR)	5.0 (3.0–9.0)		
Abbreviated Mental Test (0–10), median (IQR)	10.0 (9.0–10.0)		
Activities-specific Balance Confidence (ABC) Scale (0–100)	71.3 (31.4)		
Balance performance, median (IQR)			
Mini-BESTest (0–28)	19.0 (14.0–22.0)	27.0 (26.0–27.0)	
Berg Balance Scale (0–56)	54.0 (50.0–56.0)		
Functional Reach Test, cm	25.4 (22.9–30.5)		
One-leg standing: paretic side, s	1.3 (0.8–4.4)		
One-leg standing: nonparetic side, s	12.7 (4.4–36.0)		
Timed “Up & Go” Test, s	16.6 (12.1–35.2)		

<sup>a</sup> Values are mean ± SD unless otherwise indicated. IQR=interquartile range, LCI=lacunar circulation infarct, Mini-BESTest=Mini-Balance Evaluation Systems Test, PACI=partial anterior circulation infarct, PCI=posterior circulation infarct, TACI=total anterior circulation infarct.

in Figure 1B–F. We found that the Mini-BESTest had significantly less skewness than other balance measures ( $P \leq .001$ ), except OLS on the nonparetic side ( $P = .965$ ) (Tab. 2). The proportion of participants with the lowest and highest possible Mini-BESTest scores was 0% and 0.9%, respectively. The BBS had the most severe ceiling effect, with 32% of the individuals achieving the highest possible score.

### Reliability Analysis

Thirty individuals with stroke participated in the reliability assessment. The Mini-BESTest demonstrated good internal consistency, with Cronbach alpha values of .89, .93, and .94 for raters 1, 2, and 3, respectively. Intrarater reliability of the Mini-BESTest total score was excellent (ICC [3,1] = .97,  $P \leq .001$ ), yielding an MDC<sub>95</sub> value of 3.0 points. The Mini-BESTest total score also



**Figure.**

Score distribution of the balance tests. Frequency distributions of scores on the (A) Mini-Balance Evaluation Systems Test (Mini-BESTest), (B) Berg Balance Scale (BBS), (C) Functional Reach Test (FRT), (D) Timed “Up & Go” Test (TUG), (E) one-leg standing (OLS) (paretic side), and (F) OLS (nonparetic side) are shown. The data of 106 individuals with stroke are shown, except for the TUG, which was based on 102 participants with stroke only, as 4 participants were unable to walk without manual assistance.

showed excellent interrater reliability (ICC [2,1]=.96,  $P \leq .001$ ). When the test items were analyzed separately, adequate to excellent intrarater and interrater reliability were found for all items (Tab. 3), except for item 5 (compensatory stepping correction in backward direction), item 6 (compensatory stepping correction in lateral direction), and item 8 (stand on foam surface with eyes closed), which showed fair reliability (kappa = .30-.40).

### Validity Analysis

**Concurrent validity.** In the stroke group, significant relationships were found between the Mini-BESTest total score and the BBS ( $\rho = .83$ ,  $P \leq .001$ ), FRT ( $\rho = .55$ ,  $P \leq .001$ ), OLS on the paretic side ( $\rho = .83$ ,  $P \leq .001$ ), OLS on the nonparetic side ( $\rho = .54$ ,  $P \leq .001$ ), and TUG ( $\rho = -.82$ ,  $P \leq .001$ ).

**Convergent and discriminant validity.** In the stroke group, the Mini-BESTest total score was significantly correlated with the Chedoke-McMaster Stroke Assessment leg score ( $\rho = .53$ ,  $P \leq .001$ ) and foot score ( $\rho = .64$ ,  $P \leq .001$ ), MAS ( $\rho = -.22$ ,  $P = .02$ ), and ABC Scale ( $\rho = .50$ ,  $P \leq .001$ ), but not with the GDS ( $\rho = -.17$ ,  $P = .08$ ) and AMT ( $\rho = .08$ ,  $P = .42$ ), thus demonstrating good convergent and discriminant validity.

**Known-groups validity.** Significant differences in the Mini-BESTest total score and most individual item scores were found between the stroke and control groups and between fallers and nonfallers in the stroke group (Tab. 4).

**ROC curve analysis.** Receiver operating characteristic curves were constructed to assess the ability of the various balance measures to distinguish people with stroke with and without a history of falls (Tab. 5). The cutoff score for the Mini-BESTest

**Table 2.**

Comparison of Mini-BESTest With Other Balance Measures: Floor and Ceiling Effects<sup>a</sup>

Balance Measure	Skewness ( $\gamma_1$ )	Floor Effect (% Participants With Lowest Possible Score)	Ceiling Effect (% Participants With Highest Possible Score)
Mini-BESTest (0-28)	-0.81	0	0.9
Berg Balance Scale (0-56)	-2.69 <sup>b</sup>	0	32.1
Functional Reach Test, cm	-1.15 <sup>b</sup>	2.8	NA
One-leg standing: paretic side, s	4.06 <sup>b</sup>	13.2	0.9
One-leg standing: nonparetic side, s	0.80	7.5	14.2
Timed "Up & Go" Test, s	1.69 <sup>b,c</sup>	NA	NA

<sup>a</sup> Mini-BESTest=Mini-Balance Evaluation Systems Test, NA=not applicable.

<sup>b</sup> Significant difference in skewness compared with the Mini-BESTest ( $P \leq .001$ ).

<sup>c</sup> The analysis of skewness was based on 106 participants with stroke, except for the Timed "Up & Go" Test data, which were based on 102 people with stroke only.

was 17.5, and the ROC curve yielded an AUC of 0.64 (95% CI=0.51-0.77), a sensitivity of 64.0% (95% CI=44.5-79.7), and a specificity of 64.2% (95% CI=53.3-73.7). The associated LR+ and LR- values were 1.8 (95% CI=1.2-2.7) and 0.6 (95% CI=0.3-1.0), respectively. The AUC value of the Mini-BESTest then was compared with that of the BBS, TUG, OLS, and FRT. We found that the AUC of the Mini-BESTest was significantly smaller than that of the BBS ( $\chi^2 = 7.36$ ,  $P = .01$ ). The AUC of the Mini-BESTest was not significantly different from that of the TUG ( $\chi^2 = 0.05$ ,  $P = .82$ ), OLS on the paretic side ( $\chi^2 = 0.80$ ,  $P = .37$ ), OLS on the nonparetic side ( $\chi^2 = 0.01$ ,  $P = .90$ ), and FRT ( $\chi^2 = 0.48$ ,  $P = .49$ ).

### Discussion

In this study, the psychometric properties of the Mini-BESTest for people with chronic stroke were examined. The ceiling and floor effects and ability of the Mini-BESTest to identify fallers among individuals with chronic stroke also were systematically compared with those of 4 other balance measures for the first time. The study showed that the Mini-BESTest is a reliable and valid measure of balance performance for community-dwelling individuals with chronic stroke, with no signifi-

cant floor or ceiling effects. The association between the Mini-BESTest and fall history, however, is limited.

### Score Distribution and Ceiling and Floor Effects

Our results showed that among the various balance measures, the Mini-BESTest has the least floor or ceiling effects, as indicated by both the degree of skewness and the proportion of participants with minimum and maximum possible scores. In contrast, a significant ceiling effect was found for the BBS (32.5%). Mao et al<sup>22</sup> found a similar ceiling effect of the BBS among patients with chronic stroke (at 180 days after discharge) (28.8%). A study comparing the Mini-BESTest with the BBS in patients with PD also showed that the score distribution for the BESTest was significantly less skewed than that for the BBS.<sup>31</sup> Our data revealed that the score distribution for the TUG demonstrated substantial skewness (Tab. 2), with almost half of our participants with stroke being able to complete the task within 15 seconds (ie, ceiling effect) (Fig. 1D). The BBS consists of a good number of relatively less demanding tasks such as sitting unsupported, standing unsupported, and moving from sitting to standing, whereas the TUG is a single-item assessment involving



Table 3.

Intrarater and Interrater Reliability of the Mini-BESTest<sup>a</sup>

Mini-BESTest item score	Intrarater Reliability (n=30)								Interrater Reliability (n=30)							
	Count <sup>b</sup> (Time 1)			Count <sup>b</sup> (Time 2)			Kappa	P	Count <sup>b</sup> (Rater 2)			Count <sup>b</sup> (Rater 3)			Kappa	P
	0	1	2	0	1	2			0	1	2	0	1	2		
1. Sit to stand	1	4	25	1	4	25	1.00	≤.001 <sup>c</sup>	1	4	25	1	4	25	1.00	≤.001 <sup>c</sup>
2. Rise to toes	13	11	6	13	12	5	.58	≤.001 <sup>c</sup>	13	11	6	14	11	5	.68	≤.001 <sup>c</sup>
3a. Paretic side, stand on one leg	6	22	2	7	20	3	.64	≤.001 <sup>c</sup>	6	22	2	14	3	3	.49	≤.001 <sup>c</sup>
3b. Nonparetic side, stand on one leg	4	20	6	6	15	9	.60	≤.001 <sup>c</sup>	4	20	6	3	19	8	.67	≤.001 <sup>c</sup>
4. Compensatory stepping correction in forward direction	9	0	21	9	0	21	.84	≤.001 <sup>c</sup>	9	0	21	9	0	21	.84	≤.001 <sup>c</sup>
5. Compensatory stepping correction in backward direction	14	4	12	18	2	10	.37	.01 <sup>c</sup>	14	4	12	14	6	10	.57	≤.001 <sup>c</sup>
6a. Displacement toward the paretic side (stroke group) or left side (control group): compensatory stepping correction in lateral direction	20	2	8	22	3	5	.64	≤.001 <sup>c</sup>	20	2	8	22	4	4	.36	.01 <sup>c</sup>
6b. Displacement toward the nonparetic side (stroke group) or right side (control group): compensatory stepping correction in lateral direction	16	0	14	18	0	12	.73	≤.001 <sup>c</sup>	16	0	14	12	4	14	.36	.02 <sup>c</sup>
7. Stance, eyes open on firm and flat surface	2	2	26	2	2	26	1.00	≤.001 <sup>c</sup>	2	2	26	2	2	26	1.00	≤.001 <sup>c</sup>
8. Stance, eyes closed on foam surface	5	22	3	6	20	4	.43	.01 <sup>c</sup>	5	2	3	12	16	2	.38	.01 <sup>c</sup>
9. Stance, eyes closed on firm and inclined surface	3	1	26	3	1	26	1.00	≤.001 <sup>c</sup>	3	1	26	3	1	26	1.00	≤.001 <sup>c</sup>
10. Change in gait speed	4	2	24	4	2	24	.80	≤.001 <sup>c</sup>	4	2	24	5	5	20	.46	≤.001 <sup>c</sup>
11. Walk with horizontal head turns	5	16	9	5	17	8	.61	≤.001 <sup>c</sup>	5	16	9	5	6	19	.41	≤.001 <sup>c</sup>
12. Walk with pivot turns	5	24	1	5	25	0	.89	≤.001 <sup>c</sup>	5	24	1	5	21	4	.76	≤.001 <sup>c</sup>
13. Step over obstacle	19	8	3	19	10	1	.54	≤.001 <sup>c</sup>	19	8	3	15	8	7	.43	≤.001 <sup>c</sup>
14. TUG and TUG with dual task (cognitive)	5	19	6	5	22	3	.76	≤.001 <sup>c</sup>	5	19	6	5	18	7	.70	≤.001 <sup>c</sup>
<b>Mini-BESTest total score</b>	<b>Time 1 Median (IQR)</b>		<b>Time 2 Median (IQR)</b>		<b>ICC (3,1)</b>	<b>P</b>	<b>Rater 2 Median (IQR)</b>		<b>Rater 3 Median (IQR)</b>		<b>ICC (2,1)</b>	<b>P</b>				
	18.0 (12.0–21.0)		16.5 (13.8–21.0)		.97	≤.001 <sup>c</sup>	18.0 (12.0–21.0)		18.0 (11.0–22.0)		.96	≤.001 <sup>c</sup>				

<sup>a</sup> Mini-BESTest=Mini-Balance Evaluation Systems Test, TUG=Timed “Up & Go” Test, IQR=interquartile range, ICC=intraclass correlation coefficient.

<sup>b</sup> Count: the number of participants who received a score of 0, 1, or 2 for each item is shown.

<sup>c</sup> Statistically significant at  $P \leq .05$  (kappa for item scores or ICC for total scores).

only moving from sitting to standing, walking, and turning. The majority of our participants, however, have regained their ambulatory function, thus leading to a ceiling effect. In contrast, the inclusion of more challenging tasks such as postural responses to external perturbations (items 4–6) and walking balance tasks (items 11–14) in the Mini-BESTest may have improved the discrimination between participants.

The OLS (paretic side) showed considerable positive skewness, indicating a possible floor effect. It reveals that maintaining balance while standing on the paretic leg remains a very difficult task for many individuals with stroke, despite all of our participants being community-dwelling. Eighty-three (78%) of our participants with stroke had an OLS time of less than 5 seconds, and 14 (13%) of these individuals were even

unable to perform the task (ie, score of 0 second) (Fig. 1E).

### Reliability

The Mini-BESTest had high internal consistency (Cronbach alpha=.89–.94), indicating all of the items measure the same underlying attribute. The intrarater and interrater reliability of the Mini-BESTest also were excellent when administered to people with stroke, comparable to those

**Table 4.**  
Known-Groups Validity of the Mini-BESTest<sup>a</sup>

Mini-BESTest item score	Stroke Group (n=106)				Control Group (n=48)				P	Fallers (n=25)				Nonfallers (n=81)				P
	Count <sup>b</sup>			Mean Rank	Count <sup>b</sup>			Mean Rank		Count <sup>b</sup>			Mean Rank	Count <sup>b</sup>			Mean Rank	
	0	1	2		0	1	2			0	1	2		0	1	2		
1. Sit to stand	1	8	97	75.5	0	0	48	82.0	.04 <sup>c</sup>	1	4	20	47.3	0	4	77	55.4	.01 <sup>c</sup>
2. Rise to toes	27	43	36	62.0	0	1	47	111.7	≤.001 <sup>c</sup>	11	9	5	41.5	16	34	31	57.2	.01 <sup>c</sup>
3a. Paretic side (stroke group) or left side (control group), stand on one leg	12	86	8	58.9	0	8	40	118.7	≤.001 <sup>c</sup>	6	19	0	43.7	6	67	8	56.5	.01 <sup>c</sup>
3b. Nonparetic side (stroke group) or right side (control group), stand on one leg	7	56	43	66.5	0	7	41	101.8	≤.001 <sup>c</sup>	4	15	6	42.3	3	41	37	56.9	.02 <sup>c</sup>
4. Compensatory stepping correction in forward direction	24	20	62	67.9	0	1	47	98.7	≤.001 <sup>c</sup>	10	1	14	48.6	14	19	48	55.0	.30
5. Compensatory stepping correction in backward direction	34	29	43	63.6	0	1	47	108.3	≤.001 <sup>c</sup>	15	2	8	41.6	19	27	35	57.2	.01 <sup>c</sup>
6a. Displacement toward the paretic side (stroke group) or left side (control group): compensatory stepping correction in lateral direction	66	11	29	60.6	0	3	45	114.8	≤.001 <sup>c</sup>	17	3	5	49.8	49	8	24	54.6	.43
6b. Displacement toward the nonparetic side (stroke group) or right side (control group): compensatory stepping correction in lateral direction	41	3	62	68.8	0	4	44	96.8	≤.001 <sup>c</sup>	14	0	11	45.0	27	3	51	56.1	.07
7. Stance, eyes open on firm and flat surface	3	3	100	76.1	0	0	48	80.5	.09	1	3	21	48.1	2	0	79	55.1	.01 <sup>c</sup>
8. Stance, eyes closed on foam surface	16	69	21	59.5	0	3	45	117.2	≤.001 <sup>c</sup>	7	13	5	48.1	9	56	16	55.1	.23
9. Stance, eyes closed on firm and inclined surface	3	3	100	76.1	0	0	48	80.5	.09	2	1	22	50.0	1	2	78	54.5	.11
10. Change in gait speed	5	15	86	73.0	0	0	48	87.5	≤.001 <sup>c</sup>	2	5	18	48.5	3	10	68	55.0	.17
11. Walk with horizontal head turns	9	34	63	70.0	0	5	43	94.0	≤.001 <sup>c</sup>	5	9	11	43.5	4	25	52	56.5	.03 <sup>c</sup>
12. Walk with pivot turns	16	66	24	61.1	0	5	43	113.8	≤.001 <sup>c</sup>	6	17	2	43.2	10	49	22	56.6	.02 <sup>c</sup>
13. Step over obstacle	57	28	21	59.2	0	4	44	118.0	≤.001 <sup>c</sup>	17	5	3	45.5	40	23	18	55.9	.12 <sup>c</sup>
14. TUG and TUG with dual task (cognitive)	21	68	17	75.4	0	45	3	82.2	.13	7	14	4	49.8	14	54	13	54.6	.42
<b>Mini-BESTest total score</b>	<b>Stroke Group Median (IQR)</b>			<b>Control Group Median (IQR)</b>	<b>P</b>			<b>Fallers Median (IQR)</b>	<b>Nonfallers Median (IQR)</b>			<b>P</b>						
	19.0 (14.0–22.0)			27.0 (26.0–27.0)	≤.001 <sup>c</sup>			16.0 (10.5–21.0)	19.0 (15.5–22.0)			.03 <sup>c</sup>						

<sup>a</sup> Mini-BESTest=Mini-Balance Evaluation Systems Test, TUG=Timed “Up & Go” Test, IQR=interquartile range.

<sup>b</sup> Count: the number of participants who received a score of 0, 1, or 2 for each item is shown.

<sup>c</sup> Statistically significant difference at  $P \leq .05$  (Mann-Whitney *U* test).

of the BBS (intrarater=.92-.98, interrater=.93-.99),<sup>15,22,30,40</sup> TUG (intrarater=.96),<sup>40</sup> OLS (intrarater=.88-.92),<sup>18</sup> and FRT (interrater=.99)<sup>40</sup> previously reported in people with stroke. Our results are thus in line with those of Godi et al,<sup>30</sup> who found that the Mini-BESTest had excellent intrarater reliability (ICC=.96) and interrater reliability (ICC=.98) in a

sample of people with different balance disorders. Leddy et al<sup>32</sup> also evaluated both the intrarater and interrater reliability of the Mini-BESTest, and their results obtained from patients with PD are similar to ours (intrarater=.88-.91, interrater=.91-.96). The MDC<sub>95</sub> obtained in our study was 3.0 points, which represents the minimum difference

that would reflect a real change in the mini-BESTest total score. Godi et al<sup>30</sup> found a very similar MDC<sub>95</sub> value (3.5 points) in their sample of participants with mixed conditions. The minimal detectable change established here would be useful for future stroke clinical trials in determining whether the experimental

intervention has caused any real change in balance ability.

It is noted that item 5 (compensatory stepping correction in a backward direction), item 6 (compensatory stepping correction in a lateral direction), and item 8 (standing on a foam surface with eyes closed) showed fair reliability only. The discrepancies in scoring between the 2 testing sessions or between the 2 raters may have been partly due to the actual change in patients' performance. These 3 items represent the more challenging tasks, with the majority of participants attaining a score of only 0 or 1 at initial assessment (Tab. 4). A patient's performance of these tasks thus might be more variable with repeated testing. For the compensatory stepping reaction tests (items 5 and 6), the lower agreement in scores also might be related to the consistency of the therapist in applying the displacement. A slight increase or decrease in magnitude of the displacing force applied by the therapist might elicit a very different balance response from the patient.

### Validity

We found that the Mini-BESTest total score was significantly associated with other established balance measures (BBS, OLS, FRT, and TUG) and other measures evaluating related concepts (lower-limb motor recovery, ABC Scale), but not with measures assessing different attributes (eg, GDS, AMT), thus demonstrating good concurrent, convergent, and discriminant validity, respectively. Our results are in agreement with King et al,<sup>31</sup> who found a strong association of the Mini-BESTest with the BBS ( $r=.79$ ) and Unified Parkinson's Disease Rating Scale motor score ( $r=-.51$ ) among patients with PD. The results showed that the Mini-BESTest total score was able to separate people with different balance abilities (ie, known-groups validity),

as indicated by the significant difference in scores between the stroke and control groups and between people with stroke with and without a history of falls. Our results concord with the findings of King et al,<sup>31</sup> who showed that the Mini-BESTest can effectively distinguish between individuals with and without postural response deficits as defined by the Hoehn and Yahr scale.

When comparing the ROC curves, however, the results show that the Mini-BESTest (AUC=0.64, 95% CI=0.51-0.77), similar to the TUG (AUC=0.66, 95% CI=0.53-0.80), OLS on the paretic side (AUC=0.67, 95% CI=0.54-0.80), OLS on the nonparetic side (AUC=0.64, 95% CI=0.52-0.77), and FRT (AUC=0.67, 95% CI=0.55-0.79), has a limited association with fall history (AUC <0.7). Only the BBS showed a reasonable AUC value of 0.72 (95% CI=0.61-0.83), which was significantly greater than that of the Mini-BESTest. Whether this statistically significant difference in AUC was clinically meaningful will need further study.

The limited association of the Mini-BESTest with fall history in people with stroke may be explained by several reasons. First, it is well known that the causes of falls are multifactorial. Many factors other than balance ability, both intrinsic and extrinsic, may contribute to falls after stroke.<sup>58</sup> For example, Harris et al<sup>27</sup> found that ambulatory individuals with stroke who attained a low BBS score and used a wheelchair or walker for longer distances had lower risk for falls compared with those who had a higher BBS score and only used a cane for ambulation. Apparently, the relationship between balance and falls is not linear and involves the interplay of many other factors. This possible explanation may partly explain why balance assessment tools, when used

alone, may not be effective in predicting falls in people with stroke. Indeed, a number of previous studies have shown that various balance assessment tools commonly used in stroke rehabilitation, such as the BBS and TUG, have limited ability to predict falls after chronic stroke.<sup>25-27,59</sup> Second, the fall data were collected retrospectively, which is more susceptible to recall problems and bias than when a prospective design is used for fall data collection. For example, a fall that occurred earlier in the period (eg, 10 months previously) may not be reported compared with a fall that occurred more recently (eg, 2 weeks previously). One may not recall a fall that was relatively inconsequential compared with a fall that necessitated medical attention. Further study should assess the utility of the Mini-BESTest for predicting future falls in patients with stroke.

Our results are in contrast to the findings of Duncan et al,<sup>34</sup> who examined the relationship between the Mini-BESTest and recurrent falls during the previous 6 months (retrospective) and future 12 months (prospective) in a sample of 80 patients with PD. Their results showed a strong association of the Mini-BESTest with recurrent falls, both retrospectively and prospectively. The AUC values reported were 0.77 to 0.86, with a sensitivity of 0.62 to 0.88, a specificity of 0.74 to 0.78, an LR+ of 2.4 to 4.0, and an LR- of 0.15 to 0.52. The discordance in results between their study and ours may be explained by the different study population and research methods. Patients with PD were used in their study, whereas our sample consisted of only people with chronic stroke. In their study, the Mini-BESTest was used to predict recurrent fallers (those who experienced 2 or more falls), whereas the faller group included both single and recurrent fallers in our study. The fall

rate reported also was higher in their study. The proportion of fallers in our study was 23.6%, and only 6.6% were recurrent fallers, whereas 27.5% and 32.5% of their study participants reported recurrent falls in the previous 6 months and the 12-month follow-up period, respectively. The lower fall rate may be due to several factors. First, our sample was relatively young (mean age=57.1 years). The time since the onset of stroke was more than 6 months for all of our participants (median=2.9 years). Thus, they likely had developed compensatory strategies in their adaptation to a chronic and presumably more stable condition. In contrast, the patients with PD in the study by Duncan et al<sup>34</sup> were older (mean age=68.2 years) and were coping with a disease that was progressive in nature.

**Limitations and Future Research Directions**

This study has several limitations. First, because the participants in the stroke group were community-dwelling and most were ambulatory, the results are generalizable only to people with similar characteristics. Further research is needed to validate the Mini-BESTest in people who are in acute or subacute stages of stroke recovery, severely impaired, or institutionalized. Second, the ability to carry on a normal conversation was used as an eligibility criterion, but it may not be equivalent to being able to follow directions. Perhaps a cutoff score of a standardized assessment of cognition should have been used to determine eligibility. Third, the actual number of enrolled participants was higher than that derived from the sample size calculation described in the “Method” section. We received an overwhelming response, and a large number of people volunteered to participate in our study. As there were no substantial budgetary concerns, we decided to measure all volunteers who were eli-

**Table 5.** Comparison of Mini-BESTest With Other Balance Measures: Differentiating Between Fallers and Nonfallers in the Stroke Group<sup>a</sup>

Balance Measure	Fallers (n=25) Median (IQR)	Nonfallers (n=81) Median (IQR)	Cutoff Score	Distinguishing Fallers From Nonfallers				
				Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)	AUC (95% CI)
Mini-BESTest (0-28)	16.5 (7.5-21.0)	19.0 (15.0-22.0) <sup>b</sup>	17.5	64.0% (44.5-79.7)	64.2% (53.3-73.7)	1.8 (1.2-2.7)	0.6 (0.3-1.0)	0.64 (0.51-0.77)
BBS (0-56)	50.0 (43.0-54.0)	54.0 (51.0-56.0) <sup>b</sup>	50.5	52.0% (33.5-70.0)	80.2% (70.3-87.4)	2.6 (1.5-4.5)	0.6 (0.4-0.9)	0.72 (0.61-0.83)
FRT, cm	22.8 (19.0-27.9)	26.6 (22.8-30.4) <sup>b</sup>	24.1	52.0% (33.5-70.0)	74.0% (63.6-82.4)	2.0 (1.2-3.4)	0.6 (0.4-1.0)	0.67 (0.55-0.79)
OLS: paretic side, s	0.9 (0.0-2.3)	1.5 (1.0-5.0) <sup>b</sup>	0.9	56.0% (37.0-73.3)	77.8% (67.6-85.5)	2.5 (1.5-4.3)	0.6 (0.4-0.9)	0.67 (0.54-0.80)
OLS: nonparetic side, s	7.5 (1.0-20.1)	16.0 (5.1-40.0) <sup>b</sup>	3.6	40.0% (23.4-59.3)	84.0% (74.4-90.4)	2.5 (1.2-5.0)	0.7 (0.5-1.0)	0.64 (0.52-0.77)
TUG, s <sup>c</sup>	23.4 (13.3-50.6)	14.8 (11.6-21.1)	19.0	60.9% (40.8-77.8)	67.1% (56.2-76.4)	1.8 (1.2-2.9)	0.6 (0.3-1.0)	0.66 (0.53-0.80)

<sup>a</sup> Mini-BESTest=Mini-Balance Evaluation Systems Test, IQR=interquartile range, 95% CI=95% confidence interval, LR+=positive likelihood ratio, LR-=negative likelihood ratio, AUC=area under the curve, BBS=Berg Balance Scale, FRT=Functional Reach Test, OLS=one-leg standing, TUG=Timed “Up & Go” Test.  
<sup>b</sup> Significant difference between fallers and nonfallers at P≤.001 (Mann-Whitney U test).  
<sup>c</sup> The analysis of TUG data was based on 102 participants (23 fallers, 79 nonfallers).

gible. Although the power analysis a priori helped us to determine the minimum sample size required to detect significant findings, a larger sample size presumably would have further increased the statistical power of the study. Indeed, with the current sample size of 106 people with stroke, the power was increased to 0.95, if the alpha level (.05) and acceptable and expected AUC (0.7 and 0.9, respectively) remained the same as originally planned.

We also acknowledge that other clinical balance scales are available for patients with stroke, including the Postural Assessment Scale for Stroke Patients, Trunk Control Test, and many others,<sup>14,28,60-62</sup> but were not used for comparison with Mini-BESTest in this study. We selected only the most commonly used balance assessment tools in stroke rehabilitation and research for comparison. In addition, feasibility of the study and patient fatigue would be concerns if more balance tests were added to the assessment battery. Another interesting research question has to do with the responsiveness of the Mini-BESTest. Godi et al<sup>30</sup> found that the Mini-BESTest is more responsive to change in balance ability than the BBS in a sample consisting of patients with different balance disorders. Is the Mini-BESTest more responsive than other balance measures in detecting treatment effects among individuals with stroke at different stages of recovery? Further study is needed to address this interesting and important question.

Overall, although the association of fall history with the Mini-BESTest is limited, the Mini-BESTest remains a better option than other balance measures used in this study to assess balance function in community-dwelling people with chronic stroke who have mild to moderate neurological impairments, as it has excel-

lent reliability and validity, with no significant floor and ceiling effects. Additionally, compared with single-item measures such as the TUG and OLS, the Mini-BESTest is useful in identifying specific postural control problems and directing treatment.

Ms Tsang and Dr Pang provided concept/idea/research design and project management. Ms Tsang, Mr Liao, and Dr Pang provided writing. Ms Tsang and Mr Liao provided data collection. All authors provided data analysis. Dr Pang provided fund procurement and facilities/equipment. Ms Tsang provided institutional liaisons. Ms Tsang, Dr Chung, and Dr Pang provided consultation (including review of manuscript before submission).

Ethics approval for the study was granted by the Ethics Review Committee of the Hong Kong Polytechnic University.

The preliminary data were presented in abstract format at the 21st European Stroke Conference; May 22-25, 2012; Lisbon, Portugal.

Mr Liao was supported by a full-time research studentship granted by the Hong Kong Polytechnic University.

DOI: 10.2522/ptj.20120454

## References

- 1 The global burden of disease: 2004 update. World Health Organization Health Statistics and Information website. Available at: [http://www.who.int/healthinfo/global\\_burden\\_disease/2004\\_report\\_update/en/index.html](http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/index.html). Accessed January 10, 2012.
- 2 de Haart M, Geurts AC, Huidekoper SC, et al. Recovery of standing balance in post-acute stroke patients: a rehabilitation cohort study. *Arch Phys Med Rehabil*. 2004;85:886-895.
- 3 Perlmutter S, Lin F, Makhsous M. Quantitative analysis of static sitting posture in chronic stroke. *Gait Posture*. 2010;32:53-56.
- 4 Pang MYC, Eng JJ, Dawson AS. Relationship between ambulatory capacity and cardiorespiratory fitness in chronic stroke: influence of stroke-specific impairments. *Chest*. 2005;127:495-501.
- 5 Harley C, Boyd JE, Cockburn J, et al. Disruption of sitting balance after stroke: influence of spoken output. *J Neurol Neurosurg Psychiatry*. 2006;77:674-676.
- 6 Desrosiers J, Noreau L, Rochette A, et al. Predictors of handicap situations following post-stroke rehabilitation. *Disabil Rehabil*. 2002;24:774-785.

- 7 Lamb SE, Ferrucci L, Volapto S, et al. Risk factors for falling in home-dwelling older women with stroke: the Women's Health and Aging Study. *Stroke*. 2003;34:494-501.
- 8 Lin JH, Hsieh CL, Hsiao SF, Huang MH. Predicting long-term care institution utilization among post-rehabilitation stroke patients in Taiwan: a medical centre-based study. *Disabil Rehabil*. 2001;23:722-730.
- 9 de Haart M, Geurts AC, Dault MC, et al. Restoration of weight-shifting capacity in patients with postacute stroke: a rehabilitation cohort study. *Arch Phys Med Rehabil*. 2005;86:755-762.
- 10 Pang MY, Eng JJ, Dawson AS, et al. A community-based fitness and mobility exercise program for older adults with chronic stroke: a randomized, controlled trial. *J Am Geriatr Soc*. 2005;53:1667-1674.
- 11 Smania N, Picelli A, Gandolfi M, et al. Rehabilitation of sensorimotor integration deficits in balance impairment of patients with stroke hemiparesis: a before/after pilot study. *Neurol Sci*. 2008;29:313-319.
- 12 Shumway-Cook A, Woollacott M. Attentional demands and postural control: the effect of sensory context. *J Gerontol A Biol Sci Med Sci*. 2000;55:M10-M16.
- 13 Pollock CL, Eng JJ, Garland SJ. Clinical measurement of walking balance in people post stroke: a systematic review. *Clin Rehabil*. 2011;25:693-708.
- 14 de Oliveira CB, de Medeiros IR, Frota NA, et al. Balance control in hemiparetic stroke patients: main tools for evaluation. *J Rehabil Res Dev*. 2008;45:1215-1226.
- 15 Berg K, Wood-Dauphinée S, Williams JI. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med*. 1995;27:27-36.
- 16 Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. *J Gerontol*. 1990;45:M192-M197.
- 17 Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*. 1991;39:142-148.
- 18 Flansbjerg UB, Blom J, Brogårdh C. The reproducibility of Berg Balance Scale and the Single-leg Stance in chronic stroke and the relationship between the two tests. *PM R*. 2012;4:165-170.
- 19 Bohannon RW, Walsh S, Joseph MC. Ordinal and timed balance measurements: reliability and validity in patients with stroke. *Clin Rehabil*. 1993;7:9-13.
- 20 Franchignoni F, Horak F, Godi M, et al. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med*. 2010;42:323-331.
- 21 Leroux A, Pinet H, Nadeau S. Task-oriented intervention in chronic stroke: changes in clinical and laboratory measures of balance and mobility. *Am J Phys Med Rehabil*. 2006;85:820-830.

- 22 Mao HF, Hsueh IP, Tang PF, et al. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke*. 2002;33:1022-1027.
- 23 Tyson SF, DeSouza LH. Development of the Brunel Balance Assessment: a new measure of balance disability post stroke. *Clin Rehabil*. 2004;18:801-810.
- 24 Blum L, Korner-Bitensky. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther*. 2008;88:559-566.
- 25 Belgen B, Beninato M, Sullivan PE, Nariellwalla K. The association of balance capacity and falls self-efficacy with history of falling in community-dwelling people with chronic stroke. *Arch Phys Med Rehabil*. 2006;87:554-561.
- 26 Beninato M, Portney LG, Sullivan PE. Using the *International Classification of Functioning, Disability and Health* as a framework to examine the association of falls and clinical assessment tools in people with stroke. *Phys Ther*. 2009;89:816-825.
- 27 Harris JE, Eng JJ, Marigold DS, et al. Relationship between balance and mobility to fall incidence in people with chronic stroke. *Phys Ther*. 2005;85:150-158.
- 28 Fugl-Meyer AR, Jaasko L, Leyman I, et al. The post stroke hemiplegic patient. I: a method of evaluation of physical performance. *Scand J Rehabil Med*. 1975;7:13-31.
- 29 Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. *Phys Ther*. 2009;89:484-498.
- 30 Godi M, Franchignoni F, Caligari M, et al. Comparison of reliability, validity, and responsiveness of the Mini-BESTest and Berg Balance Scale in patients with balance disorders. *Phys Ther*. 2013;93:158-167.
- 31 King LA, Priest KC, Salarian A, et al. Comparing the Mini-BESTest with the Berg Balance Scale to evaluate balance disorders in Parkinson's disease. *Parkinsons Dis*. 2012;2012:375419. doi: 10.1155/2012/375419.
- 32 Leddy AL, Crouner BE, Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. *J Neurol Phys Ther*. 2011;35:90-97.
- 33 Duncan RP, Leddy AL, Cavanaugh JT, et al. Accuracy of fall prediction in Parkinson disease: six-month and 12-month prospective analyses. *Parkinsons Dis*. 2012;2012:237673. doi: 10.1155/2012/237673.
- 34 Duncan RP, Leddy AL, Cavanaugh JT, et al. Comparative utility of the BESTest, Mini-BESTest, and Brief-BESTest for predicting falls in individuals with Parkinson disease: a cohort study. *Phys Ther*. 2013;93:542-550.
- 35 Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
- 36 Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, NY: John Wiley & Sons Inc; 2000:156-164.
- 37 Hyndman D, Ashburn A, Stack E. Fall events among people with stroke living in the community: circumstances of falls and characteristics of fallers. *Arch Phys Med Rehabil*. 2002;83:165-170.
- 38 Jorgensen L, Engstad T, Jacobsen BK. Higher incidence of falls in long-term stroke survivors than in population controls: depressive symptoms predict falls after stroke. *Stroke*. 2002;33:542-547.
- 39 Forster A, Young J. Incidence and consequences of falls due to stroke: a systematic inquiry. *BMJ*. 1995;311:83-86.
- 40 Wolf SL, Catlin PA, Gage K, et al. Establishing the reliability and validity of measurements of walking time using the Emory Functional Ambulation Profile. *Phys Ther*. 1999;79:1122-1133.
- 41 Flansbjerg UB, Holmbäck AM, Downham D, et al. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med*. 2005;37:75-82.
- 42 Knorr S, Brouwer B, Garland SJ. Validity of the Community Balance and Mobility Scale in community-dwelling persons after stroke. *Arch Phys Med Rehabil*. 2010;91:890-896.
- 43 Gowland C, Stratford PW, Ward M, et al. Measuring physical impairment and disability with the Chedoke-McMaster Stroke Assessment. *Stroke*. 1993;24:58-63.
- 44 Gregson JM, Leathley M, Moore AP, et al. Reliability of the Tone Assessment Scale and the modified Ashworth scale as clinical tools for assessing poststroke spasticity. *Arch Phys Med Rehabil*. 1999;80:1013-1016.
- 45 Mak MK, Lau AL, Law FS, et al. Validation of the Chinese translated Activities-specific Balance Confidence scale. *Arch Phys Med Rehabil*. 2007;88:496-503.
- 46 Hsu PC, Miller WC. Reliability of the Chinese version of the Activities-specific Balance Confidence Scale. *Disabil Rehabil*. 2006;28:1287-1292.
- 47 Botner EM, Miller WC, Eng JJ. Measurement properties of the Activities-specific Balance Confidence Scale among individuals with stroke. *Disabil Rehabil*. 2005;27:156-163.
- 48 Bamford JM. The role of the clinical examination in the subclassification of stroke. *Cerebrovasc Dis*. 2000;10(suppl 4):2-4.
- 49 Lindley RI, Warlow CP, Wardlaw JM, et al. Interobserver reliability of a clinical classification of acute cerebral infarction. *Stroke*. 1993;24:1801-1804.
- 50 Selvarajah JR, Graves M, Wainwright J, et al. Classification of minor stroke: intra- and inter-observer reliability. *Cerebrovasc Dis*. 2009;27:209-214.
- 51 Chu LW, Pei CKW, Ho MH, Chan PT. Validation of the Abbreviated Mental Test (Hong Kong version) in the elderly medical patient. *Hong Kong Med J*. 1995;1:207-211.
- 52 Lam SC, Wong YY, Woo J. Reliability and validity of the Abbreviated Mental Test (Hong Kong version) in residential care homes. *J Am Geriatr Soc*. 2010;58:2255-2257.
- 53 Sheikh, JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *J Aging Ment Health*. 1986;5:165-173.
- 54 Sivrioglu EY, Sivrioglu K, Ertan T, et al. Reliability and validity of the Geriatric Depression Scale in detection of post-stroke minor depression. *J Clin Exp Neuropsychol*. 2009;31:999-1006.
- 55 Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-374.
- 56 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.
- 57 Confidence interval calculator. Available at: <http://www.pedro.org.au/wp-content/uploads/Ccalculator.xls>. Accessed February 26, 2013.
- 58 Pang MYC, Eng JJ. Falls-related self-efficacy, but not balance and mobility performance, is related to accidental falls in chronic stroke survivors with low bone mineral density. *Osteoporos Int*. 2008;19:919-927.
- 59 Eng JJ, Pang MYC, Ashe MA. Balance, falls, and bone health: Role of exercise in reducing fracture risk after stroke. *J Rehabil Res Dev*. 2008;45:297-314.
- 60 Benaïm C, Perennou DA, Villy J, et al. Validation of a standardized assessment of postural control in stroke patients: the Postural Assessment Scale for Stroke Patients (PASS). *Stroke*. 1999;30:1862-1868.
- 61 Franchignoni FP, Tesio L, Ricupero C, Martino MT. Trunk Control Test as an early predictor of stroke rehabilitation outcome. *Stroke*. 1997;28:1382-1385.
- 62 Hsieh C-L, Sheu C-F, Hsueh I-P, Wang C-H. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. *Stroke*. 2002;33:2626-2630.