The Berg Balance Scale - Determining its usefulness in the elderly.

A thesis by publication submitted for the degree of

Master of Philosophy

By

Stephen Downs
BAppSc(Physiotherapy)

The School of Health Sciences
Faculty of Health and Medicine
University of Newcastle

Submitted September 2013
Declaration of originality:

I hereby certify that this thesis is submitted in the form of a series of published papers of which I am a joint author. I have included as part of the thesis a written statement from each co-author attesting to my contribution to the joint publications. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying subject to the provisions of the Copyright Act 1968.

In addition, ethics approval from the University of Newcastle and North Coast Area Health Service Human Research Ethics Committees was granted for the clinical study presented in this thesis. Participants were required to read a participant information document and written informed consent was gained prior to data collection.

Stephen Downs
Acknowledgements:

I wish to acknowledge my supervisors Pauline Chiarelli and Jodie Marquez. Their wise mentoring and assistance has been superb. At important times they have steered this research in very useful ways.

I would also like to acknowledge the support provided by the Rural Research Capacity Building Program – run by a part of the NSW Health Service which has been named the Health Education and Training Institute, the Clinical Training Institute, and the Institute of Rural Clinical Services and Teaching.

I would also like to acknowledge to flexibility provided by my employer, the Mid North Coast Local Health District, which has been a big help in the completion of this research.
Publications and Presentations

Peer reviewed publications included in this thesis:


Conference presentations:

Downs, S. Mobility Outcomes at Bellingen and Macksville Hospitals. Australian Association of Gerontology (NSW Division) Rural Conference. Cessnock March -April 2011

Downs, S. Mobility Outcomes at 2 small hospitals in the Mid North Coast of NSW. NSW Rural Allied Health Conference. Port Macquarie November 2011

Downs, S. The Berg Balance Scale has high intra- inter-rater reliability, but absolute reliability varies across the scale: a systematic review. Australian Physiotherapy Association Conference. Melbourne October 2013

# Table of Contents

Declaration......................................................................................................................... i  
Supervisors............................................................................................................................ ii  
Acknowledgements................................................................................................................. ii  
Publications and presentations................................................................................................. iii  

Abstract: ................................................................................................................................ 1  
Chapter 1: Introduction ........................................................................................................... 3  
Chapter 2: Clinical outcomes using the Berg Balance Scale  
Introduction............................................................................................................................. 17  
Publication............................................................................................................................... 18  
Chapter 3: Reliability of the Berg Balance Scale:  
Introduction............................................................................................................................. 34  
Publication............................................................................................................................... 35  
Chapter 4: Normal values of the Berg Balance Scale  
Introduction............................................................................................................................. 52  
Publication............................................................................................................................... 54  
Chapter 5: Conclusion............................................................................................................. 70  
Appendices  
Appendix 1: Search strategy for reliability study ................................................................. 76  
Appendix 2: The Berg Balance Scale ...................................................................................... 77  
Appendix 2: Copyright permission......................................................................................... 79  
Appendix 3: Ethics approval ................................................................................................. 80  
Appendix 4: Consent and participant information forms...................................................... 86  
Appendix 5: Papers as published........................................................................................... 88
Abstract

This thesis by publication examines the Berg Balance Scale (BBS), a balance assessment tool used globally by clinicians and researchers for varied clinical and non-clinical populations. Despite its extensive application, important aspects of the utility of the BBS remain to be determined. This thesis evaluated three of these properties: validity when applied in a rural, clinical population, reliability evaluated from previously published literature, and normative values as measured in the healthy aged.

The first published study investigated the BBS in the context of two rural hospitals. It aimed to describe predicted and measured balance changes in patients receiving physiotherapy in two rural hospitals and to explore the relationship between patients’ measured balance at discharge, carer availability and patients’ discharge destination. A strong relationship was discovered between the BBS at discharge and the probability of discharge to nursing home. The degree to which the balance of patients changed from the commencement of their physiotherapy intervention to their point of discharge displayed high variability and was very difficult to predict.

The study attempted to identify information available to hospitals at admission which might predict change in balance. The presence or absence of several diagnoses and health issues was recorded at admission. Although 13 such potential predictors were recorded, none of them was shown to be useful in predicting change in balance. Treating physiotherapists were asked at commencement of therapy to predict how much change in patients’ measured balance they anticipated and their estimates were accurate.

The findings of this study are valuable because they suggest that balance, as measured by the BBS is an important predictor of the need to enter nursing home care. These results also suggest that discharge planners and aged care assessment teams should give credence to the treating physiotherapist’s estimate of probable change in balance when planning hospital discharge destinations.

The second published study used a systematic review and meta-analysis of the published literature to assess the relative and absolute reliability of the BBS. It aimed to determine the inter rater and intra rater relative reliability and the absolute reliability of the BBS. The relative reliability of the BBS was found to be high. The pooled estimate of the relative intra-rater reliability of 0.98 (95% CI 0.97 to 0.99) and the pooled estimate of the relative inter-rater reliability was 0.97 (95% CI 0.96 to 0.98). The absolute reliability of the BBS was found to vary across the scale. The minimal detectable change with 95% confidence varied between 2.8/56 and 6.6/56, with the highest reliability found when the
BBS had scores near 56/56. No data was found describing the absolute reliability in the range 0-20/56. While the BBS has acceptable reliability, it may not detect a modest, clinically relevant change. These results are important because they provide guidance for clinicians using the BBS about whether a measured change in a balance score might be confidently interpreted as real change or whether it is likely to be measurement error.

The third study, accepted for publication used a systematic review and meta-regression analysis to assess the relationship between BBS and the age of healthy, community dwelling elderly people. It aimed to find the mean BBS scores of healthy elderly people, and how much the BBS scores of healthy elderly people vary with age. The findings of this analysis are robust for healthy elderly aged 70-80 years, being based on data from 1363 participants. Data analysis for people aged over 80 years is based on only 258 participants. The BBS demonstrated a ceiling effect when used in healthy people aged 70 years and younger, but declines at the rate of 0.75/56 points per year. As the age of healthy elderly people increases the variability in their BBS scores also increases. These findings provide guidance as to when the BBS should, and should not be used and offers a perspective on the meaning of BBS scores in elderly people and provide important normative information to allow clinicians to accurately determinate balance related to age.

In conclusion, this thesis provides new and robust insight about how balance, as measured by the BBS, relates to the ability to live in the community. Furthermore, it provides a perspective on whether a measured change in BBS is meaningful or not. Finally the research provides normative data on the BBS, to guide its appropriate implementation and interpretation of scores when applied to the elderly.
Chapter 1: Introduction

Australia and other comparable countries around the world are experiencing ageing of their population, with increasing life expectancy. A negative consequence of increased life expectancy is the accompanying increased disability associated with having a higher proportion of population of advanced age. In the 2009-2010 financial year an estimated 83800 people aged 65 and older presented at Australian hospitals due to injurious falls (Bradley 2013). Structured interviews with elderly fallers found that 73.7% reported either a trip, slip or loss of balance as the reason for falling (Lord et al 1993), suggesting that poor balance is likely to be a significant risk factor for falls.

In Australia, ischaemic heart disease, stroke, dementia, chronic obstructive pulmonary disease, type II diabetes, and falls are among the 10 leading burdens of disease in Australians aged 75 and older, as measured by disability adjusted life (Australian Institute of Health and Welfare 2007). It has been estimated that in the 2006-2007 financial year 16412 hip fractures occurred in Australia, with risk of hip fracture increasing with increasing age (Crisp et al 2012). These figures highlight the current and escalating importance of monitoring and maintaining safe mobility in our older citizens which includes the evaluation and maintenance of adequate balance, one important measure of which is the Berg Balance Scale (BBS) (Berg et al 1989).

In the 2012-2013 financial year the Australian Government spent over 9.3 billion dollars on residential and community aged care (Australian Department of Health and Ageing 2013a). In addition to financial costs, quality of life costs associated with entering residential aged care must also be considered since few older people want to enter residential aged care (Salvage et al 1989). A recent study (Heydari et al 2012) compared the quality of life in nursing home residents to that of nearby community dwelling elderly. Multivariate regression analysis found a lower quality of life in nursing home residents, even after adjusting for several potential confounding factors.

The Australian government funds services in the home, including assistance with domestic tasks, and personal care tasks such as showering (Australian Department of Health and Ageing 2013b), with the goal of providing care in the community rather than in residential aged care facilities. Government policy in Australia has been to increase the proportion of this care to enable more older Australians to continue living in the community. This can be seen in the number of long term community care packages funded by the Australian government, which have increased from 2532 care packages in the community in 1995 (compared to 134810 residential care places) to 51467 care packages in the community (compared to 182850 residential care places) (Australian Institute of Health and Welfare 2011). Although substantial, this support for in home care does not include a paid carer with 24 hours per day availability. The absence of this 24 hour service in the community has the potential to
limit the ability of community based care to maintain older people in their own homes making admission to residential aged care unavoidable in many cases.

Dependency in activities of daily living has been found to be the single biggest risk factor for entry to Australian residential aged care facilities (Kendig et al 2010). The Australian aged care guidelines (Australian Department of Health and Ageing 2006) identify dependency in tasks such as mobility, bathing, grooming, toileting, continence and dressing as reasons for entry to residential aged care. All of these tasks require adequate balance ability.

The cost of falls in Australia, in both of quality of life and financial terms is high, due in part to the large number of recorded injurious falls (Crisp et al 2012). A large number of elderly Australians require community or residential aged care (Australian Institute of Health and Welfare 2011). The large amount of aged care provided to Australians suggests a high level of disability among older Australians. If even a small proportion of falls and disability related to poor balance could be prevented, the cost savings to the Australian taxpayer and the benefits of improved quality of life for elderly Australians would be substantial.

**Falls Risk**

Although there are several published studies that have investigated the relationship between the BBS and prospective risk of falling, this thesis does not aim to conduct a systematic review to investigate any such relationship. There appears to be three main complicating factors which make such a review difficult.

Firstly, well conducted studies using tools other than the BBS have found a non-linear relationship between mobility and falls (Barker et al 2012) and between balance and falls (Lord et al 2003). These studies found the lowest risk of falls in people with the highest level of disability, a moderate risk of falls in people with the lowest level of disability, and the highest risk of falls in people with a moderate level of disability. These findings suggest that people with the highest level of disability might be unable to attempt activities such as walking and might therefore experience fewer falls. A systematic review investigating the relationship between the BBS and falls would have to consider the possibility that there may be a non-linear relationship between the BBS and falls.

Secondly, studies comparing the relationship between the BBS and falls have differing measurement outcomes, including incidence of falls (Hall et al 2001), multiple fallers compared to non or single fallers (Mackintosh et al 2006), and fallers compared to non-fallers (Saverino et al 2006, Wrisley and Kumar 2010). The length of time over which falls are recorded also varies, with 12 and 6 month follow up time periods being the most common.
Thirdly, studies consider subjects with very different BBS scores, with some studies (Ersoy et al 2009, Hall et al 2001, Muir et al 2008, Wrisley and Kumar 2010) having mean balance scores above 50/56. Studies with a mean BBS close to 100% might be affected by a ceiling effect, and might potentially underestimate the ability of the BBS to predict falls when used in groups where the BBS is not affected by a ceiling effect.

Balance and Ageing

Balance is a multifaceted process involving the reception and integration of sensory inputs and the planning and execution of movement. The location of the centre of gravity, the base of support, the limit of stability, the surface conditions, the visual environment and the intentions of the task, all produce changing demands on the systems which control balance (Umphred et al 2013). Balance can be defined as the ability to maintain the centre of gravity within the base of support in a specified posture, such as standing or sitting, while voluntarily moving, and while being disturbed such as being pushed (Pollock et al 2000). It essentially involves controlling the body’s position in space for the dual purposes of stability/orientation and function.

Age related diseases can incur balance deficits requiring management by physiotherapists in different ways. Many elderly people experience multiple coexisting chronic illnesses (Marengoni et al 2011), which can potentially make the balance deficits and their physiotherapy management much more complex. Even in the absence of any disease the process of sarcopaenia (age related loss of muscle mass and strength) can be expected to reduce muscle strength with increasing age (Ryall et al 2008), and with sufficient loss of muscle strength, balance will be affected.

This thesis seeks to inform the use of the BBS, both among clinicians and researchers by providing information about the reliability, normal values, and validity of the BBS. This thesis does not include a specific section for a literature review, as two of the three already published studies are themselves systematic reviews.

Rationale

The first study examines outcomes from two small rural hospitals. The baseline and discharge BBS scores of whose patients were recorded. Potential predictors of change in BBS scores were considered and the relationship between BBS scores and discharge destination was investigated. The second study considers the reliability of the BBS by conducting a systematic review and meta-analysis. The third study considers the normal values of the BBS in healthy, community dwelling
elderly people by conducting a systematic review with meta-regression analysis, considering the relationship between age and BBS scores.

**Issues to consider when using balance assessment tools**

**Floor and Ceiling Effects**

The floor effect occurs when a significant proportion of a tested population achieves the lowest possible score on a test, similarly the ceiling effect occurs when a significant proportion of a tested population scores the highest possible score on a test (Everitt 2010). A balance test might have a floor effect when used to measure the balance of people with very poor balance or a ceiling effect when used to measure the balance of people with very good balance. A balance test with a floor or ceiling effect can be expected to have difficulty in both detecting changes in individual balance and in detecting differences in balance between individuals.

Floor and ceiling effects are an important consideration when selecting any balance test, both for clinicians and researchers. The degree to which a balance test is affected by the floor and ceiling effects depends on the balance of the people being tested.

**Validity**

The validity of a measurement has been defined as the “degree to which a useful (meaningful) interpretation can be inferred from a measurement” (Rothstein et al 1991). There are different forms of validity. The content validity of a balance test can be defined as the extent to which the test measures balance. Criterion-based validity has three forms: concurrent validity, predictive validity and prescriptive validity. Concurrent validity describes the extent to which a balance test is supported by evidence obtained at approximately the same time. Predictive validity describes how well a balance test can predict future events or conditions. Proscriptive validity describes how useful a balance measurement is in predicting the outcomes of therapy or other intervention.

The BBS was developed be asking clinicians and patients to identify testing components they considered important, which might potentially result in strong criterion-based validity, while potentially resulting in weaker content validity than designing a balance test by consulting a bio mechanist to test all aspects of balance.
A valid balance tool will correlate with outcomes relevant to patients. One outcome highly important to many older people is the ability to remain living in the community and to remain out of residential aged care facilities.

As previously discussed, the decision to enter a high care residential aged care facility (nursing home) costs the Australian taxpayer a substantial sum of money and can have major implications for a person’s life style. Demonstrating a relationship between low BBS scores and a higher probability of entering nursing home would provide important evidence of the concurrent validity of the BBS. Quantification of a relationship between BBS and probability of nursing home placement could potentially improve the targeting and measurement of therapy designed to allow people to remain living in the community by improving their balance.

Chapter 2 of this thesis compares the relationship between BBS and probability of admission to nursing home at discharge from hospital, which has important implications for the concurrent validity of the BBS.

Reliability

Reliability refers to the reproducibility of a test. The absolute reliability of a balance test refers to the how much difference is likely when the test is repeated in a subject in the absence of real change in balance. Absolute reliability is useful for clinicians interpreting changes in balance and differences in balance between patients. This is usually expressed as minimal detectable change with 95% confidence (MDC95). Intra-rater reliability describes the reproducibility of a balance score when tested and retested by the same assessor. Inter-rater reliability describes the reproducibility of a balance score when measured by different assessors. Relative reliability compares the variation in scores due to measurement error with the variation within a population. Published papers considered by this review commonly use this measure of reliability, expressed as intra-class correlation (ICC) where a score of 1 represents perfect agreement and a score of 0 represents no relationship. While relative reliability is potentially useful when comparing the reliability of the BBS with the reliability of other measurements, it is dependent on variability within the study sample and is less clinically useful than absolute reliability. Studies of populations with a high degree of variability in the measurement under consideration can be expected to find a very high relative reliability, even when the test not sufficiently reliable to detect clinically important changes (Bland and Altman 1986).
When considering reliability intra-rater and inter-rater reliability are both important. Intra-rater reliability of the BBS in the literature is established by having an assessor measure balance and then repeating their measurement on the same person after a specified time. Inter-rater reliability can be established either by repeated measurement by different assessors or by having one assessment of the BBS test and assessments by other people, in person or by videorecording. Repeated measurements have the disadvantage that a person’s underlying balance might change between two measurements and may therefore potentially underestimate the actual reliability of the BBS.

Simultaneous testing of the BBS to measure inter-rater reliability has different disadvantages. The verbal instructions for administering the BBS have been clearly defined, yet several aspects of non-verbal assessor behaviour may affect the reliability of the participants’ performance and ultimate score. For example, the assessor may provide a demonstration on how to perform balance tests and this may vary between assessors. This can be partially overcome by simultaneous BBS testing, either in person or by video. This method can assess the reliability of how different assessors interpret a subject performing the BBS, but will not detect differences in how assessors instruct subjects to perform BBS testing and may therefore overestimate the actual reliability of the BBS. Another issue is the subjective manner in which assessors determine whether to provide supervision. Most components of the BBS record a reduced score if the subject requires supervision, and having an assessor stand very close to a subject while performing balance testing might demonstrate that they believe supervision is required. Furthermore, safety considerations may lead some assessors to not even attempt components of the BBS which other assessors might consider safe.

**Normative Data**

Knowledge of normative data are important when interpreting the results of any balance tool. Knowledge that a person has worse balance than expected from healthy people of the same age can assist the identification and effective management of balance problems. Normative data can help avoid floor or ceiling effects when selecting a balance tool to screen populations or assess the effect of interventions. Interventions to improve balance can be validly assessed by comparison to normative data related to balance of healthy elderly people in specific age cohorts.

Knowledge of normal balance values allows clinicians to set appropriate goals. Clinicians accustomed to working with balance impaired people may easily underestimate normal balance values of healthy elderly on the basis of their experience with balance impaired people, and so set inadequate
rehabilitation goals. However, the normative range of balance, as measured by the BBS in healthy elderly people in specific age cohorts is yet to be determined.

**Balance assessment tools**

As discussed earlier Australia, like many other countries is experiencing an increase in the ageing population, and therefore has a greater proportion of people experiencing disability and falls, partly related to poor balance. Rehabilitation interventions can be cost effective when they enhance patients’ functional ability and capacity for independence. At the centre of this restoration of independent functioning is the ability to mobilise, transfer and perform activities of daily living – all requiring substantial ability to balance. The importance of balance to function and participation cannot be overstated. Balance impairments negatively impact function and often lead to restricted activity levels and reliance on aids and/or the support of others. Poor balance can result in falls, often causing serious injuries. To avoid these consequences it is imperative for therapists to methodically assess balance and interpret the findings before implementing a management strategy.

Due to the complexity of balance there is no single test that can adequately measure all the components of balance and different tests measure different facets of postural control. Balance tests can be classified by type and include the following: static (quiet) standing, dynamic (active) standing, and functional scales involving whole-body movement tasks. In all, there are over 40 published balance scales from which therapists must choose the most suitable for each individual client.

Fallscreen (Lord and Clark 1996) and a shortened version, Quickscreen (Tiedemann et al 2012) are tools developed in Australia, designed to predict the risk of falls and containing substantial balance components alongside measurement of other important falls risk considerations such as visual acuity. At times peoples’ ability to function independently, and remain living in the community might be a higher priority than their risk of falling. Due to the established links between balance and function, a comprehensive balance evaluation should include performance of functional tasks to identify those activities that need to be addressed in therapy. While balance and mobility tests have been developed, many have been validated using the BBS as “gold standard” (Connelly et al 2009, Creel et al 2001, Desrosiers et al 2005, Ditunno et al 2007, Linder et al 2006, Podsiadlo and Richardson 1991). A small number of balance and mobility tests predominate in the peer reviewed literature, in addition to the BBS. The Dynamic Gait Index (Shumway-Cook et al 1997) examines how a person responds to various challenges to balance such as turning one’s head and navigating
between obstacles, while walking. The functional reach test (FRT) (Duncan et al 1990) measures how far a person can reach forward while standing and maintain balance. It was developed to measure balance in the context of a person performing voluntary movement, while at the same time being quick and clinically convenient. The timed up and go test (TUG) (Podsiadlo and Richardson 1991) measures how far a person can stand up, walk at a safe and comfortable pace three metres, walk back and sit down. It was developed to be quick and practical to administer. The Tinetti test (Tinetti et al 1986) has static balance and walking balance components, each consisting of several items. The usefulness of any balance test will depend on its purpose and the population it is used to measure. As discussed earlier, floor and ceiling effects are important considerations in deciding what test to use.

The therapist must consider the level of challenge the tests provide in conjunction with the clients level of ability, the time taken to perform the test, the number of staff who must be present, the space and equipment required, and the cost and practicality of the test (Umphred et al 2013). Tests are only useful when subjects can make a meaningful attempt to complete them, so tests of balance in the context of walking such as the TUG and the Tinetti walking test which measure balance in people while walking are only useful for subjects able to walk. When considering all of these factors, the BBS rates well as the 14-items are ranked from relatively easy to more challenging tasks and it requires minimal equipment or space. The BBS is quick and easy to administer although it does require more time than extremely brief tests such as the FRT. The BBS also tests a variety of types of balance, and with different degrees of challenge, particularly for subjects unable to walk, for whom many other tests of balance will not be useful.

When comparing the BBS with other tests which measure balance one apparent limitation appears to be the lack of a walking component. Although the BBS contains tests of transferring between chairs, turning around on the spot, alternatively placing feet on a step and placing one foot in front of another, neither the speed of walking nor the quality of gait is measured. Patients with balance and mobility problems related primarily to poor dynamic motor control and less related to poor muscle strength might therefore have their balance more usefully measured by other balance tests.

Another apparent limitation of the BBS is a lack of highly challenging items. The two most challenging items in the BBS are standing on one leg, with the maximum score achieved after ten seconds, and standing with the heel directly in front of the toe, with the maximum score achieved after holding this posture for half a minute. Fallscreen and Quickscreen contain items measuring sway which record measurements of sway even in young, healthy people. A study of normative values (Isles et al 2004) measured balance using tools including the TUG and FRT in healthy women.
aged from 20 years to 70 years. This study clearly showed that balance deteriorates with age. This suggests that the FRT is unlikely to be affected by a ceiling effect and the TUG is unlikely to be affected by a floor effect when used with elderly people.

Another limitation of the BBS is that it only tests balance and does not measure other risk factors for falls such as poor visual acuity. Clinicians engaged in Falls prevention seeking to address risk factors other than poor balance might consider using tests such as Fallscreen or Quickscreen in place of the BBS or alternatively might consider testing balance with a tool such as the BBS and screening for other risk factors with separate, specific tests.

**Development of the Berg Balance Scale**

The BBS is a tool originally developed by physiotherapists to measure balance in a geriatric rehabilitation ward (Berg et al 1989). The first phase of its development was to interview professionals, including physical therapists, physicians, occupational therapists and nurses, to establish specific balance tasks they believed stressed a person’s balance and specific balance tasks they used to assess their patients. Patients were asked open ended questions about movements they felt made them unsteady, as well as circumstances of past falls. The second phase tested 33 balance items developed from the first phase on patients. Professionals identified items they believed were not helpful. Other items were discarded because they appeared not to discriminate well between patients. The third phase of the development of the BBS videotaped testing of the remaining 22 items on patients, with feedback from professionals discarding six items. The fourth phase of development assigned scores ranging from zero to four based on how patients had performed while having their balance tested in earlier phases. The remaining 16 items were tested by five experienced physical therapists not involved in development of the scale on patients selected to represent the full spectrum of balance in the ward, with the measurements videotaped. Relative inter-rater reliability was determined by having all five physical therapists score all tests. Two items with inadequate reliability were discarded, with the remaining 14 items, each scored between zero and four for a total of 0 to 56 constituting the BBS.

Since its development, the BBS has been a widely and frequently used tool by clinicians (Korner-Bitensky et al 2006, Stokes and O’Neill 2008). The extent to which the BBS has been studied and used for research is illustrated by the second of the systematic reviews presented in this thesis, which found a total of 859 articles published in the peer reviewed literature containing the term “Berg Balance Scale”.

11
The BBS was chosen as the focus of this study partly because of the wide range of balance ability it can measure. Most other tests of balance require subjects to be able to stand independently, while the BBS can be used to measure the balance of a range from subjects who are restricted to chair sitting, right up to those who are capable of single leg standing balance. The wide use of the BBS was the other reason it was the focus of this study.

Summary

This thesis seeks to describe important aspects of the validity of the BBS by describing the usefulness of the BBS in two small rural hospitals, principally describing the relationship between BBS at discharge and probability of discharge to nursing home. This thesis also uses systematic reviews to explore and describe the reliability of BBS and to establish normative data describing BBS scores of healthy elderly people.
References:


Chapter 2: Clinical outcomes using the Berg Balance Scale

Introduction

This study describes the baseline and discharge BBS scores of in-patients in two small hospitals on the Mid-North Coast of NSW. This study describes a strong relationship between discharge BBS and probability of discharge to nursing home, supporting the concurrent validity of the BBS. The highly variable changes in BBS experienced by in-patients and difficulty predicting these changes suggest that hospitals should be careful to avoid prematurely discharging patients to nursing home. Physiotherapists were asked at admission to estimate how much their patients balance would change. The accuracy of these estimates suggests that the discharge planning process from hospital should give credence to the treating physiotherapists estimates of how much their patients balance will change.

Ethical considerations

Ethics approval was granted by the North Coast Area Health Service human research ethics committee and by the University of Newcastle human research ethics committee (see appendices three and four).

Published with permission of Wiley-Blackwell
Abstract:

Objectives: To describe predicted and measured balance changes in patients receiving physiotherapy in two rural hospitals and to explore the relationship between balance at discharge, carer availability and patients’ discharge destination.

Design: Prospective measurement study.

Setting: Two rural Australian hospitals.

Participants: Eighty nine inpatients with a median age of 84.

Main outcome measures: Berg Balance Scale (BBS) on admission and the treating physiotherapist’s estimate at admission of individual patient’s discharge BBS. Follow up measures included: Discharge BBS, carer availability after discharge and patient discharge destination.

Results: Although change in measured balance of study participants had wide variability, balance measured by the BBS displayed a statistically and clinically significant improvement. A strong relationship was found between balance scores and discharge destination. However, no relationship was found between carer availability and discharge destination. Physiotherapists’ estimates of discharge BBS displayed an average error of 7/56.

Conclusions: The strong relationship between measured balance and discharge destination in these elderly study participants suggests that maximising their balance might minimize admissions to nursing home. The high variability of measured balance change suggests outcomes are difficult to predict. The study results suggest premature assessment of patient’s suitability for nursing home placement should be avoided. The accuracy of physiotherapist’s estimates of discharge BBS suggests greater weight might be placed on their input to facilitate the discharge planning process.

Keywords: Berg Balance Scale; Discharge planning; Care of the elderly

What is already known on this subject:

- An ageing population means that accurate assessment of the potential of elderly people to improve function will become increasingly important.
- Elderly people may experience better functional outcomes from small hospitals.
- Little objective data are available to inform discharge planning decisions from small hospitals.
What this paper adds:

- A strong relationship has been described between Berg Balance Score at discharge and probability of discharge to nursing home, emphasising the importance of maximising elderly people’s balance.
- Elderly people in rural hospitals are likely to have improved balance during the period of their admission.
- Change in balance is highly variable and difficult to predict. The best predictor of change in balance available at admission was the treating physiotherapist’s estimate of change.

Introduction:

The proportion of the population aged greater than 85 years in Australia is expected to increase from 1.8% in 2010 to 3.0% in 2030 (Australian Bureau of Statistics 2008). Along with this will be a significant increase in the number of elderly people with multiple health problems including poor mobility. There is an apparent need to provide healthcare for people who are unable to participate in rehabilitation programs, but who are not acutely unwell enough to require intensive medical management. In Australia much of this care, appears to be provided by hospitals, especially smaller hospitals. Such hospitals with between 10 and 50 beds are mostly in rural areas and provide 14% of Australia’s hospital beds (Australian Institute of Health and Welfare 2010). Understanding the movement of patients between hospital and the community, hostel and nursing home is important in understanding the role of these hospitals.

While the two hospitals in this study provide rehabilitation interventions, they cannot be called rehabilitation units since they do not provide access to specialist rehabilitation medical practitioners nor the amount of physiotherapy, occupational therapy and speech therapy services reasonably required of such units (Australasian Faculty of Rehabilitation Medicine 2011). In Europe, community hospitals provide care comparable to that provided by smaller rural hospitals in Australia such as those in this study, which have 32 and 42 general care beds respectively (see Table 2.1). Seamark et al surveyed 471 smaller community hospitals in the United Kingdom with a total of 18,579 beds (median of 33 beds) found over 75% of available beds were described as medical and elderly care beds. Care in 45.5% of the beds was led by general practitioners and all hospitals provided physiotherapy. Two randomised controlled trials with participants of comparable age and diagnoses to the patients in this study’s rural Australian hospitals (Garåsen et al 2007, Young et al 2007) found that elderly people assigned to care in community hospitals had better independent living status,
fewer hospital readmissions and less use of community care six months after discharge compared to similar patients assigned to acute-care hospitals

Elderly people often have multiple health issues that frequently include problems with deteriorating balance and mobility. Physiotherapists are often required to assess these facets of health. Overly pessimistic estimates of balance change may result in inappropriate recommendations for placement in a residential facility, while overly optimistic estimates may result in inappropriately prolonged hospital stays. Hence the importance of accuracy, reliability and validity in physiotherapy assessment cannot be overstated.

Berg Balance Scale (BBS)

The BBS was developed (Berg et al 1989) primarily for measuring the functional balance of elderly people. It consists of 14 components each scored between 0 and 4 to give a composite score out of 56. It has been shown to have a high relative inter-rater and intra-rater reliability (Berg et al 1995, Cattaneo et al 2007, Donoghue et al 2009, Gan et al 2008, Liaw et al 2008, Mao et al 2002, Sackley et al 2005, Wirz et al 2010). The absolute reliability of the BBS in elderly people has been found to have a minimal detectable change with 95% confidence (MDC95) for people with a BBS between 0-24 to be 4.6 and the MDC95 of people with a BBS between 25-34 to be 6.3 (Donoghue et al 2009). Most prospective studies have shown the BBS to be valuable for predicting who will fall (Hall et al 2001, Li et al 2004, Muir et al 2008).

A person’s capacity for mobility, bathing, grooming, toileting, continence and dressing are important considerations in assessing eligibility for admission to nursing home (Australian Department of Health and Ageing 2006). Adequate balance is required for all of these activities. The BBS tests multiple balance tasks, suggesting the possibility that BBS might able to predict admission to nursing home.

Aims

This study aims to answer the following questions for mobility limited patients in two small rural Australian hospitals: How does functional balance change between initial physiotherapy intervention and discharge from hospital? How accurate is the change in balance predicted by the physiotherapist? What relationship exists between balance as scored by the BBS and discharge
destination? Is discharge destination related to availability of a patient’s carer? Is change in balance related to patient’s diagnosis at admission?

**Method**

*Participants:* All inpatients receiving physiotherapy related to mobility at two rural Australians hospitals between January and November 2010 were given a copy of the study information sheet, allowed time to make an informed decision when invited to join the study.

*Exclusion criteria:* patients were excluded from the study if they were: less than 16 years old, orthopedically unable to fully weight bear on both legs, medically unfit to undergo balance testing, unable to understand instructions for the purpose of balance testing, unable to provide informed consent, or expected to have a length of stay less than two days.

*Baseline measurements:* Patient demographics, gait status and presence of clinically relevant health issues were recorded. Recorded demographic information included: age, gender and preadmission residence. Residents of aged care facilities receiving low level care were considered to be hostel residents, while those receiving high level care were considered nursing home residents. Participants not residing in an aged care facility were considered to be community dwelling. Availability of a carer able to provide at least basic support with mobility for 12 or more hours per day after discharge, and the date of initial physiotherapy intervention for mobility were also documented. Medical conditions that either affected mobility at admission, or were a significant reason for patients’ admission to hospital were considered to be clinically significant and were recorded.

Baseline and discharge BBS measurements were tested by a physiotherapist experienced in aged care physiotherapy and familiar with use of the tool but who was not the treating physiotherapist. The well documented BBS measurement protocol was followed (Berg et al 1989). When any component of the BBS appeared to risk injury to the patient, it was not undertaken and a score of 0/4 was assigned.

The BBS was measured within 2 days of the initial physiotherapy intervention for mobility. The mean difference between date of initial physiotherapy and baseline BBS measurement was less than one
day (0.6 days). The baseline BBS was shown to the treating physiotherapist who then estimated what the patient's discharge BBS was likely to be.

**Discharge measurements:**

The discharge BBS was measured within 4 days of discharge. This usually occurred on the ward but if necessary a home visit was conducted. The mean difference between date of discharge and discharge BBS measurement was less than one day (0.8 days). The discharge destination was recorded in terms of discharge to a nursing home, hostel or the community. The date of discharge and the number of physiotherapist mobility interventions were recorded based on medical record entries and discussion with the treating physiotherapists.

**Statistical analysis** Statistical analysis was performed using STATA version 11. Two tailed t-tests were used to compare change in BBS by condition. Fisher’s exact test was used to compare the proportion of people with or without a carer, who were discharged to nursing home with those who were not. Logistic regression analysis was used to investigate the relationship between BBS and probability of nursing home placement. Spearman’s rho was used to investigate the relationship between the number of days under physiotherapy care and change in BBS. A two sample Wilcoxon’s rank sum test was used to compare the days under physiotherapy care of participants discharged to nursing home and those not discharged to nursing home. The sample size of 89 was adequate to answer most study questions. However, power analysis suggested a sample size of approximately 400 would be required to determine a relationship between health conditions and change in BBS, considering a 6/56 change in BBS as clinically significant.

**Results**

Eighty nine participants completed the study (see Figure 2.1) with a median age of 84 (see Figure 2.2). Forty seven women and 42 men completed the study. Most participants had multiple clinically significant conditions (see Table 2.2). In no case was the presence of these conditions able to predict a statistically significant difference in change in balance.

The BBS changes predicted by physiotherapists were compared with the actual BBS changes at discharge. In 6/89 cases the treating physiotherapist did not make initial predictions of BBS change.
Physiotherapists did not appear to systemically underestimate or overestimate change in balance (95% CI -3.1-0.3), the average magnitude of error physiotherapists made was 7.0/56 points on the BBS (SD± 6.5).

Mean change in BBS between baseline and discharge measures was significant at 8.47 (95% CI 6.28 – 10.65 p <0.0001). The two hospitals did not have significantly different changes in BBS (p=0.45 95% CI -2.4-6.9)

Participant preadmission and discharge destinations are presented in Table 2.3. The majority of participants resided in the community prior to admission, with the majority returning to their homes on discharge (79%). While basic mobility support by a carer was available after discharge for 42 participants, support was not available for 39 patients, and 8 patients were resident in a hostel or nursing home prior to admission. There was no relationship between availability of a carer and discharge destination (p=0.47).

**Relationship between BBS and discharge destination**

At discharge the mean BBS of participants discharged to nursing home was 13.5 (SD± 9.9), while the mean BBS of participants discharged to hostel or to the community was 38.5 (SD ±10.0) and 35.3 (SD ±13.0) respectively

Logistic regression analysis exploring the relationship between BBS at discharge and discharge destination revealed that for a 1 unit increase in final BBS, the relative risk of being discharged to a nursing home decreases by 13% (RR 0.87, 95% CI 0.81-0.93, p<0.001) compared to being not being discharged to a nursing home (see Figure 2.3).

**Relationship between number of days under physiotherapy care and change in BBS**

Several participants had prolonged hospital stays while waiting for an available nursing home place, long after the decision to enter nursing home was made. Participants discharged to nursing home had more time under physiotherapy care p=0.014). To control for this factor time under physiotherapy care was compared with change in BBS for patients not discharged to nursing home. Correlation coefficient showed a moderate positive association between improvement in BBS and days under physiotherapy care (Spearman’s rho = 0.32, p= 0.007, see Figure 2.4).
Discussion

The decision to place a person in a nursing home should only be made when people are medically stable and after rehabilitation is completed (Australian Department of Health and Ageing 2006). Accurate prediction of likely changes in BBS of in-patients, combined with substantive evidence related to the BBS of people admitted to nursing home may provide a valuable tool to assist in making timely and appropriate decisions for placement in nursing homes. Such evidence might also help hospitals focus resources where they are most effective.

Physiotherapy advice is frequently sought concerning the likely functional balance and mobility status of patients at their discharge from hospital. Treating physiotherapists were asked, at baseline, to estimate each patient's final discharge BBS. The average error in their predictions was 7.0/56 points. This measurement might underestimate physiotherapists’ accuracy estimating changes in balance as in many clinical situations physiotherapists’ BBS estimates are likely to be made closer to the discharge date than the estimates provided within this study. Comparison with the test-retest minimal detectable change with 95% confidence (MDC95) provides perspective on the extant average error of 7.0/56 within this study. The MDC95 of the BBS in a similar patient cohort has been found to vary between 4.6/56 and 6.3/56 (Donoghue et al 2009). The outcomes from the current study suggest physiotherapist’s estimates of functional balance at discharge are sufficiently accurate to be useful and therefore should be considered as part of the discharge planning process. The strong association between functional balance and nursing home placement highlights the importance of maximising the functional balance of elderly people.

While the change in BBS showed a wide variability, a positive correlation was established between improved BBS and the number of days under physiotherapy care for mobility. These findings suggest that care should be taken to avoid premature assessment of people in small hospitals as requiring nursing home when more time and physiotherapy intervention may lead to a change in outcome.
The most surprising finding was that having a carer available to provide basic mobility support was not associated with avoiding discharge to a nursing home. Wee et al (1999) found that having a carer was a far stronger predictor of returning home after hospital admission than the admission measures of BBS within a stroke rehabilitation unit. This suggests that the mid north coast of NSW has effective services available to elderly people in the community which allows many of them to return home despite disability and lack of a carer.

This study has a number of limitations, the first being that it was not a randomised control trial and therefore unable to establish the causes of improved balance. Secondly the physiotherapist measuring BBS was not blinded to whether the measurements were baseline or discharge, a potential source of bias, and finally this study did not follow patients after discharge from hospital, nor did it investigate other factors such as cognitive impairment which might contribute to nursing home placement.

Acknowledgements: We wish to acknowledge Alastair Merrifield, trainee biostatistician from the NSW centre for Epidemiology and Research.
References:


Table 2.1 Physiotherapy services at the 2 hospitals under consideration

<table>
<thead>
<tr>
<th></th>
<th>Hospital A</th>
<th>Hospital B</th>
</tr>
</thead>
<tbody>
<tr>
<td>General care beds</td>
<td>32</td>
<td>42</td>
</tr>
<tr>
<td>Full time equivalent physiotherapists</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mean rate of interventions received</td>
<td>1.6</td>
<td>3.0</td>
</tr>
<tr>
<td>(number per week)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Therapy assistant</td>
<td>None</td>
<td>Available for a part of the study time</td>
</tr>
<tr>
<td>Cover for physiotherapy leave</td>
<td>Limited availability</td>
<td>Reliably available</td>
</tr>
<tr>
<td>Physiotherapy on weekends and public holidays</td>
<td>Unavailable</td>
<td>Unavailable</td>
</tr>
<tr>
<td>Physiotherapy position vacancies</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Medical Care</td>
<td>Primarily provided by local general practitioners</td>
<td>Primarily provided by local general practitioners</td>
</tr>
<tr>
<td>Clinically significant Condition</td>
<td>Number of patients with this condition</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>----------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Dementia</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Cardiac, Respiratory, Vascular</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Infection</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Other Musculoskeletal Problem</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Delirium</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Other Neurological</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Lower Limb Joint Replacement</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Fractured Proximal Femur</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Palliative Care</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Fractured Pelvis</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.3: Participant preadmission status and discharge destinations

<table>
<thead>
<tr>
<th>Pre Admission status</th>
<th>Discharged to the community</th>
<th>Discharged to hostel</th>
<th>Discharged to nursing home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community (81)</td>
<td>64 (79%)</td>
<td>4 (5%)</td>
<td>13 (16%)</td>
</tr>
<tr>
<td>Hostel (5)</td>
<td>0</td>
<td>2 (40%)</td>
<td>3 (60%)</td>
</tr>
<tr>
<td>Nursing home (3)</td>
<td>0</td>
<td>0</td>
<td>3 (100%)</td>
</tr>
</tbody>
</table>
Figure 2.1 Participant flow

173 Potential participants
Identified by hospital physiotherapists

42 did not meet criteria
2 Acutely unwell
2 end stage palliative care
15 not fully weight bearing
9 too confused to follow instructions
14 discharge expected too soon for baseline and discharge measures to be reasonably expected to change

131 Approached

30 Declined

101 Enrolled

12 Lost from study
1 transferred to acute care hospital
1 too unwell on the day of discharge to allow BBS testing
7 lost to follow up
3 withdrew

89 Completed Study
Figure 2.2 Age distribution of participants

Figure 2.3 Probability of discharge to nursing home compared with final BBS
Chapter 3 – Reliability of the Berg Balance Scale

Introduction

Reliability is the first consideration when considering any clinical measurement tool. Reliability refers to the reproducibility of a tool. Ideally when a subject’s underlying balance is stable and unchanging their performance as measured on a balance test should not change. Knowledge of reliability is important when interpreting and selecting balance tests. A balance test which changes by less than the minimal detectable change cannot provide convincing evidence of real change.

This study considers the relative and absolute reliability of the BBS by means of a systemic review. It finds the BBS to have a very high relative reliability, but an absolute reliability that varies across the scale. This means that a large change in recorded BBS is unlikely to be the result of test unreliability, although modest changes which the preceding chapter suggests might be clinically significant are not always reliably detectable, especially towards the mid-range of the BBS. These findings help clinicians to interpret BBS results by determining if observed changes are likely to be caused by test unreliability. Clinicians considering using the BBS can compare the minimum detectable change established by this research with the minimum change they consider clinically relevant to help them determine if the BBS will be useful for patients with specific levels of balance.
ABSTRACT

Questions: What is the intra-rater and inter-rater relative reliability of the Berg Balance Scale? What is the absolute reliability of the Berg Balance Scale? Does the absolute reliability of the Berg Balance Scale vary across the scale?

Design: Systematic review with meta-analysis of reliability studies.

Participants: Any clinical population that has undergone assessment with the Berg Balance Scale.

Outcome measures: Relative intra-rater reliability, relative inter-rater reliability, and absolute reliability.

Results: Eleven studies involving 668 participants were included in the review. The relative intra-rater reliability of the Berg Balance Scale was high, with a pooled estimate of 0.98 (95% CI 0.97 to 0.99). Relative inter-rater reliability was also high, with a pooled estimate of 0.97 (95% CI 0.96 to 0.98). A ceiling effect of the Berg Balance Scale was evident for some participants. In the analysis of absolute reliability, all of the relevant studies had an average score of 20 or above on the 0 to 56 point Berg Balance Scale. The absolute reliability across this part of the scale, as measured by the minimal detectable change with 95% confidence, varied between 2.8 points and 6.6 points. The Berg Balance Scale has a higher absolute reliability when close to 56 points due to the ceiling effect. We identified no data that estimated the absolute reliability of the Berg Balance Scale among participants with a mean score below 20 out of 56.

Conclusion: The Berg Balance Scale has acceptable reliability, although it might not detect modest, clinically important changes in balance in individual subjects. The review was only able to comment on the absolute reliability of the Berg Balance Scale among people with moderately poor to normal balance.

INTRODUCTION

The Berg Balance Scale was developed in 1989 via health professional and patient interviews that explored the various methods used to assess balance (Berg et al 1989). Initially, 38 balance tests were selected as potential components of the score and then refined through
further interviews and trials to 14 items. Each of these items is scored from 0 to 4, which are summed to make a total score between 0 to 56, with a higher score indicating better balance. Although the Berg Balance Scale was originally developed to measure balance in the elderly, it has since been used to measure balance in a wide variety of patients.

All clinical measurement tools need to be reliable. Absolute reliability is clinically relevant and appears to be the most useful way of describing the reliability of the Berg Balance Scale (Bland and Altman 1986). The absolute reliability of the Berg Balance Scale provides a confidence interval, within which one can be confident that a change in balance is real change. The most common way of expressing this is the minimal detectable change with 95% confidence (MDC95). With regard to balance, intra-rater reliability refers to the reproducibility of a balance score when tested and retested by the same assessor. Inter-rater reliability refers to the reproducibility of a balance score when measured by different assessors. Relative reliability provides information about the variation in a score due to measurement error relative to variation within a population. This measure of reliability appears commonly in the literature, usually expressed as intra-class correlation (ICC) where a score of 1 represents perfect agreement and a score of 0 represents no relationship. Relative reliability provides perspective of the reliability of the Berg Balance Scale compared to other measurements, but is less useful clinically and is dependent on variability within the study sample. Studies of heterogeneous populations may find a very high relative reliability, even when the test is unable to reliably detect clinically important changes (Bland and Altman 1986). Three commonly used methods of calculating ICC are used, generally referred to as type 1, type 2 and type 3 (Shrout and Fleiss 1979). If a type 1 calculation is incorrectly used, the reported ICC is likely to be an underestimate. Use of a type 3 calculation is likely to result in a higher ICC, however type 3 calculations cannot be generalised validly to assessors not involved in the study (Shrout and Fleiss 1979).

The objective of this review was to summarise the available evidence for the reliability of the Berg Balance Scale across all age groups and conditions where the Berg Balance Scale was used as a balance measurement tool.
Intra-rater reliability is measured by having an assessor measure balance and then repeat their measurement of the same person after a specified time lapse. Inter-rater reliability can be measured either by repeated measures by different assessors or by one assessor performing the test and other assessors rating the test. In the case of the Berg Balance Scale, the second rating can be done either in person or by reviewing a videorecording. Repeated measurements have the disadvantage that a person’s underlying balance might change between two measurements and therefore may underestimate the actual reliability of the Berg Balance Scale.

Simultaneous testing of the Berg Balance Scale to measure inter-rater reliability has different disadvantages. The Berg Balance Scale instructions may be interpreted and delivered in slightly different ways by different assessors. Non-verbal components such as demonstrating how to perform balance tests may vary between assessors. Safety considerations may lead some assessors to not attempt components of the Berg Balance Scale that other assessors might consider safe to attempt. An assessor might stand very close to a subject while performing balance testing, and so demonstrate that supervision is required. Simultaneous Berg Balance Scale testing, either in person or by video, can assess the reliability of how different assessors interpret a subject performing the Berg Balance Scale, but will not detect differences in how assessors instruct subjects to perform Berg Balance Scale testing and may therefore overestimate the actual reliability of the Berg Balance Scale.

It is reasonable to speculate that the reliability of the Berg Balance Scale may vary for each of the test items and for different populations. For example, in healthy community-dwelling people, reliability might be affected by disagreement about how item 14 ‘standing on one leg’ is measured, while easier items such as item 3 ‘sitting balance’ might be expected to have almost complete agreement of 4/4 among assessments. Conversely, when applied to people with stroke who are unable to stand, the reliability of ‘sitting balance’ may be more affected, while more difficult tasks such as ‘standing on one leg’ are likely to be universally assessed as 0/4. An additional factor that might cause variation in the reliability of the Berg Balance Scale is the underlying health conditions of subjects whose balance is tested. Individual studies are unlikely to be able to investigate the Berg Balance Scale over the full range of the scale and over the broad spectrum of causes of disordered balance. This review describes the range of subjects in whom the reliability of the Berg Balance Scale has been studied, reporting both their balance as well as any underlying health condition.
A previous literature review of the Berg Balance Scale (Blum and Korner-Bitensky 2008) considered the relative reliability of the Berg Balance Scale in patients with stroke and found the Berg Balance Scale to have strong reliability. The current review covers important aspects of the reliability of the Berg Balance Scale not considered by the earlier review, including absolute reliability, and the reliability of the Berg Balance Scale in patients with conditions other than stroke.

Floor or ceiling effects occurs when a significant proportion of a tested population achieve the lowest or highest possible score on a test, respectively (Everitt 2010). In groups where the mean Berg Balance Scale is close to 0 or 56, the Berg Balance Scale is unlikely to be useful in discriminating between individuals and will exhibit floor or ceiling effects. In such cases the scale is unlikely to be able to detect a change in balance, even if there is a real change. While floor and ceiling effects can potentially impair the clinical and research usefulness of the Berg Balance Scale, they are also likely to inflate its absolute reliability. A person with extremely poor balance is likely to be uniformly rated at 0/4 on most elements of the Berg Balance Scale. Conversely, a person with extremely good balance is likely to be uniformly rated 4/4 on most items of the Berg Balance Scale. Floor and ceiling effects comprise groups with lower variability, which in turn lead to lower estimates of relative reliability compared to groups with more variable scores. Therefore, absolute and relative reliability should be interpreted with reference to floor and ceiling effects.

The specific study questions for this systematic review were:

1. What is the relative intra-rater and inter-rater reliability of the Berg Balance Scale?

2. What is the absolute reliability of the Berg Balance Scale, defined as the minimal detectable difference able to be determined with 95% confidence?

3. Does the absolute reliability of the Berg Balance Scale vary across the scale?

METHOD

Identification and selection of studies
A literature search was undertaken to locate eligible published studies. Electronic searches of Medline, CINAHL, Embase, and the Cochrane Library from 1980 to August 2010 were conducted using ‘Berg Balance Scale’ as a search term. No search terms were used for intervention type or health condition and no methodological filter was used for study design. See Appendix 1 for the detailed search strategy. All potentially relevant papers were identified from abstracts and assessed for inclusion. The reference lists of included studies were searched for additional relevant papers. Data were extracted from the included studies by two authors (SD and PC) with any disagreements adjudicated by a third author (JM).

The inclusion criteria for studies are presented in Box 3.1. Studies investigating the relative reliability of the Berg Balance Scale had to supply a confidence interval around the estimate of the reliability of the scale or data allowing a confidence interval to be calculated. A minimum sample size of 10 was also applied, as recommended by Walter et al (1998). Studies examining translated versions of the scale were included if the study was reported in English. Studies examining a modified or partial version of the scale were excluded. Studies that excluded people who were unable to attempt some items of the scale were excluded. Studies that used incorrect or unclear methods to calculate the intra-class correlation coefficient (ICC) and articles not containing original data, such as letters and reviews, were also excluded. Cognitive impairment initially was not a basis for excluding papers. However, only one paper studied people who predominantly had substantial cognitive impairment, so this paper was considered separately.

Assessment of the characteristics of the studies

The following data were extracted from each included study: the number of participants and their age, diagnosis, disease severity, and distribution of scores of the Berg Balance Scale. Any exclusion criteria applied in the original studies were also recorded.

Data analysis
Meta-analyses of the relative intra-rater and inter-rater reliability were performed. Confidence intervals were assessed at 95%. Sensitivity analysis was conducted on studies examining translations of the Berg Balance Scale by individually omitting studies, repeating the analysis and determining if results were significantly different without any study. If not specifically stated, it was assumed that studies conducted in predominantly non-English speaking locations used translations.

To calculate the relationship between absolute reliability and samples of Berg Balance Scale data, samples were weighted for sample size and the mean Berg Balance Scale was plotted against the MDC95. A quadratic line of best fit was used as the floor and ceiling effects can be expected to cause increased absolute reliability as the mean Berg Balance Scale approaches 0 or 56. Meta-analysis of absolute reliability was not conducted due to the confounding effect of sample mean Berg Balance Scale on MDC95.

**RESULTS**

**Flow of studies through the review**

511 papers were identified: 510 from electronic searches and 1 from reference lists. Based on information in the title and abstract, 27 were identified as being related to reliability. Of these, 15 were excluded, primarily for having inadequate detail about the methods or insufficient data to include in the meta-analysis. Eleven studies were included in analysis of the reliability of the Berg Balance Scale. The flow of studies through the review is presented in Figure 3.1.

**Characteristics of the studies**

Table 3.1 presents the characteristics of the included studies, including a description of the participants. The included studies included elderly patients (Donoghue et al 2009), elderly residents of an aged care facility (Berg et al 1995), patients with stroke (Liaw et al 2008, Mao et al 2002, Stevenson 2001), multiple sclerosis (Cattaneo et al 2007, Paltamaa et al 2005),

**Relative reliability**

The intra-rater relative reliability of the Berg Balance Scale was estimated by meta-analysing data from 3 studies with a total of 101 subjects. The pooled estimate of the intra-rater relative reliability of the Berg Balance Scale was 0.98, with a 95% CI of 0.97 to 0.99, as presented in Figure 3.2. A further analysis was conducted to examine the inter-rater relative reliability of the Berg Balance Scale by meta-analysing data from 5 studies with a total of 345 subjects. The pooled estimate of the inter-rater reliability was 0.97, with a 95% CI of 0.96 to 0.98, as presented in Figure 3.3. These studies included participants from a variety of clinical populations with balance abilities across the full spectrum of the Berg Balance Scale, although only one study had a sizeable number of subjects with very low Berg Balance Scale scores (Berg et al 1995).

Sensitivity analyses did not find evidence that translations of the Berg Balance Scale into languages other than English have different reliability to the English version. In all cases repeating the analysis omitting translations of the Berg Balance Scale changed the relative reliability by less than 1%. All papers used Shrout and Fleiss type 2 calculation to calculate ICC except (Berg et al 1995), which used type 1.

**Absolute reliability**

Studies investigating the absolute intra-rater reliability of the Berg Balance Scale show that the MDC95 varies in relation to the Berg Balance Scale of the sample, as presented in Figure 3.4. The review did not identify data about the absolute reliability of the Berg Balance Scale within its lower range of 0 to 20. Only one study examined the absolute inter-rater reliability of the Berg Balance Scale (Cattaneo et al 2007). This found very similar results for absolute intra- and inter-rater reliability.
Sensitivity analysis was conducted individually on all papers studying the absolute reliability of the Berg Balance Scale using translations. A Swedish translation studying the reliability of the Berg Balance Scale in residential aged care facilities with substantially cognitively impaired residents found a significantly lower absolute reliability with a MDC95 of 7.7 (mean Berg Balance Scale 30.1) (Conradsson et al 2007). These study findings were not included in our analysis of the absolute reliability of Berg Balance Scale. In all other cases the line of best fit with the individual study excluded was almost identical to the analysis presented.

DISCUSSION

Our review identified substantial and consistent evidence of high intra-rater and inter-rater relative reliability of the Berg Balance Scale. Absolute reliability data were also favourable, although some people might experience moderate change in balance that would not be reliably detected by the Berg Balance Scale. Furthermore, the absolute reliability data were only available for people with Berg Balance Scores above 20.

The reliability of the Berg Balance Scale has been investigated among a wide variety of subjects, although both studies investigating the reliability of the Berg Balance Scale in patients with Parkinson’s disease used subjects with high Berg Balance Scale which incurred a ceiling effect. The results of these studies might therefore be considered invalid in terms of describing the reliability of the Berg Balance Scale for patients with Parkinson’s disease whose balance scores are in the middle or lower range of the Berg Balance Scale.

This review found little evidence describing the reliability of the English language Berg Balance Scale in people with substantial cognitive impairment, although a Swedish language Berg Balance Scale translation (Conradsson et al 2007) suggests the Berg Balance Scale may be less reliable in people with substantial cognitive impairment.

While the high relative reliability suggests the Berg Balance Scale is clinically useful, there is little specific guidance as to how confident one can be that a real change in balance has occurred between tests across time for individual patients. This review suggests that if an individual has a Berg Balance Scale of between 20 and 56 and experiences a change of between 3 and 7 (see Figure 3.4), one can be 95% confident that there has been a real change in balance. Individuals may experience clinically relevant changes in balance which cannot be
reliability detected. Downs et al (2012) found inpatients with a Berg Balance Scale of 20 have approximately a 30% probability of being discharged to a nursing home, while those with a Berg Balance Scale of 25 have approximately 20% probability of being discharged to a nursing home, suggesting that a difference in balance which is only barely detectable with 95% confidence in any individual may in fact be highly clinically relevant.

Changes in the average Berg Balance Scale of patient or research groups have a smaller minimal detectable change than individual subjects. Thus, while moderately clinically important balance changes might not always be detectable with 95% confidence in individuals, they can be expected to be reliably detectable within groups. Researchers or clinicians who find clinically important changes in the average Berg Balance Scale of a group of individuals might therefore be confident that the change was not caused by random variation.

This literature review did not find data describing the absolute reliability in groups with very low Berg Balance Scale scores, although data presented by Cattaneo et al (2007) suggest that the absolute reliability of the Berg Balance Scale might be higher in the 0 to 20 range than the 20 to 56 range. Bimodal distribution of the Berg Balance Scale has been reported previously (Berg et al 1995, Downs et al 2012), suggesting subjects might be categorised into two distinct groups: those able to stand independently and those unable to stand independently. Where people were able to stand independently, they were also able to attempt and usually achieve a score on several items, generally achieving a Berg Balance Scale greater than 20. Those unable to stand independently are unable to attempt these items and usually score less than 15. The dichotomous nature of these two groups suggests that the absolute reliability of the lower Berg Balance Scale between 0 and 20 cannot be validly inferred from data related to the higher 20 to 56 range.

This review was underpinned by very broad inclusion criteria which may have impacted the findings. Although studies published in non-English journals were excluded, most of the studies in this review were performed in countries predominantly speaking a language other than English and may have used translations of the Berg Balance Scale.

Our meta-analysis has shown that the Berg Balance Scale has a high intra-rater and inter-rater relative reliability. Several studies of absolute reliability suggest that the Berg Balance Scale
is able to detect many clinically significant changes in balance with 95% confidence, although some individuals might experience moderate change in balance which cannot be reliably detected by the Berg Balance Scale. This review found little evidence describing the absolute reliability of the Berg Balance Scale for people with a Berg Balance Scale score between 0 and 20.
References:


Table 3.1. Summary of included studies (n = 12).

<table>
<thead>
<tr>
<th>Study</th>
<th>Reliability examined</th>
<th>n</th>
<th>Setting</th>
<th>Diagnosis</th>
<th>Age (yr) mean (SD)</th>
<th>Severity</th>
<th>Exclusion Criteria</th>
<th>Berg Balance Scale^a (0 to 56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg 1995</td>
<td>Relative</td>
<td>63</td>
<td>Acute stroke ward</td>
<td>Acute stroke</td>
<td>73 (9)</td>
<td>At least some motor impairment</td>
<td>Medically unstable</td>
<td>Full range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aged care facility</td>
<td>84 (5)</td>
<td>Independently mobile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattaneo 2007</td>
<td>Both</td>
<td>25</td>
<td>Multiple Sclerosis clinic</td>
<td>Multiple Sclerosis</td>
<td>42 (13)</td>
<td>Able to walk 6 metres</td>
<td>Cognitive impairment that might hinder testing</td>
<td>47.27 (7.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Geriatric rehabilitation inpatient unit, day hospital</td>
<td>Varied</td>
<td>82 (6)</td>
<td>Able to walk</td>
<td>Significant cognitive impairment, fracture</td>
<td>44.4</td>
</tr>
<tr>
<td>Halsaa 2007</td>
<td>Relative</td>
<td>83</td>
<td>Geriatric rehabilitation inpatient unit, day hospital</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mao 2002</td>
<td>Relative</td>
<td>11</td>
<td>Hospital inpatients</td>
<td>Recent stroke</td>
<td>69 (11)</td>
<td>Able to follow commands</td>
<td>Unable to give informed consent, able to consent by proxy not excluded</td>
<td>34.8 (18.6)</td>
</tr>
<tr>
<td>Wirz 2010</td>
<td>Relative</td>
<td>40</td>
<td>Outpatient rehabilitation centre</td>
<td>Spinal cord injury</td>
<td>49 (12)</td>
<td>Able to walk 15 metres</td>
<td>Balance affecting co-morbidity</td>
<td>41.1 (15.2)</td>
</tr>
<tr>
<td>Liaw 2008</td>
<td>Both</td>
<td>52</td>
<td>Outpatient rehabilitation centre</td>
<td>Chronic stroke</td>
<td>60 (13)</td>
<td>Able to follow verbal instructions</td>
<td>Cognitive impairment, other major diseases</td>
<td>Full range</td>
</tr>
<tr>
<td>Donoghue 2009</td>
<td>Absolute</td>
<td>11</td>
<td>Outpatient rehabilitation centre</td>
<td>Various</td>
<td>81 (7)</td>
<td>n/s</td>
<td>Cognitive impairment, Parkinson’s disease, unable to consent, stroke, recent hip replacement</td>
<td>38.6 (9.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lim 2005</td>
<td>Absolute</td>
<td>26</td>
<td>Home visits</td>
<td>Parkinson’s disease</td>
<td>63 (8)</td>
<td>Able to walk without gait aid</td>
<td>Cognitive impairment, co-morbidities affecting balance</td>
<td>53.8 (2)</td>
</tr>
<tr>
<td>Steffen and Seney 2008</td>
<td>Absolute</td>
<td>37</td>
<td>University</td>
<td>Parkinson’s disease</td>
<td>N/S</td>
<td>Able to walk independently</td>
<td>Cognitive impairment, activity limiting heart disease</td>
<td>50 (7)</td>
</tr>
<tr>
<td>Paltamaa 2005</td>
<td>Absolute</td>
<td>19</td>
<td>Physiotherapy outpatient department</td>
<td>Multiple Sclerosis</td>
<td>43 (9)</td>
<td>Able to walk 20 m</td>
<td>Unable to give written informed consent</td>
<td>54.3 (2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inter-rater: 49 (9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conradsson 2007</td>
<td>Absolute</td>
<td>45</td>
<td>Aged care facility</td>
<td>Various</td>
<td>83 (7)</td>
<td>Dependant for personal care, mean MMSE^b 17.5</td>
<td>Unable to stand from chair</td>
<td>30.1 (15.6)</td>
</tr>
<tr>
<td>Stevenson 2001</td>
<td>Absolute</td>
<td>48</td>
<td>Rehabilitation inpatients</td>
<td>Sub acute stroke patients</td>
<td>74 (7)</td>
<td>n/s</td>
<td>Unable to consent</td>
<td>IQR 36.5 to 47</td>
</tr>
</tbody>
</table>

^a mean (SD) unless otherwise stated, ^b Mini Mental State Examination (Folstein et al 1975) n/s = not stated
*where the full range is not studied the mean ± standard deviation or inter-quartile range is stated
**Mini Mental State Examination
Figures

Box 3.1. Inclusion criteria

<table>
<thead>
<tr>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Reliability studies examining the Berg Balance Scale</td>
</tr>
<tr>
<td>• Published in English</td>
</tr>
<tr>
<td>• Sample size ≥10 participants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Any clinical population</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Relative intra- and inter-rater reliability</td>
</tr>
<tr>
<td>• Absolute reliability</td>
</tr>
</tbody>
</table>
Figure 3.1. Flow of studies through the review

Articles identified from electronic database searches (n = 510) and reference lists (n = 1)

Excluded by title and abstract as not assessing reliability (n = 484)

Full text articles obtained (n = 27)

Excluded (n = 15)
- insufficient data reported (n = 4)
- incorrect or unclear methods used to calculate ICC (n = 2)
- investigated a modified version of the Berg Balance Scale (n = 2)
- inadequate details of method (n = 2)
- excluded subjects unable to complete the Berg Balance Scale (n = 1)
- inadequate number of subjects (n = 1)
- not published in English (n = 1)
- review article (n = 1)
- letter (n = 1)

Included studies (n = 12)

Studies included in meta-analysis (n = 11)

Study summarised narratively due to substantial cognitive impairment of participants (n = 1)
Figure 3.2. Intra rater relative reliability of the Berg Balance Scale

<table>
<thead>
<tr>
<th>Study</th>
<th>Relative reliability (95% CI) (random effects)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg</td>
<td>97 (93 to 99)</td>
<td>9</td>
</tr>
<tr>
<td>Cattaneo</td>
<td>97 (91 to 98)</td>
<td>6</td>
</tr>
<tr>
<td>Liaw</td>
<td>98 (97 to 99)</td>
<td>83</td>
</tr>
<tr>
<td>Pooled</td>
<td>98 (97 to 99)</td>
<td>100</td>
</tr>
</tbody>
</table>

Figure 3.3. Inter rater relative reliability of the Berg Balance Scale

<table>
<thead>
<tr>
<th>Study</th>
<th>Relative reliability (95% CI) (random effects)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg</td>
<td>0.98 (0.97 to 0.98)</td>
<td>32</td>
</tr>
<tr>
<td>Cattaneo</td>
<td>0.96 (0.90 to 0.97)</td>
<td>11</td>
</tr>
<tr>
<td>Mao</td>
<td>0.95 (0.93 to 0.97)</td>
<td>20</td>
</tr>
<tr>
<td>Wirz</td>
<td>0.95 (0.91 to 0.98)</td>
<td>12</td>
</tr>
<tr>
<td>Halsaa</td>
<td>0.99 (0.97 to 1.00)</td>
<td>23</td>
</tr>
<tr>
<td>Pooled</td>
<td>0.97 (0.96 to 0.98)</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 3.4. Minimal detectable difference with 95% confidence compared to mean of sample population

\[\text{Liaw only reports Median}\]
Chapter 4: Normal Values of the Berg Balance Scale

Introduction

Earlier chapters of this thesis have provided an important perspective on the validity of the BBS and critically analysed the peer reviewed literature describing the reliability of the BBS. This chapter considers the BBS from another perspective: what BBS values can be expected when it is used with healthy older people?

Australia’s ageing population is faced with increasing burdens of disability and falls. Health promotion can be expected to become increasingly important to limit these burdens. Clinicians using balance tools in any therapy, especially health promotion should be aware of what constitutes normal balance. Knowledge of normal balance values allows clinicians working in health promotion to identify people with lower than normal balance, potentially assisting interventions to be directed to people with the greatest need. This knowledge may also help identify which patients have reached their normal levels of balance, and which patients have most need of ongoing therapy.

In this study a systematic review with meta-analysis was used to describe the balance of healthy, community dwelling people. The integrity of the results of this review was dependent on the selection of individual studies included within it. Studies specifically conducted to determine normative data have been included, although this review has not been limited to such studies. Several studies conducted for purposes other than the establishment of normative data have nevertheless provided useful data. Examples of these are studies measuring baseline BBS in healthy elderly people prior to the commencement of falls prevention exercises and measurement of BBS in subjects selected as healthy matched controls. Excluding studies likely to report on biased BBS values is central to the validity of this review. Several studies have been excluded on the basis of being likely to report better or worse balance scores than typical healthy subjects. The most common reasons for excluding studies on the basis of bias were that some studies measured the BBS of only healthy elderly people not using gait aids such as walking sticks and that some studies measured the BBS of only healthy elderly people using gait aids.

There are substantial data, with considerable agreement describing the balance of healthy people aged between the ages of 70 and 80. There is much more limited data describing the balance of healthy elderly people aged 80 years and older. This review adds to the usefulness of the BBS by
describing robust normative data for people aged 80 years and younger. This review has found that as healthy people age past the age of 80 years, normal balance values continue to deteriorate but that inadequate data exists to precisely describe this deterioration.
Abstract

Questions: What is the mean Berg Balance Scale score of healthy elderly people living in the community and how does it vary with age? How much variability in Berg Balance Scale scores is present in groups of healthy elderly people and how does this vary with age?

Design: Systematic review with meta-analysis. Participants: Any group of healthy community-dwelling people with a mean age of 70 years or greater, that has undergone assessment using the Berg Balance Scale. Outcome Measurement: Mean and standard deviations of Berg Balance Scale scores within cohorts of elderly people of known mean age.

Results: The search yielded 17 relevant studies contributing data from a total of 1363 participants. The mean Berg Balance Scale scores ranged from 37 to 55 out of a possible maximum score of 56. The standard deviation of Berg Balance Scale scores varied from 1.0 to 9.2. Although participants aged around 70 years had very close to normal Berg Balance Scale score, there was a significant decline in balance with age at a rate of 0.7 points on the 56-point Berg Balance Scale per year. There was also a strong association between increasing age and increasing variability in balance ($R^2 = 0.56$, $p < 0.001$).

Conclusion: Healthy community-dwelling elderly people have modest balance deficits, as measured by the Berg Balance Scale, although balance scores deteriorate and become more variable with age.

INTRODUCTION

The Australian Institute of Health and Welfare has found that 65-year-old Australians have increasing life expectancy, both of years lived with disability and years lived without disability (Australian Institute of Health and Welfare 2012). With the percentage of Australians aged 85 years and older expected to increase from 2% in 2013 to 3.5% in 2033 (Australian Bureau of Statistics 2008) the costs of disability in older Australians and can be expected to substantially increase unless disability can be prevented and treated more efficiently. Falls are a major contributor to injury with subsequent disability in the elderly and poor balance is associated with increased risk of injurious falls (Muir et al 2008). The development and implementation of effective and cost-efficient strategies to prevent falls in older people is therefore an urgent challenge for healthcare. Such strategies require accurate and comprehensive measurement of balance ability.

The Berg Balance Scale was developed in 1989 using health professional and patient interviews,
which explored the various methods used to assess balance (Berg et al 1989). Thirty-eight component balance tests were originally selected and then refined through further interviews and trials to 14 items, each scored from 0 to 4, making a possible total score between 0 and 56, with a higher score indicating better balance. Although the Berg Balance Scale was originally developed to measure balance in the elderly, it has since been used to measure balance in a wide variety of patients.

The convergent validity of the BBS has been established across several different domains. Hospital inpatients with a lower Berg balance score have been found to have a significantly higher chance of being discharged to nursing home accommodation (Downs et al 2012). Among community-dwelling veterans, progressively lower Berg Balance Scale scores are associated with increased risk of injurious falls (Muir et al 2008). Responsiveness to change was established in a trial enrolling sedentary older people, where those who exercised improved their Berg Balance Scale scores and reported fewer falls, compared to a control group (Li et al 2004). The Berg Balance Scale also had greater ability than four other performance measures to predict the onset of difficulty in activities of daily living in older adults (Wennie Huang et al 2010).

Normative data are important when interpreting any balance tool, both for clinicians and researchers. Knowledge that a person or a group of people has significantly worse balance than a healthy person of the same age may assist the identification and effective management of balance problems. The effect of interventions to improve balance can be assessed by comparison to normative data for balance from healthy elderly people in specific age cohorts. Knowledge of the variability of the Berg Balance Scale in groups of healthy elderly people can be used to interpret individual results and to help establish the sample sizes required for future studies.

An earlier review (Downs et al 2013) searched for the phrase “Berg Balance Scale” and, despite finding 511 articles, did not identify any published review of normative data of the Berg Balance Scale.

The study questions for the systematic review were:
1. What is the mean Berg Balance Scale score of healthy elderly people living in the community and how does it vary with age?
2. How much variability in Berg Balance Scale scores is present in groups of healthy elderly people and how does this vary with age?
METHOD

Identification and selection of studies
A literature search was undertaken to locate all relevant published studies. Electronic searches of MEDLINE, CINAHL, Embase, and the Cochrane Library databases from 1980 to September 2012 were conducted using "Berg Balance Scale" as the search term. No keywords related to intervention type or health condition were used and no methodological filters to identify particular study designs were used. All potentially relevant papers were identified by screening the abstracts, assessed for inclusion and had data extracted by two authors (SD and PC) with any disagreements adjudicated by a third author (JM).

The a priori criteria for studies to be included in the review are presented in Box 4.1. Studies were excluded if the participants were hospital inpatients or resided in an aged care facility. Studies in which subjects had health conditions likely to significantly affect their balance were also excluded, as were studies in which healthy elderly subjects with extremes of balance (either minimal or maximal deficits) were excluded, or gait aid users were excluded. Where there were inadequate details of methods or results, an email was sent to the author where possible to seek further information.

Assessment of characteristics of trials
Participants: The inclusion and exclusion criteria and the country in which the data were collected were extracted for each trial. The sample size and the mean age of the participants were also extracted, along with whether the participants were enrolled as an observational cohort, an intervention group, or a control group.

Outcome: Means and standard deviations were extracted for baseline Berg Balance Scale scores. Where variability data were presented as other statistics, these were converted to standard deviations.

Data analysis
Meta-regression analysis of the mean Berg Balance Scale scores was conducted. Where studies provided participant groups stratified by age, analysis was conducted using subgroups rather than pooled data. In studies where subjects were listed by age decade without provision of the mean age
within the data, the mean age was assumed to be the mid-point of the decade. Where studies provided data for treatment and control groups in a trial, the baseline data for each group were included in the analysis separately.

To account for differences in the statistical power of the studies included in the meta-regression analysis, samples with larger numbers and samples with homogenous balance scores are weighted more highly when calculating the overall relationship between age and Berg Balance Scale score. Conversely small samples and samples with highly variable balance scores were given less weight.

The relationship between the mean age of a sample and the standard deviation of the Berg Balance Scale scores of the sample was investigated using linear regression analysis, with weighting for sample size.

RESULTS

Flow of studies through the review

After duplicates were removed, 859 articles were found containing the term “Berg Balance Scale” in their abstract, title or keywords. Hand searches of reference lists revealed one additional relevant paper. Of these 17 were deemed relevant and included in the analysis. Figure 4.1 presents the flow of studies through the review and the reasons for exclusion. The main reasons for exclusion from the study were: the participants had significant health conditions or limited mobility, the participants were too young, the participants were hospital inpatients, and the authors reported inadequate details about the participants, methods or results. The 17 included studies contributed data on 23 study cohorts involving 1363 participants in total.

The main properties of the studies of healthy elderly are presented in Table 4.1. In cases where studies contain more than one group of subjects the groups are listed individually.

The meta-regression analysis of mean age compared to mean Berg Balance Scale score in community dwelling healthy elderly is presented in Figure 4.2. Each circle represents an individual sample, with the diameter of the circle representing the weight given to that sample because of its variability and sample size. The analysis shows the deterioration of Berg Balance Scale score with increasing age ($R^2 = 0.81, p < 0.001$). The Berg Balance Scale score of healthy people aged 70 years
and older can be estimated by the formula: Berg Balance Scale score (over 70 years) = 107.7 - (age in years * 0.75).

Linear regression analysis found a strong relationship between increasing age and increasing variability of Berg Balance Scale scores (R2 = 56%, p < 0.001). This analysis is presented in Figure 4.3. The standard deviation of the Berg Balance Scale in groups of healthy people aged 70 and older can be estimated by the formula: standard deviation of the Berg Balance Scale score (over 70 years) = (age in years * 0.328) – 20.5

**DISCUSSION**

The results of the meta-regression of mean Berg Balance Scale scores suggests that a 70-year-old community dwelling person without health conditions likely to significantly affect their balance is likely to have a Berg Balance Scale score close to the maximum possible value of 56. The estimate of the decline in Berg Balance Scale with age beyond 70 years was fairly strongly supported by a large pooled sample of data (1363 participants). Interpretation of this decline in Berg Balance Scale with age should, however, acknowledge that only three studies (four samples, 210 participants) had participants with a mean age over 80 years and that the statistical power of these studies were weakened by large standard deviations.

These findings are broadly comparable to normative measures of mobility and balance using tools other than the Berg Balance Scale, which also show deterioration with age (Isles et al 2004). The normal values of the Berg Balance Scale suggest a ceiling effect in people younger than 70 years of age. Because of limited data from participants over 80 years old, further study is warranted to explore the relationship between the Berg Balance Scale and age among healthy, community-dwelling people aged 80 years or more.

This review found variation in the relationship between average Berg Balance Scale and age in healthy, community-dwelling elderly people. Several factors might explain this variability. Studies measuring the balance of healthy, community-dwelling elderly included in this review had similar, but not identical, eligibility criteria. Two outliers in the meta-regression, with lower Berg Balance Scale scores than expected for their age, were the treatment and control groups from a study that included only healthy sedentary elderly (Li et al 2004), suggesting that sedentary elderly might have
poorer balance than active elderly.

Two other outliers in the meta-regression, with higher Berg Balance Scale than expected for age, were cohorts from studies that included only participants without a history of hip or knee joint replacement surgery (Coleman and Cliftt 2010, Jongjit et al 2003). We can speculate that patients with a history of hip or knee replacement differ from other subjects for several reasons: they are more likely to have a history of arthritis; reduced physical activity following surgery might affect the long-term balance of some people; surgery might involve loss of proprioception at the affected joint; and patients with a history of hip replacement may be more likely to have a history of falls. For these reasons the finding that studies excluding patients with history of hip or knee replacement find a higher Berg Balance Scale than studies including such patients is unsurprising.

With the exception of the outliers discussed above, all the samples included in this review reported mean Berg Balance Scale scores within 2.3 points of the line of best fit. Given that the Berg Balance Scale is scored from 0 to 56, this suggests that there is relatively little heterogeneity within the studies considered by this review. Random sampling error appears to explain at least some of this heterogeneity, particularly among studies with a small sample size and high variability (displayed in figure as a small circle). The small amount of heterogeneity also suggests that the balance of healthy, community-dwelling elderly, as measured by the Berg Balance Scale, is similar in all countries where studies included in the review have been conducted.

This review provides an important perspective on the normal values of the Berg Balance Scale. It demonstrates that with increasing age, Berg Balance Scale scores of healthy, community-dwelling people become more variable. Some people retain good balance, with very high Berg Balance Scale scores into very old age, while some demonstrate very large deficits in balance. The increasing standard deviation of the Berg Balance Scale scores with age suggests that trials involving very old but otherwise unselected participants will require larger sample sizes to allow for the greater variability compared to trials in younger participants. Alternatively, at the expense of external validity and ease of recruitment, researchers could select very old participants with a specific degree of balance deficit.

Clinicians accustomed to working with balance impaired people may easily underestimate normal balance values of healthy elderly on the basis of their experience with balance impaired people and fail to set adequate treatment goals for their patients to attain optimal balance. These pooled
normative data will help to identify the usual balance performance of healthy, community-dwelling people aged 70 years or more.

Acknowledgements:

We wish to Acknowledge Alastair Merrifield, who provided bio statistical advice while he was a trainee biostatistician with the trainee NSW Centre for Epidemiology and Research.

Conflict of interest:

There were no conflicts of interest
References:


Box 4.1: Inclusion criteria

<table>
<thead>
<tr>
<th>Box 1. Inclusion criteria.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design</strong></td>
</tr>
<tr>
<td>• Any study design reporting baseline data on an unselected cohort</td>
</tr>
<tr>
<td>• Original research report (ie, not literature review)</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
</tr>
<tr>
<td>• Community dwelling</td>
</tr>
<tr>
<td>• Free of health condition likely to affect balance</td>
</tr>
<tr>
<td>• Mean age at least 70 years</td>
</tr>
<tr>
<td><strong>Outcomes measures</strong></td>
</tr>
<tr>
<td>• Berg Balance Scale mean</td>
</tr>
<tr>
<td>• Berg Balance Scale variability</td>
</tr>
</tbody>
</table>

Graphics
Table 4.1 – Summary of included studies and samples (n=17) and samples (n=23)

<table>
<thead>
<tr>
<th>Author and year</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
<th>Subjects from</th>
<th>n</th>
<th>Mean age</th>
<th>Sample Type</th>
<th>BBS ± standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhatt 2011</td>
<td>Ambulatory</td>
<td>Serious musculo-skeletal, neurological disorders, osteoporosis</td>
<td>USA</td>
<td>59</td>
<td>71.6</td>
<td>Normal</td>
<td>53.9±2.2</td>
</tr>
<tr>
<td>Coleman 2010</td>
<td>Able to follow directions</td>
<td>Stroke, Transient ischemic attack, Parkinson’s Disease, Joint replacement.</td>
<td>USA</td>
<td>53</td>
<td>75.4</td>
<td>Intervention</td>
<td>53±4</td>
</tr>
<tr>
<td>Daubney 1999</td>
<td>Aged 65+, Ambulatory</td>
<td>Any medical condition that may affect ability to test BBS</td>
<td>Canada</td>
<td>50</td>
<td>74.8</td>
<td>Normal</td>
<td>49.7±6.6</td>
</tr>
<tr>
<td>Eyigor 2009</td>
<td>Aged 65+, Active, Independent</td>
<td>Any health or cognitive problem limiting ability to learn dance</td>
<td>Turkey</td>
<td>19</td>
<td>73.5</td>
<td>Intervention</td>
<td>54.1±2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>71.2</td>
<td>Control</td>
<td>53.6±2.1</td>
</tr>
<tr>
<td>Hatch 2003</td>
<td>Able to: Walk 6.1 metres without help &amp; follow commands</td>
<td>Depression, neurological disorder, recent lower limb fracture or surgery</td>
<td>USA</td>
<td>50</td>
<td>81.7</td>
<td>Normal</td>
<td>46.5±9.5</td>
</tr>
<tr>
<td>Hinman 2002</td>
<td>Ambulatory, able to follow structured</td>
<td>Acute neurological or orthopaedic condition, dementia</td>
<td>USA</td>
<td>28</td>
<td>73.1</td>
<td>Intervention</td>
<td>53.1±2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>72.6</td>
<td>Intervention</td>
<td>52.9±3.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Group Description</th>
<th>Condition</th>
<th>Country</th>
<th>Sample Size</th>
<th>Age Range</th>
<th>Intervention</th>
<th>Control</th>
<th>Measurement</th>
<th>Study</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jongjit</td>
<td>2003</td>
<td>Age matched controls</td>
<td>Fracture, major orthopaedic surgery</td>
<td>Thailand</td>
<td>55</td>
<td>75.7</td>
<td>Control</td>
<td>54.1±3.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>2005</td>
<td>Age 70+ Ambulatory but inactive</td>
<td>Cognitive impairment, unable to exercise at moderate intensity</td>
<td>USA</td>
<td>125</td>
<td>76.9</td>
<td>Intervention</td>
<td>45.7±3.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>131</td>
<td>78.0</td>
<td>Control</td>
<td>46.2±4.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lusardi</td>
<td>2003</td>
<td>Ambulatory</td>
<td>Symptomatic cardiac, respiratory or neurological disease, depression, major surgery in last 6 months, dementia, cancer, acute illness or injury</td>
<td>USA</td>
<td>19</td>
<td>75</td>
<td>Normal</td>
<td>52.7±2.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>85</td>
<td>Normal</td>
<td>42±9.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>95.5</td>
<td>Normal</td>
<td>37.2±9.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacIntyre</td>
<td>2010</td>
<td>Community dwelling control subjects</td>
<td>Taking medication known to alter bone metabolism, known medical condition</td>
<td>Canada</td>
<td>11</td>
<td>71</td>
<td>Control</td>
<td>55.3±1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ozdemir</td>
<td>2009</td>
<td>Volunteers</td>
<td>Debilitating cardiac, respiratory or neurological disease, musculoskeletal conditions limiting movement.</td>
<td>Turkey</td>
<td>30</td>
<td>70.4</td>
<td>Normal</td>
<td>54.5±3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pardasaney</td>
<td>2012</td>
<td>Age 65+, able to climb a flight of stairs, MMSE* 23+</td>
<td>Unstable acute or chronic disease</td>
<td>USA</td>
<td>111</td>
<td>75.9</td>
<td>Normal</td>
<td>51.5±9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sihvonen</td>
<td>2009</td>
<td>Age 60-85, control group</td>
<td>Hip fracture, neurological or progressive severe disease</td>
<td>Finland</td>
<td>31</td>
<td>73.4</td>
<td>Normal</td>
<td>52.9±3.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun</td>
<td></td>
<td>Age 65, control</td>
<td>Surgery to spine or knees, arthritis, cardiac disorder, other condition that might affect</td>
<td>Taiwan</td>
<td>50</td>
<td>73.8</td>
<td>Control</td>
<td>53.2±3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Group</td>
<td>BBS Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Wang</td>
<td>Age 65+, independent in self care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hip or knee surgery, unable to follow instructions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td>Taiwan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>268</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>73.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td>53.3±4.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Group</th>
<th>BBS Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wennie Huang</td>
<td>2010</td>
<td>Aged 60+ MMSE* 24+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Activity of daily living difficulty, terminal health condition, dementia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>49.8±4.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Group</th>
<th>BBS Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrisley</td>
<td>2010</td>
<td>Age 60-90, MMSE* above 25+, Able to stand 1 minute</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Significant neurological, orthopaedic, cardiac or respiratory disease, fall in last year</td>
</tr>
<tr>
<td></td>
<td></td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.7±4</td>
</tr>
</tbody>
</table>

Figure 4.1. Flow chart describing selection of papers

859 Articles
(858 identified from literature search +1 from reference lists)

842 Excluded
- 13 Article not English
- 22 Case studies
- 26 Considering group with better than normal balance
- 352 Considering group with health problems
- 1 Considering group who had recently completed exercise program
- 29 Considering population
- 8 Double reporting data
- 31 Includes a disproportionate number of mobility limited people
- 9 Inadequate details of methods or results
- 102 Studying inpatients
- 44 Not testing BBS
- 22 Studies of residential aged care facilities
- 3 Reviews
- 180 Too young

17 Included
Figure 4.2 Relationship between mean age and mean BBS in healthy, community dwelling elderly

![Figure 4.2](image)

4.3. Relationship between BBS Standard Deviation (BBS SD) and Mean Age

![Figure 4.3](image)
Chapter 5: Conclusion

*Importance of Balance:*

Australia has an ageing population and without effective new interventions older Australians are likely to suffer from increased levels of disability and falls. The financial costs of health care and aged care to the Australian community from falls and disability are also likely to rapidly increase. The strong association demonstrated in this thesis between poor balance and probability of nursing home placement suggests that poor balance might force some older people to enter nursing homes and therefore be an important driver of care costs.

There are many age related diseases with the potential to reduce balance. Even in the absence of specific disease the process of sarcopaenia reduces muscle strength, potentially reducing balance. This thesis has shown that the balance of healthy elderly people as measured by the BBS decreases after the age of 70 years.

*Implications for clinicians:*

Validity:

Research conducted as part of this thesis suggested a relationship between very poor balance and a higher chance of being discharged to nursing home, suggesting that the BBS is a valuable tool for clinicians when measuring the functional balance of older people who are at risk of being unable to return home to the community. These findings also suggest the possibility that if more resources were invested to improve the balance of elderly inpatients, a significant number of elderly people discharged to residential aged care facilities might instead be discharged home to the community. This study also found that the change in BBS experienced by inpatients is quite variable and hard to predict, suggesting care should be taken not to prematurely discharge elderly inpatients to nursing homes on the basis of poor measured balance, as their balance might improve more than expected with time. This study also suggests that when planning discharge from hospital, credence should be given to treating physiotherapists’ estimates of how much change in balance is likely.
Reliability:

The BBS has very high relative intra-rater and inter-rater reliability. The absolute reliability of the BBS varies across the scale between 2.8 and 6.6, being poorest toward the middle of the scale. Clinicians can be confident that any large change in BBS is not likely due to random variation, although they should not discount other potential causes, such as natural recovery leading to patient improvement overall as well as measurement bias. From the research presented within this thesis, it is clear that modest, although clinically significant changes, might not always be detectable in individual clients with 95% confidence. The BBS might be less reliable when used to measure the balance of people with significant cognitive impairment.

This review has been unable to find strong evidence describing the absolute reliability of the BBS for people with a BBS of less than 20/56, although a small amount of data from two studies (Conradsson et al 2007, Liaw et al 2008) suggests the absolute reliability for the BBS between 0 and 10 might be very high.

Normal values:

Among healthy community dwelling elderly, those aged 70 years old have BBS scores close to the maximum of 56/56. As healthy, community dwelling elderly age, their mean BBS reduces at the rate of 0.75 points per year, with variability increasing with age. The data suggest that sedentary healthy elderly might have poorer balance than non-sedentary healthy elderly people. However, data describing the BBS of healthy, community dwelling elderly people aged 80 years and older is very limited.

Clinicians might conduct early interventions to improve balance before low levels of disability worsen and begin to cause problems. The data suggest that the BBS is likely to be an inappropriate screening tool for this purpose when used to assess healthy people close to 70 years of age, since in this age group a ceiling effect can be expected.

Our findings imply that healthy elderly people with a history of lower limb joint replacement surgery might have poorer balance than those without joint replacements. Until further research is conducted, clinicians should consider the possibility that otherwise healthy older people with lower limb joint replacement might be at risk of impaired balance. Screening these older people might potentially identify a group of people who would benefit from more intensive physiotherapy.
interventions than currently provided as part of usual care intervention, to improve long term balance.

Implications for research:

Validity:

This thesis demonstrates a strong relationship between balance and probability of discharge to nursing home from hospital, suggesting the possibility that physiotherapy intervention to improve balance might reduce the need for aged care services. The Australian population can be expected to continue to age and experience a greatly increased burden of age related disability if no effective interventions are found to reduce the number of disability affected years by improving the functional abilities of older people. The cost to the Australian taxpayer of the increasing burden of disability might be substantial as the population ages. A compelling case can be established for further research into the relationship between balance, age related disability, utilisation of aged care services and whether physiotherapy intervention can improve balance and reduce the need for residential and community aged care services.

Reliability:

Further research is required to determine the absolute reliability of the BBS in the range 0-20. The findings of this thesis were that elderly people with a BBS of 0/56 on discharge from hospital had almost a 100% probability of discharge to a nursing home, while elderly people with a BBS of 20/56 had approximately a 30% probability of discharge to nursing home. The lower range of the BBS appears important for therapists attempting improve the balance of elderly people at risk of nursing home placement and for clinicians in aged care assessment teams assessing whether or not elderly people require nursing home placement. Given this importance further research into the absolute reliability of the BBS in the range 0-20/56 appears warranted.

Normal values:

Data describing the BBS of healthy, community dwelling elderly people aged 80 years and older is very limited, with further research into this age group having the potential to substantially improve our understanding of the relationship between age and normal balance.
Further research might show more clearly the relationship between joint replacement surgery and balance. Further research also appears warranted to determine if providing physiotherapy intervention more intensively than is presently normal would improve the long term balance of people who have had a lower limb joint replacement.

Summary

This thesis has investigated important aspects of the BBS, a widely used balance tool. One issue important to many older Australians and therefore important when selecting a balance test is whether inpatients return home from hospital or whether they are admitted to nursing home care (high level residential aged care). The impost on the quality of life of elderly Australians upon entry to nursing home accommodation cannot be underestimated. The disability itself that contributes to such placement also impacts substantially on quality of life. Financial costs to the Australian taxpayer are also significant. With an ageing population all of these costs can be expected to further increase. This thesis has demonstrated a strong relationship between the BBS of hospital inpatients at discharge and the probability of their entry to nursing home. This thesis provides evidence that the BBS has high concurrent validity in this important domain.

Clinicians need to feel confident that any measurement tools are accurate and reliable. This thesis provides a summary of the reliability of the BBS, based on the published literature. Figure 3.4 (page 51*) provides clinicians with a tool to determine if measured changes noted in their patients are likely to reflect real changes (above the line of best fit) or might be due to measurement error (below the line of best fit).

Importantly this thesis provides evidence of how normal balance changes in healthy elderly people as they grow older. This more accurately informs the use of the BBS in healthy ageing people. The normative data within this thesis provides a perspective on the BBS that enables a comparison of individuals against age stratified normal values. This study has found that if therapists use the BBS to measure the balance of healthy people aged close to 70 years old they can expect to find most results to be 56/56. This ceiling effect would mean that if the BBS were to be used within this group it would not be useful for detecting differences between people nor would it be useful for detecting improvements in balance following intervention. The findings of this study suggest that the BBS might be a useful screening tool for healthy people aged 75 years and older and for people with health conditions likely to cause balance impairment.
This thesis has described a relationship between balance, as measured by the BBS and probability of inpatients being discharged to nursing home care. Therapy to improve balance might potentially lower the rate of admissions to nursing home, so research into how effective these therapies are at improving balance and lowering the rate of admissions to nursing homes appears warranted. This thesis has described the reliability of the BBS, but has identified a gap of knowledge of the absolute reliability of the BBS in the range 0/56-20/56. Finally this thesis has described normal values of the BBS in healthy older adults, but has identified a paucity of data describing these values in people aged older than 80 years.
References:


Appendix 1

Full search strategy for: The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review

Downs S, Marquez J, Chiarelli P

Databases: Medline, Embase CINAHL, Cochrane Library

Using the search terms below, the full holdings of Medline, Embase, CINAHL and the Cochrane Library were searched from 1980 to August 2010. The search strategy for each database is described below.

**Medline search strategy**
1. Search title field for ‘Berg Balance Scale’
2. Search abstract field for ‘Berg Balance Scale’
3. Search Key Words field for ‘Berg Balance Scale’
4. 1 AND 2 AND 3
5. Remove duplicates

The search yield was limited to ‘English language’.

**Embase search strategy**
1. Search title field for ‘Berg Balance Scale’
2. Search abstract field for ‘Berg Balance Scale’
3. Search Key Words field for ‘Berg Balance Scale’
4. 1 AND 2 AND 3
5. Remove duplicates

In Embase, the search was limited to exclude papers from Medline.

**CINAHL search strategy**
1. Search title field for ‘Berg Balance Scale’
2. Search abstract field for ‘Berg Balance Scale’
3. Search Key Words field for ‘Berg Balance Scale’
4. 1 AND 2 AND 3
5. Remove duplicates

In Cinahl, the search was limited to exclude papers from Medline and Embase.

**Cochrane Library search strategy**
1. Search title field for ‘Berg Balance Scale’
2. Search abstract field for ‘Berg Balance Scale’
3. Search Key Words field for ‘Berg Balance Scale’
4. 1 AND 2 AND 3
5. Remove duplicates

In Cochrane library, the search was limited to exclude papers from Medline, Embase and Cinhal.
# Appendix 2: The Berg Balance Scale

<table>
<thead>
<tr>
<th>Berg Balance Scale</th>
<th>Total</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DOB</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MRN</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total /56**

Please stand up. Try not to use your hand for support.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to stand without using hands and stabilize independently</td>
</tr>
<tr>
<td>3</td>
<td>able to stand independently using hands</td>
</tr>
<tr>
<td>2</td>
<td>able to stand using hands after several tries</td>
</tr>
<tr>
<td>1</td>
<td>needs minimal aid to stand or stabilize</td>
</tr>
<tr>
<td>0</td>
<td>needs moderate or maximal assist to stand</td>
</tr>
</tbody>
</table>

Please stand for two minutes without holding on

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to stand safely for 2 minutes</td>
</tr>
<tr>
<td>3</td>
<td>able to stand 2 minutes with supervision</td>
</tr>
<tr>
<td>2</td>
<td>able to stand 30 seconds unsupported</td>
</tr>
<tr>
<td>1</td>
<td>needs several tries to stand 30 seconds unsupported</td>
</tr>
<tr>
<td>0</td>
<td>unable to stand 30 seconds unsupported</td>
</tr>
</tbody>
</table>

If a subject is able to stand 2 minutes unsupported, score full and skip

**Sitting with back unsupported but feet supported on floor or on a stool**

Please sit with arms folded for 2 minutes

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to sit safely and securely for 2 minutes</td>
</tr>
<tr>
<td>3</td>
<td>able to sit 2 minutes under supervision</td>
</tr>
<tr>
<td>2</td>
<td>able to sit 30 seconds</td>
</tr>
<tr>
<td>1</td>
<td>able to sit 10 seconds</td>
</tr>
<tr>
<td>0</td>
<td>unable to sit without support 10 seconds</td>
</tr>
</tbody>
</table>

Please sit down.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>sits safely with minimal use of hands</td>
</tr>
<tr>
<td>3</td>
<td>controls descent by using hands</td>
</tr>
<tr>
<td>2</td>
<td>uses back of legs against chair to control descent</td>
</tr>
<tr>
<td>1</td>
<td>sits independently but has uncontrolled descent</td>
</tr>
<tr>
<td>0</td>
<td>needs assist to sit</td>
</tr>
</tbody>
</table>

Arrange chairs for pivot transfer. Ask subject to transfer toward a seat with armrests and back toward a seat without armrests. You may use two chairs or a bed and a chair.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to transfer safely with minor use of hands</td>
</tr>
<tr>
<td>3</td>
<td>able to transfer safely definite need of hands</td>
</tr>
<tr>
<td>2</td>
<td>able to transfer with verbal cuing and/or supervision</td>
</tr>
<tr>
<td>1</td>
<td>needs one person to assist</td>
</tr>
<tr>
<td>0</td>
<td>needs two people to assist or supervise to be safe</td>
</tr>
</tbody>
</table>

Please close your eyes and stand still for 10 seconds.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to stand 10 seconds safely</td>
</tr>
<tr>
<td>3</td>
<td>able to stand 10 seconds with supervision</td>
</tr>
<tr>
<td>2</td>
<td>able to stand 3 seconds</td>
</tr>
<tr>
<td>1</td>
<td>unable to keep eyes closed 3 seconds but stays safely</td>
</tr>
<tr>
<td>0</td>
<td>needs help to keep from falling</td>
</tr>
</tbody>
</table>

Place your feet together and stand without holding on.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>able to place feet together independently and stand 1 minute safely</td>
</tr>
<tr>
<td>3</td>
<td>able to place feet together independently and stand 1 minute with supervision</td>
</tr>
<tr>
<td>2</td>
<td>able to place feet together independently but unable to hold for 30 seconds</td>
</tr>
<tr>
<td>1</td>
<td>needs help to attain position but able to stand 15 seconds feet together</td>
</tr>
<tr>
<td>0</td>
<td>needs help to attain position and unable to hold for 15 seconds</td>
</tr>
</tbody>
</table>

Lift arm to 90 degrees. Stretch

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>can reach forward confidently 25 cm (10 inches)</td>
</tr>
<tr>
<td>Action</td>
<td>0</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td>Out your fingers and reach forward as far as you can. When possible, ask subject to use both arms when reaching to avoid rotation of the trunk.</td>
<td>0</td>
</tr>
<tr>
<td>Pick up the shoe/slipper, which is placed in front of your feet. (STANDING)</td>
<td>0</td>
</tr>
<tr>
<td>Turn to look directly behind you over toward the left shoulder. Repeat to the right.</td>
<td>0</td>
</tr>
<tr>
<td>360 DEGREES INSTRUCTIONS: Turn completely around in a full circle. Pause. Then turn a full circle in the other direction.</td>
<td>0</td>
</tr>
<tr>
<td>Place each foot alternately on the step/stool. Continue until each foot has to touch the step/stool four times.</td>
<td>0</td>
</tr>
<tr>
<td>Place one foot directly in front of the other. If you feel that you cannot place your foot directly in front, try to step far enough ahead that the heel of your forward foot is ahead of the toes of the other foot. (To score 3 points, the step should exceed the length of the other foot and the width of the stance should approximate the subject’s normal stride width.)</td>
<td>0</td>
</tr>
<tr>
<td>Stand on one leg as long as you can without holding on.</td>
<td>0</td>
</tr>
</tbody>
</table>

**TOTAL SCORE (Maximum = 56)**
Appendix 3 Copyright Permission

From: Judy Waters [mailto:Judy.Waters@physiotherapy.asn.au]
Sent: Tuesday, July 23 4:05 PM

Hello Stephen

You must be pleased that the grind is nearly over!

We do have instances of thesis by publication, and in all cases we place no impediment in the
author’s way. Please feel free to include your JoP paper in the University’s digital repository. There
are no forms to sign, and this message is sufficient authority.

I’m not sure where matters stand with your second submission. But I’ll ask Mark Elkins and let you
know.

With best wishes

From: Griffin, Tom - Melbourne [mailto:tgriffin@wiley.com]
Sent: Tuesday, 3 September 2013 5:14 PM
To: Downs, Stephen
Subject: RE: copyright

Hi Stephen,

Happy to say that we can grant you permission based on the conditions in copyright form
attached. Essentially you can place the post peer-reviewed version of the paper in the repository –
but only the word version, not the PDF. You will also need some acknowledgements listed on the
paper and a link to the actual final PDF online. The other condition was it needed to be 12 months
after publication, which we are almost at (the article was first published on September 23\textsuperscript{rd} last year)
so you are in the clear.

Have a quick read of the attached, the relevant section is next to the dot point “after acceptance”. If
you have any questions please let me know.

Hope it all goes well!

Tom
## Notification of Expedited Approval

<table>
<thead>
<tr>
<th>To Chief Investigator or Project Supervisor:</th>
<th>Associate Professor Pauline Chiarelli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cc Co-investigators / Research Students:</td>
<td>Mrs Jodie Marquez Mr Stephen Downs</td>
</tr>
<tr>
<td>Re Protocol:</td>
<td>Mobility outcomes for people referred to physiotherapy at Bellingen and Macksville Hospital</td>
</tr>
<tr>
<td>Date:</td>
<td>14-Apr-2011</td>
</tr>
<tr>
<td>Reference No:</td>
<td>H-2010-1336</td>
</tr>
<tr>
<td>Date of Initial Approval:</td>
<td>13-Apr-2011</td>
</tr>
</tbody>
</table>

Thank you for your Response to Conditional Approval submission to the Human Research Ethics Committee (HREC) seeking approval in relation to the above protocol.

Your submission was considered under Expedited review by the Chair/Deputy Chair.

I am pleased to advise that the decision on your submission is Approved effective 13-Apr-2011.

In approving this protocol, the Human Research Ethics Committee (HREC) is of the opinion that the project complies with the provisions contained in the National Statement on Ethical Conduct in Human Research, 2007, and the requirements within this University relating to human research.

Approval will remain valid subject to the submission, and satisfactory assessment, of annual progress reports. If the approval of an External HREC has been "noted" the approval period is as determined by that HREC.

The full Committee will be asked to ratify this decision at its next scheduled meeting. A formal Certificate of Approval will be available upon request. Your approval number is H-2010-1336.

If the research requires the use of an Information Statement, ensure this number is inserted at the relevant point in the Complaints paragraph prior to distribution to potential participants. You may then proceed with the research.

### Conditions of Approval

This approval has been granted subject to you complying with the requirements for Monitoring of Progress, Reporting of Adverse Events, and Variations to the Approved Protocol as detailed below.

PLEASE NOTE:
In the case where the HREC has "noted" the approval of an External HREC, progress reports and reports of adverse events are to be submitted to the External HREC only. In the case of Variations to the approved protocol, or a Renewal of approval, you will apply to the External HREC for approval in the first instance and then Register that approval with the University’s HREC.
• Monitoring of Progress

Other than above, the University is obliged to monitor the progress of research projects involving human participants to ensure that they are conducted according to the protocol as approved by the HREC. A progress report is required on an annual basis. Continuation of your HREC approval for this project is conditional upon receipt, and satisfactory assessment, of annual progress reports. You will be advised when a report is due.

• Reporting of Adverse Events

1. It is the responsibility of the person first named on this Approval Advice to report adverse events.

2. Adverse events, however minor, must be recorded by the investigator as observed by the investigator or as volunteered by a participant in the research. Full details are to be documented, whether or not the investigator, or his/her deputies, consider the event to be related to the research substance or procedure.

3. Serious or unforeseen adverse events that occur during the research or within six (6) months of completion of the research, must be reported by the person first named on the Approval Advice to the (HREC) by way of the Adverse Event Report form within 72 hours of the occurrence of the event or the investigator receiving advice of the event.

4. Serious adverse events are defined as:
   - Causing death, life threatening or serious disability.
   - Causing or prolonging hospitalisation.
   - Overdoses, cancers, congenital abnormalities, tissue damage, whether or not they are judged to be caused by the investigational agent or procedure.
   - Causing psycho-social and/or financial harm. This covers everything from perceived invasion of privacy, breach of confidentiality, or the diminution of social reputation, to the creation of psychological fears and trauma.
   - Any other event which might affect the continued ethical acceptability of the project.

5. Reports of adverse events must include:
   - Participant's study identification number;
   - date of birth;
   - date of entry into the study;
   - treatment arm (if applicable);
   - date of event;
   - details of event;
   - the investigator's opinion as to whether the event is related to the research procedures; and
   - action taken in response to the event.

6. Adverse events which do not fall within the definition of serious or unexpected, including those reported from other sites involved in the research, are to be reported in detail at the time of the annual progress report to the HREC.

• Variations to approved protocol

If you wish to change, or deviate from, the approved protocol, you will need to submit an Application for Variation to Approved Human Research. Variations may include, but are not limited to, changes or additions to investigators, study design, study population, number of participants, methods of recruitment, or participant information/consent documentation. Variations must be approved by the (HREC) before they are implemented except when Registering an approval of a variation from an external HREC which has been designated the lead HREC, in which case you may proceed as soon as you receive an acknowledgement of your Registration.

Linkage of ethics approval to a new Grant
HREC approvals cannot be assigned to a new grant or award (i.e. those that were not identified on the application for ethics approval) without confirmation of the approval from the Human Research Ethics Officer on behalf of the HREC.

Best wishes for a successful project.

Professor Alison Ferguson
Chair, Human Research Ethics Committee

For communications and enquiries:
Human Research Ethics Administration

Research Services
Research Integrity Unit
HA148, Hunter Building
The University of Newcastle
Callaghan NSW 2308
T +61 2 492 18988
F +61 2 492 17164
Human-Ethics@newcastle.edu.au

Linked University of Newcastle administered funding:

<table>
<thead>
<tr>
<th>Funding body</th>
<th>Funding project title</th>
<th>First named investigator</th>
<th>Grant Ref</th>
</tr>
</thead>
</table>

3 of 3
8 November 2009

Stephen Downs
Physiotherapist
Transitional Aged Care
Bellingen River District Hospital
PO Box 21
BELLINGEN NSW 2454

Dear Stephen

RE: NCAHS HREC NO. 478N
Mobility outcomes for physiotherapy patients: Mobility Outcomes for patients referred to physiotherapy at Bellingen Hospital and Macksville Hospital

Thank you for your correspondence by email dated 4 November 2009 to the North Coast Area Health Service (NCAHS) Human Research Ethics Committee (HREC) in response to clarifications requested by the NCAHS HREC, at its meeting held Thursday, 22 October 2009. The following documents were received:

- Updated Ethics Application form
- Updated Patient Information Sheet and Consent Form

Also received separately was the Management Approval.

The Chair, Paul Corben, reviewed and approved the above documents.

Final approval to commence the 'Mobility outcomes for physiotherapy patients' study has now been granted and includes the following documents approved at the NCAHS HREC meeting of 22 October 2009:

- Flowchart
- Berg Balance Scale

The NCAHS HREC is constituted and operates in accordance with the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research (National Statement - 2007).

As part of this approval, the following must be provided to the NCAHS HREC:

Amendments and Reporting of Serious Adverse Events

Researchers should immediately report anything to the Research Ethics Committee which might warrant review of ethical approval of the protocol, including:

- Serious or unexpected adverse effects on local participants (reports to be de-identified);

Human Research Ethics Committee
Clinical Governance Unit
North Coast Area Health Service
PO Box 126, Port Macquarie NSW 2444
Tel (02) 6588 2941 Fax (02) 6588 2942
Website www.ncahs.nsw.gov.au
ABN 37 940 606 983

Page 83
Proposed changes in the protocol or any other material given to the participants in the study must be known prior to being actioned, including patient information and consent forms; and unforeseen events that might affect continued ethical acceptability of the project.

Study Progress Reports

At least annually, reports from principal researchers should be submitted to the Research Ethics Committee on matters including:

- Progress to date or outcome in the case of completed research;
- Maintenance and security of records;
- Compliance with the approved protocol;
- Compliance with any conditions of approval;
- If the research project is discontinued before the expected date of completion;
- Published abstracts/reports resulting from the research;
- Upon completion of the research the HREC Clearance form to be submitted to HREC.

It is requested that updated Patient Information Consent Forms that are approved by the HREC, to be forwarded to all patients on the Trial.

Please quote HREC No. 478N, short and full study name in all correspondence and ensure all documentation relating to this study is forwarded, original with twelve copies (total 13) being double-sided and 2-hole punched, to:

Research Ethics Officer
Human Research Ethics Committee
North Coast Area Health Service
PO Box 126
PORT MACQUARIE NSW 2444

On behalf of the NCAHS HREC I wish you all the best with your research.

If you wish to discuss any matters further, please contact me on 02 6588 2941.

Yours sincerely,

Val Johnstone
Research Ethics Officer
Human Research Ethics Committee
31 May 2010

Stephen Downs
Physiotherapist
Transitional Aged Care
Bellinger River District Hospital
PO Box 21
BELLINGEN NSW 2454

Dear Stephen

RE: NCAHS HREC NO. 478N
Mobility outcomes for physiotherapy patients: Mobility Outcomes for patients referred to physiotherapy at Bellingen Hospital and Macksville Hospital

Thank you for your correspondence dated 13 May 2010 to the North Coast Area Health Service (NCAHS) Human Research Ethics Committee (HREC). The following documents were received:

- Cover letter of explanation
- Amendment form

The above documents were reviewed by the NCAHS HREC at its meeting held on Thursday, 27 May 2009. The Committee approved the change to classifications of patients and additional information collected from the medical records.

Please quote HREC No. 478N, short and full study name in all correspondence and ensure all documentation relating to this study is forwarded, with required number of copies being doublesided and 2-hole punched, to:

Human Research Ethics Committee
North Coast Area Health Service
PO Box 126
PORT MACQUARIE NSW 2444

If you wish to discuss any matters further, please contact me on 02 6588 2941.

Yours sincerely

Val Johnstone
Research Ethics Officer
Human Research Ethics Committee
Consent form

For Research into Mobility outcomes for physiotherapy patients. I agree to participate in this research project and give my consent freely.

1. I agree to have a balance test (The Berg Balance Scale) lasting about 10 minutes at the start and at the end of my physiotherapy program
2. I agree to have a copy of this balance test put in my medical record so that my physiotherapist has access to it
3. I understand, have read and explained to my satisfaction the Study Information Statement, a copy of which I have retained. I have also retained a copy of this Consent Form
4. I understand I can withdraw from the project at any time and do not have to give any reason for withdrawing.
5. I understand that my decision to participate in the study or to withdraw or not participate in the study will not affect my physiotherapy treatment
6. I understand that my personal information will remain confidential to the researchers.
7. I have the opportunity to have questions answered to my satisfaction

Name of participant

Signature

Name of person signing (if different to participant)

Relationship of person signing to participant

North Coast Area Health Service
Locked Bag 11 Lismore NSW 2480
Tel (02) 6620 2100 Fax (02) 6621 7088
Website www.ncahs.nsw.gov.au
ABN 37 940 606 983

Page 86
**Mobility outcomes for physiotherapy patients**

**Patient Information Sheet**

There is a new research project at Bellingen and Macksville Hospitals. We will be collecting information between January 2010 and November 2010. The purpose of the research is to better understand:

- How peoples' balance changes when they are in hospital
- How balance impacts on peoples' ability to continue living at home

The research is being conducted by Stephen Downs who is a physiotherapist with 11 years experience. Most of this experience is in Aged Care, especially helping people maintain their mobility and preventing falls. Stephen is conducting the research as part of the NSW Institute of Rural Clinical Services and Teaching (IR CST) "Rural Research Capacity Building Program" which aims to increase the number of rural and remote health workers with knowledge and skills in evaluation and research methods.

If you choose to participate in this study Stephen Downs will measure your balance using a series of tests called the "Berg Balance Scale" which take a total of about 10 minutes. Your balance will be tested once when you first have physiotherapy in hospital and once when you finish having physiotherapy. Only balance tests which are indicated to be safe will be done. However even with careful testing there is a very slight risk of falling over. The Berg Balance Scale has information that might help your physiotherapist, so a copy will be put in your medical record where your physiotherapist can access it.

Participation in this study is entirely voluntary. If you participate or don't participate in this study you will get the same physiotherapy while they you are in hospital. If you participate in this study you are free to withdraw from the study at any time, without any change to you treatment. You do not have to give any reason if you choose not to be part of the study.

All personal information collected in the study will be confidential. No personal information including your name or any other information which could identify you will be released. The findings of the study may be published in professional journals and discussed at professional conferences. Reports will be given to health professionals and health service managers with the intention of improving Hospital practices.

If you would like any information about the study you can contact Stephen Downs:
Phone 66 595 873
Mobile 0400 150 521
Email Stephen.Downs@ncahs.health.nsw.gov.au

This research project has been approved by the North Coast Area Health Service Human Research Ethics Committee. If you have any complaints or concerns about the conduct of this study you can contact:

The Executive Officer
Human Research Ethics Committee
North Coast Area Health Service
Phone 02 65 882 941
Appendix 5 – Papers As Published

The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review

Stephen Downs¹, Jodie Marquez² and Pauline Chiarelli²

¹Traditional Aged Care, Bellingen Hospital, ²Discipline of Physiotherapy, University of Newcastle, Australia

Questions: What is the intra-rater and inter-rater relative reliability of the Berg Balance Scale? What is the absolute reliability of the Berg Balance Scale? Does the absolute reliability of the Berg Balance Scale vary across the scale? Design: Systematic review with meta-analysis of reliability studies. Participants: Any clinical population that has undergone assessment with the Berg Balance Scale. Outcome measures: Relative intra-rater reliability, relative inter-rater reliability, and absolute reliability. Results: Eleven studies involving 668 participants were included in the review. The relative intra-rater reliability of the Berg Balance Scale was high, with a pooled estimate of 0.98 (95% CI 0.97 to 0.99). Relative inter-rater reliability was also high, with a pooled estimate of 0.97 (95% CI 0.96 to 0.98). A ceiling effect of the Berg Balance Scale was evident for some participants. In the analysis of absolute reliability, all of the relevant studies had an average score of 20 or above on the 0 to 56 point Berg Balance Scale. The absolute reliability across this part of the scale, as measured by the minimal detectable change with 95% confidence, varied between 2.8 points and 6.6 points. The Berg Balance Scale has a higher absolute reliability when close to 56 points due to the ceiling effect. We identified no data that estimated the absolute reliability of the Berg Balance Scale among participants with a mean score below 20 out of 56. Conclusion: The Berg Balance Scale has acceptable reliability, although it might not detect modest, clinically important changes in balance in individual subjects. The review was only able to comment on the absolute reliability of the Berg Balance Scale among people with moderately poor to normal balance. [Downs S, Marquez J, Chiarelli P (2013) The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. Journal of Physiotherapy 59: 93–99]

Key words: Postural Balance, Meta-Analysis, Reproducibility of Results

Introduction

The Berg Balance Scale was developed in 1989 via health professional and patient interviews that explored the various methods used to assess balance (Berg et al 1989). Initially, 38 balance tests were selected as potential components of the score and then refined through further interviews and trials to 14 items. Each of these items is scored from 0 to 4, which are summed to make a total score between 0 and 56, with a higher score indicating better balance. Although the Berg Balance Scale was originally developed to measure balance in the elderly, it has since been used to measure balance in a wide variety of patients.

All clinical measurement tools need to be reliable. Absolute reliability is clinically relevant and appears to be the most useful way of describing the reliability of the Berg Balance Scale (Bland and Altman 1986). The absolute reliability of the Berg Balance Scale provides a confidence interval, within which one can be confident that a change in balance is real change. The most common way of expressing this is the minimal detectable change with 95% confidence (MDC95). With regard to balance, intra-rater reliability refers to the reproducibility of a balance score when tested and retested by the same assessor. Inter-rater reliability refers to the reproducibility of a balance score when measured by different assessors. Relative reliability provides information about the variation in a score due to measurement error relative to variation within a population. This measure of reliability appears commonly in the literature, usually expressed as intra-class correlation (ICC) where a score of 1 represents perfect agreement and a score of 0 represents no relationship. Relative reliability provides perspective of the reliability of the Berg Balance Scale compared to other measurements, but is less useful clinically and is dependent on variability within the study sample. Studies of heterogeneous populations may find a very high relative reliability, even when the test is unable to detect clinically important changes reliably (Bland and Altman 1986). Three commonly used methods of

What is already known on this topic: The Berg Balance Scale rates balance from 0 (very poor) to 56 (normal) and is widely used in many clinical populations. The reliability of the scale in people with stroke has previously been reviewed but its reliability across all clinical populations has not been summarised.

What this study adds: Relative intra- and inter-rater reliability of the Berg Balance Scale are high. Absolute reliability was assessable between 20 and 56 on the scale. Absolute reliability varied within this range.
calculating ICC are used, generally referred to as Type 1, Type 2, and Type 3 (Shrout and Fleiss 1979). If a Type 1 calculation is incorrectly used, the reported ICC is likely to be an underestimate. Use of a Type 3 calculation is likely to result in a higher ICC, however Type 3 calculations cannot be generalised validly to assessors not involved in the study (Shrout and Fleiss 1979).

The objective of this review was to summarise the available evidence for the reliability of the Berg Balance Scale across all age groups and conditions whereas the Berg Balance Scale was used as a balance measurement tool.

Intra-rater reliability is measured by having an assessor measure balance and then repeat the measurement of the same person after a specified time lapse. Inter-rater reliability can be measured either by repeated measures by different assessors or by one assessor performing the test and other assessors rating the test. In the case of the Berg Balance Scale, the second rating can be done either in person or by reviewing a videorecording. Repeated measurements have the disadvantage that a person’s underlying balance might change between two measurements and therefore may underestimate the actual reliability of the Berg Balance Scale.

Simultaneous testing of the Berg Balance Scale to measure inter-rater reliability has different disadvantages. The Berg Balance Scale instructions may be interpreted and delivered in slightly different ways by different assessors. Non-verbal components such as demonstrating how to perform balance tests may vary between assessors. Safety considerations may lead some assessors not to attempt components of the Berg Balance Scale that other assessors might consider safe to attempt. An assessor might stand very close to a subject while performing balance testing, and so demonstrate that supervision is required. Simultaneous Berg Balance Scale testing, either in person or by video, can assess the reliability of how different assessors interpret a subject performing the Berg Balance Scale, but will not detect differences in how assessors instruct subjects to perform Berg Balance Scale testing and may therefore overestimate the actual reliability of the Berg Balance Scale.

It is reasonable to speculate that the reliability of the Berg Balance Scale may vary for each of the test items and for different populations. For example, in healthy community-dwelling people, reliability might be affected by disagreement about how Item 14 ‘standing on one leg’ is measured, while easier items such as Item 3 ‘sitting balance’ might be expected to have almost complete agreement of 4/4 among assessments. Conversely, when applied to people with stroke who are unable to stand, the reliability of ‘sitting balance’ may be more affected, while more difficult tasks such as ‘standing on one leg’ are likely to be universally assessed as 0/4. An additional factor that might cause variation in the reliability of the Berg Balance Scale is the underlying health conditions of subjects whose balance is tested. Individual studies are unlikely to be able to investigate the Berg Balance Scale over the full range of the scale and over the broad spectrum of causes of disordered balance. This review describes the range of subjects in whom the reliability of the Berg Balance Scale has been studied, reporting both their balance as well as any underlying health condition.

A previous literature review of the Berg Balance Scale (Blum and Korner-Bitensky 2008) considered the relative reliability of the Berg Balance Scale in patients with stroke and found it to have strong reliability. The current review covers important aspects of the reliability of the Berg Balance Scale not considered by the earlier review, including absolute reliability, and the reliability of the Berg Balance Scale in patients with conditions other than stroke.

Floor or ceiling effects occur when a significant proportion of a tested population achieve the lowest or highest possible score on a test, respectively (Everitt 2010). In groups where the mean Berg Balance Scale score is close to 0 or 56, the scale is unlikely to be useful in discriminating between individuals and will exhibit floor or ceiling effects. In such cases the scale is unlikely to be able to detect a change in balance, even if there is a real change. While floor and ceiling effects can potentially impair the clinical and research usefulness of the Berg Balance Scale, they are also likely to inflate its absolute reliability. A person with extremely poor balance is likely to be uniformly rated at 0/4 on most elements of the Berg Balance Scale. Conversely, a person with extremely good balance is likely to be uniformly rated 4/4 on most items of the Berg Balance Scale. Floor and ceiling effects involve groups with lower variability, which in turn lead to lower estimates of relative reliability compared to groups with more variable scores. Therefore, absolute and relative reliability should be interpreted with reference to floor and ceiling effects.

The specific study questions for this systematic review were:
1. What is the relative intra-rater and inter-rater reliability of the Berg Balance Scale?
2. What is the absolute reliability of the Berg Balance Scale, defined as the minimal detectable difference able to be determined with 95% confidence?
3. Does the absolute reliability of the Berg Balance Scale vary across the scale?

Method

Identification and selection of studies

A literature search was undertaken to locate eligible published studies. Electronic searches of Medline, CINAHL, Embase, and the Cochrane Library from 1980 to August 2010 were conducted using ‘Berg Balance Scale’ as a search term. No search terms were used for intervention type or health condition and no methodological filter was used for study design. See Appendix 1 on the eAddenda for the detailed search strategy. All potentially relevant papers were identified from abstracts and assessed for inclusion. The reference lists of included studies were searched for additional relevant papers. Data were extracted from the included studies by two authors (SD and PC) with any disagreements adjudicated by a third author (JM).

The inclusion criteria for studies are presented in Box 1. Studies investigating the relative reliability of the Berg Balance Scale had to supply a confidence interval around the estimate of the reliability of the scale or data allowing a confidence interval to be calculated. A minimum sample size of 10 was also applied, as recommended by Walter et al (1998). Studies examining translated versions of the scale were included if the study was reported in English. Studies examining a modified or partial version of the scale were excluded. Studies that excluded people who were
Box 1. Inclusion criteria.

**Design**
- Reliability studies examining the Berg Balance Scale
- Published in English
- Sample size ≥ 10 participants

**Participants**
- Any clinical population

**Outcomes**
- Relative intra- and inter-rater reliability
- Absolute reliability

Articles identified from electronic database searches (n = 510) and reference lists (n = 1)

Excluded by title and abstract as not assessing reliability (n = 484)

Full text articles obtained (n = 27)

Excluded (n = 15)
- insufficient data reported (n = 4)
- incorrect or unclear methods used to calculate ICC (n = 2)
- investigated a modified version of the Berg Balance Scale (n = 2)
- inadequate details of method (n = 2)
- excluded subjects unable to complete the Berg Balance Scale (n = 1)
- inadequate number of subjects (n = 1)
- not published in English (n = 1)
- review article (n = 1)
- letter (n = 1)

Included studies (n = 12)

Studies included in meta-analysis (n = 11)

Study summarised narratively due to substantial cognitive impairment of participants (n = 1)

Figure 1. Flow of studies through the review.

**Results**

**Flow of studies through the review**

Of the 511 papers identified (510 from electronic searches and 1 from reference lists), 27 were identified as being related to reliability based on information in the title and abstract. We excluded 15 studies, primarily for having inadequate detail about the methods or insufficient data to include in the meta-analysis. Eleven studies were included in analysis of the reliability of the Berg Balance Scale. The flow of studies through the review is presented in Figure 1.

**Characteristics of the studies**

Table 1 presents the characteristics of the included studies, including a description of the participants. These studies included elderly patients (Donoghue et al 2000), elderly residents of an aged care facility (Berg et al 1995), and patients with stroke (Liaw et al 2008, Mao et al 2002, Stevenson 2001), multiple sclerosis (Cattaneo et al 2007, Paltamaa et al 2005), spinal cord injury (Wirz et al 2010), and Parkinson’s disease (Limb et al 2005, Steffen and Seney 2008).
<table>
<thead>
<tr>
<th>Study</th>
<th>Reliability</th>
<th>n</th>
<th>Setting</th>
<th>Diagnosis</th>
<th>Age (yr) mean (SD)</th>
<th>Severity</th>
<th>Exclusion criteria</th>
<th>Berg Balance Scale scorea (0–56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg 1995</td>
<td>Relative</td>
<td>63</td>
<td>Acute stroke ward</td>
<td>Acute stroke</td>
<td>73 (9)</td>
<td>At least some motor impairment</td>
<td>Medically unstable</td>
<td>Full range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aged care facility</td>
<td>Not stated</td>
<td>84 (5)</td>
<td>Independently mobile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cattaneo 2007</td>
<td>Both</td>
<td>25</td>
<td>Multiple Sclerosis clinic</td>
<td>Multiple Sclerosis</td>
<td>42 (13)</td>
<td>Able to walk 6 metres</td>
<td>Cognitive impairment that might hinder testing</td>
<td>47.3 (7.7)</td>
</tr>
<tr>
<td>Halsaa 2007</td>
<td>Relative</td>
<td>83</td>
<td>Geriatric rehab. inpatient unit</td>
<td>Varied</td>
<td>82 (6)</td>
<td>Able to walk</td>
<td>Significant cognitive impairment, fracture</td>
<td>44.4</td>
</tr>
<tr>
<td>Mao 2002</td>
<td>Relative</td>
<td>112</td>
<td>Inpatients</td>
<td>Recent stroke</td>
<td>69 (11)</td>
<td>Able to follow commands</td>
<td>Unable to give informed consent, unable to consent by proxy not excluded</td>
<td>34.8 (18.6)</td>
</tr>
<tr>
<td>Wirz 2010</td>
<td>Relative</td>
<td>40</td>
<td>Outpatient rehabilitation centre</td>
<td>Spinal cord injury</td>
<td>49 (12)</td>
<td>Able to walk 15 metres</td>
<td>Balance affecting co-morbidity</td>
<td>41.1 (15.2)</td>
</tr>
<tr>
<td>Liaw 2008</td>
<td>Both</td>
<td>52</td>
<td>Outpatient rehab. centre</td>
<td>Chronic stroke</td>
<td>60 (13)</td>
<td>Able to follow verbal instructions</td>
<td>Cognitive impairment, other major diseases</td>
<td>Full range</td>
</tr>
<tr>
<td>Donoghue 2009</td>
<td>Absolute</td>
<td>118</td>
<td>Outpatient rehab. centre</td>
<td>Various</td>
<td>81 (7)</td>
<td>n/s</td>
<td>Cognitive impairment, Parkinson's disease, unable to consent, stroke, recent hip replacement</td>
<td>36.6 (9.6)</td>
</tr>
<tr>
<td>Lim 2005</td>
<td>Absolute</td>
<td>26</td>
<td>Home visits</td>
<td>Parkinson's disease</td>
<td>63 (8)</td>
<td>Able to walk without gait aid</td>
<td>Cognitive impairment, co-morbidities affecting balance</td>
<td>53.8 (2)</td>
</tr>
<tr>
<td>Steffen &amp;</td>
<td>Absolute</td>
<td>37</td>
<td>University</td>
<td>Parkinson's disease</td>
<td>n/s</td>
<td>Able to walk independently</td>
<td>Cognitive impairment, activity limiting heart disease</td>
<td>50 (7)</td>
</tr>
<tr>
<td>Seney 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paltamaa 2005</td>
<td>Absolute</td>
<td>19</td>
<td>Physiotherapy outpatient</td>
<td>Multiple sclerosis</td>
<td>Intra-rater: 43 (9)</td>
<td>Able to walk 20 m</td>
<td>Unable to give written informed consent</td>
<td>54.3 (2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>department</td>
<td></td>
<td>Inter-rater: 49 (9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conradsson 2007</td>
<td>Absolute</td>
<td>45</td>
<td>Aged care facility</td>
<td>Various</td>
<td>83 (7)</td>
<td>Dependant for personal care, mean MMSE 17.5</td>
<td>Unable to stand from chair</td>
<td>30.1 (15.6)</td>
</tr>
<tr>
<td>Stevenson 2001</td>
<td>Absolute</td>
<td>48</td>
<td>Rehab. inpatients</td>
<td>Sub acute stroke</td>
<td>74 (7)</td>
<td>n/s</td>
<td>Unable to consent</td>
<td>IQR 36.5 to 47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a mean (SD) unless otherwise stated, rehab. = rehabilitation, MMSE = Mini Mental State Examination (Folstein et al 1975), n/s = not stated, IQR = interquartile range
Table 1: Intra-rater relative reliability of the Berg Balance Scale.

<table>
<thead>
<tr>
<th>Study</th>
<th>Relative reliability (95% CI) (random effects)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg</td>
<td>97 (93 to 99)</td>
<td>9</td>
</tr>
<tr>
<td>Cattaneo</td>
<td>97 (91 to 98)</td>
<td>6</td>
</tr>
<tr>
<td>Llaw</td>
<td>98 (97 to 99)</td>
<td>83</td>
</tr>
<tr>
<td>Pooled (I² = 0%, p = 0.73)</td>
<td>98 (97 to 99)</td>
<td>100</td>
</tr>
</tbody>
</table>

Figure 2. Intra-rater relative reliability of the Berg Balance Scale.

Table 2: Inter-rater relative reliability of the Berg Balance Scale.

<table>
<thead>
<tr>
<th>Study</th>
<th>Relative reliability (95% CI) (random effects)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berg</td>
<td>0.98 (0.97 to 0.98)</td>
<td>32</td>
</tr>
<tr>
<td>Cattaneo</td>
<td>0.96 (0.90 to 0.97)</td>
<td>11</td>
</tr>
<tr>
<td>Mao</td>
<td>0.95 (0.93 to 0.97)</td>
<td>20</td>
</tr>
<tr>
<td>Wirz</td>
<td>0.95 (0.91 to 0.98)</td>
<td>12</td>
</tr>
<tr>
<td>Halsaa</td>
<td>0.99 (0.97 to 1.00)</td>
<td>23</td>
</tr>
<tr>
<td>Pooled (I² = 69%, p = 0.012)</td>
<td>0.97 (0.96 to 0.98)</td>
<td>100</td>
</tr>
</tbody>
</table>

Figure 3. Inter-rater relative reliability of the Berg Balance Scale.

Relative reliability

The intra-rater relative reliability of the Berg Balance Scale was estimated by meta-analysing data from three studies with a total of 101 subjects. The pooled estimate of the intra-rater relative reliability of the Berg Balance Scale was 0.98 (95% CI 0.97 to 0.99), as presented in Figure 2. A further analysis was conducted to examine the inter-rater relative reliability of the Berg Balance Scale by meta-analysing data from five studies with a total of 345 subjects. The pooled estimate of the inter-rater reliability was 0.97 (95% CI 0.96 to 0.98), as presented in Figure 3. These studies included participants from a variety of clinical populations with balance abilities across the full spectrum of the Berg Balance Scale, although only one study had a sizeable number of subjects with very low Berg Balance Scale scores (Berg et al 1995).

Sensitivity analyses did not find evidence that translations of the Berg Balance Scale into languages other than English have different reliability to the English version. In all cases repeating the analysis omitting translations of the Berg Balance Scale changed the relative reliability by less than 1%. All papers used Shrout and Fleiss Type 2 calculation to calculate ICC except Berg et al (1995), which used Type 1.

Absolute reliability

Studies investigating the absolute intra-rater reliability of the Berg Balance Scale show that the MDC05 varies in relation to the mean Berg Balance Scale scores of the sample, as presented in Figure 4. The review did not identify data about the absolute reliability of the Berg Balance Scale within its lower range of 0 to 20. Only one study examined the absolute inter-rater reliability of the Berg Balance Scale (Cattaneo et al 2007). This found very similar results for absolute intra- and inter-rater reliability.
Sensitivity analysis was conducted individually on all papers studying the absolute reliability of the Berg Balance Scale using translations. A Swedish translation studying the reliability of the Berg Balance Scale in residential aged care facilities with substantially cognitively impaired residents found a significantly lower absolute reliability with a MDC95 of 7.7 (mean Berg Balance Scale 30.1) (Conradsson et al 2007). These study findings were not included in our analysis of the absolute reliability of Berg Balance Scale. In all other cases the line of best fit with the individual study excluded was almost identical to the analysis presented.

**Discussion**

Our review identified substantial and consistent evidence of high intra-rater and inter-rater relative reliability of the Berg Balance Scale. Absolute reliability data were also favourable, although some people might experience moderate change in balance that would not be reliably detected by the scale. Furthermore, the absolute reliability data were only available for people with Berg Balance Scores above 20.

The reliability of the Berg Balance Scale has been investigated among a wide variety of subjects, although both studies investigating the reliability of the Berg Balance Scale in patients with Parkinson’s disease used subjects with high Berg Balance Scale scores which incurred a ceiling effect. The results of these studies might therefore be considered invalid in terms of describing the reliability of the Berg Balance Scale for patients with Parkinson’s disease whose balance scores are in the middle or lower range of the Berg Balance Scale.

This review found little evidence describing the reliability of the English language Berg Balance Scale in people with substantial cognitive impairment, although a Swedish language Berg Balance Scale translation (Conradsson et al 2007) suggests the Berg Balance Scale may be less reliable in people with substantial cognitive impairment.

While the high relative reliability suggests the Berg Balance Scale is clinically useful, there is little specific guidance as to how confident one can be that a real change in balance has occurred between tests across time for individual patients. This review suggests that if an individual has a Berg Balance Scale score of between 20 and 56 and experiences a change of between 3 and 7 (see Figure 4), one can be 95% confident that there has been a real change in balance. Individuals may experience clinically relevant changes in balance that cannot be reliably detected. Downs et al (2012) found hospital inpatients with a Berg Balance Scale of 20 have approximately a 30% probability of being discharged to a nursing home, while those with a Berg Balance Scale of 25 have approximately 20% probability of being discharged to a nursing home, suggesting that a difference in balance which is only barely detectable with 95% confidence in any individual may in fact be highly clinically relevant.

Changes in the average Berg Balance Scale score of patient or research groups have a smaller minimal detectable change than individual subjects. Thus, while moderately clinically important balance changes might not always be detectable with 95% confidence in individuals, they can be expected to be reliably detectable within groups. Researchers or clinicians who find clinically important changes in the average Berg Balance Scale score of a group of individuals might therefore be confident that the change was not caused by random variation.
This literature review did not find data describing the absolute reliability in groups with very low Berg Balance Scale scores, although data presented by Cattaneo et al (2007) suggest that the absolute reliability of the Berg Balance Scale might be higher in the 0 to 20 range than the 20 to 56 range. Bimodal distribution of the Berg Balance Scale has been reported previously (Berg et al 1995, Downs et al 2012), suggesting subjects might be categorized into two distinct groups: those able to stand independently and those unable to stand independently. Where people were able to stand independently, they were also able to attempt and usually achieve a score on several items, generally achieving a Berg Balance Scale score greater than 20. Those unable to stand independently are unable to attempt these items and usually score less than 15. The dichotomous nature of these two groups suggests that the absolute reliability of the lower Berg Balance Scale between 0 and 20 cannot be validly inferred from data related to the higher 20 to 56 range.

This review was underpinned by very broad inclusion criteria which may have impacted the findings. Although studies published in non-English journals were excluded, most of the studies in this review were performed in countries predominantly speaking a language other than English and may have used translations of the Berg Balance Scale.

Our meta-analysis has shown that the Berg Balance Scale has high intra- and inter-rater relative reliability. Several studies of absolute reliability suggest that the Berg Balance Scale is able to detect many clinically significant changes in balance with 95% confidence, although some individuals might experience moderate change in balance that cannot be reliably detected by the Berg Balance Scale. This review found little evidence describing the absolute reliability of the Berg Balance Scale for people with a Berg Balance Scale score between 0 and 20.

**eAddenda:** Appendix 1 available at jop.physiotherapy.asn.au

**Support:** Research was conducted as part of a Master’s degree with the University of Newcastle.

**Acknowledgement:** We thank Alastair Merrifield from the NSW Centre for Epidemiology and Research for his assistance with the project.

**Correspondence:** Stephen Downs, Transitional Aged Care, Bellingen Hospital, Church Street, Bellingen, 2454, NSW, Australia. Email: Stephen.downs@ncahs.health.nsw.gov.au

**References**


Research

Normative scores on the Berg Balance Scale decline after age 70 years in healthy community-dwelling people: a systematic review

Stephen Downs a, Jodie Marquez b, Pauline Chiarello b

a Transitional Aged Care Service, Bellbridge Hospital, Bellbridge, Australia; b Discipline of Physiotherapy, School of Health Sciences, University of Newcastle, Callaghan, Australia

Keywords

Berg Balance Scale, Normal values, Literature review, Meta-analysis, Age

Abstract

Questions: What is the mean Berg Balance Scale score of healthy elderly people living in the community and how does it vary with age? How much variability in Berg Balance Scale scores is present in groups of healthy elderly people and how does this vary with age? Design: Systematic review with meta-analysis. Participants: Any group of healthy community-dwelling people with a mean age of 70 years or greater that has undergone assessment using the Berg Balance Scale. Outcome measures: Mean and standard deviations of Berg Balance Scale scores within cohorts of elderly people of known mean age. Results: The search yielded 17 relevant studies contributing data from a total of 1363 participants. The mean Berg Balance Scale scores ranged from 37 to 55 out of a possible maximum score of 56. The standard deviation of Berg Balance Scale scores varied from 1.0 to 9.2. Although participants aged around 70 years had very close to normal Berg Balance Scale scores, there was a significant decline in balance with age at a rate of 0.7 points on the 56-point Berg Balance Scale per year. There was also a strong association between increasing age and increasing variability in balance ($R^2 = 0.56, p < 0.001$). Conclusion: Healthy community-dwelling elderly people have modest balance deficits, as measured by the Berg Balance Scale, although balance scores deteriorate and become more variable with age. [Downs S, Marquez J, Chiarello P (2014) Normative scores on the Berg Balance Scale decline after age 70 years in healthy community-dwelling people: a systematic review. Journal of Physiotherapy 60, 83–89] © 2014 Published by Elsevier B.V. on behalf of Australian Physiotherapy Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Introduction

The Australian Institute of Health and Welfare has found that 65-year-old Australians have increasing life expectancy, both of years lived with disability and years lived without disability.1 With the percentage of Australians aged 85 years and older expected to increase from 2% in 2013 to 5.5% in 2033,2 the costs of disability in older Australians can be expected to substantially increase unless disability can be prevented and treated more efficiently. Falls are a major contributor to injury with subsequent disability in the elderly, and poor balance is associated with increased risk of injurious falls.3 The development and implementation of effective and cost-efficient strategies to prevent falls in older people is therefore an urgent challenge for health care. Such strategies require accurate and comprehensive measurement of balance ability.

The Berg Balance Scale was developed in 1989 using health professional and patient interviews, which explored the various methods used to assess balance.4 Thirty-eight component balance tests were originally selected and then refined through further interviews and trials to 14 items, each scored from 0 to 4, making a possible total score between 0 and 56, with a higher score indicating better balance. Although the Berg Balance Scale was originally developed to measure balance in the elderly, it has since been used to measure balance in a wide variety of patients. The convergent validity of the Berg Balance Scale has been established across several different domains. Hospital inpatients with a lower Berg balance score have been found to have a significantly higher chance of being discharged to nursing home accommodation.5 Among community-dwelling veterans, progressively lower Berg Balance Scale scores are associated with increased risk of injurious falls.6 Responsiveness to change was established in a trial enrolling seniory older people, where those who exercised improved their Berg Balance Scale scores and reported fewer falls, compared to a control group.7 The Berg Balance Scale also had greater ability than four other performance measures to predict the onset of difficulty in activities of daily living in older adults.8

Normative data are important when interpreting any balance tool, both for clinicians and researchers. Knowledge that a person or a group of people has significantly worse balance than a healthy person of the same age may assist the identification and effective management of balance problems. The effect of interventions to improve balance can be assessed by comparison to normative data.
for balance from healthy elderly people in specific age cohorts.
Knowledge of the variability of the Berg Balance Scale in groups of healthy elderly people can be used to interpret individual results
and to help establish the sample sizes required for future studies.
An earlier review searched for the phrase ‘Berg Balance Scale’
and, despite finding 511 articles, did not identify any published
review of normative data of the Berg Balance Scale.
The study questions for the systematic review were:
1. What is the mean Berg Balance Scale score of healthy elderly
people living in the community and how does it vary with age?
2. How much variability in Berg Balance Scale scores is present in
groups of healthy elderly people and how does this vary with age?

Method
Identification and selection of studies
A literature search was undertaken to locate all relevant
published studies. Electronic searches of MEDLINE, CINahl,
Embase, and the Cochrane library databases from 1980 to
September 2012 were conducted using ‘Berg Balance Scale’ as
the search term. No keywords related to intervention type or
health condition were used and no methodological filters to
identify particular study designs were used. All potentially
relevant papers were identified by screening the abstracts and
assessed for inclusion. Data were extracted by two authors (SD and
PC) with any disagreements adjudicated by a third author (IM).
The a priori criteria for studies to be included in the review are
presented in Box 1. Studies were excluded if the participants were
hospital inpatients or resided in an aged care facility. Studies in
which subjects had health conditions likely to significantly affect
their balance were also excluded, as were studies in which healthy
elderly subjects with extremes of balance (either minimal or
maximal deficits) were excluded, or gait aid users were excluded.
Where there were inadequate details of methods or results, an
email was sent to the author where possible to seek further
information.

Assessment of characteristics of trials
Participants: The inclusion and exclusion criteria and the
country in which the data were collected were extracted for each
trial. The sample size and the mean age of the participants were
also extracted, along with whether the participants were enrolled
as an observational cohort, an intervention group, or a control
group.
Outcome: Means and standard deviations were extracted for
baseline Berg Balance Scale scores. Where variability data were
presented as other statistics, these were converted to standard
deviations.

Data analysis
Meta-regression analysis of the mean Berg Balance Scale scores
was conducted. Where studies provided participant groups
stratified by age, analysis was conducted using subgroups rather
than pooled data. In studies where subjects were listed by age
decade without provision of the mean age within the data, the
mean age was assumed to be the mid-point of the decade. Where
studies provided data for treatment and control groups in a trial,
the baseline data for each group were included in the analysis
separately.
To account for differences in the statistical power of the studies
included in the meta-regression analysis, samples with larger
numbers and samples with homogenous balance scores are
weighted more highly when calculating the overall relationship
between age and Berg Balance Scale score. Conversely, small
samples and samples with highly variable balance scores were
given less weight.
The relationship between the mean age of a sample and the
standard deviation of the Berg Balance Scale scores of the sample
was investigated using linear regression analysis, with weighting
for sample size.

Results
Flow of studies through the review
After duplicates were removed, 859 articles were found
containing the term ‘Berg Balance Scale’ in their abstract, title,
or keywords. Hand searches of reference lists revealed one
additional relevant paper. Of these, 17 were deemed relevant
and included in the analysis. Figure 1 presents the flow of studies
through the review and the reasons for exclusion. The main
reasons for exclusion from the study were: the participants had
significant health conditions or limited mobility; the participants
were too young; the participants were hospital inpatients; and
the authors reported inadequate details about the participants.

Box 1. Inclusion criteria.

Design
• Any study design reporting baseline data on an
  unselected cohort
• Original research report (i.e., not literature review)
Participants
• Community dwelling
• Free of health condition likely to affect balance
• Mean age at least 70 years
Outcomes measures
• Berg Balance Scale mean
• Berg Balance Scale variability

Articles included in the review (n = 17)

Articles excluded after evaluation (n = 942)
• participants had health problems (n = 352)
• participants were young (n = 160)
• participants were inpatients (n = 102)
• outcome not Berg Balance Scale (n = 44)
• study cohort had disproportionate number of
  participants with mobility limitation (n = 31)
• sample population was likely to have included
  some people with impaired balance (n = 29)
• participants had supra-normal balance (n = 26)
• participants in residential care facilities (n = 22)
• case reports (n = 22)
• non-English (n = 13)
• inadequate detail of methods or results (n = 9)
• duplicate report of data (n = 8)
• literature reviews (n = 3)
• participated recently completed an exercise
  intervention (n = 1)

Figure 1. Flow of studies through the review.
methods, or results. The 17 included studies contributed data on 23 study cohorts involving 1363 participants in total. The main properties of the studies of healthy elderly are presented in Table 1. In cases where studies contain more than one group of subjects, the groups are listed individually.

The meta-regression analysis of mean age compared to mean Berg Balance Scale scores in community-dwelling healthy elderly is presented in Figure 2. Each circle represents an individual sample, with the diameter of the circle representing the weight given to that sample because of its variability and sample size. The analysis shows the deterioration of Berg Balance Scale score with increasing age ($R^2 = 0.81, p < 0.001$). The Berg Balance Scale score of healthy people aged 70 years and older can be estimated by the formula: Berg Balance Scale score (age 70 years) = (age in years × 0.028) – 20.5.

**Discussion**

The results of the meta-regression of mean Berg Balance Scale scores suggest that a 70-year-old community-dwelling person with no health conditions likely to significantly affect their balance is likely to have a Berg Balance Scale score close to the maximum possible value of 56. The estimate of the decline in Berg Balance Scale with age beyond 70 years was fairly strongly supported by a large pooled sample of data (1363 participants). Interpretation of this decline in Berg Balance Scale with age should, however, acknowledge that only three studies (four samples, 210 participants) had participants with a mean age over 80 years, and that the statistical power of these studies was weakened by large standard deviations.

These findings are broadly comparable to normative measures of mobility and balance using tools other than the Berg Balance Scale.
Figure 2. Relationship between mean age and mean Berg Balance Scale (BBS) score in healthy, community-dwelling elderly people.

Figure 3. Relationship between standard deviation of the Berg Balance Scale (BBS) score and mean age.

Scale, which also show deterioration with age. The normal values of the Berg Balance Scale suggest a ceiling effect in people younger than 70 years of age. Because of limited data from participants over 80 years old, further study is warranted to explore the relationship between the Berg Balance Scale and age among healthy, community-dwelling people aged 80 years or more.

This review found variation in the relationship between average Berg Balance Scale and age in healthy, community-dwelling elderly people. Several factors might explain this variability. Studies measuring the balance of healthy, community-dwelling elderly included in this review had similar, but not identical, eligibility criteria. Two outliers in the meta-regression, with lower Berg Balance Scale scores than expected for their age, were the treatment and control groups from a study that included only healthy sedentary elderly, suggesting that sedentary elderly might have poorer balance than active elderly.

Two other outliers in the meta-regression, with higher Berg Balance Scale than expected for age, were cohorts from studies that included only participants without a history of hip or knee joint replacement surgery. We can speculate that patients with a history of hip or knee replacement differ from other subjects for several reasons: they are more likely to have a history of arthritis; reduced physical activity following surgery may affect the long-term balance of some people; surgery might involve loss of proprioception at the affected joint; and patients with a history of hip replacement may be more likely to have a history of falls. For these reasons, the finding that studies excluding patients with history of hip or knee replacement find a higher Berg Balance Scale than studies including such patients is unsurprising.

With the exception of the outliers discussed above, all the samples included in this review reported mean Berg Balance Scale scores within 2.3 points of the line of best fit. Given that the Berg Balance Scale is scored from 0 to 56, this suggests that there is relatively little heterogeneity within the studies considered by this review. Random sampling error appears to explain at least some of this heterogeneity, particularly among studies with a small sample size and high variability (displayed in figure as a small circle). The small amount of heterogeneity also suggests that the balance of healthy, community-dwelling elderly, as measured by the Berg Balance Scale, is similar in all countries where studies included in the review have been conducted.

This review provides an important perspective on the normal values of the Berg Balance Scale. It demonstrates that with increasing age, Berg Balance Scale scores of healthy, community-dwelling people become more variable. Some people retain good balance, with very high Berg Balance Scale scores into very old age, while some demonstrate very large deficits in balance. The increasing standard deviation of the Berg Balance Scale scores with age suggests that trials involving very old but otherwise unselected participants will require larger sample sizes to allow for the greater variability compared to trials in younger participants. Alternatively, at the expense of external validity and ease of recruitment, researchers could select very old participants with a specific degree of balance deficit.

Clinicians accustomed to working with balance-impaired people may easily underestimate normal balance values of healthy elderly on the basis of their experience with balance-impaired people and fail to set adequate treatment goals for their patients to attain optimal balance. These pooled normative data will help to identify the usual balance performance of healthy, community-dwelling people aged 70 years or more.

What is already known on this topic: The Berg Balance Scale scores balance from 0 (very poor to 56 (normal)) and is widely used in many clinical populations. It has well-established, favourable discriminant properties.

What this study adds: Normative data from community-dwelling people aged 70 years or more suggest a normal Berg Balance Scale score. With each subsequent year, however, mean scores decrease by about 0.7 points, and variability in the scores increases.

Ethics: Not applicable.
Competing interests: Nil.
Support: This research was conducted as part of a master’s degree by Stephen Downs with the University of Newcastle. The University provided academic supervision and use of the library, including electronically accessing papers and the use of 'get-it' to access papers not electronically available. Support has also been provided to attend conferences to present research findings. No direct financial support has been provided.

Acknowledgements: The authors acknowledge Alastair Meredith, who provided biostatistical advice while he was a trainee biostatistician with the NSW Centre for Epidemiology and Research.

Correspondence: Stephen Downs, Transitional Aged Care Service, Bellingen Hospital, Bellingen 2454, Australia. Email: Stephen.down@nswhealth.health.nsw.gov.au