



The Internet Journal of Allied Health Sciences and Practice

<http://ijahsp.nova.edu>

A Peer Reviewed Publication of the College of Allied Health & Nursing at Nova Southeastern University

Dedicated to allied health professional practice and education

<http://ijahsp.nova.edu> Vol. 5 No. 4 ISSN 1540-580X

Functional Balance Assessment of Older Community Dwelling Adults: A Systematic Review of the Literature

Felicity Anne Langley

Shylie F.H. Mackintosh, PhD, MSc, B.ApplSc

University of South Australia, City East Campus, North Terrace, Adelaide, South Australia

CITATION: Langley, FA., Mackintosh, SFH. Functional balance assessment of older community dwelling adults: a systematic review of the literature. *The Internet Journal of Allied Health Sciences and Practice*. Oct 2007, Volume 5 Number 4.

ABSTRACT

Background: For allied health professionals wishing to assess the functional balance of older adults living in the community, the vast number of functional balance tests available makes it difficult to decide which assessment is most appropriate. Objective: To identify the reliability, concurrent validity and clinical practicality of functional balance tests with community dwelling older adults. Methods: A systematic review of published literature relevant to 17 functional balance tests was undertaken. The 17 functional balance tests were identified by a preliminary literature search and through consultation with an expert in the field of functional balance assessment. Studies published in English before January 2007, assessing the use of these functional balance tests with community dwelling adults aged 65 years or above were included. The CINAHL, MEDLINE, Ageline, Amed, PubMed, Cochrane library, PEDro and Joanna Briggs Institute databases were searched. The methodological quality of studies was assessed using a checklist criteria adapted from the Cochrane Working Group for Screening and Diagnostic Tests. Results: Eight databases were searched and 21 studies were included. The majority of studies demonstrated low to moderate methodological quality scores. Despite limitations reported for clinical application with community dwelling older adults, the Berg Balance Scale and the Timed Up and Go Test have been most rigorously tested. Reliability and concurrent validity of the Balance Screening Tool and the Fullerton Advanced Balance Scale had also been established in this population, however only one study was retrieved for each. Conclusion: The Berg Balance Scale and Timed Up and Go Test have published reliability and validity with community dwelling older adults. Further testing of other functional balance tests is required to establish their reliability and validity in this target population.

BACKGROUND

The performance of all activities of daily living requires good balance control while at rest or when moving from one position to another.^{1,2} Maintenance of balance requires the coordination of sensory, neural and musculoskeletal systems.^{1,2} Many of these systems undergo deterioration as people age.^{3,4} This has the potential to affect balance, restrict safe mobility, increase the likelihood of a fall and adversely affect quality of life.^{3,5} Therefore, the assessment of balance with older people is important to direct appropriate interventions to improve balance performance and to monitor changes in balance over time.⁶

Various approaches to measure balance have been developed.^{3,7} Questionnaires such as the Rivermead Mobility Index and the Activities Specific Balance Scale provide self-report information regarding functional status.⁸ Laboratory measures such as computerized force platforms provide accurate measurements of postural sway.^{3,9} Functional performance-based tests such as the Berg Balance Scale and the Timed Up and Go Test may be defined as tests which objectively measure a person performing balance or walking tasks.^{7,10,11} An advantage of functional balance tests is their practicality for assessment in a variety of settings because of their low cost, lack of complex equipment and time efficiency.^{10,11}

To substantiate the clinical use of a functional balance test, it must be established to be reliable and valid when used with the target population.⁹ A reliable test should provide similar results when performed on the same person and under the same conditions both on different occasions and by different raters.³ Results of a test with concurrent validity should correlate with results of another “gold standard” test when performed on the same person.³ For functional balance tests, the Berg Balance Scale has been reported to be the “gold standard”, and will be considered the gold standard for this systematic review.^{7, 21} To detect balance impairments of older adults, functional balance tests should also challenge the various sensory, neural and musculoskeletal systems required for postural control. To be a useful outcome measure, functional balance tools also need to be sensitive to change in the elderly community dwelling population.^{3,9}

As there are a vast number of functional balance assessment tests available, it is often difficult for a therapist or researcher to decide which to use. Psychometric properties and the practical application of some tests have been summarized in previous literature reviews and a systematic review for people with stroke.^{7,11,12} Such reviews include results of studies involving a wide range of populations. To date, no systematic reviews have been conducted to assess the psychometric properties and practical application of functional balance tests in community dwelling older people. Such a review would aid therapists and researchers to make decisions about which tool may be most useful when assessing functional balance in this target population.

Therefore, the aims of the following systematic review were; to determine the reliability, concurrent validity and the reported practicality for application of functional balance measures when used with community dwelling older adults.

METHOD FOR REVIEW

Study eligibility criteria

Functional Balance Assessment Tools

A preliminary search and consultation with an expert in balance assessment identified 33 tests of functional balance for older adults. Assessment tools were included for review if they fulfilled each of the following criteria;

- A test of functional balance appropriate for people living in the community (includes standing tasks).
- Practical and able to be tested in a variety of settings (requires minimal and easily accessible equipment and take less than 20 minutes to administer).
- Challenges at least two of the following balance components;
 - Static balance
 - Dynamic balance
 - Change in sensory conditions (visual, vestibular, somatosensory)

Seventeen of the 33 functional balance tests identified met these criteria. Each of these tests are listed in table 1 as search terms.

Types of Studies

Studies were included if they assessed any of the 17 functional balance tools identified with regards to either; inter-rater reliability, intra-rater reliability or concurrent validity. Investigations of construct validity, sensitivity or specificity with regards to falls status were excluded due to the wide scope of this area which does not pertain exclusively to functional balance status. To meet the inclusion criteria for the review, all 17 functional balance tests that were selected demonstrated face and content validity. Studies were excluded if they were not reported in English and no studies were sought after January 2007.

Types of Participants

Studies which assessed community dwelling participants aged 65 years or older were included. Studies which involved participants who were not community dwelling or aged less than 65 years were included if the majority of participants were community dwelling older adults, or a separate analysis of this group was provided. Study populations including participants with conditions which may affect balance, such as stroke, Parkinson's disease, Multiple Sclerosis, vestibular dysfunction and acute medical conditions were excluded.

Search Strategy

Studies were identified using a systematic search of computer databases. CINAHL, MEDLINE, Ageline, Amed (Allied and Complementary Medicine) and PubMed were searched using combinations of the keywords listed in table 1 (two or more terms). The third keyword, 'balance assessment' was used where more than 50 results were returned by the initial search using the first two terms. The Cochrane library, PEDro, and the Joanna Briggs Institute database were also searched however these sources did not yield any further relevant literature. A manual search was conducted using the reference lists of relevant articles and the researchers' library.

Table 1: Search keywords

1. Functional Balance Tool	2. Psychometric property	3. Balance Assessment
Berg Balance Scale (BBS)	Reliab*	Functional bal*
Functional Reach Test (FRT)	Valid*	Equilib*
Lateral Reach Test (LRT)	Sensitivity*	Geriatric* Assess*
Timed Up and Go Test (TUG)	Response*	Balance, Postural
Step test	Ceiling effect*	Functional Assess*
Tinetti Performance Orientated Mobility Assessment (POMA)		Musculoskeletal Equilib*
Four square step test (FSST)		
Elderly Mobility Scale (EMS)		
Balance Screening Tool (BST)		
Physical Mobility Scale (PMS)		
Sensory Orientated Mobility Assessment Instrument (SOMAI)		
Clinical Test of Sensory Interaction and Balance (CTSIB)		
Short Form Berg Balance Scale (SF-BBS)		
Hierarchical Assessment of Balance and Mobility (HABAM)		
Fullerton Advanced Balance Scale (FAB)		

Boolean operators relevant to the database searched (*)

Critical Appraisal

Critical appraisal was undertaken using an adaptation of the checklist from the Cochrane Methods Working Group for Screening and Diagnostic Tests cited in the National Health and Medical Research Council (NHMRC) guidelines (1999) (Table 2).¹³ While the reliability and validity of this critical appraisal tool is not reported, no other validated tool was located that addressed the issues of potential bias in test-retest design reliability and concurrent validity studies.

All studies were assessed by two independent reviewers. Any disagreements were resolved by discussion. Had a resolution not been able to be reached, a third independent reviewer was available, however this was not required. In studies which investigated both reliability and concurrent validity, these psychometric properties were critically appraised separately. Critical appraisal scores were calculated by summing up the number of 'Yes' responses that were obtained and dividing them by the total number possible (excluding the number of 'Not Applicable' responses). This produced a percentage score, ranging from 0% (no quality) to 100% (excellent methodological quality).

Table 2: Critical Appraisal Checklist

	Item	Score
1	Have main features of subjects been clearly described?	Y/N/NA
2	Were subjects recruited consecutively?	Y/N/NA
3	Were all subjects included that should have been included?	Y/N/NA
4	Have reasons for exclusion been described?	Y/N/NA
5	Has the assessment method been clearly described?	Y/N/NA
6	Was there a valid reference standard?	Y/N/NA
7	Are test and reference standards measured independently?	Y/N/NA
8	Are tests measured independently of other clinical and test information?	Y/N/NA
9	Do table and figures support the text?	Y/N/NA
10	Have appropriate statistical methods been used?	Y/N/NA
11	Are results supported by existing evidence?	Y/N/NA
12	Have all possible confounding variables been identified?	Y/N/NA
13	Were all subjects included in analysis?	Y/N/NA
14	Has attrition been recorded?	Y/N/NA
	Total:	Yes/14-NA

Yes (Y), No (N), Not Applicable (NA)

RESULTS

Search Results

The initial list of results was screened for eligibility by one evaluator using abstracts and titles, leaving 21 studies for review. Most studies were excluded because they did not investigate either reliability or concurrent validity (figure 1).

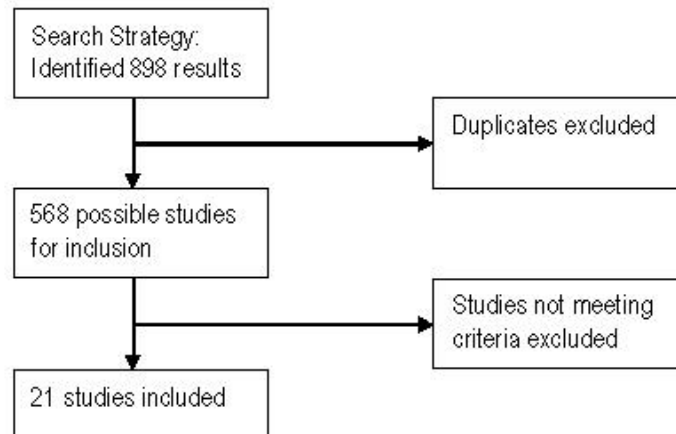


Figure 1. Search Results

Methodological Quality

Thirteen studies assessing both reliability and validity of a functional balance test were retrieved. Two studies investigating validity and six investigating reliability alone were retrieved. The average quality score for reliability and validity investigations were 47.47 percent and 51.60 percent respectively.

Poor reporting of participant characteristics was common to reliability studies. In 11 studies, reliability data was obtained through testing of a subset of the total sample, of which descriptive characteristics and selection methods were not provided. Methodological procedures regarding blinding, test order, and test protocol were also poorly reported in most studies. Only one concurrent validity study reported that raters were blinded to the participants initial test results and in numerous inter-rater reliability studies, it was not clear whether raters were blinded to each others scores. For intra-rater reliability studies, memory will impact on results as it is impossible to blind the assessor to results of the initial test. Most studies however did not identify this as a possible confounding variable. Reporting of methodological details are important in ensuring that rater bias and confounding variables are minimized in reliability and validity studies.

The characteristics of the raters involved in a study, such as their level of experience in the administration of functional balance tests, may contribute to the generalisability of results to clinical situations. Only ten out of the 21 studies reported these rater characteristics. In the majority of these studies, physiotherapists with one to fifteen years experience administered the functional balance tests. Poor reporting of attrition rates further contributed low quality scores. This is important as the omission of outlying values, or scores of participants unable to complete the test may significantly affect results.

Strengths of both reliability and validity studies were the use of appropriate statistical methods and presentation of results. Most reliability studies reported Intra-class Correlation Coefficients (ICC) however did not report the model used. This test is considered appropriate for ordinal data where data sets followed similar and normal distributions. Concurrent validity was assessed using either the Pearson Product or Spearman Rank correlation coefficients. With the exception of three studies, all provided results presented clearly in tables and graphs to support the text.

Individual Functional Balance Test Results

Results of eligible studies have been presented with respect to each individual functional balance tool. Correlation coefficient values above 0.90 were considered excellent correlation, values above 0.75 good correlation, values above 0.50 moderate correlation, values above 0.25 fair correlation and values below 0.25 poor correlation.¹⁴ Tests reporting moderate reliability coefficients and moderate correlations with other established tests of functional balance were considered to be reliable and valid tests.

Berg Balance Scale (BBS)

The BBS consists of 14 functional items which are scored using an ordinal scale. It takes approximately 15 minutes to complete and requires a step, two chairs, a stopwatch, a 40cm ruler and minimal space.¹² Reliability of the BBS has been assessed with community dwelling older adults in three studies, and validity in five studies.^{1, 15-20} Good to excellent inter-rater reliability was reported for the BBS in all studies (ICC=0.88-0.98). Intra-rater reliability demonstrated greater variability (ICC=0.68-0.99).^{1,15,16} Excellent intra-rater reliability obtained by Berg et al (1989a) (ICC=0.99) may be attributed to the scoring of videotaped sessions.¹ This method reduces the confounding effects of change in patient condition, fatigue and learning, which may be present when tests are re-administered on separate occasions, as in Bennie et al (2003) (ICC=0.68).¹⁵

Fair to moderate correlations with the Timed Up and Go test (TUG) ($r=-0.47$ - -0.69) and tests of reach ($r=0.36$ - 0.48), and good to excellent correlations with the Fullerton Advanced Balance Scale (FAB) ($r_s=0.75$) and the Balance Screening Tool (BST) ($r_s=-0.87$ - -0.92) have been reported for the BBS.¹⁵⁻²⁰ Correlation with the FAB and BST may be expected as similar functional balance components are assessed. Conversely, lower correlation of the BBS which involves predominantly static balance tasks is expected with the TUG which tests dynamic balance. Results indicate that the BBS is a reliable and valid test of functional balance for older community dwelling adults. The BBS is often used as a gold standard to validate other functional balance measures.^{7,21}

Timed Up and Go Test (TUG)

The TUG involves timing a person as they rise from a chair, walk three meters, turn and return to the chair.¹⁷ The TUG takes approximately 1-2 minutes to complete and requires only a chair and a stopwatch.¹² Excellent inter-rater reliability has been reported for the TUG in three studies (ICC=0.98-0.99).^{15,17,22} Three studies reported excellent intra-rater reliability for the TUG (ICC=0.97-0.98).^{15, 17, 23} The moderate intra-rater reliability reported by Rockwood et al (2000) (ICC=-0.56), using a test-retest method, may be attributed to the average period of 112 ± 72.4 days between the test administrations and that each test was potentially administered in varied environments.²⁴

Six studies have reported fair to moderate correlations of the TUG with studies largely assessing static balance, such as the BBS ($r=-0.47$), Clinical Test of Sensory Interaction and Balance (CTSIB) ($r=-0.44$), Tinetti Performance Orientated Mobility Assessment (POMA) ($r=-0.55$) and Multi-Directional Reach Test (MDRT) ($r=-0.26$ - -0.42).^{15,17,20,25-27} While gait speed and the TUG both measure dynamic balance, the moderate correlations reported ($r=0.55$ - 0.66) may be attributed to the added tasks involved with the TUG of mobilizing from sit to stand and turning around.^{17,25}

Across all studies, test protocols of the TUG varied in relation to the amount of trials allowed, and how these were recorded. In Shumway-Cook et al (2000) the average of three trials was recorded, while in Rockwood et al (2000), following one practice trial, the best of two test trials was recorded.^{22,24} The use of walking aids to complete the TUG also varied between studies. These varied or poorly reported elements suggest the need for a standardized test protocol for the TUG, appropriate for use with community dwelling older people. Despite varied test protocols, reliability and validity of the TUG has been established with older community dwelling adults.

Functional Reach Test (FRT)

The FRT involves measuring the distance a person can reach forwards while standing. The only equipment required for the FRT is a yardstick fixed to the wall at shoulder height.⁷ Four studies have reported good to excellent reliability for the FRT in older community dwelling populations (ICC= 0.75-0.99).^{15,24,28,29} The concurrent scoring method used in Bennie et al (2003), which eliminates the confounding effects of fatigue and learning may have contributed to the higher inter-rater reliability value reported (ICC=0.97), compared with Giorgetti et al (1998) (ICC=0.75) with which a test-retest protocol was adopted.^{15,29}

Only fair associations of the FRT with the BBS ($r=0.42$) and POMA ($r=0.48$) have been established.^{15,25} However more dynamic tests (tandem walk, gait speed) have shown moderate correlations with the FRT ($r=0.67$ - 0.71).²⁸ Results indicate that while the FRT is reliable for use with community dwelling older adults, it has not yet been validated in this population with respect to another established measure of functional balance.

Tinetti Performance Orientated Mobility Assessment (POMA), balance subscale

The POMA consists of balance and gait subscales. Like the BBS, the balance subscale assesses a person's ability to perform numerous functional tasks. The inclusion of one task which involves an externally provided sternal nudge defines it from the BBS. The test takes 5-10 minutes to complete and requires a stopwatch, a chair and an object to pick up from the floor.^{30,31}

No studies investigating the reliability of the POMA in older community dwelling adults were retrieved. Concurrent validity has been established through moderate correlations with the TUG ($r=-0.55$) and fair correlations with the FRT and Rapid Step Test (RST).^{25,27} For use with community dwelling older adults, the POMA has not yet been established as a reliable test, or validated for use against a gold standard measure of functional balance such as the BBS.

Balance Screening Tool (BST)

The BST consists of six static and dynamic functional balance tasks and takes less than five minutes to administer.¹⁹ The only equipment required for the BST is a stopwatch.¹⁹ Being a newly developed test, the psychometric properties of the BST have only been assessed in one study with 30 community dwelling older adults.¹⁹ Excellent inter-rater ($r_s=0.89$) and intra-rater ($r_s=0.89$) reliability for total BST scores have been established.¹⁹ Chance corrected agreements of all individual items were consistently high for intra-rater reliability ($\kappa=0.64-1.00$).¹⁹ Moderate agreement for single limb stance items between raters may be attributed to differences in scoring interpretations and changes in balance performance between the two test administrations (up to one week).¹⁹

Mackintosh et al (2006) established good to excellent concurrent validity of the BST with the BBS on two separate occasions ($r_s=-0.87$, $r_s=-0.92$).¹⁹ While the tests assess similar balance components, the BST is designed to provide an efficient screen of balance abilities, whereas the BBS is a more comprehensive assessment of the sources of impairment. Results of this study indicate that the BST is a reliable and valid screening test of functional balance impairment in community dwelling older adults, however these results have not yet been confirmed by further research.

Multi-Directional Reach Test (MDRT) and Lateral Reach Test (LRT)

The incidence of older people falling in backwards and lateral directions prompted the development of the MDRT and the LRT as modifications of the original FRT.^{20,32} The MDRT involves a person reaching in four directions without losing their balance or taking a step.²⁰ The LRT involves a person reaching as far to the left and to the right without losing balance or taking a step.³² The only equipment required for each of these tests is a yardstick.^{20,32}

The MDRT has been assessed with community dwelling older adults in one study.²⁰ With the exception of forward reach (ICC=0.94), poor intra-rater reliability was reported for all directions of the MDRT (ICC=0.09).²⁰ Concurrent validity was assessed, reporting fair correlations of the MDRT with both the TUG ($r=-0.26$ - -0.44) and the BBS ($r=0.36$ - 0.48).²⁰ The LRT has been assessed with 18 community dwelling older adults by Brauer (1999) and a larger sample of 383 community dwelling older adults by Takahashi (2006).^{32,33} While excellent intra-rater reliability for the LRT (ICC=0.90-0.99) is reported, inter-rater reliability was not investigated in either of these studies.^{32, 33} Concurrent validity of the LRT has only been established by moderate correlation with an electronic measure of functional reach ($r=0.65$, $p<0.001$).³² Studies retrieved indicate that while intra-rater reliability of the LRT has been established, neither the LRT or MDRT has been established to be reliable both between raters and test administrations, and valid for use for community dwelling older adults.

Step Tests

The Step Test, Rapid Step Test (RST) and Four Square Step Test (FSST) have been developed as assessments of dynamic single limb stance.³⁴⁻³⁶ The Step Test involves stepping onto a block for 30 seconds using the left then the right leg.³⁴ The RST and FSST involve timing a person as they step in multiple directions.^{35,36} In the FSST participants are required to step over a low obstacle whereas in the RST, the stepping directions are repeated in a random order. All tests require a stop watch for administration while the Step Test also requires a 7.5cm step and the FSST requires four cones to step over.³⁴⁻³⁶

Only one study has assessed the FSST with community dwelling older adults.³⁶ Excellent inter-rater (ICC=0.99) and intra-rater (ICC=0.98) was established with 20 participants.³⁶ Concurrent validity was assessed with 81 community dwelling older adults.³⁶ Good correlation was reported with the TUG ($r_s=0.88$), while correlation with the FRT was fair ($r_s=-0.47$).³⁶ The Step Test only retrieved one study assessing psychometric properties with community dwelling older adults.³⁴ In a sample of 14 participants, excellent intra-rater reliability was reported (ICC=0.90).³⁴ Good correlations with measures of gait (Gait Speed: $r=0.82$; Stride Length: $r=0.83$) and a moderate correlation with the FRT ($r=0.68$ - 0.73) were reported.³⁴ Two studies which assessed the RST with community dwelling older adults were retrieved.^{27, 35} Lower intra-rater reliability of the RST reported by Cho et al (2004) (ICC=0.42) compared with Medell and Alexander (2000) (ICC=0.71-0.91), may be attributed to the homogeneity of the sample group, resulting from participants being recruited for a balance exercise program.¹⁴ Participants who were too able or too unstable were excluded.²⁷ Inter-rater reliability of the RST has not been assessed. Concurrent validity of the RST has been assessed with measures of gait (Tandem walk: $r=0.33$ - 0.61 ; TUG=0.33).^{27,35}

For use in older community dwelling populations, results indicate that neither the Step Test or RST have been validated as measures of functional balance, nor has inter-rater reliability been established. The FSST has been established as a reliable and valid test in this older population, however further investigation using a gold standard test is required to validate the FSST as a test of functional balance.

Fullerton Advanced Balance Scale (FAB)

The FAB scale consists of 10 functional tasks which assess both static and dynamic balance under varying sensory conditions.¹⁸ Like the BBS, a four point ordinal scale is used to score performance for each task. It is reported that the FAB requires less time than the BBS to administer, however requires more equipment, including a stopwatch, a ruler, a six inch high bench, masking tape, a square of foam and a metronome.¹⁸

To date, psychometric properties of the FAB scale have only been assessed as part of the original study which described its development.¹⁸ Excellent reliability ($r_s=0.93-1.00$) of the FAB scale has been assessed using video-taped assessments to reduce the impact of confounding variables such as learning, fatigue, and change in condition. Test-retest reliability in which the 31 community dwelling older adults completed two administrations at an interval of two to four days also indicated excellent reliability ($r_s=0.96$).¹⁸ Concurrent validity of the FAB scale was established through good correlation with BBS scores ($r_s=0.75$).¹⁸ Results of this initial study indicate that the FAB scale is a reliable and valid test of functional balance in older community dwelling adults, however these results have not yet been confirmed by further research.

Sensory Orientated Mobility Assessment Instrument (SOMAI)

The SOMAI is an ordinal scale which includes 10 functional balance tasks which are performed under two visual conditions; normal vision and peripherally occluded vision.³⁷ Equipment required includes a chair, two cushions to walk around and a visual conflict dome.³⁷ Our search retrieved only one study which assessed the SOMAI in 27 community dwelling older adults.³⁷ This study did not address reliability. Concurrent validity was assessed through comparison with the Clinical Test of Sensory Interaction and Balance (CTSIB), with only fair correlations reported ($r_s=0.21-0.53$).³⁷ Results indicate that the reliability and validity of the SOMAI has not yet been established for use as a functional balance test with community dwelling older adults.

Clinical Test of Sensory Interaction and Balance (CTSIB)

The CTSIB assesses the reliance on the various sensory systems in maintaining balance.³⁸ Equipment required include a stopwatch, a visual-conflict dome and a square of foam approximately 8cm thick.³⁸ Only one study was retrieved that assessed reliability of the CTSIB in older community dwelling adults.²⁶ Good test-retest reliability ($r_s=0.75$) was reported using a small sample ($n=10$). The concurrent validity of the CTSIB has not been clearly established in a group of older community dwelling adults. Poor to moderate correlations of the CTSIB with the TUG and the SOMAI have been reported.^{26, 37} Results of the studies retrieved do not establish the CTSIB as a reliable or valid functional measure for use with older community dwelling populations.

Tests returning no studies assessing community dwelling older adults

No studies assessing the reliability and concurrent validity with community dwelling older adults were returned for the Hierarchical Assessment of Balance and Mobility (HABAM), Physical Mobility Scale (PMS), Short-Form Berg Balance Scale (SF-BBS) or Elderly Mobility Scale (EMS). Most of these tests were not developed for use in this population.

DISCUSSION

Results of this systematic review suggest that while numerous tests have been developed to measure functional balance, the reliability and validity of a number of the identified tests have not yet been established for use with older community dwelling people (Table 3).

Table 3: Summary of Functional Balance Test Properties

Test	Reliable	Valid	Practical
BBS	Yes	Yes	Yes
TUG	Yes	Yes	Yes
FRT	Yes	NE	Yes
POMA	NE	NT	Yes
BST	Yes	Yes	Yes
MDRT	NE	NE	Yes
LRT	NE	NT	Yes
Step Test	NE	NT	Yes
RST	NE	NT	Yes
FSST	Yes	NT	Yes
FAB	Yes	Yes	Yes
SOMAI	NT	NT	Yes
CTSIB	NE	NT	Yes
HABAM	NT	NT	Yes
PMS	NT	NT	Yes
SF-BBS	NT	NT	Yes
EMS	NT	NT	Yes

Adequate levels of reliability and concurrent validity not established (NE), Concurrent validity with gold standard functional balance measure, or reliability (inter or intra-rater) not tested (NT), Berg Balance Scale (BBS), Timed Up and Go (TUG), Functional Reach Test (FRT), Performance Orientated Mobility Assessment (POMA), Balance Screening Tool (BST), Multi-Directional Reach Test (MDRT), Lateral Reach Test (LRT), Rapid Step Test (RST), Four Square Step Test (FSST), Fullerton Advanced Balance Scale (FAB), Sensory Orientated Mobility Assessment Instrument (SOMAI), Clinical Test of Sensory Interaction and Balance (CTSIB), Hierarchical Assessment of Balance and Mobility (HABAM), Physical Mobility Scale (PMS), Short Form Berg Balance Scale (SF-BBS), Elderly Mobility Scale (EMS).

For most of the functional balance tests which retrieved studies involving community dwelling older adults, reliability had been assessed to some extent. Intra-rater reliability was most commonly assessed in tests which take minimal time to administer or involve numerous trials (FRT, LRT, RST, Step Test). While good reliability has been reported between successive trials of these tests, further investigation with tests administered at greater time intervals are required to minimize the effects of learning and memory. Further research is also required to establish inter-rater reliability for the SOMAI, RST, Step Test, LRT, MDRT and POMA in community dwelling older populations.

While concurrent validity had been assessed for most of the investigated tests, this had only been established using a 'gold standard' functional balance measure such as the BBS, with the TUG, BST and FAB. Concurrent validity of the FSST with the TUG was reported, establishing the FSST as a measure of dynamic balance. Furthermore concurrent validity of the POMA had only been demonstrated for community dwelling older adults with respect to the TUG. Further investigation in community dwelling older populations and using 'gold standard' criterion measures is required to establish tests as measures of functional balance.

Excellent reliability and concurrent validity has however been established for use of the BBS, TUG, BST and FAB scale for older community dwelling people. A limitation of the TUG and BST is that they only provide information on a few aspects of balance. The TUG assesses dynamic balance and was initially designed as a test of basic mobility for older people.¹⁷ The BST was developed to provide an efficient screen of static and dynamic balance abilities.¹⁹ Consequently, while the BST removes visual input for one task, the TUG, and BST provide little information about the source of a balance problem. Furthermore, meaningful scores can not be recorded for the TUG where participants are physically unable to rise from a chair or walk independently.²⁴ Consequently, while the TUG and BST provide an efficient, practical, reliable and valid screen of functional balance for community dwelling older adults, they do not have the depth of information to discriminate between the various sources of impairment.

Despite good reliability and validity reported for the BBS, the presence of a ceiling effect when used with community dwelling older adults, limits the use of this scale to detect balance impairments.³⁹ With few items which test dynamic balance, the BBS may not provide a great enough challenge to older adults who live independently.¹⁸ Furthermore, if community dwelling older adults score highly initially on the BBS, its use as an outcome measure is compromised.¹⁸ A further limitation is that the BBS has one of the longest administration times of functional balance tests.¹² Difficulties interpreting scoring criteria, resulting in inconsistencies with score allocation have also been reported.^{16, 40} Consequently, modifications of the BBS are required for use with older community dwelling people.

The FAB scale is a newly developed functional balance test which has a shorter administration time than the BBS (10-12 minutes) and includes more challenging tasks that test dynamic balance and challenge sensory components (tandem walk, two-footed jump, walk with head turns and standing on foam with eyes closed).¹⁸ To date, psychometric properties of the FAB scale have only been assessed in the original study which described its development.¹⁸ Interpretation of results is limited, as descriptive statistics indicating the spread of scores and presence of floor or ceiling effects, are not presented. Furthermore, the use of video-taped performances to assess reliability is not consistent with clinical application, limiting the generalisability of results. While preliminary results are promising, further research of higher methodological quality, is required to establish the FAB scale as a reliable and valid measure of functional balance.

The search strategy used was limited by time and resource constraints resulting in not seeking unpublished studies and excluding studies which were not published in English. Possible bias may exist as only one reviewer was involved in the search and selection of studies.

Clinical and Research Implications

Of the numerous functional balance tests available, only the BBS, TUG, BST and FAB scale have established reliability and validity with community dwelling older people. The TUG and BST provide an efficient screen of a persons balance abilities, however do not offer enough detail to determine the source of impairment. While being the most rigorously developed functional balance test, the BBS may be more appropriate for use for frail older adults due to its limited assessment of dynamic balance. The FAB scale may be more applicable for older people living independently, however further research is required to establish its reliability and validity. Consequently, there remains a need for a valid and reliable functional balance test that appropriately challenges balance of older community dwelling people.

A further recommendation that arises from this systematic review relates to the low overall quality of reliability and validity studies. There is a need for a critical appraisal tool which is more relevant to reliability and validity studies of outcome measures. Furthermore, development of a set of guidelines describing details that should be included for such studies may help to improve quality of reporting in studies.

REFERENCES

1. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada*. 1989a; 41(6): 304-11.
2. Huxham F, Goldie PA, Patla, AE. Theoretical considerations in balance assessment. *Australian Journal of Physiotherapy*. 2001; 47: 89-100.
3. Berg K. Balance and its measure in the elderly: a review. *Physiotherapy Canada*. 1989b; 41(5): 240-5.
4. Alexander NB. Postural Control in Older Adults. *Journal of the American Geriatrics Society*. 1994; 42(1): 93-108.
5. Patla A, Frank JS, Winter DA. Balance control in the elderly: implications for clinical assessment and rehabilitation. *Canadian Journal of Public Health*. 1992; 83 Suppl 2: S29-S33.
6. Lord S, Dayhew J. Visual risk factors for falls in older people. *Journal of the American Geriatrics Society*. 2001; 49(5): 508-15.
7. Tyson S, DeSouza L. A systematic review of methods to measure balance and walking post stroke. *Physical Therapy Reviews*. 2002; 7: 17-86.
8. Forlander D, Bohannon RW. Rivermead mobility index: a brief review of research to date. *Clinical Rehabilitation*. 1999; 13: 97-100.
9. Horak F. Clinical measurement of postural control in adults. *Physical Therapy*. 1987; 67(12): 1881-5.
10. Berg K, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural control in an elderly population. *Archives of Physical Medicine and Rehabilitation*. 1992; 73: 1073-80
11. Farrell M. Using functional assessment and screening tools with frail older adults. *Topics in Geriatric Rehabilitation*. 2004; 20(1): 14-20.

12. Whitney S, Poole JL, Cass SP. A review of balance instruments for older adults. *The American Journal of Occupational Therapy*. 1998; 52(8): 666-71.
13. Cochrane Methods Working Group on Diagnostic and Screening Tests cited in, *How to review the evidence: systematic identification and review of the scientific literature*. National Health and Medical Research Council. 1999: 62.
14. Portney L, Watkins MP. *Foundations of Clinical Research*. Prentice-Hall Health; 2000
15. Bennie S, Bruner K, Dizon A, Fritz H, Goodman B, Peterson S. Measurements of balance: comparison of the timed "up and go" test and the functional reach test with the berg balance scale. *Journal of Physical Therapy Science*. 2003; 15(2):93-7.
16. Bogle-Thorban L., Newton RA. Use of the berg balance test to predict falls in elderly persons. *Physical Therapy*. 1996; 76(6): 576-85.
17. Podsiadlo D, Richardson S. "The timed "up & go", a test of basic functional mobility for frail elderly persons." *Journal of the American Geriatrics Society*. 1991; 39: 142-8.
18. Rose D, Lucchese N, Wiersma LD. Development of a multidimensional balance scale for use with functionally independent older adults. *Archives of Physical Medicine and Rehabilitation*. 2006; 87: 1478-85.
19. Mackintosh S, Datson N, Fryer C. A balance screening tool for older people: reliability and validity. *International Journal of Therapy and Rehabilitation*. 2006; 13(12): 558-61.
20. Newton R. Validity of the multi-directional reach test: a practical measure for limits of stability in older adults. *Journal of Gerontology: Medical Sciences*. 2001; 56A(4): M248-M252.
21. Whitney S, Wrisley D, Furman J. Concurrent validity of the berg balance scale and the dynamic gait index in people with vestibular dysfunction. *Physiotherapy Research International*. 2003; 8(4): 178-86.
22. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the timed up and go test. *Physical Therapy*. 2000; 80(9): 896-903.
23. Steffen TM, Hacker TA, Mollinger L. Age- gender- related test performance in community dwelling elderly people. Six-minute walk test, berg balance scale, timed up and go test and gait speeds. *Physical Therapy*. 2002; 82:128-37
24. Rockwood K, Awalt E, Carver D, MacKnight C. Feasibility and measurement properties of the functional reach and the timed up and go tests in the canadian study of health and aging. *Journal of gerontology: Medical Sciences*. 2000; 55A(2): M70-M73.
25. Lin MR, Hwang HF, Hu MH, Wu HDI, Wang YW, Huang FC. Psychometric Comparisons of the Timed Up and Go, One-Leg Stand, Functional Reach, and Tinetti Balance Measures in Community-Dwelling Older People. *Journal of the American Geriatrics Society*. 2004; 52(8): 1343-48.
26. Anacker S, DiFabio RP. Influence of sensory inputs on standing balance in community-dwelling elders with a recent falls history. *Physical Therapy*. 1992; 72(8): 575-84.
27. Cho BL, Scarpace D. Test of stepping as indicators of mobility, balance and falls-risk in balance impaired older adults. *Journal of the American Geriatrics Society*. 2004; 52(7):1168-73.
28. Weiner DK, Duncan PW, Chandler J, Studenski SA. Functional reach: a marker of physical frailty. *Journal of the American Geriatric Society*. 1992; 40(3) 203-7.
29. Giorgetti MM, Harris BA, Jette A. Reliability of clinical balance outcome measures in the elderly. *Physical Therapy Research International*. 1998; 3(4): 274-83.
30. Nakamura M, Holm MB, Wilson A. Measures of balance and fear of falling in the elderly: a review. *Physical and Occupational Therapy in Geriatrics*. 1998; 15(4): 17-33.
31. Tinetti M. Performance-orientated assessment of mobility problems in elderly patients. *Journal of the American Geriatrics Society*. 1998; 34(2): 119-26.
32. Brauer S, Burns Y, Galley P. Lateral reach: a clinical measure of medio-lateral postural stability." *Physiotherapy Research International*. 1999; 4(2): 81-8.
33. Takahashi T, Ishida K, Yamamoto H, Takata J, Nishinaga M, Doi Y, et al. Modification of the functional reach test: analysis of lateral and anterior functional reach in community-dwelling older people. *Archives of Gerontology and Geriatrics*. 2006; 42: 167-73.
34. Hill K, Bernhardt J, McGann AM, Maltese D, Berkovits D. A new test of dynamic standing balance for stroke patients: reliability, validity and comparison with healthy elderly. *Physiotherapy Canada*. 1996; Fall: 257-62.
35. Medell J, Alexander NB. A clinical measure of maximal and rapid stepping in older women. *Journal of Gerontology: Medical Sciences*. 2000; 55A(8): M429-M433.
36. Dite W, Temple VA A clinical test of stepping and change of direction to identify multiple falling older adults." *Archives of Physical Medicine and Rehabilitation*. 2002; 83: 1566-71.
37. Tang P, Moore S, Woollacott MH. Correlation between two clinical balance measures in older adults: functional mobility and sensory organisation test. *Journal of gerontology: Medical Sciences*. 1998; 53A(2): 140-6.

38. Shumway-Cook A, Horak F. Assessing the influence of sensory interaction on balance. *Physical Therapy*. 1986; 66(10): 1548-50.
39. Newton R. Balance screening of an inner city older adult population. *Archives of Physical Medicine and Rehabilitation*. 1997. 78: 587-91.
40. Kornetti D, Fritz SL, Chiu YP, Light KE, Velozo CA. Rating scale analysis of the berg balance scale. *Archives of Physical Medicine and Rehabilitation*. 2004; 85: 1128-35.