Reliability and validity of Persian versions of Mini-BESTest and Brief-BESTest in persons with Parkinson’s disease

Amin Nakhostin-Ansari MD, MPH, Noureddin Nakhostin Ansari PT, PhD, Milad Mellat-Ardakani MD, MPH, Mostafa Nematizad MD, Soofia Naghdi PT, PhD, MohammadJavad Babaki MD, Mehdi Farhangian MD, Amir Hassan Habibi MD, Abbas Tafakhori MD, and Scott Hasson PT, EdD

“Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences, Tehran, Iran; Department of Physiotherapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran; Neumusculoskeletal Research Center, Iran University of Medical Sciences, Tehran, Iran; Department of Neurology, Rasoul Akram Hospital, Iran University of Medical Sciences, Tehran, Iran; Department of Neurology, School of Medicine, Imam Khomeini Hospital and Iranian Center of Neurological Research, Tehran University of Medical Sciences, Tehran, Iran; Department of Physical Therapy, Augusta University, Augusta, GA, USA

ABSTRACT

Background: Mini-BESTest and Brief-BESTest are used to assess balance in patients with a wide range of balance disorders. While there are Persian versions of Mini-BESTest and Brief-BESTest, the psychometric properties have not been thoroughly evaluated. This study aimed to assess the reliability and validity of the Persian versions of Mini-BESTest and Brief-BESTest in persons with Parkinson’s disease (PD).

Methods: Three medical students rated videotaped performances of 49 individuals with PD on the Persian Mini-BESTest, Persian Brief-BESTest, and Berg balance scale (BBS). Healthy adults were matched with persons having PD in terms of age and gender.

Results: There were no floor and ceiling effects. Inter- and intra-rater reliability was excellent (ICC = 0.965–0.973). The minimal detectable changes were 2.37 and 3.47 for Persian versions of Mini-BESTest and Brief-BESTest, respectively. The Persian versions of Mini-BESTest and Brief-BESTest had very good correlations with BBS (r > 0.7) confirming construct validity. There was a very good correlation between the Mini-BESTest and the Brief-BESTest total scores (r = 0.78). There were significant differences between the persons with PD and healthy adults on both tests supporting discriminant validity. Significant differences in balance performances across Hoehn and Yahr stages were found which supported known-groups validity.

Conclusion: The Persian versions of Mini-BESTest and Brief-BESTest are reliable and valid instruments for balance evaluation in persons with PD. Further study to determine the reliability and validity of both tests when examining patients in real-time in the clinic is warranted.

ARTICLE HISTORY

Received 23 April 2020
Revised 20 August 2020
Accepted 7 September 2020

KEYWORDS

Balance; Mini-BESTest; Brief-BESTest; Parkinson’s disease; reliability; validity

Introduction

Parkinson’s disease (PD) affects about 4.1 million people globally. It has been estimated that about 8.7 million people will be affected by 2030. Most of the persons with PD will be from developing countries (Dorsey et al., 2007). The incidence of PD increases by age, especially after the age of 60 (Van Den Eeden et al., 2003). Postural instability, freezing of gait, impaired anticipatory and reactive balance, impaired cognition, reduced leg muscle strength, reduced proprioception, and frontal lobe impairment are causes of balance impairment and falls in persons with PD (Kim, Allen, Canning, and Fung, 2013; Latt, Lord, Morris, and Fung, 2009; Paul et al., 2014). About 68.3% of persons with PD fall annually and about 50.5% of them are recurrent fallers (Wood, Bilclough, Bowron, and Walker, 2002). Falling puts persons with PD at a higher risk of fractures, soft tissue injury, immobilization, depression, daily activity restriction, and mortality (Adkin, Frank, and Jog, 2003; Bloem et al., 2001; Bloem, Hausdorff, Visser, and Giladi, 2004; Melton et al., 2006). As a result, balance assessment has a major role in diagnosing persons with PD who are at risk of falling and also the determination of intervention outcomes.

There are several tools to assess balance dysfunction. The Balance Evaluation System Test (BESTest) was developed by Horak, Wrisley, and Frank (2009) and consists of 36 items in 6 sections assessing: (1) Biomechanical constraints; (2) Stability limits/
Vertically; (3) Anticipatory postural adjustment; (4) Postural response; (5) Sensory orientation; and (6) Stability in gait to identify the system responsible for the balance deficit. BESTest is a reliable and valid tool for assessing individuals with PD (Leddy, Crowner, and Earhart, 2011a, 2011b). The BESTest takes about 30–45 minutes to administer; this limits feasibility in the clinic and for research, hence a shorter version named Mini-BESTest was developed (Franchignoni et al., 2010).

Mini-BESTest, developed by Franchignoni et al. (2010), assesses dynamic balance and consists of 14 items that take 10–15 minutes to administer. The Mini-BESTest does not assess the biomechanical constraints and stability limits from the six sections of the BESTest. Each item is scored on a 3 level from 0 to 2 (Total score equals 28 points). Mini-BESTest has high reliability and validity in the balance evaluation of persons with PD disease (Duncan et al., 2012, 2013; King et al., 2012; Leddy, Crowner, and Earhart, 2011b). Despite the reduced time for administration, it is still too lengthy for use in the clinic. In addition, the Mini-BESTest has excluded items on the mechanical constraints and limits of stability, and the remaining items represent the Mini-BESTest as a singular construct of dynamic balance assessing only postural control (Padgett, Jacobs, and Kasser, 2012).

Brief-BESTest, developed by Padgett, Jacobs, and Kasser (2012) is also a shorter version of the BESTest which assesses all sections of the BESTest using the most representative item of each section. Thus, while the Brief-BESTest keeps multiple constructs from the original BESTest, it is one-dimensional and takes less time to administer (Bravini et al., 2016; Padgett, Jacobs, and Kasser, 2012). Brief-BESTest scores are highly correlated with Mini-BESTest and BESTest scores in balance evaluation of persons with PD (Duncan et al., 2013).

While there are Persian versions of Mini-BESTest (http://bestest.us/test_copies, Appendix) and Brief-BESTest (Komijani et al., 2018, Appendix), the psychometric properties of the Persian versions of Mini-BESTest and Brief-BESTest are not thoroughly evaluated. Therefore, the objective of the present study is to determine the inter- and intra-rater reliability, construct validity, discriminant validity, and known-groups validity of Persian versions of the Mini-BESTest and Brief-BESTest in balance evaluation of persons with PD.

**Methods**

**Design**

A prospective study design was used to validate the Persian version of the Mini-BESTest and Brief-BESTest in persons with PD. The study protocol was reviewed and approved by the Research Council of Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences. Ethical approval was granted by the Ethical Committee, Tehran University of medical sciences (IR.TUMS.IKHC.REC.1396.2275). Informed consent was obtained from each patient before entering the study.

**Participants**

Fifty persons diagnosed with PD participated in this study. Persons were recruited from the University Rasoul Akram Hospital and University Imam Khomeini Hospital Complex in Tehran, Iran. All persons with PD and ability to follow commands were included in the study. The exclusion criteria were: (1) history of other neurological disorders; (2) musculoskeletal damage which affects walking ability; (3) presence of serious medical condition; (4) history of surgical treatments for PD; and (5) not willing to participate in the study. Hoehn and Yahr staging was used to determine the severity of the disease. All patients were using medications prescribed by their neurologists to control the disease symptoms.

To assess the discriminant validity of Mini-BESTest and Brief-BESTest, the healthy adult matched for age and gender included in the study with the following inclusion criteria: (1) living independently in the community; (2) able to speak and read Persian; (3) able to follow commands; (4) able to walk 6 m independently; and (5) willing to participate in the study. The exclusion criteria for healthy subjects were: (1) history of faint, vertigo, dizziness, or taking medications which can cause dizziness; and (2) history of any medical condition which can affect balance.

**Raters**

Three 5th year medical students randomly administered and rated the persons’ balance. Before the study initiation, all raters were trained in using the Mini-BESTest and Brief-BESTest according to the instructional DVD of the BESTest provided by Prof. Horak, developer of BESTest, as well as the Berg Balance Scale (BBS) under the supervision of a physical therapist faculty member (Author SN) experienced in neurological rehabilitation. The training session included video demonstrations of tests, scoring method, and practice with feedback.
Medical students were included to assess the reliability and validity of Persian versions of Mini-BESTest and Brief-BESTest as they are usually involved in the clinic and research settings with an independent examination of patients with balance dysfunctions. Further, three raters were included in the present study.

**Procedure**

**Inter-rater reliability**
Persons were assessed in a quiet room with adequate light. All persons were supervised during testing for their safety and to avoid falling. For inter-rater reliability of the Persian versions of Mini-BESTest and the Brief-BESTest, the performances of persons were video recorded in the test session and scored independently by each of three raters. First, persons were assessed for the items of the Brief-BESTest of Hip/Trunk Lateral Strength and Functional Reach Forward; these tests were excluded in the Mini-BESTest. Then, the other items, common between the Mini-BESTest and the Brief BEST were assessed. Persons were allowed to rest for 2 minutes between tests if needed.

**Intra-rater reliability**
For intra-rater reliability, rater 2 repeated the same procedure 2 weeks later, approximately at the same time, and rated the videotaped performances of persons according to the Mini-BESTest and Brief-BESTest.

**Construct validity**
The BBS was the last test applied for the evaluation of construct validity. Rater 2 independently examined and rated the performance of individuals with PD based on the BBS for construct validity.

**Discriminant validity**
Rater 2 also assessed the videotaped performances of all healthy subjects based on the Persian versions of Mini-BESTest and the Brief-BESTest for discriminant validity.

**Outcome measures**
The Mini-BESTest consists of 14 items in 4 sections: (1) Anticipatory postural adjustment; (2) Reactive postural control; (3) Sensory orientation; and (4) Dynamic gait. Two items are assessed in both right and left side and the lower score is used for total score calculation. Each item is scored from 0 to 2; higher score means better balance performance. All items’ scores are summed to obtain the total score which ranges from 0 to 28 (Franchignoni et al., 2010).

The Brief-BESTest consists of eight items from six sections of (1) Biomechanical constraints: hip strength; (2) Stability limits: reach forward; (3) Transitions–anticipatory postural adjustment: stand on one limb with each left and right scored; (4) Reactive postural response: compensatory stepping with each left and right scored; (5) Sensory orientation: stance on foam with eyes closed; and 6) Stability in gait: get up and go test. Each item is scored from 0 to 3; higher score means better balance performance. All items’ scores are summed to obtain the total score which ranges from 0 to 24 (Padgett, Jacobs, and Kasser, 2012).

Berg Balance Scale (BBS) is a balance test which has been translated to Persian and validated (Babaei-Ghazani et al., 2017). BBS consists of 14 items. Each item is scored from 0 to 4 in which higher scores belong to better performances. All item’s scores are summed to obtain the total score, which ranges from 0 to 56 (Berg, Wood-Dauphinee, Williams, and Maki, 1992).

**Statistical analysis**
We used SPSS version 22.0 for data analysis. We set the level of significance at 0.05. Descriptive statistics of mean and standard deviation (SD) or percentages were calculated for demographic and outcome data. We used the Kolmogorov–Smirnov (KS) test to determine the normal distribution of data.

**Floor and ceiling effects**
Frequency of the total scores of the Mini-BESTest and Brief-BESTest was calculated for floor and ceiling effects; it was considered significant if ≥15% of the highest or lowest possible scores were achieved.

**Reliability**
The Intraclass Correlation Coefficient (ICC) [ICC$ _{2,1}$, two-way random effects model, consistency definition] was used to assess the inter-rater reliability and intra-rater reliability. The ICC formula for two-way random effects model (Liljequist, Elfving, and Skavberg Roaldsen, 2019) is: $MSBS-MSE/(MSBS + (k-1) MSE)$; MSBS, mean square between subjects; MSE, mean square error; $k$, number of measurements.

The ICC values were interpreted as excellent (≥0.75), good (0.60–0.75), fair (0.40–0.59) (Fleiss, Levin, and Paik, 2003; Naghdi et al., 2015). The standard error of measurement (SEM) and Minimal Detectable Change (MDC) were calculated using the following formulas: $SEM = SD\sqrt{1-ICC}$; $MDC = 1.96\times\sqrt{2 \times SEM}$. 
Construct validity
Pearson correlation coefficient was used to analyze construct validity. Correlation coefficients were interpreted as follows: 1.0–0.81 (excellent); 0.8–0.61 (very good); 0.6–0.41 (good); 0.4–0.21 (fair); and 0.20–0.00 (poor) (Feise and Menke, 2001).

Discriminant validity
Mann–Whitney U test was used to determine whether there is any significant difference between persons with PD and healthy adults on the Persian versions of Mini-BESTest and the brief-BESTest total score.

Known-groups validity
The mean of scores provided by three raters at test was averaged for analysis of variance (ANOVA) to determine whether there was a significant difference between total scores of Mini-BESTest and Brief-BESTest across the Hoehn and Yahr scale; Bonferroni analysis was performed to determine the differences between stages. Further analysis was performed with the pooled data from persons at stages 1 and 2 (early stage of PD based on Hoehn and Yahr scale) and compared with pooled data from those in stages 3 and 4 (mid-stage of PD based on Hoehn and Yahr scale) using the Independent t-Test.

Results
Characteristics of persons
In this study, 50 persons with PD were recruited and 1 was excluded because of inability in completing the assessments. Therefore, data from 49 persons with PD and 49 healthy adults were included for analyses. As shown in Table 1, the mean (SD) age of persons with PD was 60.8 (13.9) [men: 67.3%; disease duration: 5.27 (4.62) years]. Most persons with PD were in the Hoehn and Yahr stage 2 (36.7%) followed by stage 3 (34.7%), stage 1 (16.3%), and stage 4 (12.2%). No patient was in the Hoehn and Yahr stage of 5.

Table 1. Demographic data of persons with Parkinson’s disease (PD) and healthy adults.

|                          | Persons with PD (n = 49) | Healthy adults (n = 49) |
|--------------------------|--------------------------|------------------------|
|                          | Mean (SD) | Range     | Mean (SD) | Range     |
| Age (years)              | 60.8 (13.9) | 25.0–86.0 | 60.6 (13.3) | 25.0–80.0 |
| Weight (Kg)              | 70.8 (12.7) | 39.0–95.0 | 73.9 (12.6) | 53.0–110.0 |
| Height (cm)              | 166.0 (10.1) | 140.0–185.0 | 169.2 (8.8) | 154.0–188.0 |
| BMI†                     | 25.6 (3.9) | 18.3–34.7 | 25.8 (3.6) | 19.7–34.3 |

* Standard deviation; †body mass index.

Floor and ceiling effect
No floor and ceiling effects were detected. No patient had the lowest possible total score of the Persian versions of Mini-BESTest and Brief-BESTest. The percentage of persons who achieved the highest possible total score on the Mini-BESTest and the Brief-BESTest was 2% and 2.7%, respectively (Table 2).

Reliability
Both Persian versions of the Mini-BESTest and the Brief-BESTest had excellent inter- and intra-rater reliability with ICC2,1 = 0.965–0.973 (Table 2).

Standard error of measurement and minimal detectable change
For Mini-BESTest, the SEM and MDC were 0.85 and 2.37, respectively. For Brief-BESTest, the SEM and MDC were 1.25 and 3.47, respectively.

Construct validity
The mean BBS total score for persons with PD was 48.85 (SD = 6.42, CI 95% = 47.01–50.7). The Pearson correlation test revealed a significant very good association between the Brief-BESTest and the BBS total scores as well as between the Mini-BESTest and the BBS total scores (both r = 0.74, p < .001).

Discriminant validity
In healthy adults, the mean of total scores for the Persian versions of Mini-BESTest and Brief-BESTest was 24.22 (SD 3.2; CI95% 23.31–25.13) and 19.04 (SD 4.7; CI95% 17.7–20.4), respectively. The data from the healthy adults were compared with those of rater 2 in the test stage of the study and showed significant poor performance of persons with PD (P < .001).

Known-groups validity
Descriptive statistics for the Persian versions of Mini-BESTest and Brief-BESTest total scores across Hoehn and Yahr stages of persons with PD are displayed in Table 3. The ANOVA showed significant differences in the Mini-BESTest and the Brief-BESTest total scores across Hoehn and Yahr stages of persons with PD. However, there was not a significant difference between Brief-BESTest total scores between the stages 2 and 3 (p = 1.0). There were no significant differences between the Mini-BESTest total scores of
stages 1 and 2 ($p = .65$) and stages 2 and 3 ($p = 1.0$). The differences between the persons in the early stage (pooled stages 1 and 2) with those in mid-stage (pooled stages 3 and 4) were significant on both the Persian versions of Mini-BESTest ($t = 2.99$, $df = 47$, $p = .004$) and the Brief-BESTest ($t = 3.49$, $df = 47$, $p = .001$).

### Discussion

This study is the first that examined the inter-rater and intra-rater reliability and validity of Persian Mini-BESTest and Brief-BESTest scores when they were administered to persons with PD. Results showed that the Persian versions of Mini-BESTest and Brief-BESTest are reliable and valid tools in balance evaluation of persons with PD.

#### Floor and ceiling effects

We found no floor and ceiling effects for Persian Mini-BESTest in balance evaluation of persons with PD which is in line with those reported for Swedish and English versions of the Mini-BESTest (Bergstrom, Lenholm, and Franzen, 2012; King et al., 2012). We also did not observe floor and ceiling effects for the Persian Brief-BESTest. There was no study assessing the floor and ceiling effects of Brief-BESTest in persons with PD to compare our findings. However, studies in persons with cervical spondylotic myelopathy and chronic stroke have reported no ceiling effect (Chiu and Pang, 2017; Huang and Pang, 2017). The lack of floor and ceiling effects can be an advantage for the Persian versions of the Mini-BESTest and Brief-BESTest as this indicates sensitivity to detect changes after interventions in persons with PD (Naghdi et al., 2016). This implies that if the Persian versions of the Mini-BESTest and Brief-BESTest are administered at follow-up after the conclusion of treatment, there is room to measure change in the patients’ health (improvement or worsening) if they have truly occurred (Nakhostin Ansari et al., 2016).

#### Standard error of measurement and minimal detectable change

Minimal detectable change is the smallest change that falls beyond the measurement error using a tool to measure a phenomenon (e.g. balance). The MDC is important in interpreting the results about the clinical relevance of change that occurs following treatment in a patient and also for studies examining the treatment effectiveness. We found an MDC of 2.37 and 3.47 for the Persian versions of Mini-BESTest and Brief-BESTest, respectively. These results suggest that improvement or worsening after treatment in balance of persons with PD must be above 2.5 or 3.5 on Persian versions of Mini-BESTest and Brief-BESTest, respectively, to be interpreted as clinically relevant. We did not find any report regarding the SEM or MDC for the tests evaluated in the current study. However, the present results on the MDC values are in line with previous reports on the Mini-BESTest (range 3.5 to 5.6) (Godi et al., 2013; Jacome et al., 2018; Lampropoulou et al., 2019) and the Brief-BESTest in persons with end-stage renal disease (i.e. 5.6) (Jacome et al., 2018). In fact, the size of MDC values in this study is somewhat similar to those found across studies conducted with different populations in different languages, cultures, and countries. The consistency of our findings on Persian versions of Mini-BESTest and Brief-BESTest with those reported in different patient populations might be interpreted as supporting the validity of the MDC values obtained from the present study.

#### Construct validity

Construct validity is one of the measures used in the validation of tests. Construct validity is used to determine how well the Persian versions of Mini-BESTest and Brief-BESTest measure what is claimed to be measured. In other words, it was assumed that the Mini-BESTest and Brief-BESTest were constructed to successfully test balance. In the current study, the construct validity of the Persian Mini-BESTest and the

| Test     | Rater 1 | Rater 2 | Rater 3 |
|----------|---------|---------|---------|
|          | Mean    | SD      | 95% CI† | Mean    | SD      | 95% CI† | Mean    | SD      | 95% CI† |
| Brief-BESTest | 14.9    | 5.33    | 13.3–16.4 | 14.5    | 4.4     | 13.3–15.8 | 14.5    | 4.5     | 13.2–15.8 |
| Mini-BESTest  | 21.6    | 4.1     | 20.4–22.8 | 20.4    | 3.6     | 19.3–21.4 | 19.9    | 3.5     | 18.9–20.9 |

*Intraclass correlation coefficient; † standard deviation; ‡ confidence interval.

Table 2. Descriptive statistics and ICC* values for inter- and intra-rater reliability (n = 49).
Table 3. Descriptive statistics of Persian versions of Mini-BESTest and Brief-BESTest across Hoehn and Yahr stages of Parkinson’s disease.

|                  | Stage 1 (n = 8) | Stage 2 (n = 18) | Stage 3 (n = 17) | Stage 4 (n = 6) | ANOVA† |
|------------------|-----------------|------------------|------------------|-----------------|--------|
|                  | Mean            | SD*              | 95% CI           | Mean            | SD*    | 95% CI         | Mean            | SD*    | 95% CI         | Mean            | SD*    | 95% CI         | df   | F    | p      |
| Brief-BESTest    | 20.1            | 2.5              | 18.1–22.2        | 15.3            | 4.7              | 13.0–17.7 | 13.8            | 3.7              | 11.9–15.7 | 8.6              | 6.9–10.2 | 48    | 11.2          | <0.001 |
| Pooled           | 16.8 ± 4.7, range 6.0–23.7 | 12.4 ± 4.0, range 6.7–20.3 | 9.6 ± 3.8, range 6.7–20.3 | 12.1 ± 3.8, range 11.7–26.3 | t = 3.49, df = 47, p = .001 |
| Mini-BESTest     | 24.8            | 1.6              | 23.5–26.1        | 21.6            | 3.2              | 20.0–23.2 | 21.1            | 2.7              | 19.7–22.5 | 15.4              | 12.1–18.8 | 48    | 12.7          | <0.001 |
| Pooled           | 22.6 ± 3.1, range 17.3–27.3 | 19.6 ± 3.8, range 11.7–26.3 | 16.8 ± 4.7, range 6.0–23.7 | 20.3 ± 3.8, range 16.3–24.3 | t = 2.99, df = 47, p = .004 |

* Standard deviation; 95% confidence interval; † analysis of variance.
Persian Brief-BESTest were evaluated by comparing each to the BBS, a widely accepted valid measure of balance, to examine the level of correlation. Both Persian versions of Mini-BESTest and Brief-BESTest showed very good correlations with BBS ($r = 0.74$). These results are consistent with those from studies of English ($r = 0.79$) and Swedish ($r = 0.94$) versions of Mini-BESTest (Bergstrom, Lenholm, and Franzen, 2012; King et al., 2012).

**Discriminant validity**

The discriminant validity for the Persian versions of Mini-BESTest and Brief-BESTest was established through comparison between persons with PD and healthy adults to examine if their scores are different and unrelated. The significant difference between the two groups indicates that the Persian versions of Mini-BESTest and Brief-BESTest are able to discriminate between persons with PD and healthy adult people. A study used the English version of Brief-BESTest and found the Brief-BESTest was able to significantly differentiate people with neurological diseases ($n = 20$; 4 Parkinson disease, 1 stroke, 4 multiple sclerosis, 1 peripheral neuropathy due to diabetes, 1 tremor) from healthy individuals ($n = 9$) (Padgett, Jacobs, and Kasser, 2012). The same study further analyzed the data from the cohort of multiple sclerosis patients ($n = 13$) and demonstrated the discriminant validity of Brief-BESTest in distinguishing participants who fell from those with no fall history (Padgett, Jacobs, and Kasser, 2012). A study of patients with stroke examined the discriminant validity of the Brief-BESTest by correlation with dissimilar measures of Geriatric Depression Scale (GDS) and Montreal Cognitive Assessment (MoCA) and found low correlations that were interpreted as a good discriminant validity characteristic of the Brief-BESTest (Huang and Pang, 2017). In persons with Parkinson’s disease ($n = 80$), the Mini-BESTest mean score of fallers ($14.3 \pm 6.2$) was significantly lower than those of non-fallers ($22.9 \pm 5.5$) which indicates the discriminant validity of Mini-BESTest (Leddy, Crowner, and Earhart, 2011b). Demonstrating both discriminant and construct validity for the Persian versions of Mini-BESTest and Brief-BESTest in line with previous investigations (King et al., 2012; Leddy, Crowner, and Earhart, 2011b; Padgett, Jacobs, and Kasser, 2012) means that both tests measure similar balance constructs and yet at the same time discriminate constructs that are dissimilar.

**Known-groups validity**

Known-groups validity of the Persian versions of Mini-BESTest and Brief-BESTest refers to the ability of the tests in discriminating between subgroups of PD known to differ on the variable of balance. We compared the pooled data from the persons at stages 1 and 2 who are at the earliest stages of PD to those who are at the mid-stages of PD (stages 3 and 4). The differences between these two groups were significant on both tests with better scores on balance for the persons at the early stage, which supports the known-groups validity of the Persian versions of the Mini-BESTest and Brief-BESTest. The Known-groups validity of the Persian versions of Mini-BESTest and Brief-BESTest demonstrated in the current study is in agreement with previous investigations. In persons with PD, King et al. (2012) compared the Berg balance test and the Mini-BESTest and found the Mini-BESTest was more effective than the Berg balance test for differentiating between those with and without balance disorder classified according to the Hoehn and Yahr stages. Another study reported that the Brief-BESTest total scores and individual item scores showed significant differences between persons with chronic stroke and control group confirming good known-groups validity (Huang and Pang, 2017). Taken together, the findings support the known-groups validity of the Mini-BESTest and the Brief-BESTest in distinguishing people with different mean scores on the balance tests, individuals with and without a balance deficit, or those with a different level of the balance dysfunction. Known-groups validity found in the current study further supports the construct validity of the Mini-BESTest and the Brief-BESTest.

**Study limitations**

This validation study of the Persian versions of the Mini-BESTest and Brief-BESTest in assessing the balance of individuals with PD has several limitations. For inter-rater reliability and intra-rater reliability, raters independently viewed the videotapes and rated the balance performance of persons with PD. The level of reliability of measurements was not determined in the clinic, and the findings may differ when examining patients in real-time. The responsiveness to detect changes in balance performance after intervention has not been explored in the current study and thus needs investigation. The
balance performance of persons with PD may change with on-off medication. It is suggested that clinicians test individuals with PD under different medication conditions (i.e. both on and off) using the Persian versions of the Mini-BESTest and Brief-BESTest. This study did not evaluate the ability of the tests to predict falls in individuals with PD. Further studies are needed to determine whether the Persian versions of the Mini-BESTest and Brief-BESTest can in fact predict falls in individuals with PD. This study may lack adequate power to discriminate the different stages of Hoehn and Yahr such that in the stages 1 and 4 there were less than 10 subjects. More subjects, at least 30 for each stage of Hoehn and Yahr, should be included to reach an adequate sample. Further study with a larger sample size is needed to confirm the findings on the known-groups validity of the Persian versions of Mini-BESTest and Brief-BESTest.

In conclusion, the Persian versions of the Mini-BESTest and Brief-BESTest are reliable and valid tools for assessing the balance status of individuals with PD. The Persian versions of the Mini-BESTest and Brief-BESTest are valuable balance tools that can be used by Persian-speaking health professionals to examine balance performance and may assist in developing appropriate interventions for individuals with PD. Further study to investigate the measurement characteristics of both tests in the clinical context when examining patients in real-time is needed.

**Acknowledgments**

The authors wish to thank Prof. Fay B. Horak for her kind cooperation during the translation process of BESTest and Mini-BESTest and providing us with the DVD in using the BESTest. The grant support from Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences is greatly acknowledged (Grant number: 96-01-53-33984).

**Declaration of interest**

The authors report no declarations of interest.

**References**

Adkin AL, Frank JS, Jog MS 2003 Fear of falling and postural control in Parkinson’s disease. Movement Disorders 18(5): 496–502. 10.1002/mds.10396

Babaei-Ghazani A, Mohammadi H, Shahidi GA, Habibi SA, Forogh B, Ahadi T, Eftekharasadat B 2017 Reliability and validity of the Persian translation of Berg balance scale in Parkinson disease. Aging Clinical and Experimental Research 29(5): 857–862. 10.1007/s40520-016-0682-7

Berg KO, Wood-Dauphinee SL, Williams JI, Maki B 1992 Measuring balance in the elderly: Validation of an instrument. Canadian Journal of Public Health 83(Suppl 2): S7–11.

Bergstrom M, Lenholm E, Franzen E 2012 Translation and validation of the Swedish version of the mini-BESTest in subjects with Parkinson’s disease or stroke: A pilot study. Physiotherapy Theory and Practice 28(7): 509–514. 10.3109/09593985.2011.653707

Bloom BR, Grimbergen YA, Cramer M, Willemsen M, Zwierdeman AH 2001 Prospective assessment of falls in Parkinson’s disease. Journal of Neurology 248(11): 950–958. 10.1007/s004150170047

Bloom BR, Hausdorff JM, Visser JE, Giladi N 2004 Falls and freezing of gait in Parkinson’s disease: A review of two interconnected, episodic phenomena. Movement Disorders 19(8): 871–884. 10.1002/mds.20115

Bravini E, Nardone A, Godi M, Guglielmetti S, Franchignoni F, Giordano A 2016 Does the Brief-BESTest meet classical test theory and Rasch analysis requirements for balance assessment in people with neurological disorders? Physical Therapy 96(10): 1610–1619. 10.2522/ptj.20150550

Chiu AY, Pang MY 2017 Assessment of psychometric properties of various balance assessment tools in persons with cervical spondylotic myelopathy. Journal of Orthopaedic and Sports Physical Therapy 47(9): 673–682. 10.2519/jospt.2017.7283

Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Holloway RG, Kieburtz K, Marshall FJ, Ravina BM, Schiffito G, Siderowf A, et al. 2007 Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology 68(5): 384–386. 10.1212/01.wnl.0000247740.47667.03

Duncan RP, Leddy AL, Cavanaugh JT, Dibble LE, Ellis TD, Ford MP, Foreman KB, Earhart GM 2012 Accuracy of fall prediction in Parkinson disease: Six-month and 12-month prospective analyses. Parkinson’s Disease 2012: 237673.

Duncan RP, Leddy AL, Cavanaugh JT, Dibble LE, Ellis TD, Ford MP, Foreman KB, Earhart GM 2013 Comparative utility of the BESTest, mini-BESTest, and brief-BESTest for predicting falls in individuals with Parkinson disease: A cohort study. Physical Therapy 93(4): 542–550. 10.2522/ptj.20120302

Feise RJ, Menke J 2001 Functional rating index: A new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. Spine 26(1): 78–86. 10.1097/00007632-200101010-00015

Fleiss JL, Levin B, Paik MC 2003 Statistical methods for rates and proportions. 3rd. Hoboken, NJ: John Wiley and Sons.

Franchignoni F, Horak F, Godi M, Nardone A, Giordano A 2010 Using psychometric techniques to improve the balance evaluation systems test: The mini-BESTest. Journal of Rehabilitation Medicine 42(4): 323–331. 10.2340/16501977-0537

Godi M, Franchignoni F, Caligari M, Giordano A, Turcato AM, Nardone A 2013 Comparison of reliability, validity, and responsiveness of the mini-BESTest and Berg balance scale in patients with balance disorders. Physical Therapy 93(2): 158–167. 10.2522/ptj.20120171

Horak FB, Wrisley DM, Frank J 2009 The balance evaluation systems test (BESTest) to differentiate balance deficits. Physical Therapy 89(5): 484–498. 10.2522/ptj.20080071
Huang M, Pang MY 2017 Psychometric properties of brief-balance evaluation systems test (Brief-BESTest) in evaluating balance performance in individuals with chronic stroke. Brain and Behavior 7(3): e00649. 10.1002/brb3.649

Jacome C, Flores I, Martins F, Castro C, McPhee CC, Shepherd E, Demain S, Figueiredo D, Marques A 2018 Validity, reliability and minimal detectable change of the balance evaluation systems test (BESTest), mini-BESTest and brief-BESTest in patients with end-stage renal disease. Disability and Rehabilitation 40(26): 3171–3176. 10.1080/09638288.2017.1375034

Kim SD, Allen NE, Canning CG, Fung VS 2013 Postural instability in patients with Parkinson's disease. Epidemiology, pathophysiology and management. CNS Drugs 27: 97–112.

King LA, Priest KC, Salarian A, Pierce D, Horak FB 2012 Comparing the Mini-BESTest with the Berg Balance Scale to evaluate balance disorders in Parkinson’s disease. Parkinson’s Disease 2012: 375419.

Komijani P, Naghdi S, Nakhostin Ansari N, Bolhasani F, Nakhostin Ansari A 2018 Inter- and intra-rater reliability of brief BESTest in balance evaluation of patients with stroke: Brief report. Tehran University Medical Journal 76: 67–73.

Lampropoulou SI, Billis E, Gedikoglou IA, Michailidou C, Nowicky AV, Skrinou D, Michailidi F, Chandrinou D, Meligkoni M 2019 Reliability, validity and minimal detectable change of the Mini-BESTest in Greek participants with chronic stroke. Physiotherapy Theory and Practice 35(2): 171–182. 10.1080/09593985.2018.1441931

Latt MD, Lord SR, Morris JG, Fung VS 2009 Clinical and physiological assessments for elucidating falls risk in Parkinson’s disease. Movement Disorders 24(9): 1280–1289. 10.1002/mds.22561

Leddy AL, Crown BE, Earhart GM 2011a Functional gait assessment and balance evaluation system test: Reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall. Physical Therapy 91(1): 102–113. 10.2522/ptj.20100113

Leddy AL, Crown BE, Earhart GM 2011b Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. Journal of Neurologic Physical Therapy 35(2): 90–97. 10.1097/NPT.0b013e31821a620c

Liljequist D, Elfving B, Skavberg Roaldsen K 2019 Intraclass correlation – A discussion and demonstration of basic features. PLoS One 14(7): e0219854. 10.1371/journal.pone.0219854

Melton LJ, Leibson CL, Achenbach SJ, Bower JH, Maraganore DM, Oberg AL, Rocca WA 2006 Fracture risk after the diagnosis of Parkinson’s disease: Influence of concomitant dementia. Movement Disorders 21(9): 1361–1367. 10.1002/mds.20946

Naghdi S, Nakhostin Ansari N, Farhadi Y, Ebadi S, Entezary E, Glazer D 2016 Cross-cultural adaptation and validation of the injury-psychological readiness to return to sport scale to Persian language. Physiotherapy Theory and Practice 32(7): 528–535. 10.1080/09593985.2016.1221146

Naghdi S, Nakhostin Ansari N, Rustaie N, Akbari M, Ebadi S, Senobari M, Hasson S 2015 Simple shoulder test and oxford shoulder score: Persian translation and cross-cultural validation. Archives of Orthopaedic and Trauma Surgery 135(12): 1707–1718. 10.1007/s00402-015-2330-7

Padgett PK, Jacobs JV, Kasser SL 2012 Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. Physical Therapy 92: 1197–1207.

Paul SS, Sherrington C, Canning CG, Fung VS, Close JC, Lord SR 2014 The relative contribution of physical and cognitive fall risk factors in people with Parkinson’s disease: A large prospective cohort study. Neurehabilitation and Neural Repair 28(3): 282–290. 10.1177/1545968313508470

Van Den Eeden SK, Tanner CM, Bernstein AL, Fross RD, Leimpepert A, Bloch DA, Nelson LM 2003 Incidence of Parkinson’s disease: Variation by age, gender, and race/ethnicity. American Journal of Epidemiology 157: 1015–1022. 10.1093/aje/kwg068

Wood BH, Biclawh JA, Bowron A, Walker RW 2002 Incidence and prediction of falls in Parkinson’s disease: A prospective multidisciplinary study. Journal of Neurology, Neurosurgery, and Psychiatry 72(6): 721–725. 10.1136/jnnp.72.6.721