UNDERSTANDING APPROACHES TO BALANCE ASSESSMENT IN PHYSICAL THERAPY PRACTICE WITH VARIOUS AVAILABLE AND VALID SCALES

Shubha Akadas *1, Pramod Kshirasagar 2, Ganukumar Wadhokar 3.

*1 MPT Student, SDM College of Physiotherapy, Sattur, Dharwad – 580009, Karnataka, India.
2 Professor, SDM College of Physiotherapy, Sattur, Dharwad – 580009, Karnataka, India.
3 Lecturer, SDM College of Physiotherapy, Sattur, Dharwad – 580009, Karnataka, India.

ABSTRACT

Background: For allied health professionals, especially physiotherapists wishing to assess the functional balance of individuals living in the community, the vast number of functional balance tests available makes it difficult to decide which assessment is most appropriate for each target population.

Objective: To identify their reliability, concurrent validity and clinical implications of functional balance tests for paediatric, adult and geriatric age group focusing mainly on post surgery or post operative individuals.

Methods: A systematic review of published literature relevant to 10 functional balance tests was undertaken. The 10 functional balance tests were identified by a preliminary literature search and through consultation with an expert in the field of functional balance assessment. Among all the scales, the Berg Balance Scale and the Timed Up and Go Test have been most rigorously tested, in acute care set up the Morse Fall Scale was most reliable and in the paediatric age group the valid and most tested scale was WeeFIM.

Conclusion: The Berg Balance Scale, Timed Up and Go Test and Morse Fall scale for inpatient or hospitalised individuals have published reliability and validity. Further testing of other functional balance tests is required to establish their reliability and validity in the target population of post operative individuals.

KEY WORDS: Balance, Berg Balance Scale, Balance assessment, Morse Fall Scale, Timed Up And Go Test, WeeFIM, John Hopkins Fall Risk Assessment Tool.

BACKGROUND

The performance of all daily living activities requires good balance control at rest or when moving from one position to another. Maintenance of balance needs the coordination of sensory, neural and musculoskeletal systems [1,2]. Many of these mentioned systems undergo deterioration with age, diseased conditions, surgical procedures, medications and prolonged bedridden state [3,4]. This may have an effect on balance, restrict safe mobility, increase the likelihood of a fall and adversely affect quality of life [3,5]. Therefore, the assessment of balance is important to direct required interventions to improve balance and to monitor changes in balance in patients over time [6].

Various approaches to evaluate balance, been developed [3,7]. Questionnaires like Rivermead Mobility Index and the Activities Specific Balance Scale provide self-report information
regarding functional position [8]. Laboratory instrument such as computerized force platforms gives correct measurements of postural sway [3,9]. Functional performance-based tests such as the Berg Balance Scale and the Timed Up and Go Test is defined as tests which evaluate a person doing balance or walking tasks [7,10,11]. Functional balance tests have an advantage as they carry advantages such as their practicality for assessment in a variety of settings because of their cost effectiveness, lack of complex equipment and time efficiency [10,11].

To substantiate the clinical need of a functional balance test, it should be reputable to be reliable and valid when it has to be implemented in the target population [9]. A reliable test should provide similar results when performed on the same person and under the same conditions different occasions [3]. Results of a test with equivalent validity should correlate with results of another standard test when done on the same person [3]. For functional balance tests, the Berg Balance Scale has been found to be most valid and reliable, “gold standard” scale, and will be considered the gold standard for this systematic review [7,19].

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 for the given target population. To date, no reviews have been conducted to assess the psychometric properties and practical application of functional balance tests in community dwelling people especially for both pediatric, adult, geriatric population as well as hospitalised individuals. Such review would help the therapists and researchers to make decisions about which tool may be most appropriate when assessing functional balance of given target population.

Therefore, the aims of the following systematic review were; to determine the reliability, concurrent validity and the reported common sense for use of functional balance measures when performed with community dwelling adults, pediatric population, hospitalised, and post operative individuals.

**METHODOLOGY**

Studies were identified using a systematic search of computer databases. CINAHL, MEDLINE, Amed (Allied and Complementary Medicine) and PubMed were searched using combinations of the above mentioned keywords. The third keyword, ‘balance assessment’ was used where more than 40 results were obtained by the initial search using the first two terms. The Cochrane library, PEDro, and the Joanna Briggs Institute record were also found that these sources not yet given any further relevant literature.

**RESULT**

The initial list of results was screened for eligibility using abstracts and titles, leaving around 16 studies for review. Most of the studies were excluded from the study because they did not investigate either reliability or concurrent validity. Following 10 scales were shortlisted considering the assessment in different age groups and various set ups.

**Berg Balance Scale:** The BBS consists of 14 functional items which are scored using an ordinal scale. It requires 15 minutes to complete and requires a step, two chairs, a stopwatch, a 40cm ruler and minimal space. (12) Good to excellent inter-rater reliability was reported for the BBS in all studies (ICC=0.88-0.98). Intra-rater reliability verified greater variability (ICC=0.68-0.99) (1) (13) (14). Excellent intra-rater reliability obtained by Berg et al (1989a) (ICC=0.99) may be attributed to the scoring of videotaped sessions (1). BBS consists of 14 items that are scored on a scale of 0 to 4. A score of 0 is given if the participant is not able to complete the task, and a score of 4 is given if the participant is completed the task based on the criteria that has been assigned to it. The total maximum score on the test that can be achieved is 56. The items include simple mobility tasks (e.g. transfers, standing unsupported and sit-to-stand) and more difficult tasks (e.g., tandem standing, turning 360°, and single-leg stance).

Fair to moderate correlations of BBS with the Timed Up and Go test (TUG) and tests of reach and good to excellent correlations with the Fulerton Advanced Balance Scale (FAB) and the Balance Screening Tool (BST) have been reported (13-18). Correlation of BBS with the FAB and BST may be expected as similar functional
balance criteria are assessed. Conversely, lower correlation of the BBS which involves mainly the static balance tasks is expected with the TUG test which tests dynamic balance. Results indicate that the BBS is a reliable and valid test of functional balance for adult as well as geriatric age group but there are no studies which have shown reliability with balance assessment in post operative cases or post surgery individuals. The BBS is often used as a gold standard to validate other functional balance measures [7,19].

**Timed Up and Go Test (TUG):** The TUG involves timing the task of a person as they rise from a chair, walk three meters, turn and return to the chair [15]. The TUG takes around 1-2 minutes to complete and the only requirement is a chair and a stopwatch [12]. Excellent inter-rater reliability has been reported for the TUG in three studies [13,15,20]. Three studies reported excellent intra-rater reliability for the TUG [13,15,21]. The moderate intra-rater reliability reported by Rockwood et al (2000) using a test-retest method, may be attributed to the average period of 112+72.4 days between the administered test and that each test was potentially administered in various environments [22].

Six studies have reported fair to moderate correlations of the TUG test with studies mainly assessing the static balance, such as the BBS, Clinical Test of Sensory Interaction and Balance (CTSIB), Tinetti Performance Orientated Mobility Assessment (POMA) and the Multi-Directional Reach Test (MDRT) [13,15,18,23-25]. While gait speed and the TUG both measure dynamic balance, the moderate correlations that have been found may be attributed to the added tasks involved with the TUG of ambulation from sit to stand and turning around [15,23].

Across all studies, test protocols of the TUG varied in relation to the amount of trials allowed, and how these were recorded [20,22]. 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. 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. Despite varied test protocols, reliability and validity of the TUG has been established with older community dwelling adults, but not with diseased individuals.

**Functional Reach Test (FRT):** The FRT involves measurement of the distance a person can reach forward while standing. The equipment required for the FRT is a yardstick fixed to the wall at shoulder level (7). The 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 compared with Giorgetti et al (1998) with which a test-retest protocol was adopted [13, 26].

Association of the FRT with the BBS and POMA have been established [13,23]. However more dynamic tests (tandem walk, gait speed) have shown moderate correlations with the FRT. Results indicate that while the FRT is reliable for use with adults and geriatric age group, it has not yet been validated in post operative individuals or population with respect to another established measure of functional balance.

**Tinetti Performance Orientated Mobility Assessment (POMA), balance subscale:** The POMA comprises of balance and gait subscales. Like the BBS, the balance subscale assesses a person’s ability to perform various functional tasks. The inclusion of the task which involves an externally provided nudge at the sternum 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 [27,28].

Concurrent validity has been established through moderate correlations with the TUG and fair correlations with the FRT and Rapid Step Test (RST) [23,25]. 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 comprises of six static and dynamic functional balance tasks that takes less than five minutes to administer [17]. The only equipment required for the BST is a stopwatch [17]. As the BST is a newly developed test, the reliability and validity of the BST have only been assessed in one study [17]. Excellent inter-rater and intra-rater reliability for
total BST scores have been established. 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) [17]. Mackintosh et al (2006) established good to excellent concurrent validity of the BST with the BBS on two separate occasions [17]. While both the tests assess similar balance components, the BST is designed so as 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 concludes that the BST is a reliable and valid screening test for functional balance impairment in community dwelling older adults, however these results have not yet been confirmed by further research.

**Fullerton Advanced Balance Scale (FAB):** The FAB scale consists of 10 functional tasks which are for assessing both static and dynamic balance under varying sensory conditions [16]. 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 in terms of administration, however requires more equipment, which includes, a ruler, around six inch high bench, a stopwatch, masking tape, a square of foam and a metronome [16].

Till date, the psychometric properties of FAB scale have been only assessed as part of the original study which describes its development [16]. Excellent reliability of the FAB scale has been assessed with the help of video-taped evaluation to reduce the influence of confounding variables such as learning, fatigue, and change in condition. Concurrent validity of the FAB scale was established through good correlation with BBS scores [16]. Results of this initial study indicate that the FAB scale is a reliable and valid test of functional balance; however these results have not yet been confirmed by further research.

**Morse Fall Scale:** The MFS was developed by Morse (1986). It consists of six variables which are as follows; history of falling (0 and 15 points), secondary disease (0 and 15 points), ambulatory aid (0, 15, and 30 points), intravenous therapy or heparin lock (0 and 20 points), gait (0, 10, and 20 points), and mental status (0 and 15 points) (Morse, Tylko, & Dixon, 1989; E. A. Kim et al., 2007). The total score can range varying from 0 to 125 points. A total score below 25 points is classified as the low risk group, a score that falls between 25 and 30 points is classified as the intermediate risk group, and any score above 51 points is regarded as high risk (Morse et al., 1989). The intertester reliability at the time of tool development was 96%. This tool is used in assessing the balance of individuals in acute care hospitals and developed for elderly patients in long term care hospitals and community or home set ups and not much information is available to support validity and reliability of this scales for use in post operative individuals.

**Bobath Memorial Hospital Fall Risk Assessment Scale (BMFRAS):** The BMFRAS was developed for elderly hospitalized patients at the Bobath Memorial Hospital in Korea in the year 2003. It consists of eight categories in which the assessment of following things are done; age (0–3 points), history of falling (0–3 points), gait (0–8 points), cognition (0–8 points), communication (0–3 points), number of risk factors (sleep disturbance, urination problems, diarrhoea, visual disturbance, dizziness, depression, agitation, and anxiety) (0–3 points), number of related diseases (stroke, hypertension, hypotension, dementia, parkinsonism, osteoporosis, musculoskeletal disease, and seizure) (0–3 points), and number of medication (antihypertensive, diuretics, digitalis, sedatives, antidepressants, antipsychotics, anti parkinson drugs, and anticonvulsants) (0–3 points) .Total score above 15 points represent a high risk, and it is recommended that patients with total scores above 20 points need intensive monitoring. This study was also used in acute care hospitals for elderly. Not much information is available to support validity and reliability of these scales for use in post operative patients.

**Johns Hopkins Hospital Fall Risk Assessment Tool (JHFRAT):** The JHFRAT was developed by the Johns Hopkins Hospital in 2005, and was supplemented in 2007 with opinions experts in clinical practice. JHFRAT consists of eight main evaluation areas of fall risk factor categories
which are as follows: age (0–3 points), fall history (0–5 points), elimination (0–4 points), medications (0–7 points), patient care equipment if used (0–3 points), mobility (0–2 points), and cognition (0–4 points). A total score between 6 and 13 points represents an intermediate fall risk, and a total score above 13 points indicates a high fall risk. Fall risk assessment is performed during the first eight hours of hospitalization, once in a day, and when there is any change in a patient’s condition or risk condition (Poe et al., 2007). There is no valid study using the scale in post operative patients.

**WeeFIM:** The Functional Independence Measure (WeeFIM) for children is a simple scale for assessing independence across 3 aspects in American children. WeeFIM was based on a conceptual framework by the World Health Organization (1980) of pathology, impairment, disability and handicap, and the “burden of care.”(30) WeeFIM is used in assessment of functional independence in children aged 6 months to 7 years. It can be used for children with developmental disabilities aged 6 months to 21 years. WeeFIM is an 18-item, 7-level ordinal scale instrument that measures a child’s consistent performance in essential daily functional skills. Three main categories (self-care, mobility, and cognition) are assessed by interviewing or by observing a child’s performance of a task to criterion standards. WeeFIM is categorized into 2 main functional streams: “Dependent” (i.e., requires helper: scores 1–5) and “Independent” (i.e., requires no helper: scores 6–7). Scores 1 (total assistance) and 2 (maximal assistance) belongs to the category of “Complete Dependence”. Scores 3 (moderate assistance), 4 (minimal contact assistance), and 5 (supervision or set-up) belongs to the “Modified Dependence” category. Scores 6 (modified independence) and 7 (complete independence) belong to the “Independent” category. The WeeFIM is a 7-level criterion-specific ordinal scale. Level 7 is where there is no requirement of assistance for the child and the child completes the task independently without requiring a device. During the task, there is no concern about safety or taking an inordinate amount of time. Level 6 depicts modified independence and includes use of an assistive device or not completing the task in a timely or safe manner. The administration of the WeeFIM requires less time and provides direct information which is relevant for evaluating functional outcomes for children with disabilities. The advantages of WeeFIM include its conciseness (simple scoring of 1–7), extensive (covers all developmental aspects), uniformity (certified training), and discipline-free requirements (can be administered by trained health, developmental, or educational professionals). The reliability and validity of this scale have been studied in healthy as well as disabled children.

**DISCUSSION**

This systematic review suggests that while various 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 in post operative patients. 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. Excellent reliability and concurrent validity has however been established for use of the BBS, TUG, BST and FAB scale for older community dwelling people (19). A limitation of the TUG and BST is that they only provide information on a few aspects of balance. The BST was developed to provide an efficient screen of static and dynamic balance abilities. 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 cannot be recorded for the TUG where participants are physically unable to rise from a chair or walk independently [22].

Despite good reliability and validity reported for the BBS, the ceiling effect was observed when used with community dwelling older adults and post operative individuals which limits the use of this scale to detect balance impairments [28]. With few items which test dynamic balance, the BBS may not provide a great enough challenge to elderly population who are independent [16]. Furthermore, if elderly or geriatric individual score high initially on the BBS, its use as an outcome measure is compromised [16]. Another
limitation is that the BBS has one of the longest administration times of functional balance tests. Consequently, modifications of the BBS are required for use with geriatric or elderly individuals and who have undergone a surgery and balance assessment is required for them for further physiotherapy rehabilitation.

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 both, dynamic balance and sensory components (tandem walking, two footed jump, walk with head turns and standing on foam with eyes closed) [16]. To date, the reliability and validity of the FAB scale have only been assessed in the original study which described its development [16]. Interpretation of results is limited, as descriptive statistics indicates that the spread of scores and presence of floor or ceiling effects, are not present. Furthermore, the use of video-taped performances to assess reliability is not consistent with clinical application. While preliminary results are promising, further study of higher methodological quality, is required to establish the FAB scale as a reliable and valid measure of functional balance. The search strategy used in the study was limited by few components like time and resource constraints resulting in not seeking unpublished studies and excluding studies which were not published in English. Possible bias may exist as there was involvement of only one reviewer in the search and selection of studies.

Among the three mentioned scales for assessing balance in hospitalised individuals; BMFRAS, JHFRAT and MFS, the highest predictive validity for identifying patients at high risk for falls was achieved by the MFS.

**Clinical and Research Implications:** Of the numerous functional balance tests available, only the BBS, BST, TUG and FAB scale have shown to have established reliability and validity with elderly population but not the post operative patients. The TUG and BST provide an efficient screen of a person balance abilities, but they do not seem to offer enough detail to find the source of impairment. The BBS is the most rigorously developed functional balance test and most appropriate scale for assessment, but it lacks in the field of limited assessment of dynamic balance.

The FAB scale may be more applicable for elderly population living independently, however further research is required to establish its psychometric properties. Consequently, there remains a need for a valid and reliable functional balance assessment scale that appropriately challenges the balance of both adults and also geriatric age group. And also there is a lack of studies performed in the field of post operative patient balance assessment. Therefore the gold standard scale for balance assessment remains BBS and in the acute care hospital is MFS and for the paediatric age group the most valid and reliable scale is the WeeFIM.

A further recommendation that arises from this systematic review relates to the low overall quality of reliability and validity studies. 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.

**Conflicts of interest:** None

**REFERENCES**


