# Clinical Rehabilitation

http://cre.sagepub.com


## How to measure balance in clinical practice. A systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions


SF Tyson and LA Connell



*Clin Rehabil* 2009; 23; 824 originally published online Aug 5, 2009;
DOI: 10.1177/0269215509335018

The online version of this article can be found at:
http://cre.sagepub.com/cgi/content/abstract/23/9/824

Published by:
SAGE
http://www.sagepublications.com

*Clinical Rehabilitation* 2009; **23:** 824–840

# How to measure balance in clinical practice. A systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions

**SF Tyson** Centre for Rehabilitation and Human Performance Research, University of Salford and **LA Connell** Department of Physiotherapy Education, Clinical Sciences Building, University of Nottingham, on behalf of the Greater Manchester Outcome Measures project steering group, UK

Received 15th October 2008; returned for revisions 2nd March 2009; revised manuscript accepted 2nd March 2009.

**Objective**: To identify psychometrically robust and clinically feasible measurement tools of balance activity in people with neurological conditions to recommend for use in clinical practice.
**Data sources**: MEDLINE, CINAHL, EMBASE, PEDro and AMED.
**Review methods**: Independent reviewers selected and extracted data from articles that assessed the reliability, validity, sensitivity to change and clinical utility of measures of balance activity in adult neurological conditions. Quality assessment was based on Jorstad *et al*. Measures with 'good' psychometrics and ≥9/10 clinical utility scores were recommended.
**Results**: Nineteen measurement tools were selected. Of these, the Brunel Balance Assessment, Berg Balance Scale, Trunk Impairment Scale, arm raise and forward reach tests in sitting and standing, weight shift, step/tap and step-up tests reached the required standards and are usable in clinical practice. The Brunel Balance Assessment and its associated functional performance tests have the additional advantages of being a hierarchical scale with established lack of redundancy.
**Conclusion**: The measurement tools identified above are psychometrically robust and feasible to use in clinical practice. Future objective measure development should consider the theoretical construct of the measure, the minimal detectable change and use in clinical populations other than stroke.

## Introduction

Using objective measurement tools is an accepted part of evidence-based practice, to the extent that their use is included in clinical guidelines and core standards of practice.[1–3] Despite this, adoption into clinical practice is haphazard. A huge range of objective measurement tools have been developed but most are designed for research purposes and many are impractical, irrelevant or inadequately developed for generalized use with clinical patient populations.[4–8] There is little consistency about which measurement tools are used and health care professionals report that lack of resources

Address for correspondence: SF Tyson, Centre for Rehabilitation and Human Performance Research, University of Salford, Frederick Road Campus, Salford M6 6PU, UK.
e-mail: s.tyson@salford.ac.uk



10.1177/0269215509335018

(time, information, money and expertise) makes it difficult to identify suitable measurement tools and to learn how to use them.[4–6,8] This experience was shared by neurological physiotherapists in Greater Manchester. Since 2003, they have worked together to develop a comprehensive assessment process using robust, objective measurement tools that could be implemented across the region to enable equity of treatment planning and allocation.[9] Early work identified the domains that needed to be measured[9] and then the measurement tools that measure each domain were systematically reviewed. This paper reports the first of these systematic reviews and assesses measures of balance.

'Balance' is defined as 'the ability to maintain a position within the limits of stability or base of support'[9–11] and so would be classified by the International Classification of Functioning, Disability and Health[12] as an activity. There are also a number of measurement tools, such as measures of weight distribution or postural sway that measure balance impairments (defined as a problem in body function, system or structure causing a significant deviation or loss[12]). A review of these measurement tools will be reported separately. There is currently no recognized gold standard measure for balance activity in patients with neurological conditions. Previous reviews have focused on a single measurement tool,[13–18] or a limited selection of measurement tools.[19–21] While others have only considered a single condition[22] or only the psychometric properties. Information on clinical utility is rarely available,[23] despite being a factor that significantly impacts on use in clinical practice.[24] The current review will evaluate the psychometric properties and clinical utility of all available measures of balance activity in all neurological conditions.

## Method

### Study identification and selection

Electronic databases (MEDLINE, CINAHL, EMBASE, PEDro and AMED) were searched from their earliest date to June 2008 using the following keywords: 'assessment or measurement or outcome measure or measurement tool' and 'balance or equilibrium or postural control or posture or sit or sitting or seat or seating or seated' and 'stroke or cerebrovascular accident or hemiplegia or multiple sclerosis or Parkinson's Disease or Parkinsonism or head injury or brain injury or Motor Neurone Disease or Guillain Barre syndrome'.

All searches were limited to English language and human adults. We excluded articles that involved people with non-cerebral lesions (such as spinal cord injuries or peripheral nerve lesions). We also excluded the following from the analysis:

- articles that only measured psychometric properties other than those listed in the method section below (such as predictive validity or sensitivity/responsiveness using methods other than the minimal detectable change);
- measurement tools in which only one psychometric property had been assessed and so clearly did not have sufficient information to recommend for clinical use;
- measurement tools that included balance activity as part of wider assessment of general motor or mobility abilities in which data that was specific to balance abilities could not be extracted; and
- any instrumented measurement tool/device that had no information about how the device could be obtained or was clearly not commercially available.

### Data extraction

Two reviewers (ST and LC) independently assessed the titles and abstracts of articles identified by the search. The selected articles were then given to volunteers who extracted the necessary data. They were neurological physiotherapists working in National Health Service (NHS) Trusts in the North West of England identified by previous involvement in the project,[9] other research projects undertaken by the authors and word of mouth. They worked in pairs, each reviewing a maximum of six papers. Using standardized written instructions and data extraction forms, and with training and support from the authors, they extracted the following data from the articles:

- the type of measurement tool and construct measured;

- participants on whom it was tested;
- test procedure and analysis methods used;
- psychometric properties tested and the results of their analysis; and
- clinical utility (cost, time taken to complete, need for specialist training or equipment and the portability).

**Analysis and strategy for generating recommendations**

After reviewing each paper individually, the physiotherapists met with their partner to agree a consensus about the data extracted. The data for each locality was collated by a member of the project team (who provided support and advice throughout) and was analysed by ST and LC. Two aspects were analysed; the psychometric properties and the clinical utility. The psychometric properties reviewed were reliability (inter-rater and test–retest), concurrent or criterion-related validity and sensitivity to change. The accepted methods to assess the psychometric properties were:

- for reliability: intraclass correlations (ICC), kappa statistics;
- for validity: correlation coefficients; and
- for sensitivity: effect sizes or measures of the minimum detectable change.

The strength of the psychometric properties were assessed as recommended by Jorstad *et al.*[25]:

• weak reliability or validity = scores of 0.4–0.6; weak sensitivity = effect size 0.2–0.5

• moderate reliability or validity = scores of 0.6–0.8; moderate sensitivity = effect size 0.5–0.8

• good reliability or validity = scores of 0.8 and above; good sensitivity = effect size >0.8

A measurement tool needed to obtain 'good' scores for reliability, validity and sensitivity before it could be recommended for use in clinical practice.

For ordinal scales, the scaling properties of a measure were also considered through an assessment of the hierarchy (coefficients of scalability or reproducibility) or a Rasch analysis. For the coefficient of reproducibility, a coefficient of 0.9 or above was acceptable, for the coefficient of scalability a coefficient of 0.6 or above was acceptable and for Rasch analysis an 'item fit' and 'person fit' with a mean of zero and a standard deviation of less than 1.4 was accepted.

The assessment of clinical utility related to the practical details of using an measurement tool in clinical practice and is detailed below:

- Time taken to administer, analyse and interpret the measurement tool: 3 = Less than 10 minutes, 2 = 10–30 minutes, 1 = 30–60 minutes, 0 = >1 hours.
- Cost: 3 = <£100, 2 = £100–500, 1 = £500–1000, 0 = >£1000 or unknown.
- Does the measurement tool need specialist equipment and training to use? 2 = no, 1 = yes but simple and clinically feasible, 0 = yes and not clinically feasible/unknown.
- Is the measurement tool portable? Can it be taken to the patient? 2 = yes easily (can go in pocket), 1 = yes, in a briefcase or trolley, 0 = no or very difficult.

Adding the scores gave a maximum score of 10. A score of 9 or more was required before a measurement tool could be recommended for clinical use.

## Results

The searches identified 30 measurement tools that measured balance activity in the selected clinical groups, however 11 were subsequently rejected as only one psychometric property had been tested or inappropriate tests had been used. The excluded measures were:

- Trunk Control Test[13,26,27]
- Sitting balance Scale[28]
- Sitting balance scale for hemiplegia[29]
- Advanced Mobility and Balance Scale[30]
- Postural Stress Test[31]
- Equiscale[13]
- Functional Standing Balance Scale[32]
- PLM test[33]
- Maximal load test[34]
- Sensory Organisation Test, also known as the Clinical Sensory Integration test, and the 'Foam and Dome'[35]
- Retropulsion Test.[31,36]

Nineteen remaining measurement tools were assessed on their psychometric properties and clinical utility. There were 12 functional performance tests, which measure patients' ability to perform a functional sitting, standing or stepping task, and seven ordinal scales.

Each of the measurement tools are briefly described below, including, where available, details of the scaling properties (unidimensionality, hierarchy, redundancy, floor and ceiling effects) of the ordinal scales. The data extracted are shown in Tables 1 and 2 and the assessment of the psychometric properties and clinical utility are shown in Tables 3 and 4.

## Functional performance tests

The functional performance tests fell into three main groups: timed tests, reaching and stepping tests.

### *Timed tests*

These tests time how long a patient can maintain a static sitting or standing position and have been used in people with stroke and Parkinson's disease. The length of time tested varies from 15 seconds[37] to 30 seconds[38,39] and 60 seconds.[40–44] Different standing positions have also been used: Feet in parallel (together and hip distance apart); one in front of the other (step or tandem stance) (e.g. sharpened Rhomberg)[16] and standing on one leg (single stance)[34,38,39,45] and with eyes open and closed.[38,45]

### *Reaching tests*

Most reaching tests are based on the functional or forward reach test,[46] which was developed to assess dynamic standing balance. It measures the maximum distance the patient can reach forward beyond arm's length (to the limits of stability) without moving their feet, using a ruler fixed at shoulder height. The test was originally designed for, and extensively tested with, the frail elderly[46–51] but has also been tested with people with stroke, Parkinson's disease and multiple sclerosis although less vigorously. Several variations have been reported. The repeated reach test measures the number of times the patient can reach to a set target within 60 seconds, the backward functional reach test assesses the distance a patient can reach backwards, while the forward bending reach tests how far the patient can reach to the floor in front of them. The sitting forward reach test repeats the forward reach test but in a sitting, rather than standing position.

The arm raise tests aim to assess static balance; they measure the patient's ability to maintain their position while undergoing a destabilizing force (raising and lowering the arm). The sound arm is used for people with stroke and the dominant arm for bilateral conditions. The arm raise test counts the number of times the patient can raise and lower their arm in 15 seconds, while the repeated arm raise test assesses the number of times the patient can reach to a standardized target in 60 seconds. The sitting arm raise is the same as the standing test, but with the patient in a sitting position.

### *Stepping tasks*

Three stepping tests have been developed. The step test evaluates dynamic single stance and involves stepping the stronger (or dominant) foot on and off a block as often as possible within 15 or 30 seconds.[37,52] In this review the test will be referred to as the tap test to differentiate it from the step-up test, which assesses how often the patient can step on to a block (with both feet, leading with the weaker) and down again within 15 seconds.[37] The weight shift test assesses the ability to transfer weight from one leg to the other in tandem stance (the weaker leg in front of, and parallel to, the stronger leg).[37,53] In an early version of this test, a variable load monitor was fitted inside the subject's shoe and made an audible signal when 50% of body weight was transferred on to the weak leg and the number of repetitions within 60 seconds counted in parallel and step stance. More recently, Tyson and De Souza[37] have described a more clinically feasible version, which counts the number of times the patient can transfer his/her weight onto their weaker leg and back again within 15 seconds in step standing, without using specialist equipment.

**Table 1** The psychometric properties of the functional performance tests of balance activity

| Test and reference | Psychometric property tested | Subjects | Results |
|---|---|---|---|
| Timed standing balance tests (feet apart, together, tandem stance, step stance, single-limb stance), arm raise, forward reach, bend reach, tap tests[39] | Test–retest reliability | Patients with idiopathic Parkinson's disease, age range 60–80 years | ICCs: tandem stand = 0.4–0.76; single-leg stance = 0.5–0.94; arm raise = 0.51–0.61; forward reach = 0.42–0.93; bend reach = 0.43–0.84; tap tests = 0.73–0.86 |
| Sitting arm raise, standing arm raise, weight shift, tap test, step-up, sitting forward reach, standing forward reach[34] | Inter-rater reliability, Test–retest reliability, MDC and validity wrt sitting section of MAS, Berg and RMI | 84 stroke patients in total, mean age 66.7 years, median time since stroke = 11 weeks<br>35 patients for reliability and MDC testing and 48 for validity testing | *Sitting arm raise*: test–retest reliability = 0.96; inter-rater reliability = 0.99; MDC = 3 lifts<br>Validity wrt sit-MAS $r = 0.33$; wrt RMI $r = 0.53$; wrt Berg $r = 0.54$<br><br>*Standing arm raise*: test–retest reliability = 0.93; inter-rater reliability = 0.99; MDC = 3 lifts<br>Validity wrt RMI $r = 0.32$; wrt Berg $r = 0.36$<br>*Weight shift*: test–retest reliability = 0.91; inter-rater reliability = 0.88; MDC = 3 shifts<br>Validity wrt RMI $r = 0.52$; wrt Berg $r = 0.26$<br>*Tap test*: test–retest reliability = 0.96; inter-rater reliability = 1.00; MDC = 2 taps<br>Validity wrt RMI $r = 0.46$; wrt Berg $r = 0.74$<br>*Step-up*: test–retest reliability = 0.96; Inter-rater reliability = 1.00; MDC= 1 step-up<br>Validity wrt RMI $r = 0.19$; wrt Berg $r = 0.66$<br>*Sitting forward reach*: test–retest reliability = 0.93; inter-rater reliability = 0.99; MDC = 11 cm<br>Validity wrt sit-MAS $r = 0.54$; wrt RMI $r = 0.54$; wrt Berg $r = 0.61$<br>*Standing forward reach*: test–retest reliability = 0.95; inter-rater reliability = 0.99; MDC = 7.5 cm<br>Validity wrt RMI $r = 0.57$; wrt Berg $r = 0.70$ |

(Continued)

**Table 1** Continued

| Test and reference | Psychometric property tested | Subjects | Results |
|---|---|---|---|
| Tap test[52] | Test–retest reliability and validity wrt forward reach, gait speed and stride length | 41 mobile rehab strokes. Mean age 72.5 years | ICC = 0.88–0.97 Validity: wrt forward reach $r$ = 0.68–0.73; wrt speed $r$ = 0.83; wrt stride length $r$ = 0.82–0.83 |
| Arm raise, forward reach, tap tests[66] | Test–retest reliability | 14 mobile patients with multiple sclerosis. Mean age = 39 years | ICC only reported for some tests: arm raise = 0.72–0.77; forward reach = 0.89; tap test = 0.91–0.93 |
| Sitting forward reach[67] | Test–retest reliability and validity wrt FIM at discharge and gait speed | 26 acute strokes, 7–10 days post stroke, mean age 73.5 years | ICC = 0.79<br>Validity wrt FIM = 0.57; wrt gait speed $r$ = 0.52 |
| Berg, forward reach, backward reach[68] | Validity wrt to UPDRS and each other | 25 Parkinson's disease patients, mean age = 76 years. Mean Hoehn & Yahr = 2 | UPDRS total score wrt Berg $r$ = 0.64; Berg wrt FR $r$ = 0.5; wrt backward reach $r$ = 0.51; wrt TUG $r$ = 0.71 |
| Berg, forward reach, backward reach, Romberg test, Sharpened Romberg test[45] | Test–retest reliability and MDC | 37 adults with Parkinson's disease. Mean age = 71 years. Median Hoehn & Yahr = 2 | Berg: ICC = 0.94; MDC = 5 points<br>*Forward reach*: ICC = 0.73; MDC = 9 cm<br>*Backward reach*: ICC = 0.67; MDC = 7 cm<br>*Romberg*: eyes open ICC = 0.84–0.86, MDC = 10 s; eyes closed; ICC = 0.86, MDC = 19 s<br>*Sharpened Romberg*: eyes open ICC = 0.70, MDC = 39 s; eyes closed ICC = 0.91, MDC = 19 s |

ICC, interclass correlation; wrt, with respect to; MDC, minimum detectable change; MAS, Motor Assessment Scale; sit-MAS, sitting section of the Motor Assessment Scale; RMI, Rivermead Mobility Index; FIM, Functional Independence Measure; UPDRS, Unified Parkinson's Disease Rating Scale; TUG, Timed Get Up & Go test.

**Table 2** The psychometric properties of the ordinal scales of balance activity

| Test and reference | Psychometric property tested | Subjects | Results |
|---|---|---|---|
| BBA[37] | Inter-rater reliability and test–retest reliability | 92 strokes; 37 for reliability and 55 for validity | Validity: BBA wrt sit-MAS $r=0.83$; wrt Berg $r=0.97$; wrt RMI $r=0.95$; MDC = 1 point |
| FMA-B[69] | Inter-rater reliability | 12 rehab strokes. Mean age 66 years | ICC = 0.93 |
| FMA-B[59] | Test–retest reliability, measurement error | 49 chronic strokes | ICC = 0.34; error = 4 points |
| FMA-B[35] | Validity wrt balance activity | 10 acute strokes. Mean times since strokes 15 weeks | $r=0.77$ |
| Berg[70] | Validity wrt STREAM | 63 acute strokes, mean age 67 years | $r=0.53$–$0.78$ |
| Berg[71] | Inter-rater reliability and test–retest reliability | 36 elderly residential home residents with stroke | Inter-rater reliability: ICC = 0.98; test–retest reliability: ICC (stroke) = 0.99 |
| Berg[72] | Validity wrt Tinetti (bal), TUG, BI, FMA | 70 acute strokes in hospital | Berg wrt BI $r=0.87$–$0.93$; BI (mobility) $r=0.67$–$0.94$; FMA $r=0.7$–$0$-$82$; FMA-B $r=0.84$–$0.94$; TUG $r=0.76$; Tinetti (bal) $r=0.91$ |
| Berg[68] | Validity wrt forward reach | 25 mobile people with Parkinson's disease. Mean age 76 | $r=0.54$–$0.69$ |
| Berg[73] | Validity wrt whole body reaching | 23 mobile rehabilitation strokes; age 59 years | $r=0.01$–$0.47$ |
| Short-form Berg (Berg-3P)[56] | Responsiveness of short-form Berg | 81 strokes; age = 68 years | Effect size of full Berg = 0.85; 7-item 3-point Berg = 0.78 |
| Berg[74] | Validity wrt postural sway | 15 chronic stroke, mean duration = 2 years | wrt Sway (CoP-CoM) $r=0.53$–$0.56$ |
| Berg[75] | Sensitivity to change | 61 inpatient rehabilitation strokes. Mean age = 65.2 years | Effect size = 1.01 |
| Berg[76] | Validity wrt postural sway | 20 chronic ambulant hemiplegics | wrt Dynamic activities $r=0.48$–$0.61$ |
| Berg, FMA-B[19] | Test–retest reliability; responsiveness, validity wrt Barthel and MAS (walking subscale) | 123 acute strokes <2/52 post stroke | Test–retest reliability: FMA-B $k=0.79$; Berg $k=0.92$<br>Validity: FM-Bal wrt BI $r=0.86$–$0.89$, wrt MAS $r=0.8$–$0.87$<br>Berg wrt BI $r=0.89$–$0.94$, wrt MAS $r=0.82$–$0.91$<br>Effect size: FM-B=0.33–0.1.14; Berg = 0.4–1.11 |
| Berg[77] | Validity wrt postural sway | 30 strokes, able to stand for 20 s unaided | $r=0.5$–$0.57$ for AP directions only |
| Berg[78] | Validity wrt FMA, 10mWT, 90mWT, self-rated exertion, FAC | 20 rehab, mobile strokes. Mean duration = 20 days. | wrt Walking exertion –0.73; wrt FAC $r=0.63$; wrt FMA $r=0.62$; wrt FMA-B $r=0.77$ |
| Berg[79] | Test–retest reliability and inter-rater reliability | 19 ambulatory people with stable multiple sclerosis | Test–retest reliability ICCs= 0.85; inter-rater reliability ICC = 0.99 |
| Berg[55] | Validity wrt UPDRS, Hoehn & Yahr staging scale, Modified Schwab & England Capacity for ADL | 38 mobile men with Parkinson's disease = 2.5. Mean age = 71 years | wrt UPDRS $r=-0.58$; wrt Hoehn & Yahr $r=-0.45$, wrt Schwab & England Capacity for ADL $r=-0.55$ |

(Continued)

**Table 2** Continued

| Test and reference | Psychometric property tested | Subjects | Results |
|---|---|---|---|
| Berg[80] | Validity wrt forward reach | 75 strokes >1/52 post CVA | $r = 0.78$ |
| Berg[81] | MDC; test–retest reliability | 48 acute strokes (mean 30.3 days), mean age 73.5 | Error: 5–7 points; test–retest reliability ICC = 0.92 |
| Berg[57] | Validity of modified Berg wrt to full assessment and BI | 1st part: $N = 77$ rehabilitation strokes, age 59.8 years, 2nd part: $n = 226$ retrospective note search. Mean age 69 years | Validity: Berg 3-P at 14 and 30 days wrt BI at 3/12 $r = 0.75–0.87$ |
| Standing balance scale[44] | Test–retest reliability; validity wrt gait speed and FIM | 38 acute strokes, mean time since stroke 21 days | Test–retest reliability $k = 0.86$ Validity wrt Gait speed $r = 0.53–0.66$; wrt FIM $r = 0.51–0.53$ |
| Standing balance scale[43] | Inter-rater reliability; validity wrt FIM | 52 acute rehabilitation strokes | Inter-rater reliability $k = 0.97$ Validity wrt FIM $r = 0.44–77$ |
| Sit-MAS and FMA-B[82] | Validity wrt each other | 32 strokes; mean age = 60 years; mean time since stroke = 65 days | FMA-B wrt sit-MAS $r = 0.1$; wrt total FMA = $r = 0.03$; wrt FMA (leg score) $r = 0.06$; sit-MAS wrt total MAS $r = 0.64$; wrt walking MAS $r = 0.54$ |
| Sit-MAS[83] | Inter-rater reliability | 7 rehab strokes; mean age = 74 years | $k = 0.56$ |
| Sit-MAS and FMA-B[84] | Validity of sit-MAS and FMA-B | 30 strokes; mean age 63 years; mean times since stroke 12 months | sit-MAS wrt FMA-B $r = 0.28$ |
| Sandin and Smith sitting scale[62] | Validity wrt recovery BI | 25 rehabilitation strokes | wrt recovery BI $r = 0.7–0.933$. |
| Sandin and Smith sitting scale[85] | Inter-rater reliability | 53 acute strokes; mean age 73 years | $k = 0.8$ |
| Trunk Impairment Scale[16] | Inter-rater reliability and test–retest reliability; measurement error; validity wrt BI and TCT | 28 subacute rehabilitation strokes; median age = 63 years | Inter-rater reliability: ICC = 085–0.99; test–retest relaiability ICC = 087–0.96 Measurement error = 4 points (16%) Validity: wrt BI $r = 0.86$; wrt TCT $r = 0.83$ |
| Trunk Impairment Scale[17] | Inter-rater reliability and test–retest reliability, validity wrt FIM and EDSS | 30 multiple sclerosis patients; 15 male, 15 female; mean duration = 16 years | Inter-rater reliability: ICC = 0.82–0.97; test–retest reliability ICC = 0.85–0.95 Validity: wrt FIM $r = 0.81$; wrt EDSS $r = 0.85$ |
| Trunk Impairment Scale[86] | Inter-rater reliability and test–retest reliability; validity wrt BI | 30 patients with traumatic brain injury. Median duration = 84 days, median age = 32 years | Test–retest reliability: $k = 0.34–1$; inter-rater reliability $k = 0.45–1$. Validity: wrt $r = 0.59$ |

ICC, interclass correlation; wrt, with respect to; MDC, minimum detectable change; MAS, Motor Assessment Scale; sit-MAS, sitting section of the Motor Assessment Scale; RMI, Rivermead Mobility Index; FIM, Functional Independence Measure; UPDRS, Unified Parkinson's Disease Rating Scale; TUG, Timed Get Up & Go test; BI, Barthel Index; STREAM, Stroke Rehabilitation Assessment of Movement; Tinetti (bal), balance section of Tinetti Balance and Gait Scale; TCT, trunk control test; TBI, traumatic brain injury; EDSS, Expanded Disability Status Scale; BBA, Brunel Balance Assessment; FMA-B, balance section of the Fugl-Meyer Motor Assessment; FAC, functional ambulation categories; ADL, activities of daily living.

**Table 3** Clinical utility of the measures of balance activity

| Measurement tool | Time to complete | Cost | Portability | Specialist equipment | Total (max = 10) |
|---|---|---|---|---|---|
| Timed tests (including Romberg and Sharpened Romberg) | 3 | 3 | 2 | 2 | 10 |
| Sitting forward reach | 3 | 3 | 2 | 2 | 10 |
| Standing forward reach | 3 | 3 | 2 | 2 | 10 |
| Other standing reach tests (repeated forward reach, backwards reach, bending forward reach, repeated arm raise) | 3 | 3 | 2 | 2 | 10 |
| Sitting arm raise | 3 | | 2 | 2 | 10 |
| Standing arm raise | 3 | 3 | 2 | 2 | 10 |
| Tap (aka step test) | 3 | 3 | 1 | 2 | 9 |
| Step-up | 3 | 3 | 1 | 2 | 9 |
| Weight shift | 3 | 3 | 1 | 2 | 9 |
| Berg | 2 | 3 | 2 | 2 | 9 |
| FMA-B | 3 | 3 | 2 | 2 | 10 |
| Brunel | 3 | 3 | 2 | 2 | 10 |
| Standing Balance Scale | 3 | 3 | 2 | 2 | 10 |
| Sit-MAS | 3 | 3 | 2 | 2 | 10 |
| Trunk Impairment Scale | 3 | 3 | 2 | 2 | 10 |
| Sandin and Smith | 3 | 3 | 2 | 2 | 10 |

sit-MAS, sitting section of the Motor Assessment Scale; FMA-B, balance section of the Fugl-Meyer Motor Assessment.

## Ordinal scales

Seven ordinal scales were selected, of which three involve a mixture of sitting, standing and/or stepping activities; one specifically measured standing balance and three specifically measured sitting balance. One was designed to measure 'trunk impairment' but on examination the items related to the ability to perform sitting balance activities and so is included in this review.

### *Berg Balance Scale*

The Berg Balance Scale[54] is the best-known balance measurement tool. It was originally designed to measure balance in the elderly but has also been tested with patients with stroke, multiple sclerosis and Parkinson's disease[55] (see Table 2). It consists of 14 tasks quantified on a 5-point scale. The items evaluate the ability to maintain positions of increasing difficulty by decreasing the base of support, from sitting to standing, to close standing, to tandem standing and finally to standing on one leg. Other items evaluate the ability to perform specific tasks, such as transfer between positions, reaching forward, turning round and picking up an object from the floor. Scoring is based on the ability to perform the items independently and to meet certain time or distance requirements. It takes 15–20 minutes to complete. Several recent publications have modified the assessment by reducing the number of items or the levels in the subscales[56,57] and revising the rating structure using Rasch analysis.[58] Considerable redundancy within the items and rating structure was demonstrated but it is not clear which modifications are most efficacious.

### *Balance section of the Fugl-Meyer Motor Assessment*

The Fugl-Meyer Motor Assessment[59] was developed to assess recovery of motor function in stroke patients. The balance subscale examines static sitting, 'parachute reactions in sitting', standing and single stance. The seven items are each rated on a 3-point ordinal scale. The scaling properties have not been reported and it is unlikely to form a hierarchy as automatic reactions to external perturbations in sitting (parachute reactions) are placed before static standing balance, which does not reflect the relative difficulty of the tasks.

**Table 4** Summary of the psychometric properties of the measures of balance activity

| | Groups tested | Validity | Test–retest reliability | Inter-tester reliability | MDC |
|---|---|---|---|---|---|
| Timed standing balance tests | Parkinson's disease | | ++/+++ | | 10–39 s |
| Sitting forward reach | Stroke | + Sit MAS<br>++ RMI, Berg, FIM | +++ | +++ | 11 cm |
| Standing forward reach | Stroke, Parkinson's disease, multiple sclerosis | Stroke: ++ RMI, Berg<br>Parkinson's disease ++ Berg | Parkinson's disease: ++/+++<br>Stroke: +++<br>Multiple Sclerosis: ++ | Stroke: +++ | MDC: PD: 9 cm; Stroke: 7.5 cm, effect size +++ |
| Other standing reach tests (repeated forward reach, backward reach, bending forward reach, repeated arm raise) | Stroke Parkinson's disease | ++ Berg | ++/+++ | | Backward reach 7 cm |
| Sitting arm raise | Stroke | + Sit-MAS, Berg, RMI | +++ | +++ | 3 lifts (stroke) |
| Standing arm raise | Stroke Parkinson's disease Multiple sclerosis | Stroke: +RMI, Berg | Parkinson's disease: ++/+++<br>Multiple sclerosis: ++<br>Stroke: +++ | Stroke: +++ | 3 lifts (stroke) |
| Tap (aka step test) | Stroke Parkinson's disease Multiple sclerosis | Stroke: ++RMI<br>+++ Berg, FR | Parkinson's disease: ++/+++<br>Multiple sclerosis: ++<br>Stroke: +++ | Stroke: +++ | 2 taps |
| Step-up | Stroke | +++ Berg<br>+ RMI | +++ | +++ | 1 step up |
| Weight shift | Stroke | + Berg<br>++ RMI | +++ | +++ | 3 shifts |
| Berg Balance Scale | Stroke Parkinson's disease Multiple sclerosis | Stroke: +++ BI, TUG, MAS, FAC, gait exertion<br>++/+++ FMA, FMA-B, STREAM, FR<br>++ Postural sway, ADL<br>Parkinson's disease: +/++ disease severity | Stroke: +++<br>Parkinson's disease: +++ | Stroke: +++<br>Multiple sclerosis: +++ | MDC 6/56 points (stroke)<br>+++ effect size (stroke) |
| FMA-B | Stroke | +++ SOT<br>+ MAS-sit, FMA | + | +++ | 4/14 points |
| Brunel Balance Assessment | Stroke | +++ MAS, Berg RMI | +++ | +++ | MDC = 1/12 point (stroke) |
| Bohannon | Stroke | ++/+++ gait, FIM | +++ | +++ | |

(Continued).

**Table 4** Continued

| | Groups tested | Validity | Test–retest reliability | Inter-tester reliability | MDC |
|---|---|---|---|---|---|
| Sit-MAS | Stroke | + FMA-B<br>++ MAS (total) | | ++ | |
| Dutch Trunk Impairment Scale | Stroke Multiple sclerosis Parkinson's disease Traumatic brain injury | Stroke: +++ BI, TCT<br>Multiple sclerosis: +++ FIM, EDSS<br>Parkinson's disease: ++ BI, UPDRS<br>Traumatic brain injury: ++ BI | +++<br>Traumatic brain injury: +++ | +++ | +++ MDC = 4 points (stroke) |
| Sandin and Smith | Stroke | ++/+++ recovery of ADL | | +++ | |

MAS, Motor Assessment Scale; sit-MAS, sitting section of the Motor Assessment Scale; RMI, Rivermead Mobility Index; FIM, Functional Independence Measure; UPDRS, Unified Parkinson's Disease Rating Scale; TUG, Timed Get Up & Go test; BI, Barthel Index; STREAM, Stroke Rehabilitation Assessment of Movement; Tinetti (bal), balance section of Tinetti Balance and Gait Scale; TCT, trunk control test; TBI, traumatic brain injury; EDSS, Expanded Disability Status Scale; BBA, Brunel Balance Assessment; FMA-B, balance section of the Fugl-Meyer Motor Assessment; FAC, functional ambulation categories; ADL, activities of daily living.

*Brunel Balance Assessment*

The Brunel Balance Assessment[37] operationalizes the hierarchy of balance tasks that UK physiotherapists use when assessing balance activity; progressing from sitting to standing and stepping balance. In an innovative design, it combines a 12-point ordinal scale of balance activity with functional performance tests at each level of the ordinal scale. The ordinal scale is arranged into three subscales (sitting, standing and stepping balance) that also can be used individually. It forms a true Guttman scale, indicating that it is unidimensional, a true hierarchy and there is no redundancy in the included items.

*Standing Balance Scale*

The Standing Balance Scale[40–44] was designed to quantify independent standing balance, based on timed measurements. It is a 7-point ordinal scale, categorizing the ability to maintain different standing postures for 60 seconds (e.g. 0 – unable to stand with feet apart, 3 – able to stand with feet together for up to 60 seconds, 6 – able to stand unilaterally for 60 seconds). The scale was later modified to use times of 30 seconds only and to use single-leg stance of the weak or sound leg.[43] Scaling properties have not been investigated.

*Sitting balance section of the Motor Assessment Scale*

The Motor Assessment Scale[60] was developed to assess everyday motor function in patients with stroke.[60] It comprises eight subscales, only one of which relates to balance (sitting balance). Only data specific to the sitting balance section is presented in this review. The sitting balance section is scored on a 7-point scale, with patients assessed on their ability to perform progressively harder sitting balance activities. A score of 0 indicates a patient is only able to sit with assistance of another person, whereas 6 indicates they can sit on a stool unsupported, reach sideways to touch the floor and return to the starting position. Scaling properties have not been addressed.

*Trunk Impairment Scale*

The Trunk Impairment Scale[16] aims to measure motor impairment of the trunk after stroke, however all the items relate to the ability to maintain sitting balance during different activities, which fulfils the definition of balance used in this review. It consists of 17 tasks in three subscales: static sitting balance, dynamic sitting balance and coordination. These range from sitting on the edge of the bed, sitting on the edge with a narrower base of support, and various selective movements of the trunk (lateral flexion and rotation). It has been subjected to Rasch analysis.[61] The coordination section fitted the Rasch model, as did the dynamic sitting balance section once one of the items had been adjusted. The static sitting balance section did not fit the model, probably as it needs to be tested with more people with very severe stroke.

**Sandin and Smith (1990)**

Sandin and Smith[62] assessed sitting balance using a 4-point scale evaluating the amount of assistance required during static and dynamic activities. The dynamic tests assessed whether the patient could withstand external displacement forces of 'approximately 5–10 lbs force applied anteriorally, laterally and posteriorally'. No further details were given about how this was performed or standardized, nor have the scaling properties been assessed.

The data extracted from the selected papers are shown in Tables 1 and 2 (for functional performance tests and ordinal scales respectively), the clinical utility is shown in Tables 3 and the summary of the psychometric properties are shown in Table 4. This analysis of the psychometric properties and clinical utility revealed three ordinal scales and seven functional performance tests that were psychometrically robust and feasible for use in clinical practice. These were the forward reach and arm raise tests (in sitting and standing), step/tap test, weight shift test and step-up test, Brunel Balance Assessment, Berg Balance Scale and Trunk Impairment Scale. All have been primarily tested in people with stroke, although the Berg Balance Scale, Trunk Impairment Scale and functional reach have been tested in other neurological groups (people with Parkinson's disease, multiple sclerosis and traumatic brain injury), although incompletely. All of the tests scored highly on clinical utility but for many the psychometric testing was incomplete. Although all had some assessment of validity and reliability, sensitivity to change had been less thoroughly assessed.

## Discussion

The results of this systematic review revealed 10 tests that were psychometrically robust and feasible to use in clinical practice. They were the forward reach and arm raise tests (in sitting and standing), step/tap test, weight shift test and step-up test, Brunel Balance Assessment, Berg Balance Scale and Trunk Impairment Scale. This is the first review of balance measures to systematically assess the quality of all measures of balance disability and to consider the clinical utility of the measure, it is therefore not possible to compare with previous research.

The robust and feasible tools employed diverse designs, included a wide range of items and target different patient groups. This makes it difficult to identify the most effective measurement tool for any individual patient group. The Brunel and Berg Assessments measure sitting, standing and stepping balance activities but are constructed differently, while the Trunk Impairment Scale contains items that measure sitting balance alone. The functional performance tests measure more specific aspects of balance such as maintaining balance while moving to the limits of stability or while withstanding a destabilizing force.

To make a sensible choice about which measurement tool to use, one needs a good understanding of what the measurement tool measures and why those measures are important to relate it to one's patients and clinical service. This is a challenge for balance activity measurement tools as the literature on their development gives scant consideration to any underlying theoretical construct. In most cases, where any explanation is given, the choice of items is based on clinicians' opinions with no justification or rationale for why they were included or how they were operationalized.

Although this atheoretical approach has some face and ecological validity, it limits evidence-based decisions about which measurement tool to use, how to interpret the results and how to relate them to treatment planning. Further research is needed to identify develop and articulate the theoretical construct behind the clinical balance assessment and to build a consensus. This is a challenge, as although discussions of the theoretical basis of balance problems have been published,[63–65] consistency in the language and definitions has not been achieved and none have been widely adopted.

A further limitation of the ordinal scales is that few have considered the scaling properties, particularly the hierarchy and redundancy of items. The presence of a true hierarchy has three advantages. It means that the whole measurement tool does not need to be performed each time it is used as one can assume that the patient would pass all the preceding items once one item had been 'passed', or would fail all subsequent items once an item has been 'failed'. Consequently one can start the testing at a level that is appropriate for the patient and a few of the items would need to be tested. Clearly, this would save time and effort for both clinician and patient. A true hierarchy also has no redundant items, which artificially inflate a score thereby reducing test accuracy and wasting time and effort by asking the patient to perform activities that do not contribute to the overall result. The final advantage is that the score gives an indication of the subject's ability. For example, in the Brunel Balance Assessment, which is hierarchical, a score of 6/12 indicates someone who has dynamic standing balance (and will have passed all sitting balance and static balance tests) but not stepping or single stance balance. In contrast, a score of 46/56 on the Berg Balance Scale, which is not hierarchical, gives no indication of which items were passed or failed. A hierarchical scale is advantageous as the score is more informative and relevant to clinical practice than one from a non-hierarchical design. Further research is needed to assess the hierarchy of the ordinal scales, and to adjust the construction of the scale whenever it is found to be lacking. Development of future ordinal measurement tools should include an examination of the hierarchy and redundancy of items at an early stage as part of the development of the construct of the scale.

An alternative type of measurement tool are functional performance tests, which have many attractive features. As well as being reliable and valid, they are very quick to use (most take 2 or 3 minutes), require no specialist equipment and can easily be incorporated into a full clinical assessment or treatment session. In addition the results are meaningful to clinicians and patients; it is easy to understand the significance being able to reach a few more centimetres or complete more repetitions. The disadvantage of these tests is that each individual test is only relevant for a relatively narrow band of patients and will exhibit floor and ceiling effects for people outside this band. For example, patients who cannot sit unaided would be unable to perform the sitting forward reach test (demonstrating a floor effect) but this test would also show a ceiling effect for more able patients whose dynamic sitting balance may be within normal limits. It is therefore important to marry up the appropriate test to the patient's level of ability. To do this, one needs to understand what the test is measuring: its construct. The functional performance tests recommended in this review form a battery of tests which progress by increasing the demands and complexity of the task and by decreasing the base of support.[9,34] They operationalize the hierarchy of balance activities that physiotherapists use in their clinical assessment,[9] which gives a framework for the choice of which measurement tool to use.

The psychometric property that has received least attention was the minimum detectable change, yet this is vital for the clinical interpretation of scores as it indicates whether a change in score is due to 'true change' in performance or whether it is merely due to the normal fluctuations of scoring. Furthermore, it is simple to calculate from the same data that is used to assess reliability. Future measurement tool development should include assessment of the minimum detectable change and consider application to clinical groups other than stroke.

Like all systematic reviews, the quality of the review is dependent on the papers identified.

Although we had thorough search strategies, we only included publications in English. There may have been relevant publications in other languages that we missed. We also did attempt to identify unpublished data or the grey literature, so there may have been a publication bias in the data that we identified.

To the authors' knowledge, this is the first systematic review to specifically assess the clinical utility of measurement tools. The system we developed to assess the utility was based on our clinical experience and the judgements of quality were arbitrary. Such judgements can not be assumed to be appropriate for other health care systems or other areas of clinical practice. Nevertheless, they have strong face validity and were acceptable to neurological physiotherapists working across the north-west of England, who served a population of about 6.8 million people, so we feel they are reasonably generalizable.

**Clinical messages**

- A range of psychometrically robust, clinically feasible measures of balance activity has been identified. These are the forward reach and arm raise tests (in sitting and standing), step/tap test, weight shift test and step-up test, Brunel Balance Assessment, Berg Balance Scale and Trunk Impairment Scale.
- Future research to develop measurement tools of balance activity for patients with neurological conditions should clearly state the theoretical construct on which they are based, consider the scale development of ordinal scales, include assessment of the minimal detectable change, and investigate their use in all neurological populations.

## Competing interests

The lead author (SFT) was the principal investigator on the project to develop one of the measurement tools included in this review; the Brunel Balance Assessment.

## References

1 Chartered Society of Physiotherapy. *Core standards of physiotherapy practice*. London, Chartered Society of Physiotherapy, 2005.
2 Intercollegiate Stroke Working Party. *National clinical guidelines for stroke*, second edition. London, Royal College of Physicians, 2004.
3 Rudd A, Pearson M, Georgiou A. *Measuring clinical outcome in stroke (acute care): clinical effectiveness and evaluation unit*. London, Royal College of Physicians, 2000.
4 Department of Health. *National services framework for long-term conditions*. London, Department of Health, 2005.
5 Greenhalgh J, Tyson S, Long A & Flynn R. How do multidisciplinary teams use clinical outcome measures in clinical practice? *Proceedings of Society for Research in Rehabilitation in Clinical Rehabilitation*, 2006; **20**(12): 1105.
6 Grimmer K, Bialocerkowski A, Kumar S. Implementing evidence in clinical practice: the therapies dilemma. *Physiotherapy* 2004; **90**: 189–94.
7 National Audit Office. *Reducing brain damage: faster access to better stroke care*. London, National Audit Office, 2005.
8 Pollock A, Legg L, Langhorne P. Barriers to achieving evidence-based stroke rehabilitation. *Clin Rehabil* 2000; **14**: 611–17.
9 Tyson S, Watson A, Moss S *et al*. Development of a framework for the evidence-based choice of outcome measures in neurological physiotherapy. *Disabil Rehabil* 2008; **30**: 142–49.
10 Bronstein A, Brandt T, Woolacott M. *Clinical disorders of balance, posture and gait*. London, Arnold, 1996.
11 Shumway-Cook A, Horak FB. Assessing the influence of sensory interaction on balance: suggestion from the field. *Phys Ther* 1986; **66**: 1548–50.
12 World Health Organization. *ICF-Introduction, the International Classification of Functioning Disability and Health*. Geneva, WHO, 2001.
13 Tesio L, Perucca L, Franchignoni FP, Battaglia MA. A short measure of balance in multiple sclerosis: validation through Rasch analysis. *Funct Neurol* 1997; **12**: 255–65.
14 Fujiwara T, Liu M, Tsuji T *et al*. Development of a new measure to assess trunk impairment after stroke (Trunk Impairment Scale): its psychometric properties. *Am J Phys Med Rehabil* 2004; **83**: 681–88.

15 Liu M, Chino N, Tuji T, Masakado Y, Hase K, Kimura A. Psychometric properties of the Stroke Impairment Assessment Set (SIAS). *Neurorehabil Neural Repair* 2002; **16**: 339–51.
16 Verheyden G, Nieuwboer A, Mertin J, Preger R, Kiekens C, De Weerdt W. The Trunk Impairment Scale: a new tool to measure motor impairment of the trunk after stroke. *Clin Rehabil* 2004; **18**: 326–34.
17 Verheyden G, Nuyens G, Nieuwboer A, Van Asch P, Ketelaer P, De Weerdt W. Reliability and validity of trunk assessment for people with multiple sclerosis. *Phys Ther* 2006; **83**: 66–76.
18 Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther* 2008; **88**: 559–66.
19 Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke* 2002; **33**: 1022–27.
20 Whitney SL, Poole JL, Cass SP. A review of balance instruments for older adults. *Am J Occup Ther* 1998; **52**: 666–71.
21 Anemaet WK, Moffa-Trotter ME. Functional tools for assessing balance and gait impairments. *Topics Geriatr Rehabil* 1999; **15**: 66–83.
22 van Peppen RP, Hendriks HJ, van Meeteren NL, Helders PJ, Kwakkel G. The development of a clinical practice stroke guideline for physiotherapists in the Netherlands: a systematic review of available evidence. *Disabil Rehabil* 2007; **29**: 767–83.
23 Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; **2**: 1–74.
24 Yoward LS, Doherty P, Boyes C. A survey of outcome measurement of balance, walking and gait amongst physiotherapists working in neurology in the UK. *Physiotherapy* 2008; **94**(2): 125–132.
25 Jørstad EC, Hauer K, Becker C, Lamb SE. Measuring the psychological outcomes of falling: a systematic review. *J Am Geriatr Soc* 2005; **53**: 501–10.
26 Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. *J Neurol Neurosurg Psychiatry* 1990; **53**: 576–79.
27 Duarte E, Marco E, Muniesa JM, Belmonte R, Diaz P, Tejero MEF. Trunk control test as a functional predictor in stroke patients. *J Rehabil Med* 2002; **34**: 267–72.
28 Feigin L, Sharon B, Czaczkes B, Rosin AJ. Sitting equilibrium 2 weeks after a stroke can predict the walking ability after 6 months. *Gerentology* 1996; **42**: 348–53.
29 Nieuwboer A, Feys H, De Weerdt W, Nuyens G, De Corte E. Developing a clinical tool to measure sitting balance after stroke: a reliability study. *Physiotherapy* 1995; **81**: 439–45.
30 Kairy D, Paquet N, Fung J. A postural adaptation test for stroke patients. *Disabil Rehabil* 2003; **25**: 127–35.
31 Harburn KL, Hill KM, Kramer JF, Noh S, Vandervoort AA, Teasell R. Clinical applicability and Test–retest reliability of an external perturbation test of balance in stroke subjects. *Arch Phys Med Rehabil* 1995; **76**: 317–23.
32 Pyoria O, Era P, Talvitie U. Relationships between standing balance and symmetry measurements in patients following recent strokes (<= 3 weeks) or older strokes (>=6 months). *Phys Ther* 2004; **84**: 128–36.
33 Kokko S, Paltamaa J, Ahola E, Malkia E. The assessment of functional ability in patients with Parkinson's disease: the PLM-test and three clinical tests. *Physiother Res Int* 1997; **2**: 29–45.
34 Tyson SF, DeSouza LH. Reliability and validity of functional balance tests post stroke. *Clin Rehabil* 2004; **18**: 916–23.
35 Di Fabio RP, Badke MB. Relationship of sensory organization to balance function in patients with hemiplegia. *Phys Ther* 1990; **70**: 542–48.
36 Visser M, Marinus J, Bloem BR, Kisjes H, van den Berg BM, van Hilten JJ. Clinical tests for the evaluation of postural instability in patients with Parkinson's disease. *Arch Phys Med Rehabil* 2003; **84**: 1669–74.
37 Tyson SF, DeSouza LH. Development of the Brunel Balance Assessment: a new measure of balance disability post stroke. *Clin Rehabil* 2004; **18**: 801–10.
38 Bernhardt J, Ellis P, Denisenko S, Hill K. Changes in balance and locomotion measures during rehabilitation following stroke. *Physiother Res Int* 1998; **3**: 109–22.
39 Smithson F, Morris ME, Iansek R. Performance on clinical tests of balance in Parkinson's disease. *Phys Ther* 1998; **78**: 577–92.
40 Bohannon R. Correlation of lower limb strengths and other variables with standing performance in stroke patients. *Physiother Can* 1989; **41**: 198–202.
41 Bohannon RW. Selected determinants of ambulatory capacity in patients with hemiplegia. *Clin Rehabil* 1989; **3**: 47–53.
42 Bohannon RW. Standing balance, lower extremity muscle strength, and walking performance of patients referred for physical therapy. *Percept Motor Skills* 1995; **80**: 379–85.

43 Bohannon RW, Leary KM. Standing balance and function over the course of acute rehabilitation. *Arch Phys Med Rehabil* 1995; **76**: 994–96.
44 Bohannon RW, Walsh S, Joseph MC. Ordinal and timed balance measurements: reliability and validity in patients with stroke. *Clin Rehabil* 1993; **7**: 9–13.
45 Steffen T, Seney M. Test–retest reliability and minimal detectable change on balance and ambulation tests, the 36-item Short-Form Health Survey, and the Unified Parkinson Disease Rating Scale in people with parkinsonism. *Phys Ther* 2008; **88**: 733–46.
46 Duncan P, Weiner D, Chandler Jea. Functional reach: a new clinical measure of balance. *J Gerontol* 1990; **45**: M192–M197.
47 Duncan PW, Studenski S, Chandler J, Prescott B. Functional reach: predictive validity in a sample of elderly male veterans. *J Gerontol* 1992; **47**: M93–M98.
48 Giorgetti M, Harris B, Jette A. Reliability of clinical balance outcome measures in the elderly. *Physiother Res Int* 1998; **3**: 274–83.
49 Hill K, Ellis P, Bernhardt J, Maggs P, Hull S. Balance and mobility outcomes for stroke patients: a comprehensive audit. *Aust J Physiother* 1997; **43**: 173–80.
50 Weiner D, Duncan P, Chandler J, Studenski S. Functional reach: a marker of physical frailty. *J Am Geriatr Soc* 1992; **40**: 203–207.
51 Weiner DK, Bongiorni DR, Studenski SA, Duncan PW, Kochersberger GG. Does functional reach improve with rehabilitation? *Arch Phys Med Rehabil* 1993; **74**: 796–800.
52 Hill KD, 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. *Physiother Can* 1996; **48**: 257–62.
53 Goldie PA, Matyas TA, Spencer KI, McGinley RB. Postural control in standing following stroke: Test–retest reliability of some quantitative clinical tests. *Phys Ther* 1990; **70**: 234–43.
54 Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can* 1989; **41**: 304–31.
55 Qutubuddin AA, Pegg PO, Cifu DX, Brown R, McNamee S, Carne W. Validating the Berg Balance Scale for patients with Parkinson's disease: a key to rehabilitation evaluation. *Arch Phys Med Rehabil* 2005; **86**: 789–92.
56 Chou CY, Chien CW, Hsueh IP, Sheu CF, Wang CH, Hsieh CL. Developing a short form of the Berg Balance Scale for people with stroke. *Phys Ther* 2006; **86**: 195–204.
57 Wang CH, Hsueh IP, Sheu CF, Yao G, Hsieh CL. Psychometric properties of 2 simplified 3-level balance scales used for patients with stroke. *Phys Ther* 2004; **84**: 430–38.
58 Kornetti DL, Fritz SL, Chiu Y-P, Light KE, Velozo CA. Rating scale analysis of the Berg balance scale. *Arch Phys Med Rehabil* 2004; **85**: 1128–35.
59 Beckerman H, Vogelaar TW, Lankhorst GJ, Verbeek ALM. A criterion for stability of the motor function of the lower extremity in stroke patients using the Fugl-Meyer Assessment Scale. *Scand J Rehabil Med* 1996; **28**: 3–7.
60 Carr JH, Shepherd RB, Nordholm L, Lynne D. Investigation of a new motor assessment scale for stroke patients. *Phys Ther* 1985; **65**: 175–80.
61 Verheyden, G. and Kersten, P. Evaluating the construct validity of the trunk impairment scale: a Rasch analysis of its subscales. *Clinical Rehabilitation*, **22**(7), 664.
62 Sandin KJ, Smith BS. The measure of balance in sitting in stroke rehabilitation prognosis. *Stroke* 1990; **21**: 82–86.
63 King M, Judge J, Wolfson L. Functional base of support decreases with age. *J Gerontol* 1994; **49**: M258–63.
64 Nashner L. Adaptation of human movement to altered environments. *Trends Neurosci* 1982; **5**: 358–61.
65 Pollock AS, Durward BR, Rowe PJ, Paul JP. What is balance? *Clin Rehabil* 2000; **14**: 402–406.
66 Frzovic D, Morris ME, Vowels L. Clinical tests of standing balance: performance of persons with multiple sclerosis. *Arch Phys Med Rehabil* 2000; **81**: 215–21.
67 Tsang YL, Mak MK. Sit-and-reach test can predict mobility of patients recovering from acute stroke. *Arch Phys Med Rehabil* 2004; **85**: 94–98.
68 Brusse KJ, Zimdars S, Zalewski KR, Steffen TM. Testing functional performance in people with Parkinson disease. *Phys Ther* 2005; **85**: 134–41.
69 Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer Assessment for testing motor performance in patients following stroke. *Phys Ther* 1993; **73**: 447–54.
70 Ahmed S, Mayo N, Higgins J, Salbach N, Finch L, Wood-Dauphinee S. The Stroke Rehabilitation Assessment of Movement (STREAM): a comparison with other measures used to evaluate effects of stroke and rehabilitation. *Phys Ther* 2003; **83**: 617–30.

71 Berg K, Wood-Dauphinee S, Williams J. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med* 1995; **27**: 27–36.
72 Berg K, Wood-Dauphinee S, Williams J, Maki B. Measuring balance in the elderly: validation of an instrument. *Can J Publ Health* 1992; **83**(suppl 2): S7–11.
73 Chern JS, Yang SW, Wu CY. Whole-body reaching as a measure of dynamic balance in patients with stroke. *Am J Phys Med Rehabil* 2006; **85**: 201–208.
74 Corriveau H, Hebert R, Raiche M, Prince F. Evaluation of postural stability in the elderly with stroke. *Arch Phys Med Rehabil* 2004; **85**: 1095–101.
75 English C, Hillier S, Stiller K, Warden-Flood A. The sensitivity of three commonly used outcome measures to detect change amongst patients receiving inpatient rehabilitation following stroke. *Clin Rehabil* 2006; **20**: 52–55.
76 Liston RA, Brouwer BJ. Reliability and validity of measures obtained from stroke patients using the Balance Master. *Arch Phys Med Rehabil* 1996; **77**: 425–30.
77 Niam S, Cheung W, Sullivan PE, Kent S, Gu X. Balance and physical impairments after stroke. *Arch Phys Med Rehabil* 1999; **80**: 1227–33.
78 Nilsson L, Carlsson J, Grimby G, Nordholm L. Assessment of walking, balance and sensorimotor performance of hemiparetic patients in the acute stage after stroke. *Physiother Theory Pract* 1998; **14**: 149–57.
79 Paltamaa J, West H, Sarasoja T, Wikstrom J, Malkia E. Reliability of physical functioning measures in ambulatory subjects with MS. [erratum appears in Physiother Res Int 2006; **11**: 123]. *Physiother Res Int* 2005; **10**: 93–109.
80 Smith PS, Hembree JA, Thompson ME. Berg Balance Scale and Functional Reach: determining the best clinical tool for individuals post acute stroke. *Clin Rehabil* 2004; **18**: 811–18.
81 Stevenson TJ. Detecting change in patients with stroke using the Berg Balance Scale. *Aust J Physiother* 2001; **47**: 29–38.
82 Malouin F, Pichard L, Bonneau C, Durand A, Corriveau D. Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Arch Phys Med Rehabil* 1994; **75**: 1206–12.
83 Loewen SC, Anderson BA. Reliability of the Modified Motor Assessment Scale and the Barthel Index. *Phys Ther* 1988; **68**: 1077–81.
84 Poole JL, Whitney SL. Motor assessment scale for stroke patients: concurrent validity and interrater reliability. *Arch Phys Med Rehabil* 1988; **63**: 195–97.
85 Morgan P. The relationship between sitting balance and mobility outcome in stroke. *Aust J Physiother* 1994; **40**: 91–96.
86 Verheyden G, Hughes J, Jelsma J, Nieuwboer A, De Weerdt W. Assessing motor impairment of the trunk in patients with traumatic brain injury: reliability and validity of the Trunk Impairment Scale. *S Afr J Physiother* 2006; **62**: 23–28.