Psychometric properties of the Arabic version of the Mini-Balance Evaluation Systems Test in patients with neurological balance disorders

N.I. ALYOUSEF1, A.A.M. SHAHEEN1,2, W. ELSAYED3, A.M. ALSUBIHEEN1, A. FARRAG2

1Department of Health Rehabilitation Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
2Basic Science Department, Faculty of Physical Therapy, Cairo University, Cairo, Egypt
3Department of Physical Therapy, College of Applied Medical Sciences, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Abstract. – OBJECTIVE: This study aimed to translate and cross-culturally adapt the Mini-Balance Evaluation Systems Test (Mini-BESTest) to Arabic (Mini-BESTest-Ar) and evaluate its psychometric properties in patients with neurological balance disorders.

PATIENTS AND METHODS: The translation and adaptation followed the established guidelines. Validity, internal consistency, test-retest reliability, standard error of measurement (SEM), minimal detectable change (MDC95), and limits of agreement (LOA) were examined in 56 patients. The sensitivity was investigated using the receiver operating characteristic curve.

RESULTS: The Mini-BESTest-Ar significantly correlated with the Berg balance scale (BBS) (r = 0.80; p < 0.001) and dynamic gait index (DGI) (rho = 0.75; p < 0.001). All domains showed moderate to very good correlations with BBS (r = 0.62-0.81; p < 0.001) and fair to very good correlations with DGI (rho = 0.4 -0.79; p < 0.05). The internal consistency and test-retest reliability of the total score and all domains were excellent (Cronbach’s α = 0.96-0.81, ICC = 0.95-0.81, and r = 0.92-0.68). The SEM, MDC95, and MDC% for total score and domains were 1.19-0.31, 3.29-0.86 points, and 16.5%-66.8% respectively. The LOA revealed no systematic error. A cut-off point of 21.5/28 (Area under the curve = 0.85, sensitivity = 75%, specificity = 75%) was specified.

CONCLUSIONS: The Mini-BESTest-Ar has appropriate psychometric properties supporting its usefulness for research and clinical purposes.

Key Words:
Balance, Mini-BESTest, Psychometric properties, Neurological balance disorders, Assessment.

Introduction

Balance – an essential component of daily movement and activity – is the ability to maintain the center of gravity within the base of support under dynamic and static situations. It is required to preserve stability while shifting from one position to another in order to function independently in the community. In addition to being complex, it is influenced by vestibular, visual, somatosensory, biomechanical, nervous, and cognitive systems, which can be affected by various disorders.

Balance evaluation is a critical component of physical therapy assessment. Several balance assessment tools have been used such as the Timed Up and Go (TUG) test, one leg stance (OLS) test, Berg balance scale (BBS), activities-specific balance confidence scale (ABC), dynamic gait index (DGI), and functional reach test (FRT). The TUG, FRT, and OLS are single-task tests that cannot provide comprehensive information on which postural control subsystem is dysfunctional and have a limited role in directing treatment. Important aspects of dynamic balance control that reflect balance challenges during activities of daily living are missing in the BBS.

The Balance Evaluation Systems Test (BESTest) is a multitask balance assessment tool developed to detect balance deficits and identify the flawed underlying subsystem (biomechanical constraints, stability limits, postural responses, anticipatory postural adjustments, sensory orientation, dynamic balance during gait, and cognitive effects). However, this 36-item test is slightly extensive and limited in clinical settings. Hence, a short version – the Mini-Balance Evaluation Systems Test (Mini-BESTest) – was developed.

The Mini-BESTest provides well-defined and acknowledged clinical evaluation criteria for patients with balance deficits resulting from neurologic disorders. It has evident psychometric properties. Therefore, it was accepted as a balance...
assessment tool for both clinical and research purposes. The test was translated and cross-culturally adapted to Swedish, Brazilian Portuguese, Greek, Norwegian, Japanese, French, Spanish, Persian, German, and Turkish languages. Yet, no prescribed Arabic version of the test is available.

Most the Arabic-speaking physical therapists understand written English; nevertheless, the Arabic translation of the Mini-BESTest is necessary to maximize its validity and reliability and unify this assessment tool’s language. Therefore, this study aimed to translate and cross-culturally adapt the Mini-BESTest to Arabic and examine its validity, reliability, measurement error, and sensitivity in patients with various neurological disorders.

We hypothesized that the Arabic version of the Mini-BESTest (Mini-BESTest-Ar) would have appropriate cultural adaptation and excellent content validity. Using the hypothesis testing method, we predicted that the test and its four domains would be moderate-to-strongly correlated with the A-BBS (≥ 0.5) and have a fair-to-strong correlation with A-DGI (≥ 0.4). Additionally, the test would have excellent reliability, low measurement error, acceptable minimal detectable change (MDC), and proper limit of agreement (LOA) with no proportional error and floor or ceiling effect. Finally, we predicted that the Mini-BESTest-Ar would have a moderate accuracy with the area under the curve (AUC) of 0.70 or higher.

Patients and Methods

Study Design and Setting

In this cross-sectional and psychometric testing study, 56 patients were recruited from the Neurological Department, Sultan bin Abdul-Aziz Humanitarian City. The study was approved by the Ethics Review Boards of Sultan bin Abdul-Aziz Humanitarian City (No. 25/MSc/2020) and King Saud University Medical College (No. E-20-4835). All participants received oral and written information about the study and signed the informed consent form. The study was performed in accordance with the Helsinki Declaration of 1964 and its later revisions. The data were collected between September 2020 to July 2021.

Participants

The participants were included if they were at least 18 years of age with different neurological conditions to ensure the heterogeneity of balance deficits, could walk 6 meters with or without a cane, could follow a three-step command, and could live independently in the community. In addition, they must be sober from any medications affecting balance in the preceding 48 hours. The exclusion criteria were non-ambulatory patients and medical conditions that could interfere with or affect the balance evaluation procedures (e.g., history of pathologic vertigo, fainting, and musculoskeletal problems that affect the balance) or that could affect the results of the evaluation for reasons other than impaired balance (e.g., pregnancy, recent surgery of the lower limbs).

The sample size was calculated according to the classic rule established by Kline of using 2 to 20 subjects for each item. The authors used 4 subjects for each item (4:1) with a total sample of 56 participants. For test-retest reliability, a subsample of 29 participants was calculated as proposed by Bonett in 2002.

Translation and Cross-Cultural Adaptation

The original version of Mini-BESTest was translated and cross-culturally adapted based on the 6 steps set by Beaton et al; two independent translators translated the English Mini-BESTest version into Arabic. The two forward translations were compared, and single-consensus Arabic Mini-BESTest was then constructed and any discrepancies were resolved with the translators. The back-translations were done by two separate native English speakers. Five experts (methodologist, health professionals (neurologist and physiotherapist), language professional, and translators) reviewed all reports and approved the pre-final version of Mini-BESTest-Ar. The main developer of the English version of the Mini-BESTest was a consultant.

The pre-final version was piloted on 10 physiotherapists for face and content validity. They gave their general impression of the clarity of the items, the relevance of the content, and the comprehensiveness of the instructions. They could make suggestions if necessary. Furthermore, the expert committee assessed the content validity using the content validity index (CVI). Finally, a debriefing summary — including all participant interviews — and the final decisions grid were sent to the developer for comments. Based on the results of the pre-final version of Mini-BESTest-Ar, the equivalence was reached with the original version in 4 areas: semantic, idiomatic, experiential, and conceptual equivalence. The expert committee made minor changes and accomplished the final version of Mini-BESTest-Ar.
Psychometric properties of Mini-BEST test - Ar in patients with neurological balance disorders

Equipment
The following equipment was used for the application of the Mini-BESTest: a foamy material [Airex® foam (Sins, Switzerland), ~ 5 cm thick, medium density], a chair without armrests or wheels, a step of average height, a 10-degree incline ramp (at least 60 cm x 100 cm) to stand on, stopwatch, a box (23 cm height), and a 3-meter distance measured out and marked on the floor with tape (from the chair). For the BBS application, a chair with armrests and a chair without armrests, a timer, a step of average height, and a ruler of 5, 12 and 25 cm were used.

Procedures
In the first session, the demographic and clinical information were collected by interviewing the participants. Then, the clinical balance tests were conducted randomly to ensure that fatigue did not affect the test results. Every other participant started with the A-BBS, the Mini-BESTest-Ar, and the A-DGI and then vice versa. After the completion of each scale assessment, there was a 10-min break before the commencement of the next one. The total duration of the session was 60-90 minutes.

At the second session, after 7-10 days, the participants performed the Mini-BESTest-Ar once again and were asked about any relevant changes in the self-perceived balance state since the first test session using the Global Rating of Change (GRC). The second session took approximately 40 min to administer. The two sessions were conducted in the same room, time of the day, and environment. The participants were instructed to take their medication according to their normal regime during the test period and wear comfortable clothing and flat shoes. All participants used the same shoes at both sessions.

Outcome Measures

The Mini-BESTest
The test is used to assess the risk of imbalance and falling. It comprises 14-items focusing on dynamic balance that are divided into 4 sub-scales: anticipatory postural adjustments, reactive postural control, sensory orientation, and dynamic gait. All these sub-scales of balance control are significant in reflecting balance difficulties in daily life activities in patients with neurological disorders. Some of the tasks include sit to stand, standing on toes, single-leg stance, compensatory reactions forwards, backwards and sideways, tasks with the eyes closed, on a foam surface, on an inclined surface, and tasks involving gait with a change of speed, head rotations, walking over obstacles and timing. Each task was rated on a 3-level ordinal scale ranging from 0 (severe balance impairment) to 2 (no balance impairment) with 28 as the maximal test score. The higher scores indicate better balance. The test takes 10 to 15 minutes to administer.

The BBS
The BBS is the quantitative and widely used assessment tool for evaluating balance and risk of falling in patients with neurological conditions and the elderly. The scale contains 14 items to test either static or dynamic balance. Sitting to standing, standing unsupported, sitting unsupported, standing to sitting, transfers, standing with eyes closed, standing with feet together, reaching forward with an outstretched arm, retrieving an object from the floor, turning to look behind, turning 360 degrees, placing an alternate foot on a stool, standing with one foot in front, and standing on one foot are among the functions that are evaluated. The scoring of each item is from 0 to 4 with a global score of 56 points, and higher scores reflect better balance. The total score is interpreted as follows: 0-20, high fall risk and balance disorder; 21-40, medium fall risk and the existence of an acceptable balance; and 41-56, low fall risk and presence of good balance. It requires 10-20 minutes for administration. The Arabic version (A-BBS) is valid and reliable.

The DGI
The DGI is a common gait assessment tool utilized to assess the functional stability and risk of falling in patients with different conditions. The 8-items DGI assesses walking on flat surfaces with different velocity, horizontal, vertical and pivot turn, stepping over and around obstacles, and navigation of stairs. The DGI is a 4-point ordinal scale from 0 to 3 with a total score of 24. It takes 15 minutes to be completed. The Arabic version of DGI (A-DGI) showed good psychometric properties.

The GRC
The GRC is a scale used to assess patients’ overall perceptions of improvement or deterioration over time. At the time of the second session, the participants and the physical therapist completed the GRC. The scale composed of 15 points ranging from -7 to 0 to +7, where -7 represents...
Statistical Analysis

All the analyses were conducted using the Statistical Package of Social Science (SPSS Inc., Chicago, IL, USA) version 18.0. The level of significance was set at $p < 0.05$. Descriptive statistics were calculated for the demographic variables. The data were investigated for normal distribution using Kolmogorov-Smirnov test.

Content validity was calculated using the CVI and modified kappa coefficient ($K^*$) according to Polit et al.$^{21}$ Based on the normality of data, Pearson’s correlation coefficient ($r$) and/or Spearman’s ranked correlation ($\rho$) were calculated to test the construct (criterion) validity by correlating the Mini-BESTest-Ar with A-BBS and A-DGI. Cronbach’s alpha ($\alpha$) was calculated to test the internal consistency of Mini-BESTest-Ar.

A value higher than 0.70 indicated the internal consistency of the scale. Test-retest reliability was conducted using a 2-way analysis of variance random-effect, absolute agreement intraclass-correlation ($ICC_{2,1}$) with 95% confidence intervals (CIs). A correlation between 0 and 0.25 was interpreted as little or no relationship, between 0.25 and 0.5 as a fair relationship, between 0.5 0.75 as moderate, and above 0.75 as a very good to excellent relationship.$^{29}$ The limit of agreement (LOA) between the Mini-BESTest-Ar scores at baseline and the subsequent administration was visually examined using a Bland-Altman plot.$^{30}$

The standard error of the mean (SEM) was calculated using the formula: $SEM = \frac{SD}{\sqrt{N}}$. The $MDC_{95}$ was estimated using the formula $MDC_{95} = SEM \times 1.96 \times \sqrt{2}$. Furthermore, MDC% was calculated: $MDC\% = \frac{MDC_{95}}{mean} \times 100$, where “mean” is the mean of the scores acquired in the second testing session. An MDC% under 30% was acceptable.$^{31}$

By creating a ROC curve from the change scores between the 7-10 days follow-up and the baseline, the sensitivity to change of the Mini-BESTest-Ar was investigated. The ability of Mini-BESTest-Ar to distinguish patients who improved from those who stayed stable based on the GRC was measured using the AUC. The sensitivity (patients who improved, true positive) and specificity (patients who stayed stable, true negative) values were determined.$^{32}$ The AUC ranges between 0.5 (no diagnostic accuracy beyond chance) to 1.0 (perfect diagnostic accuracy). The cut-off values were calculated by maximizing sensitivity and specificity by selecting the smallest value of $(1- sensitivity)^2 + (1- specificity)^2$.

Floor and ceiling effects were present if more than 20% of the participants achieved the lowest or respectively the highest possible score.$^{5,32}$

Results

Participants

As shown in Table I, the mean age of the 56 participants was 36.11 ± 13.11 years and approximately 67.9% were men. The mean ± SD of BMI was 26.6 ± 5.6 kg/cm$^2$. Most participants (35.7%) were diagnosed with stroke (7: subacute, 13: chronic) followed by traumatic brain injury (28.6%, 15: chronic, 1: subacute), and spinal cord injury (14.3%, chronic paraplegia). All the data are normally distributed except the A-DGI score ($p > 0.05$).

Translation and Cross-Cultural Adaptation

Most semantic differences between the two Arabic versions produced during step 1 were caused by the use of synonyms or phrases with comparable meanings. The test’s units were all changed to the worldwide metric system (feet to meters, inches to centimeters). The final version was deemed to be unambiguous, and it was thought to be appropriate and applicable for Arabic speakers. The main developer revised and approved the back-translated English version and commented to change Temper foam to Airex foam.

Content Validity

The results of content validity elaborated excellent I-CVI for all items (I-CVI = 1). The S-CVI was excellent, with values of 1 for both (S-CVI/Ave) and (S-CVI-UA). The value of the probability of a chance ($P_c$) was 0.031 for all items and $K^*$ for each item was excellent (> 0.74) (Table II). In addition, the instructions and scoring descriptions of the Mini-BESTest-Ar are comparable to those of the original version.

Construct Validity

There were strong statistically significant positive correlations between the total score of Mini-BESTest-Ar and both A-BBS ($r = 0.80; p <$
Psychometric properties of Mini-BESTest-Ar in patients with neurological balance disorders

The four domains of Mini-BESTest-Ar showed moderate to very good correlations with A-BBS (r ranged from 0.62 to 0.81; p < 0.001). Fair to very good correlations were reported between the domains and A-DGI (rho = 0.4; p < 0.05 for sensory orientation domain and 0.79; p < 0.001 for dynamic gait domain) (Table III).

Reliability

The Mini-BESTest-Ar total score (Cronbach’s α = 0.96) and all domains represented excellent internal consistency (α ranged from 0.81 for Reactive postural control domain to 0.94 for Anticipatory domain). With 29 stable participants, based on GRC, test-retest reliability was excellent for the total score Mini-BESTest-Ar (ICC\(_{2,1}\) = 0.95). The highest ICC was for the Anticipatory domain (ICC\(_{2,1}\) = 0.94 and r = 0.90, p = 0.00) (Table IV).

The SEM for the total Mini-BESTest-Ar was 1.19 with an acceptable MDC\(_{95}\) of 3.29 points and MDC% of 16.45%. The SEM for the domains varied from 0.31 (sensory orientation) to 0.88 (Reactive postural control). The MDC\(_{95}\) ranged from 0.86 (sensory orientation) to 2.43 (Reactive postural control). As shown in Table V, the MDC% was acceptable for all domains except anticipatory and reactive postural control.

Figure 1 shows that the mean difference of the Mini-BESTest-Ar total score between the sessions was -0.17 ± 2.4 (-2.4 to 2.0). The distribution of differences was within the limits of agreement and did not differ significantly from zero (r = -0.817, p = 0.421). The line of equality (zero) was contained in the 95% CI (-0.60 to 0.26) of the mean

### Table I. Participants’ characteristics (n = 56).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean ± SD (min-max)</th>
<th>Median (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>36.11 ± 13.11</td>
<td>38</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>167.9 ± 10.4</td>
<td>167.0</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>75.5 ± 15.9</td>
<td>75.0</td>
</tr>
<tr>
<td>BMI (kg/cm(^2))</td>
<td>26.6 ± 5.6</td>
<td>26.5</td>
</tr>
<tr>
<td>Sex</td>
<td>Men</td>
<td>38</td>
</tr>
<tr>
<td>Residence</td>
<td>Central region</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Southern region</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Northern region</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Eastern region</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Western region</td>
<td>8</td>
</tr>
<tr>
<td>Medical condition</td>
<td>Stroke</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>SCI</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>PD</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>GBS</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Spinal Bifida</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Meningitis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Lateral Sclerosis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cranioza</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CMT</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>20 ± 5.4 (6-28)</td>
</tr>
<tr>
<td></td>
<td>Mini-BESTest1-Ar (n = 56)</td>
<td>20 ± 5.4 (6-28)</td>
</tr>
<tr>
<td></td>
<td>Mini-BESTest2-Ar (n = 29)</td>
<td>20 ± 5.8 (8-27)</td>
</tr>
<tr>
<td></td>
<td>A-BBS (n = 56)</td>
<td>47.5 ± 7.3 (13-56)</td>
</tr>
<tr>
<td></td>
<td>A-DGI (n = 56)</td>
<td>22 (3-24)</td>
</tr>
</tbody>
</table>

All data represented as mean and standard deviation SD, or frequency and percentage %, except for DGI represented as median and range. min-max: Minimum-maximum, n number of participants, BMI: body mass index, TBI: traumatic brain injury, SCI: spinal cord injury, MS: multiple sclerosis, PD: Parkinson’s disease, GBS: Guillain Barre Syndrome, CP: Cerebral palsy, CMT: Charcot-Marie-Tooth, Mini-BESTest1-Ar: total score of Arabic Mini-BESTest at baseline, Mini-BESTest 2-Ar: total score of Arabic Mini-BESTest after 7-10 days, A-BBS: total score of Arabic Berg Balance scale, A-DGI: total score of Arabic Dynamic Gait Index.
difference, showing that there was no systematic error or proportional bias between sessions.

**Sensitivity of the Mini-BESTest-Ar**

An AUC value of 0.85 (standard error 0.05; 95% CI, 0.74-0.96) was obtained after constructing the ROC curve (Figure 2). This value demonstrated moderate accuracy and was significant at 0.05 alpha level. The cut-off point value was 21.5 with a sensitivity of 75% and specificity of 75%.

**Floor/Ceiling Effect**

There was no floor/ceiling effect for the total score of Mini-BESTest-Ar. The four domains did not reach the ceiling or floor limits.
Discussions

Discussion

This study aimed to translate and cross-culturally adapt the Mini-BESTest into Arabic and examine its psychometric properties in patients with various neurological balance disorders. In agreement with previous studies, the Mini-BESTest-Ar is comprehensible, valid, reliable, and accurate with low measurement error with no floor or ceiling effect. These results support our prior hypotheses.

Measuring a scale’s content validity is crucial for enhancing the construct validity of an instrument and indicates how appropriate the total scale is. It was evaluated with a standard scale, which...

Table IV. Internal consistency, test-retest reliability, and floor and ceiling effects for the Mini-BESTest-Ar (total and domains).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Internal Consistency (α)</th>
<th>Test re-test (ICC)</th>
<th>95% CI</th>
<th>Test re-test (r)</th>
<th>Floor No. (%)</th>
<th>Ceiling No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anticipatory domain</td>
<td>0.94</td>
<td>0.94</td>
<td>0.86-0.97</td>
<td>0.90**</td>
<td>1 (1.8)</td>
<td>4 (7.1)</td>
</tr>
<tr>
<td>Reactive postural control domain</td>
<td>0.81</td>
<td>0.81</td>
<td>0.60-0.91</td>
<td>0.68**</td>
<td>8 (14.3)</td>
<td>12 (21.4)</td>
</tr>
<tr>
<td>Sensory orientation domain</td>
<td>0.93</td>
<td>0.93</td>
<td>0.85-0.97</td>
<td>0.87**</td>
<td>1 (1.8)</td>
<td>32 (57.1)</td>
</tr>
<tr>
<td>Dynamic gait domain</td>
<td>0.93</td>
<td>0.93</td>
<td>0.84-0.97</td>
<td>0.88**</td>
<td>1 (1.8)</td>
<td>9 (16.1)</td>
</tr>
<tr>
<td>Mini-BESTest -Ar Total</td>
<td>0.96</td>
<td>0.95</td>
<td>0.88-0.98</td>
<td>0.92**</td>
<td>1 (1.8)</td>
<td>1 (1.8)</td>
</tr>
</tbody>
</table>

N: number of stable participants included in test-retest reliability, α: Cronbach alpha, ICC: Inter class correlation; CI: 95% confidence interval; r: Pearson’s correlation. *p-value is significant at < 0.05; **p-value is significant at < 0.001

Table V. Standard error and Minimal detectable change for Mini-BESTest -Ar.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SEM</th>
<th>MDC 95</th>
<th>MDC%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anticipatory</td>
<td>0.38</td>
<td>1.05</td>
<td>30.6%</td>
</tr>
<tr>
<td>Reactive postural control</td>
<td>0.88</td>
<td>2.43</td>
<td>66.8%</td>
</tr>
<tr>
<td>Sensory orientation</td>
<td>0.31</td>
<td>0.86</td>
<td>16.6%</td>
</tr>
<tr>
<td>Dynamic gait</td>
<td>0.49</td>
<td>1.35</td>
<td>17.3%</td>
</tr>
<tr>
<td>Mini-BESTest-Ar Total</td>
<td>1.19</td>
<td>3.29</td>
<td>16.5%</td>
</tr>
</tbody>
</table>

SEM: standard error of measurement, MDC: Minimal detectable change, MDC%: Minimal detectable change percentage.

not have a floor effect. On the other hand, the ceiling effect was found in two domains [Reactive postural control (21.4%) and Sensory orientation (57.1%)] (Table IV).

Figure 1. Graphical representation of discrepancies between test 1 and test 2 using the Bland-Altman technique. The mean difference is represented by the solid gray line in the middle. The black strong solid lines reflect the highest and lowest limits of agreement. The 95% CI of the mean difference between the two testing sessions is represented by the upper and lower dashed lines. The line above 0 (the line of equality) is represented by the thin solid black line in the middle.

4343
was previously developed in the same field. The results revealed excellent content validity of Mini-BESTest-Ar. The value of S-CVI was higher than that of the Spanish version (S-CVI = 0.98)\textsuperscript{14}. In accordance with our prior hypothesis, the correlation between the Mini-BESTest-Ar and A-BBS was strong (r = 0.80). Earlier studies\textsuperscript{5,6,10,12,16,23,32,33} have reported moderate to high correlations between the Mini-BESTest and the BBS. Additionally, in terms of the test domains, moderate to strong correlations with A-BBS were observed (r ranged from 0.63 (Anticipatory and Reactive Postural Control) to 0.81 (Sensory orientation domain)).

Meanwhile, fair to strong correlations between the total score, domains’ scores of the Mini-BESTest-Ar, and A-DGI were confirmed. The fair correlation was observed between the Sensory orientation domain and A-DGI (rho = 0.40) because the DGI does not include items to test the static balance, such as standing with eyes open or closed, to evaluate the risk of falling\textsuperscript{34}. Only one study\textsuperscript{34} tested the correlation between the Mini-BESTest and DGI in elders with r = 0.83, 0.84, and 1 for patients with paresis, parkinsonism, and vertigo respectively, which are much higher than ours. The difference between our results and the previous studies may be attributed to the recruitment of a heterogeneous sample (patients with different neurological balance disorders) to ensure the test’s applicability in a wide range of abnormalities\textsuperscript{1,12,13} in addition to the variety of the studies’ techniques.

The internal consistency of Mini-BESTest-Ar was excellent. The Cronbach’s α for total score was 0.96 which confirms the homogeneity of the test. This finding correlates with the previous studies\textsuperscript{3} with different samples where α ranging from 0.89 to 0.96. Moreover, the value of Cronbach’s α was higher than that of the Greek (α = 0.88), French (α = 0.895- 0.929), German (α = 0.90), Spanish (α = 0.85) versions\textsuperscript{12,14,16}. The four domains of the Mini-BESTest demonstrated very good internal consistency, which is similar to previous studies\textsuperscript{35,36} with α ranged from 0.77 to 0.94.

Regarding the test-retest reliability, the Mini-BESTest-Ar showed excellent reliability with ICC\textsubscript{2,1} = 0.95, 95% CI = 0.88-0.98 for the total score, which is higher than or nearly equal to the previous studies\textsuperscript{12,37,38}. Reproducibility of the test domains had ICC values between 0.81 and 0.94, which are higher than those reported in Norwegian study (0.53-0.87)\textsuperscript{10}, lower than ICCs reported by Winairuk et al\textsuperscript{32}, Göktaş et al\textsuperscript{10}, and Naghdí et al\textsuperscript{32}, (0.95-0.99), and nearly equal to ICCs estimated by Potter et al\textsuperscript{37} (0.85-0.92). The differences in the results may be explained by the differences in the protocols of the studies: two or more raters independently rated each patient’s performance\textsuperscript{15} while others video-recorded the performance\textsuperscript{13}.

The absolute reliability, presented by SEM and MDC\textsubscript{95} in actual scale units, is possibly the most important reliability measure for clinical purposes\textsuperscript{1}. In this study, the values of SEM ranged from 0.31 (for Sensory orientation domain) to 1.19 (for total score). The SEMs were similar to or lower than those previously reported with values ranging from 0.86 to 1.99\textsuperscript{1,11,13,37,39}. MDC\textsubscript{95} for the Mini-BESTest-Ar total score falls in the range (MDC\textsubscript{95} = 2.0-5.2)\textsuperscript{1,11,13,37,39}. For the tests’ domains, our results were similar to Potter et al\textsuperscript{37}, higher than Dominguez-Olivan et al\textsuperscript{36}, and slightly lower than Beauchamp et al\textsuperscript{39}. The small SEM of the total score of the Mini-BESTest-Ar suggests that it is a precise test with low measurement error\textsuperscript{37}. The MDC\textsubscript{95} reported in this study ranged from 0.8 to 3.29 points, which represents the minimum difference that would reflect a real change in the Mini-BESTest total and domains scores. MDC\textsubscript{95} values fall within the range reported by previous studies\textsuperscript{5,33}. One of the essential features of a clinical tool is its ability to determine actual changes in the patients’ condition. The MDC established in this study would be useful for future clinical trials in determining whether the experimental intervention has caused any real change in balance.

![Figure 2. Receiver operating characteristic curve for Mini-BESTest-Ar.](image-url)
Psychometric properties of Mini-BESTest-Ar in patients with neurological balance disorders

ability. The MDC values provide clinicians with useful parameters for interpreting changes in patients' functions and setting goals. The MDC% (16.45%) was acceptable and higher than the reported MDC% in patients with COPD (14.9%) and with spinal cord injury (13.7%)31,40. The homoscedasticity of variations between scoring seen on the Bland-Altman plot proves the absence of systematic errors and supports the use of an absolute versus relative MDC. Meanwhile, the Bland-Altman plot was performed since that high correlation does not necessarily imply an agreement16. Our results are compatible with previous studies13,16,31.

Despite the limitations of our procedure, these results confirm the high reliability and precision of the Mini-BESTest-Ar for evaluating patients with balance deficits undergoing rehabilitation and support its use among Arabic-speaking rehabilitation professionals and researchers. The sensitivity to change of the Mini-BESTest-Ar was indicated using the AUC. Its value falls well within the range reported with other samples (AUC ranged from 0.64 to 0.91, sensitivity ranged from 62% to 93%, and specificity ranged from 64% to 81%)31,32,39,41. Our results demonstrate the ability of the Mini-BESTest-Ar to distinguish patients who improved from those who stayed stable specified at a cut-off score of 21.5/28. Clinically speaking, this score gives valuable information considering that below 21, people present a higher balance deficit. Thus, healthcare professionals can use the Mini-BESTest-Ar cut-off score of 21.5 to identify individuals with and without balance deficits and adjust their interventions accordingly.

Additionally, we found no floor effect for the total score and domains' scores which is consistent with previous studies11,12,14,16. The Mini-BESTest, on the other hand, demonstrated a floor effect on patients with subacute stroke, and authors argued that the Mini-BESTest contains some items that are more difficult for those patients32. Our results are in line with those of Potter et al17, 2019 as they revealed ceiling effects in two domains: reactive postural control (37.5%) and sensory orientation (53.1%) in patients with MS37. They suggested that the clinicians should apply the whole test on the patients because few patients did well in some domains but not in the entire test17.

Limitations
Several limitations of this study need to be acknowledged. The sample size was relatively small, even though it was higher than previous studies1,3,4,6,15,16,31,35,37,42. Due to the relatively small sample size, further studies should focus on other psychometric properties (confirmatory factor analysis, item-response theory or Rasch analysis) of the Mini-BESTest-Ar in a larger sample. The participants were ambulatory, which may have contributed to the lack of a floor effect; additionally, the results could not be generalized to patients with cognitive and severe neurological disabilities.

Conclusions
In conclusion, the Mini-BESTest-Ar is comprehensible and clear. The instructions and scoring descriptions of the Mini-BESTest-Ar are comparable to those of the original version. The scale was translated into the classical Arabic language to enhance its use in all Arabic-speaking countries. In accordance with previous studies, the Mini-BESTest-Ar demonstrates strong validity and excellent relative and absolute reliability. It is a sensitive and accurate tool to provide information on which particular balance systems were the underlying cause of balance impairments in the individuals with different neurological balance disorders. Therefore, it can be recommended for use in clinical practice and research.

Conflict of Interest
The authors declare that there is no conflict of interest.

Acknowledgments
The authors thank all participants for being a part of this study.

Authors' Contributions
A.A.M.S and N.I.A suggested the research idea, helped in data collection, analyzed, and interpreted the data. W.E, A.M.A, and AF revised the data analysis and participated in data interpretation. AS and NA were the major contributors in the writing process. All authors critically read and approved the final version of the manuscript.

ORCID ID
Afaf A.M. Shaheen: 0000-0002-9265-0806

Availability of Data and Materials
The data collected and analyzed in the present study are available from the corresponding author upon request.
Informed Consent
Informed consent was obtained from all participants included in the study.

Ethics Approval
The study was approved by the Ethics Review Boards of Sultan bin Abdul-Aziz Humanitarian City (No. 25/MSc/2020) and King Saud University Medical College (No. E-20-4835).

References


22) Yusoff MSB. ABC of content validation and content validity index calculation. EIMJ 2019; 11: 49-54.


Psychometric properties of Mini-BESTest -Ar in patients with neurological balance disorders


34) Miqdad, Pawar SK. Comparison of mini-bestest and dynamic gait index for prediction of fall susceptibility in old individuals. IJRMS 2017; 5: 3018-3022.


