European Review for Medical and Pharmacological Sciences 2023; 27: 4337-4347

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

N.I. ALYOUSEF[1], A.A.M. SHAHEEN[1,2], W. ELSAYED[3], A.M. ALSUBIHEEN[1], A. FARRAG[2]

[1]Department of Health Rehabilitation Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
[2]Basic Science Department, Faculty of Physical Therapy, Cairo University, Cairo, Egypt
[3]Department 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 (MDC_{95}), 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, MDC_{95}, 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[1]. It is required to preserve stability while shifting from one position to another in order to function independently in the community[2]. In addition to being complex, it is influenced by vestibular, visual, somatosensory, biomechanical, nervous, and cognitive systems, which can be affected by various disorders[3].

Balance evaluation is a critical component of physical therapy assessment[4]. 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)[5,6]. 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[5]. Important aspects of dynamic balance control that reflect balance challenges during activities of daily living are missing in the BBS[7].

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)[7,8]. However, this 36-item test is slightly extensive and limited in clinical settings[4]. Hence, a short version – the Mini-Balance Evaluation Systems Test (Mini-BESTest) – was developed[9].

The Mini-BESTest provides well-defined and acknowledged clinical evaluation criteria for patients with balance deficits resulting from neurologic disorders[10]. It has evident psychometric properties[11]. Therefore, it was accepted as a balance

Corresponding Author: Afaf A.M. Shaheen, MD; e-mail: ashaheen@ksu.edu.sa

assessment tool for both clinical and research purposes[9,12]. The test was translated and cross-culturally adapted to Swedish[4], Brazilian Portuguese[3], Greek[12], Norwegian[1], Japanese6, French[13], Spanish[14], Persian[15], German[16], and Turkish[10] 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[17] 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[18]. For test-retest reliability, a subsample of 29 participants was calculated as proposed by Bonett in 2002[19].

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[20]; 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)[21,22]. 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.

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[23].

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[7]. 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[13].

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[24]. The scoring of each item is from 0 to 4 with a global score of 56 points, and higher scores reflect better balance[24]. 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[23]. The Arabic version (A-BBS) is valid and reliable[25].

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[26]. 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[27].

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 a very great deterioration; 0 represents about the same; and +7 represents a very great improvement. Participants with an average rating > +3 were considered improved; those with an average rating between +3 and -3 were considered stable; and those with average ratings of < -3 were categorized as experiencing a deterioration in their clinical status[28].

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 (Cronbach's α) 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 = (SD×[√(1-ICC)]), where SD was the sample standard deviation. The MDC_{95} was estimated using the formula MDC_{95} = SEM ×1.96 ×√2[5]. Furthermore, MDC% was calculated: MDC% = (MDC_{95} / mean) x 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 $(1\text{- sensitivity})^2 + (1\text{- specificity})^2$. 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 <

Table I. Participants' characteristics (n = 56).

Variables	Mean	SD	
Age	36.11	13.11	
Height (cm)	167.9	10.4	
Weight (kg)	75.5	15.9	
BMI (kg/cm²)	26.6	5.6	
		Frequency (n)	**%**
Sex	Men	38	67.9
	Women	18	32.1
Residence	Central region	21	37.5
	Southern region	16	28.6
	Northern region	3	5.4
	Eastern region	8	14.3
	Western region	8	14.3
Medical condition			
	Stroke	20	37.5
	TBI	16	8.6
	SCI	8	14.3
	MS	4	7.1
	PD	1	1.8
	GBS	1	1.8
	Spinal Bifida	1	1.8
	Meningitis	1	1.8
	Lateral Sclerosis	1	1.8
	Cranioma	1	1.8
	CP	1	1.8
	CMT	1	1.8
	Mean ± SD (min-max)		
Mini-BESTest1-Ar (n = 56)	20 ± 5.4 (6-28)		
Mini-BESTest2-Ar (n = 29)	20 ± 5.8 (8-27)		
A-BBS (n = 56)	47.5 ± 7.3 (13-56)		
	Median (range)		
A-DGI (n = 56)	22 (3-24)		

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.

0.001) and A-DGI (rho = 0.75; $p < 0.001$) using Pearson's and Spearman's correlation respectively. 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 $\alpha = 0.96$) 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 (t = -.817, $p = 0.421$). The line of equality (zero) was contained in the 95% CI (-.60 to 0.26) of the mean

Table II. Evaluation of I-CVIs with expert's agreement, scales' items.

Items	Experts' rating of relevance									
	E1	E2	E3	E4	E5	Experts in agreement	I-CVI	P_c	K*	Evaluation
1.	4	4	4	4	4	5	1	.031	1	Excellent
2.	4	4	4	4	4	5	1	.031	1	Excellent
3.	4	4	4	4	4	5	1	.031	1	Excellent
4.	4	4	4	4	4	5	1	.031	1	Excellent
5.	4	4	4	4	4	5	1	.031	1	Excellent
6.	3	4	4	4	4	5	1	.031	1	Excellent
7.	4	4	4	4	4	5	1	.031	1	Excellent
8.	4	4	4	4	4	5	1	.031	1	Excellent
9.	4	4	4	4	4	5	1	.031	1	Excellent
10.	3	4	4	4	4	5	1	.031	1	Excellent
11.	4	4	4	4	4	5	1	.031	1	Excellent
12.	4	4	4	4	4	5	1	.031	1	Excellent
13.	4	4	4	4	4	5	1	.031	1	Excellent
14.	4	4	4	4	4	5	1	.031	1	Excellent
Proportion relevance	1	1	1	1	1	S-CVI/Ave = 1				

I-CVI = item-level content validity index; P_c = probability of a chance computed by formula: P_c = [N/A(N - A)]*0.5N where N = number of experts and A = Number agreeing on good relevance; K* = modified kappa coefficient: k* = (I-CVI - P_c)/(1 - P_c), about the criteria for K*, 0.40 to 0.59 = fair, 0.60 to 0.74 = good, and > 0.74 excellent.

Table III. Construct Validity.

Mini-BESTest-Ar domains	Instrument for correlation	Correlation coefficient (r, rho)	Hpotheses confirmed?
Anticipatory domain	**A-BBS**	Expected: ≥ 0.5 Actual: r = 0.63**	Yes
Reactive postural control domain		Expected: ≥ 0.5 Actual: r = 0.63**	Yes
Sensory orientation domain		Expected: ≥ 0.5 Actual: r = 0.81**	Yes
Dynamic gait domain		Expected: ≥ 0.5 Actual: r = 0.62**	Yes
Mini-BESTest-Ar Total		Expected: ≥ 0.5 Actual r = 0.80**	Yes
Anticipatory domain	**A-DGI**	Expected: ≥ 0.4 Actual: rho = 0.53**	Yes
Reactive postural control domain		Expected: ≥ 0.4 Actual: rho = 0.62**	Yes
Sensory orientation domain		Expected: ≥ 0.4 Actual: rho=0.40*	Yes
Dynamic gait domain		Expected: ≥ 0.4 Actual: rho = 0.79**	Yes
Mini-BESTest-Ar Total		Expected: ≥ 0.4 Actual : rho = 0.75**	Yes

The correlation between Mini-BESTest -Ar' domains and total score with A-BBS and A-DGI, r: Pearson's correlation, rho: Spearman's correlation, * *p*-value is significant at < 0.05; ** *p*-value is significant at < 0.001.

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

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

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

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.

Variable	SEM	MDC_{95}	MDC%
Anticipatory	0.38	1.05	30.6%
Reactive postural control	0.88	2.43	66.8%
Sensory orientation	0.31	0.86	16.6%
Dynamic gait	0.49	1.35	17.3%
Mini-BESTest-Ar Total	1.19	3.29	16.5%

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).

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[12-14,16], 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[22]. It was evaluated with a standard scale, which

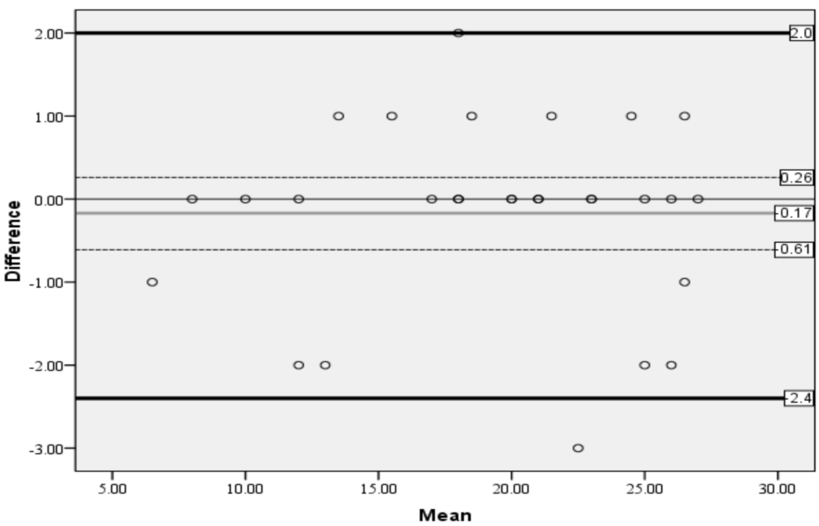

Figure 1. Graphical representation of discrepancies between test 1 and test 2 using the Bland-Altman technique. The mean difference is represented by the solid gray line in the middle. The black strong solid lines reflect the highest and lowest limits of agreement. The 95% CI of the mean difference between 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.

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)[14].

In accordance with our prior hypothesis, the correlation between the Mini-BESTest-Ar and A-BBS was strong (r = 0.80). Earlier studies[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[34]. Only one study[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[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[11] 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[12-14,16]. The four domains of the Mini-BESTest demonstrated very good internal consistency, which is similar to previous studies[35,36] with α ranged from 0.77 to 0.94.

Regarding the test-retest reliability, the Mini-BESTest-Ar showed excellent reliability with $ICC_{2,1}$ = 0.95, 95% CI = 0.88-0.98 for the total score, which is higher than or nearly equal to the previous studies[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)[1], lower than ICCs reported by Winairuk et al[32], Göktaş et al[10], and Naghdi et al[15], (0.95-0.99), and nearly equal to ICCs estimated by Potter et al[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[1,15] while others video-recorded the performance[13].

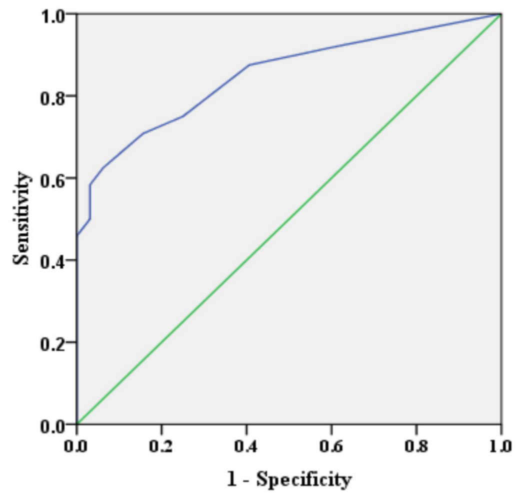

Figure 2. Receiver operating characteristic curve for Mini-BESTest-Ar.

The absolute reliability, presented by SEM and MDC_{95} in actual scale units, is possibly the most important reliability measure for clinical purposes[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 1.9[1,11,13,37,39].

MDC_{95} for the Mini-BESTest-Ar total score falls in the range (MDC_{95} = 2.0-5.2)[1,11,13,37,39]. For the tests' domains, our results were similar to Potter et al[37], higher than Dominguez-Olivan et al[36], and slightly lower than Beauchamp et al[39]. The small SEM of the total score of the Mini-BESTest-Ar suggests that it is a precise test with low measurement error[37].

The MDC_{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_{95} values fall within the range reported by previous studies[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 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_{95}[13]. Meanwhile, the Bland-Altman plot was performed since that high correlation does not necessarily imply an agreement[16]. Our results are compatible with previous studies[13,16,31].

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%)[11,32,39,41]. 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 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 studies[1,12-14,16]. 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 patients[32].

Our results are in line with those of Potter et al[37], 2019 as they revealed ceiling effects in two domains: reactive postural control (37.5%) and sensory orientation (53.1%) in patients with MS[37]. 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 test[37].

Limitations

Several limitations of this study need to be acknowledged. The sample size was relatively small, even though it was higher than previous studies[1,3,4,6,15,16,31,35,37,42]. 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

1) Hamre C, Botolfsen P, Tangen GG, Helbostad JL. Interrater and test-retest reliability and validity of the Norwegian version of the BESTest and mini-BESTest in people with increased risk of falling. BMC Geriatr 2017; 17: 1-8.
2) Büyükavcı R ŞF, Sağ S, Doğu B, Kuran B. The impact of additional trunk balance exercises on balance, functional condition and ambulation in earlystroke patients: Randomized controlled trial. Turk J Phys Med Rehab 2016; 62: 248-256.
3) Maia AC, Rodrigues-de-Paula F, Magalhaes LC, Teixeira RL. Cross-cultural adaptation and analysis of the psychometric properties of the Balance Evaluation Systems Test and MiniBESTest in the elderly and individuals with Parkinson's disease: application of the Rasch model. Braz J Phys Ther 2013; 17: 195-217.
4) Bergstrom M, Lenholm E, Franzen E. Translation and validation of the Swedish version of the mini-BESTest in subjects with Parkinson's disease or stroke: a pilot study. Physiother Theory Pract 2012; 28: 509-514.
5) Tsang CS, Liao L-R, Chung RC, Pang MY. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. Phys Ther 2013; 93: 1102-1115.
6) Oyama C, Otaka Y, Onitsuka K, Takagi H, Tan E, Otaka E. Reliability and Validity of the Japanese Version of the Mini-balance Evaluation Systems Test in Patients with Subacute Stroke. Prog Rehabil Med 2018; 3: 20180015.
7) Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation System's Test: the mini-BESTest. J Rehabil Med 2010; 42: 323.
8) Horak FB, Wrisley DM, Frank J. The balance evaluation systems test (BESTest) to differentiate balance deficits. Phys Ther 2009; 89: 484-498.
9) Lampropoulou SI, Billis E, Gedikoglou IA, Michailidou C, Nowicky AV, Skrinou D, Michailidi F, Chandrinou D, Meligkoni M. Reliability, validity and minimal detectable change of the Mini-BESTest in Greek participants with chronic stroke. Physiother Theory Pract 2019; 35: 171-182.
10) Göktaş A, Çolak FD, Kar İ, Ekici G. Reliability and validity of the turkish version of the mini-BESTest balance scale in patients with stroke. Türk Nöroloji Dergisi 2020; 26: 303-310.
11) Di Carlo S, Bravini E, Vercelli S, Massazza G, Ferriero G. The Mini-BESTest: a review of psychometric properties. Int J Rehabil Res 2016; 39: 97-105.
12) Sofia L, Ingrid GA, Christina M, Evdokia B. Cross Cultural Validation of the Mini-BESTest into Greek. WJRR 2016; 3: 61-67.
13) Lemay JF, Roy A, Nadeau S, Gagnon DH. French version of the Mini BESTest: A translation and transcultural adaptation study incorporating a reliability analysis for individuals with sensorimotor impairments undergoing functional rehabilitation. Ann Phys Rehabil Med 2019; 62: 149-154.
14) Bustamante-Contreras C, Ojeda-Gallardo Y, Rueda-Sanhueza C, Rossel PO, Martínez-Carrasco C. Spanish version of the mini-BESTest: a translation, transcultural adaptation and validation study in patients with Parkinson's disease. Int J Rehabil Res 2020; 43: 129-134.
15) Naghdi S, Nakhostin Ansari N, Forogh B, Khalifeloo M, Honarpisheh R, Nakhostin-Ansari A. Reliability and Validity of the Persian Version of the Mini-Balance Evaluation Systems Test in Patients with Stroke. Neurol Ther 2020; 9: 567-574.
16) Cramer E, Weber F, Faro G, Klein M, Willeke D, Hering T, Zietz D. Cross-cultural adaption and validation of the German version of the Mini-BESTest in individuals after stroke: an observational study. Neurol Res Pract 2020; 2: 27.
17) Kline R. edition 3rd Ed. ed. Principles and practice of structural equation modeling. New York. Guilford Press; 2011.
18) Anthoine E, Moret L, Regnault A, Sebille V, Hardouin JB. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes 2014; 12: 176.
19) Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med 2002; 21: 1331-1335.
20) Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000; 25: 3186-3191.
21) Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007; 30: 459-467.
22) Yusoff MSB. ABC of content validation and content validity index calculation. EIMJ 2019; 11: 49-54.
23) Lampropoulou SI, Billis E, Gedikoglou IA, Michailidou C, Nowicky AV, Skrinou D, Michailidi F, Chandrinou D, Meligkoni M. Reliability, validity and minimal detectable change of the Mini-BESTest in Greek participants with chronic stroke. Physiother Theory Pract 2019; 35: 171-182.
24) Berg K. Measuring balance in the elderly: Development and validation of an instrument: McGill University; 1992.

25) El-Gilany A-H, Hatata E, Aayob NS, Refaat R. Validation of the Arabic version of the Berg Balance Scale (A-BBS) among elderly residents in a rural community. ME-JAA 2012; 9: 3-7.
26) Dye DC, Eakman AM, Bolton KM. Assessing the validity of the dynamic gait index in a balance disorders clinic: an application of Rasch analysis. Phys Ther 2013; 93: 809-818.
27) Alghwiri AA. Reliability and validity of the Arabic Dynamic Gait Index in people poststroke. Top Stroke Rehabil 2014; 21: 173-179.
28) Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther 2001; 81: 776-788.
29) Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018; 18: 91-93.
30) Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015; 25: 141-151.
31) Jácome C, Cruz J, Oliveira A, Marques A. Validity, reliability, and ability to identify fall status of the Berg Balance Scale, BESTest, Mini-BESTest, and Brief-BESTest in patients with COPD. Phys Ther 2016; 96: 1807-1815.
32) Winairuk T, Pang MY, Saengsirisuwan V, Horak FB, Boonsinsukh R. Comparison of measurement properties of three shortened versions of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. J Rehabil Med 2019; 51: 683.
33) Godi M, Franchignoni F, Caligari M, Giordano A, Turcato AM, Nardone A. 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.
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.
35) Lofgren N, Lenholm E, Conradsson D, Stahle A, Franzen E. The Mini-BESTest--a clinically reproducible tool for balance evaluations in mild to moderate Parkinson's disease? BMC Neurol 2014; 14: 235.
36) Dominguez-Olivan P, Gasch-Gallen A, Aguas-Garcia E, Bengoetxea A. Validity and reliability testing of the Spanish version of the BESTest and mini-BESTest in healthy community-dwelling elderly. BMC Geriatr 2020; 20: 444.
37) Potter K, Bowling R, Kavanagh L, Stone A, Witt B, Wooldridge A. Reliability, validity, and responsiveness of the mini-balance evaluation systems test in ambulatory individuals with multiple sclerosis. Physiother Can 2019; 71: 327-334.
38) Anson E, Thompson E, Ma L, Jeka J. Reliability and Fall Risk Detection for the BESTest and Mini-BESTest in Older Adults. J Geriatr Phys Ther 2019; 42: 81-85.
39) Beauchamp MK, Niebuhr R, Roche P, Kirkwood R, Sibley KM. A prospective study to establish the minimal clinically important difference of the Mini-BESTest in individuals with stroke. Clin Rehabil 2021; 35: 1207-1215.
40) Roy A, Higgins J, Nadeau S. Reliability and minimal detectable change of the mini-BESTest in adults with spinal cord injury in a rehabilitation setting. Physiother Theory Pract 2021; 37: 126-134.
41) Duchesne E, Hebert LJ, Mathieu J, Cote I, Roussel MP, Gagnon C. Validity of the Mini-BESTest in adults with myotonic dystrophy type 1. Muscle Nerve 2020; 62: 95-102.
42) Leddy AL, Crowner 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.