Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Geriatr Phys Ther. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:

J Geriatr Phys Ther. 2019 ; 42(2): 81–85. doi:10.1519/JPT.0000000000000123.

Reliability and Fall risk Detection for the BESTest and mini-BESTest in Older Adults

Eric Anson, PT, PhD[a], **Elizabeth Thompson, PT, DPT**[b,c], **Lei Ma, MS**[b], and **John Jeka, PhD**[b]

[a]Department of Otolaryngology Head & Neck Surgery, Johns Hopkins School of Medicine, Baltimore, MD

[b]Department of Kinesiology, Temple University, Philadelphia, PA

[c]Department of Physical Therapy, Temple University, Philadelphia, PA

Abstract

Background & Purpose—Test stability and test-retest reliability have not previously been reported for either the BESTest or mBEST in a population of older adults with non-specific balance limitations. Furthermore, no criterion for identifying change greater than chance has been reported in older adults with non-specific balance problems using either BESTest or mBEST scores. The purposes of this study were to determine test stability over time, test-retest reliability, to identify minimum detectable change for the BESTest and mini-BEST test in a population of older adults with non-specific balance problems. Additionally, the ability of the BESTest and mBEST to identify past fallers was characterized.

Methods—Observational study with 58 adults 65 years old or older with a history of falls or self-reported balance problem. BESTest and mini-BEST were administered to all participants at the beginning and end of 4 weeks. Test-retest reliability was calculated with intraclass correlations, and minimum detectable change was calculated at the 95% confidence level (MDC_{95}). Receiver operating characteristics were used to characterize the sensitivity and specificity of the BESTest and mBEST to identify older adults who had previously fallen.

Results—Balance scores did not significantly change over a 4 week period. Test-retest reliability for the BESTest (.86) and mini-BEST (.84) was good to excellent. MDC_{95} scores were identified for the BESTest (8.9) and mini-BEST (4).

Conclusions—The BESTest and mini-BEST scores were stable and reliable over a period of 4 weeks for a population of older adults with self-reported balance problems or a history of falling. MDC_{95} scores allow interpretation of change in BESTest and mBEST scores following rehabilitation.

Keywords

BESTest; mini-BEST; reliability; Fall risk; Older Adults

Corresponding Author Eric Anson, 601 N Caroline Street, Ste 6030B, Johns Hopkins Medical Institutes, Baltimore, MD 21287, Phone: 410-955-7381, eanson1@jhmi.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

INTRODUCTION

In the U.S., an estimated 2.5 million non-fatal falls and 24 thousand fatal falls occur every year in individuals over age 65, with an annual cost of 19 to 40 billion dollars.[1–4] In addition to the financial impact, falls have a negative impact on health and quality of life for older adults.[5] Many clinical balance assessment tools exist to characterize current fall risk and change in fall risk following treatment.[6] The Balance Evaluation Systems Test (BESTest) and the mini-BESTest (mBEST) have received attention as tests with the potential to characterize balance ability across multiple sub-domains during standing and walking.[6–10] Assessments like these may allow a more targeted approach to rehabilitation by identifying specific systems that are limited.[6] Despite recent research attention, less than 1% of clinicians regularly use the BESTest for evaluating balance ability and identifying fall risk; but 79% of those respondents test single leg stance routinely.[11] The mBEST, a shortened version of the BESTest, was designed to reduce the time burden while focusing on dynamic balance assessment.[7,12]

The BESTest is a 27 item physical performance test with items distributed among 6 sub-systems that underlie static and dynamic balance: biomechanical constraints, stability limits/verticality, anticipatory postural adjustments, postural responses, sensory orientation, and stability in gait.[6] The BESTest has excellent inter-rater reliability (ICC = .91) with a mixed population of individuals with neurological disorders and balance limitations,[6] and it has excellent test re-test reliability (ICC = .88) for individuals with Parkinson's Disease (PD).[9] Test-retest reliability for the mBEST has also been shown to be excellent (ICC = .92–.97) for individuals with a variety of neurological disorders and balance limitations.[8,13,14] The mBEST correlated highly with the Berg Balance Scale (BBS), without the ceiling effects that limit the utility of the BBS.[10]

Previous reports addressing test-retest reliability for BESTest and mBEST in older adults had a limited number of older adults with non-specific balance limitations grouped with adults who had specific neurological disorders that contributed to their balance limitations.[8,13,14] Neurological conditions may contribute to falls in older adults but not all older fallers have specific neurological conditions.[15] However, test-retest reliability has not previously been reported for either the BESTest or mBEST in a population of older adults with non-specific balance limitations. Furthermore, no discriminatory criteria for identifying minimum detectable change (MDC_{95}) or fall risk in older adults with non-specific balance problems using either BESTest or mBEST scores have been reported. BESTest cut-off criteria previously developed for individuals with PD (69%) may not be appropriate for older adults with non-specific balance limitations.[9,16] It is not known if BESTest or mBEST scores can identify individuals in this population with elevated fall risk, or whether previously reported cut off scores are appropriate for older adults with non-specific balance limitations. The foci of this study were to describe the test stability over time, test-retest reliability, and MDC_{95} of the BESTest and mBEST, and to characterize the ability of the BESTest and mBEST to identify fallers in a population of older adults with non-specific balance limitations.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

METHODS

Participants

Fifty eight participants provided informed consent prior to participating in this study. Participants were consecutively recruited to participate in a pilot feasibility study which was subsequently transitioned to a randomized clinical trial (ClinicalTrials.gov ##366151-1). Recruitment was facilitated via advertisements in a newspaper for older adults in the greater Washington D.C. area and via fliers placed at North City Congress (Philadelphia, PA), Dr. Grace Ma's Center for Asian Health (Philadelphia, PA) and Collington Episcopal Life Care Community (Mitchellville, MD). Results reported here do not reflect results from the randomized clinical trial. All participants were adults over 65 with self-reported balance problems without neurological or musculo-skeletal diseases who scored at least 24 on the Mini-Mental State Exam and were able to walk continuously on a treadmill for 2 minutes. The inclusion criteria reported here were designed to ensure eligibility for the clinical trial which was a treadmill based intervention study that required sufficient cognition to attend to and follow multi-step instructions. This study was approved by the Institutional Review Boards at the University of Maryland and Temple University. Data presented here represent a 4 week control period prior to the training portion of the clinical trial. Subject testing occurred at The University of Maryland, Temple University, and Collington Episcopal Life Center. EA, ET, and RR performed the testing. Four participants withdrew for personal reasons prior to the second testing session resulting in 55 subjects for test stability and test-retest comparisons and 58 subjects for the fall risk analysis.

Protocol

All participants were tested at the beginning and end of a 4 week period with instructions not to change their physical activity level. Balance and walking ability were characterized with the BESTest[6] and mBEST[7], which are clinical assessments of gait and balance. Since all of the items on the mBEST are included in the BESTest, participants performed items on the BESTest once and the performance was scored separately according to BESTest and mBEST criteria to avoid repeating test items and causing unnecessary fatigue.[8] Examples of the differences in scoring criteria are presented in Table 1 for select test items from the mBEST.

Participants were classified as fallers or non-fallers according to their self-report of at least one fall in the previous 12 months. Actual falls were tracked but not used to classify participants as fallers.

Data Analysis

Mixed model repeated measures ANOVAs with 2 levels (time: test 1 and test 2) and 2 groups (fallers vs. non-fallers) were performed separately to test the overall hypothesis that there was no effect of time, group, or an interaction between time and group for the BESTest and mBEST. Fifty five participants were used for intraclass correlation coefficients (ICC) (2,1) to determine the test-retest reliability for the BESTest and mBEST. Standard error of measurement (SEM) was calculated to determine the response stability for the BESTest and mBEST. The 95% minimum detectable change (MDC_{95}) was determined for the BESTest

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and mBEST, to identify the magnitude of improvement greater than chance.[17] Receiver operating curve (ROC) analysis was performed to evaluate cut points for identifying increased fall risk for the mBEST and BESTest using scores from the first testing point with all 58 participants. Hypothesis testing was conducted with $\alpha = 0.05$.

RESULTS

The average age of the participants was 78.1 (SD 7.01, range 66–92) and 72% were female. Thirty eight out of the 58 subjects reported at least 1 fall in the previous 12 months. Three participants fell during the 4 week period between tests. There was no significant effect of time for the BESTest ($z = 1.28$, $p = 0.202$, 95% CI [−0.7, 3.3]) or mBEST ($z = 0.84$, $p = 0.399$, 95% CI [−0.5, 1.3]), see Table 2. There was not significant difference in scores between fallers and non-fallers for the BESTest ($z = 0.15$, $p = 0.880$, 95% CI [−3.9, 4.5]) or mBEST ($z = 0.02$, $p = 0.988$, 95% CI [−1.8, 1.8]). There was no interaction between group and time for either the BESTest ($z = -0.11$, $p = 0.909$, 95% CI [−2.6, 2.3]) or mBEST ($z = 0.21$, $p = 0.832$, 95% CI [−1.0, 1.3]).

Test-retest ($n = 56$) reliability ICC(2,1) values are presented in Table 2. The mBEST and BESTest both demonstrated excellent test-retest reliability (ICCs(2,1) .84–.86).[18] The SEM and MDC_{95} for the first test time point for the BESTest was 8.9 and for the mBEST was 4.0, see Table 2.

The area under the curve was similar for the BESTest (0.53, 95% CI [0.37–0.69]) and the mBEST (0.54, 95% CI [0.39–0.69]). Cut-off values to identify fallers and likelihood ratios were not calculated since the area under the curve was not meaningfully greater than chance, since the 95% CI included 0.5 for both curves. An area under the curve equal to 0.5 corresponds to a test that is not better than chance at discriminating those with a condition from those without a condition.

DISCUSSION

The primary purpose of this study was to determine test stability and test-retest reliability of the BESTest and mBEST in a population of older adults with non-specific balance limitations. Test-retest reliability for the BESTest (ICC(2,1) = 0.86) and mBEST (ICC(2,1) = 0.84) in a population of older adults with non-specific balance limitations was consistent with previous reports for participants with PD.[7,9,13,14] The 4 week time period used in this study was longer than the time periods in previous work on test-retest reliability for the BESTest and mBEST.[9,10,13,14] This extended time period could have resulted in performance variability in our participants; however, there were no significant changes over time in BESTest or mBEST test scores. Further, in this cohort participants with 1 or more past falls performed similarly compared to non-fallers (see Table 2) and neither group demonstrated significant changes over time. The longer period between each test also served to minimize the potential that scoring may be biased because the testers remembered previous test performance. Additionally, this time period reflects the reality of current clinical practice, where 4 week intervals often define the time period between functional performance assessments consistent with the Centers for Medicare & Medicaid Services

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

requirement for reassessments at least every 30 calendar days.[19] Changes in functional ability are more likely to be detected over a longer time period compared to a shorter time period.[20,21] Scores on the BESTest and mBEST were stable over this 4 week time window, regardless of whether they had a history of falling, during which participants were asked to maintain their current activity levels and are appropriate for evaluating balance in older adults with balance limitations.

Participants were instructed not to change their daily activities; however, the extended time between assessments afforded the possibility of changes in functional ability. Despite some participants being actively involved in regular exercise programs, the overall test scores did not change significantly during the 4 week period of this study. The present results demonstrate that maintaining current levels of activity, even if that included regular exercise, did not result in a meaningful improvement in balance ability after 4 weeks as measured using the BESTest and mBEST in this study. The implication is that older adults with balance problems will not improve their balance solely by *maintaining* their current level of activity. In this light, any individual with elevated fall risk as identified using the BESTest or mBEST should receive appropriate interventions to modify physical ability, environmental risks for falling, and possibly even changes in medication as appropriate.[22] Changes in activity level and overall health status during the 4 week period were not measured; but 3 participants reported non-injurious falls between the 2 assessments. Two participants withdrew from the study due to new onset knee or back pain. It is conceivable that there were complementary fluctuations in balance performance such that some participants improved while others worsened due to changes in their health status. This would be hidden in the current analysis since we could not include activity level or health status as a covariate.

The secondary purpose was to characterize the ability of BESTest and mBEST to detect change in balance ability and identify fallers for older adults with non-specific balance problems. Neither the BESTest or mBEST scores were able to identify participants with a history falling better than chance. Recruitment bias may have contributed to this effect since all eligible subjects must have had at least 1 fall in the past year or a self-reported balance problem. As such it remains to be seen whether a cut-off score for identifying fallers using BESTest or mBEST scores in a population of adults with non-specific balance limitations differs from those previously reported for individuals with PD (BESTest = 69%, mBEST = 23).[9,10,23] Interestingly, there were 3 participants who fell between the first and second tests. Their BESTest scores were: 63.9%, 69.4%, 71.3%; and mBEST scores were: 11, 17, 18. Only 1 of those participants had a BESTest score below the reported threshold for individuals with PD, while all 3 scored below the threshold for detecting fallers using the mBEST.[8] The impairments and associated balance problems may be more severe for individuals with PD compared to older adults with non-specific balance problems. This suggests that the BESTest cut-off values identified for individuals with PD may not be valid for older adults with non-specific balance problems. Future prospective studies are needed to determine whether there is an ideal test or test battery for identifying future fallers in a population of older adults with non-specific balance problems without relying on past fall history as a gold standard.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MDC_{95} score changes for BESTest (8.9) and mBEST (4) provide useful criteria for goal setting and interpretation of change in balance for older adults participating in balance training or rehabilitation programs. The identified MDC_{95} scores are consistent with those reported from testing individuals with PD[8] and other mixed populations including individuals with PD, hemiparesis, multiple sclerosis, vestibular disease, ataxia, polyneuropathy, neoplasm, and non-specific balance limitations.[14] With good to excellent test-retest reliability, BESTest and mBEST can be useful tools for evaluating balance ability and fall risk for older adults with non-specific balance problems.

LIMITATIONS

Despite identifying meaningful change scores for the BESTest and mBEST in a cohort of older adults with self-reported balance problems, cut-off scores were not calculated to identify older adults with increased fall risk. Participants were identified as fallers based on self-report of a fall in the previous 12 months. Relying on only past fall history rather than multiple falls or prospective fall tracking over an extended time period may have contributed to poor discriminatory ability to detect fallers in this cohort. Activity level was not measured and it is possible that participants who are more active may have experienced improvements while less active individuals experienced decline. Additional studies are needed to elucidate the relationship between activity level and balance function in older adults. Prospective studies are needed to determine the predictive ability of the BESTest and mBEST for identifying individuals who actually fall in the future.

CONCLUSIONS

Both BESTest and the mBEST scores demonstrate good to excellent test re-test reliability over a 4 week period in older adults with non-specific balance problems. Progress in rehabilitation can be determined by changes in BESTest (9) and mBEST (4) scores for older adults with self-reported balance problems. Balance scores on the BESTest and mBEST were stable over 4 weeks indicating that balance does not improve or decline when older adults with self-reported balance problems maintain their current activity level.

Acknowledgments

The authors would like to thank Roshita Rathore BPT, MS, for assistance with data collection.

FUNDING SOURCES

This manuscript was supported in part by NIH grant 7R21AG041714 (John Jeka, PI); NIH grant NIDCD T32 DC000023 (Eric Anson); PODS I and PODS II Scholarships from the Foundation for Physical Therapy (Eric Anson), Inc.; and a Wylie Dissertation Fellowship from the University of Maryland Graduate School (Eric Anson).

References

1. Stevens JA, Corso PS, Finkelstein EA, Miller TR. The costs of fatal and non-fatal falls among older adults. Inj Prev. 2006; 12(September 2007):290–295. [PubMed: 17018668]
2. Nurmi I, Lüthje P. Incidence and costs of falls and fall injuries among elderly in institutional care. Scand J Prim Health Care. 2002; 20(2):118–122. [PubMed: 12184711]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3. Centers for Disease Control and Prevention National Center for Injury Prevention and Control. [Accessed June 30, 2012] No Title. Web–based Inj Stat Query Report Syst. 2015. http://www.cdc.gov/injury/wisqars/index.html
4. Michael YL, Whitlock EP, Lin JS, Fu R, O'Connor EA, Gold R. Primary care-relevant interventions to prevent falling in older adults: a systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med. 2010; 153(12):815–825. [PubMed: 21173416]
5. Fuller G. Falls in the Elderly. Am Fam Physician. 2000; 61(7):2159–2168. [PubMed: 10779256]
6. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. Phys Ther. 2009; 89(5):484–498. [PubMed: 19329772]
7. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: The Mini-BESTest. J Rehabil Med. 2010; 42(4):323–331. [PubMed: 20461334]
8. Leddy AL, Crowner B, Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinsons Disease. J Neurol Phys Ther. 2011; 35(2): 90–97. [PubMed: 21934364]
9. Leddy A, Crowner B, Earhart G. Functional Gait Assessment and Balance Evaluation System Test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson Disease who fall. Phys Ther. 2011; 91(1):102–113. [PubMed: 21071506]
10. King LA, Priest KC, Salarian A, Pierce D, Horak FB. Comparing the Mini-BESTest with the Berg Balance Scale to evaluate balance disorders in Parkinson's Disease. Parkinsons Dis. 2012; 2012:375419. [PubMed: 22135761]
11. Sibley KM, Straus SE, Inness EL, Salbach NM, Jaglal SB. Balance assessment practices and use of standardized balance measures among Ontario physical therapists. Phys Ther. 2011; 91(11):1583–1591. [PubMed: 21868613]
12. King L, Horak F. On the Mini-BESTest: scoring and the reporting of total scores. Phys Ther. 2013; 93(4):571–575. [PubMed: 23547173]
13. Tsang CSL, Liao L, Raymond CK, Pang MYC. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. Phys Ther. 2013; 93(8):1102–1115. [PubMed: 23559522]
14. Godi M, Franchignoni F, Caligari M, Giordano A, Turcato AM. Comparison of reliability, validity, and responsiveness of the Mini-BESTest and Berg Balance Scale in patients with balance disorders. Phys Ther. 2013; 93(2):158–167. [PubMed: 23023812]
15. Tinetti ME, Speechley M, Ginter SF. Risk Factors for Falls Among Elderly Persons Living in the Community. N Engl J Med. 1988; 319(26):1701–1707. [PubMed: 3205267]
16. Mak MKY, Auyeung MM. The mini-BESTest can predict parkinsonian recurrent fallers: a 6-month prospective study. J Rehabil Med. 2013; 45(6):565–571. [PubMed: 23673397]
17. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes. 2006; 5:54.
18. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994; 6(4):284–290.
19. Services C for M& M. [Accessed January 23, 2015] CMS announces payment changes for Medicare home health agencies for 2015. 2014. http://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2014-Fact-sheets-items/2014-10-30.html
20. Lesinski M, Hortobágyi T, Muehlbauer T, Gollhofer A, Granacher U. Effects of balance training on balance performance in healthy older adults: A systematic review and meta-analysis. Sports Med. 2015; 45(12):1721–1738. [PubMed: 26325622]
21. Howe TE, Rochester L, Neil F, Skelton DA, Ballinger C. Exercise for improving balance in older people. Cochrane database Syst Rev. 2011; (11):CD004963. [PubMed: 22071817]
22. Pfortmueller CA, Lindner G, Exadaktylos AK. Reducing fall risk in the elderly: risk factors and fall prevention, a systematic review. Minerva Med. 2014; 105(4):275–281. [PubMed: 24867188]
23. Duncan RP, Leddy AL, Cavanaugh JT, et al. Accuracy of fall prediction in Parkinson Disease: six-month and 12-month prospective analyses. Parkinsons Dis. 2012; 2012(237673):1–7.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Exemplar differences in scoring criteria between the full BESTest and the mini-BESTest.

Test Items	BESTest Scoring Criteria	Mini-BESTest Scoring Criteria
Sit To Stand	(3) Normal: Comes to stand without the use of hands and stabilizes independently (2) Comes to stand on the first attempt with the use of hands (1) Comes to stand after several attempts or requires minimal assist to stand or stabilize or requires touch of back of leg or chair (0) Requires moderate or maximal assist to stand	(2) Normal: Comes to stand without use of hands and stabilizes independently. (1) Moderate: Comes to stand WITH use of hands on first attempt. (0) Severe: Unable to stand up from chair without assistance, OR needs several attempts with use of hands.
Compensatory Stepping Correction-Forward	(3) Recovers independently a single, large step (second realignment step is allowed) (2) More than one step used to recover equilibrium, but recovers stability independently OR 1 step with imbalance (1) Takes multiple steps to recover equilibrium, or needs minimum assistance to prevent a fall (0) No step, OR would fall if not caught, OR falls spontaneously	(2) Normal: Recovers independently with a single, large step (second realignment step is allowed). (1) Moderate: More than one step used to recover equilibrium. (0) Severe: No step, OR would fall if not caught, OR falls spontaneously.
Stance (Feet Together); Eyes Closed, Foam Surface	(3) 30s stable (2) 30s unstable (1) < 30s (0) Unable	(2) Normal: 30 s. (1) Moderate: < 30 s. (0) Severe: Unable.
Walk with Head Turns - Horizontal	(3) Normal: performs head turns with no change in gait speed and good balance (2) Mild: performs head turns smoothly with reduction in gait speed, (1) Moderate: performs head turns with imbalance (0) Severe: performs head turns with reduced speed AND imbalance AND/OR will not move head within available range while walking.	(2) Normal: performs head turns with no change in gait speed and good balance. (1) Moderate: performs head turns with reduction in gait speed. (0) Severe: performs head turns with imbalance.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Comparison of BESTest and mBEST Scores During 2 Test Sessions.

Clinical Balance and Walking Tests	Test 1 Fallers Mean (SD)	Test 1 Non-Fallers Mean (SD)	Test 2 Fallers Mean (SD)	Test 2 Non-Fallers Mean (SD)	ICC (95% CI)	SEM Test 1	MDC_{95} Test 1
BESTest (%)	72.3 (9.0)	72.9 (8.2)	73.9 (7.9)	75.0 (9.0)	.86 (.77–.92)	3.2	8.9
mBEST	19.0 (3.8)	19.3 (3.1)	19.7 (3.7)	19.8 (3.8)	.84 (.73–.90)	1.4	4.0

ICC – Intraclass correlation coefficient, BESTest – Balance Evaluation Systems Test, mBEST – mini Balance Evaluation Systems Test, SEM – standard error of measurement, MDC_{95} – 95% minimum detectable change, SEM and MDC_{95} were calculated after combining the groups (fallers and non-fallers).

J Geriatr Phys Ther. Author manuscript; available in PMC 2020 April 01.