Use of the Berg Balance Scale for Predicting Multiple Falls in Community-Dwelling Elderly People: A Prospective Study

Susan W Muir, Katherine Berg, Bert Chesworth, Mark Speechley

Background and Purpose

Falls are a significant public health concern for older adults; early identification of people at high risk for falling facilitates the provision of rehabilitation treatment to reduce future fall risk. The objective of this prospective cohort study was to examine the predictive validity of the Berg Balance Scale (BBS) for 3 types of outcomes—any fall (≥1 fall), multiple falls (≥2 falls), and injurious falls—by use of sensitivity, specificity, receiver operating characteristic (ROC) curves, area under the curve, and likelihood ratios.

Subjects and Methods

A sample of 210 community-dwelling older adults received a comprehensive geriatric assessment at baseline, which included the BBS to measure balance. Data on prospective falls were collected monthly for a year. The predictive validity of the BBS for the identification of future fall risk was evaluated.

Results

The BBS had good discriminative ability to predict multiple falls when ROC analysis was used. However, the use of the BBS as a dichotomous scale, with a threshold of ≤45, was inadequate for the identification of the majority of people at risk for falling in the future, with sensitivities of 25% and 45% for any fall and for multiple falls, respectively. The use of likelihood ratios, maintaining the BBS as a multilevel scale, demonstrated a gradient of risk across scores, with fall risk increasing as scores decreased.

Discussion and Conclusion

The use of the BBS as a dichotomous scale to identify people at high risk for falling should be discouraged because it fails to identify the majority of such people. The predictive validity of this scale for multiple falls is superior to that for other types of falls, and the use of likelihood ratios preserves the gradient of risk across the whole range of scores.

SW Muir, BSc, BScPT, is a PhD candidate, Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, University of Western Ontario, Kresge Building, London, Ontario, Canada N6A 5C1. Address all correspondence to Ms Muir at: susan.muir@uwo.ca.

K Berg, PT, PhD, is Chair and Associate Professor, Department of Physical Therapy, University of Toronto, Toronto, Ontario, Canada.

B Chesworth, BA, BScPT, MClScPT, PhD, is Assistant Professor, School of Physical Therapy, Faculty of Health Sciences, University of Western Ontario.

M Speechley, PhD, is Associate Professor, Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, University of Western Ontario.

[Muir SW, Berg K, Chesworth B, Speechley M. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people: a prospective study. *Phys Ther*. 2008;88:449–459.]

The Berg Balance Scale (BBS) was developed as a clinical measure of functional balance specifically in older people. Suggested applications were comparing balance between groups of people, describing balance in an individual, monitoring status over time, and evaluating treatment effectiveness.[1] The psychometric properties of reliability and validity for the scale have been well demonstrated.[1–3] Balance is critical in the performance of normal physical activities, and balance impairment is a key risk factor for falls in older people. Although balance impairment is neither a necessary nor a sufficient cause of falls and therefore is not present in all people who have fallen or who will fall, it has been found to increase the risk for falling, with a mean relative risk of 2.9 across 11 studies.[4,5]

It was a natural extension from the epidemiologic studies that identified an association between balance impairment and falls to evaluate the application of specific clinical measures of balance in the identification of people at risk for falling. Falls in older people are a significant public health issue, for which the evidence clearly suggests that risk can be reduced through the identification and treatment of modifiable risk factors.[6] The key challenges in the prevention of falls in older people include successfully identifying people at risk, understanding the influence and interaction of risk factors, and appropriately targeting meaningful treatment to improve or maintain functional autonomy and quality of life. The identification of people at risk facilitates the establishment of both preventive and rehabilitative therapies to ameliorate the adverse consequences of functional decline.

Although the BBS was not advocated by the original authors for use as a dichotomous scale, the value of 45 has been quoted by other authors as the recommended threshold value.[7,8] The use of the BBS score in a dichotomous format as a predictive tool for identifying people at an increased risk for falling has produced mixed results for community-dwelling older people.[2,9–14] This heterogeneity reflects differences across the samples in the prevalence of balance impairment and other risk factors for falls, frequencies of falls, and outcome definitions for falls (single, multiple, and injurious). Many investigators have used the retrospective identification of people with a past history of falls for validity studies; few prospective studies have evaluated predictive validity. It is important in establishing the validity of measurement scales to demonstrate the success of the use of the scales in multiple populations.

Because falls are multifactorial, the goal of screening is to successfully integrate information from the assessment of several domains to create an overall risk score. Given the low sensitivity of functional assessment scales for balance impairment in retrospective study designs, a lack of studies predicting the risk for falling in community-dwelling older people, inconsistent findings within the literature regarding the best score at which to dichotomize the BBS, and the need to integrate the BBS score with other information for determining fall risk, further evaluation of methods of applying the BBS is warranted. The use of likelihood ratios and prospective data for the prediction of fall risk has not been evaluated and would be a valuable addition to clinical practice if found relevant.

The purpose of this study was to examine the predictive validity of the BBS for 3 types of falls—any fall ($\geq$1 fall), multiple falls ($\geq$2 falls), and injurious falls—in community-dwelling older people by use of conventional data analysis techniques of sensitivity, specificity, receiver operating characteristic (ROC) curves, and area under the curve (AUC) and by use of likelihood ratios.

Method

Study Design

The BBS was evaluated with data on the occurrence of prospective falls that were collected during a field trial of prevention of falls, "The Project to Prevent Falls in Veterans" (PPFV). The PPFV was carried out at the University of Western Ontario and was funded jointly by Veterans Affairs Canada and Health Canada. The sampling and data collection procedures for the initial phase of the PPFV have been described in more detail elsewhere.[15]

The first phase of the PPFV was a cross-sectional mailed survey of a simple random sample generated by Veterans Affairs Canada of World War II and Korean War veterans residing in 3 regions of southwestern Ontario, Canada. Information was obtained from veterans and their caregivers who lived in the community. Respondents to the questionnaire living in London, Ontario, and the immediate surrounding region (n=1,192) were eligible to participate in the second phase of the PPFV, a prospective field intervention trial of the modification of fall risk factors.

In the risk factor modification trial, eligible candidates were divided into groups based on the number of modifiable risk factors reported on the mailed questionnaire. The use of mailed questionnaires has been demonstrated to be reliable and valid in screening community-dwelling older adults.[16] The modifiable risk factors enumerated on the questionnaire were as follows: greater than 4 prescription medications; lower-extremity muscle weakness; and balance, foot, and vision problems. The people who reported no modifiable risk factors on the questionnaire formed the control arm of the trial. Candidates who reported 1 to 5 modifiable risk factors were stratified into groups based on the number of risk factors and were then randomized to 1 of 2 treatment groups within each stratum: a regional geriatric care program or community-based primary care. The intervention, specific information on the prevention of falls, was provided to participants on the basis of the risk factors identified from the questionnaire and a geriatric assessment. The group randomized to the regional geriatric care program and the people who reported no risk factors received a one-time comprehensive geriatric assessment performed by a geriatrician using a preliminary version of the interRAI CHA (Community Health Assessment).[17,18] A total of 339 people participated in the second phase of the PPFV; 210 of them, including participants in the control and intervention arms, received the comprehensive geriatric assessment. All participants who received the comprehensive geriatric assessment formed the subgroup included in the analysis in the present study (Fig. 1).

Study Participants

Of the 210 people eligible for this analysis (ie, received the comprehensive geriatric assessment), 187 (89%) had complete data at the end of the 1-year follow-up period. For these 187 people, the mean age was 79.47 years (SD=5.83, range=47–90), and 65% were men. Participant characteristics are shown in Table 1. There were few differences between people included in the analysis (n=187) and those lost to follow-up (n=23), with only "decreased physical activity in the last 3 days" (P=.02) was statistically significant; BBS score (P>.05), history of falls (P=.08), lower-extremity weakness (P=.26), ≥4 prescription medications (P=.12), vision problems (P=.18), use of mobility aids (P=.19), cognitive impairment (P=.42), fear of falling (P=.39), decline in activities of daily living in the preceding 3 months (P=.33), and dependence in activities of daily living (P=.17) were not statistically significant. Of the 187 people with complete follow-up information, 80 (42.8%) sustained any fall, of whom 33 (17.6%) sustained multiple falls and 55 (29.4%) sustained at least 1 injurious fall.

Assessments

The basic version of the CHA is a subset of the Minimum Data Set for Home Care, version 2.0.[17,18] The reliability and validity of all items of the Minimum Data Set for Home Care in community settings have been reported.[19,20] The interRAI CHA was conceived as a modular assessment for relatively well people living in the community. People identified as having functional problems or mental health issues on the basic assessment could receive further evaluation through a specialized subassessment, the functional module or the mental health module, respectively. In the present study, the functional module and the mental health module were not used; only the basic version of the interRAI CHA was used. An additional study-specific supplement included a full list of medications and the BBS. The BBS consists of 14 tasks that are each scored on a scale of 0 to 4, for a total possible score of 56, indicating no identified balance difficulties. The scoring is graded such that a score of 0 is assigned if a person is unable to perform a task and a score of 4 is assigned when the task is performed independently.

Prospective information on daily falls was collected for a year by use of monthly mailed "fall calendars." Participants were instructed to record all falls and to mail in the calendars at the end of each month. People who indicated falling in a given month were contacted by telephone and interviewed to obtain detailed information about the specifics of the fall, including the location and activity at the time of each fall, whether they were injured, and whether they consulted a physician or went to a hospital because of the fall. A fall was defined as coming to rest unintentionally on the floor or ground. An injurious fall was defined as a fall resulting in an injury that required a person to see a physician. Informed consent was obtained from all study participants. Data collection for the baseline BBS was started in May 2002, and collection of 1-year follow-up information on prospective falls was completed in January 2004.

The present study was a secondary analysis of data from the PPFV; therefore, the *post hoc* sample size calculation for the analysis of sensitivity and specificity reflects the precision available from the number of fall events recorded over the follow-up period, with the number of people experiencing multiple falls constituting the minimum sample size available (n=33). Calculations were based on an expected sensitivity and specificity of 0.85 for the BBS according to the range of values found in the literature.[9–12,14] On the basis of the methods and sample size tables of Flahault et al[21] and an expected sensitivity and specificity of 0.85 in a population with a yearly fall risk of 30%, a sample of 33 people who sustained a fall during the study period ("fallers") and 67 people who did not sustain a fall during the study period ("nonfallers") will ensure that the lower limit of the 95% confidence interval will be greater than 0.65 with a probability of .95.[21]

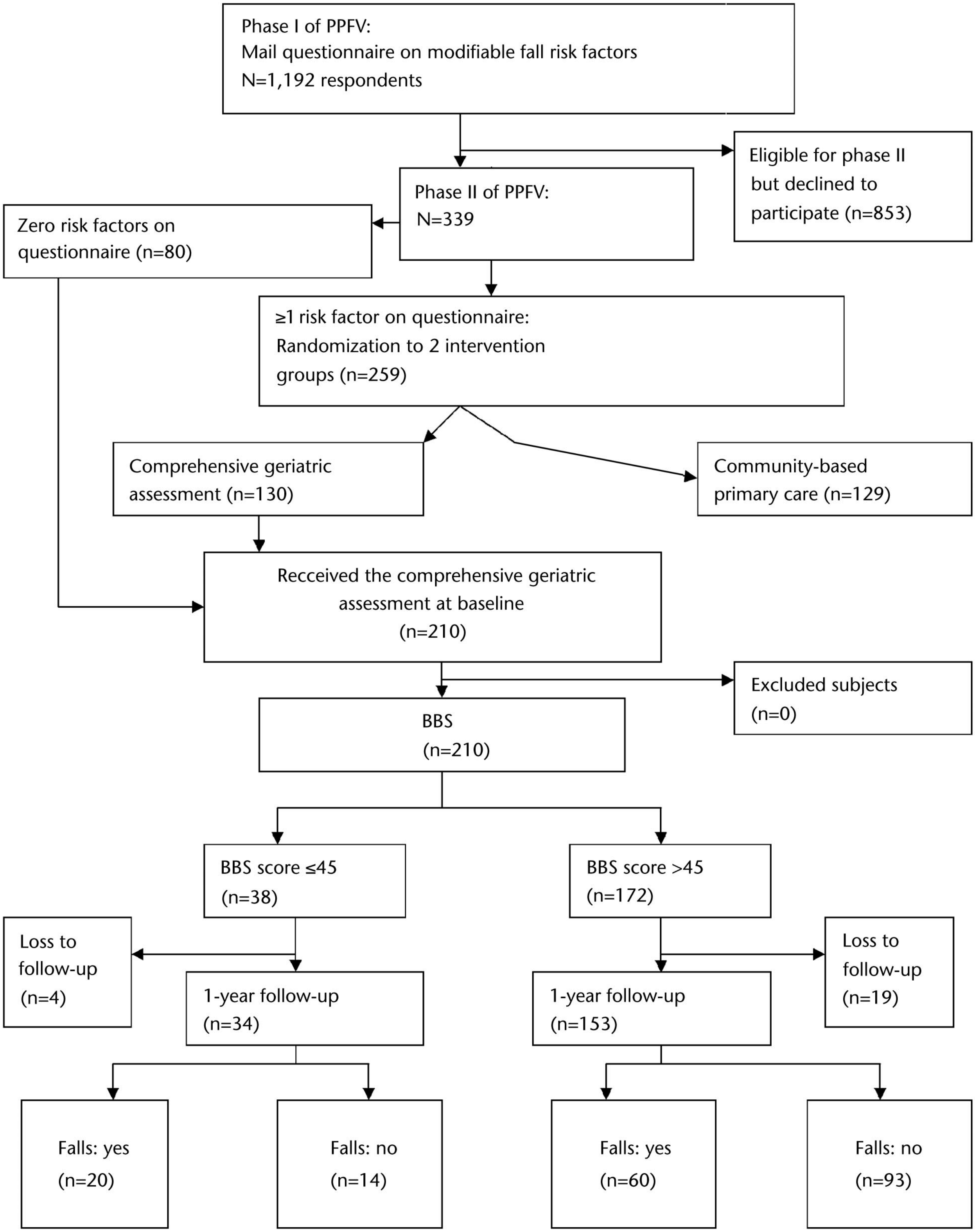

Figure 1.
Flow diagram for the first and second phases of The Project to Prevent Falls in Veterans (PPFV) and the subsample from the second phase that was included in the data analysis for the outcome any fall in the present study. BBS=Berg Balance Scale.

Table 1.
Study Participant Characteristics by Baseline Berg Balance Scale (BBS) Score (n=187)[a]

Variable	Value for Participants With BBS Scores of:	
	≤45 (n=34)	>45 (n=153)
BBS score, $\overline{X}$ (SD)	34.8 (10.8)	53.7 (2.8)
History of falling in the preceding 90 d	17 (50)	30 (20)
Lower-extremity weakness	20 (59)	7 (5)
Prescription medications (≥4)	25 (74)	111 (73)
Vision problems	7 (21)	13 (9)
Use of an assistive device	15 (44)	11 (7)
Unsteady gait	24 (71)	9 (6)
Cognitive impairment	5 (15)	5 (3)
Fear of falling	12 (35)	8 (5)
Decline in ADL status over preceding 3 mo	2 (6)	0 (0)
Physical activity in last 3 d (<2 h in total)	11 (32)	39 (26)
Any dependence in ADL	13 (38)	5 (3)

[a] Values are reported as number (percentage) unless otherwise indicated. Participant characteristics were obtained from the interRAI basic assessment form and the study-specific supplement for prescription medications and the BBS. More specifically, lower-extremity weakness was defined as the inability to stand up from a chair without using the arms of the chair; vision problems were defined as the inability to read fine print in adequate light with glasses if needed; an unsteady gait was assessed with an observational analysis of gait quality; cognitive impairment was defined as anything less than participants being able to make independent decisions that were consistent, reasonable, and safe in organizing their day; fear of falling was indicated by a “yes” response when a participant was asked whether he or she limited going outdoors because of a fear of falling; dependence in activities of daily living (ADL) (mobility in bed, locomotion in home, dressing of the upper and lower body, eating, toilet use, personal hygiene, and bathing) was defined as any state other than complete independence. The number of prescription medications was obtained from a detailed inventory that included dose and frequency of drug administration. All other information included in this table is from the participants’ responses to questions. The hours of physical activity included tasks such as walking, cleaning house, and exercising. A decline in ADL status was indicated by a participant feeling more impaired in self-performance at the time of assessment than at 90 days previously.

Data Analysis

Statistical analyses were performed with SAS, version 8.2.* The numbers of fallers and nonfallers indicated by the “fall calendar” data for prospective falls were obtained at each BBS score (possible scores of between 0 and 56) for each of the 3 fall outcomes: any fall, multiple falls, and injurious falls. Sensitivities and specificities were calculated from the numbers at each BBS score and plotted as an ROC curve for each of the 3 fall outcomes. The AUC was obtained as the “c statistic” from separate univariate logistic regression analyses of each binary fall outcome regressed on the BBS score as a continuous value. Sensitivities and specificities were reported in the present study for the dichotomized BBS for each fall outcome category with 2 cutoff values: ≤45, as recommended in previous articles, and the optimal score, the BBS score that was closest to the upper left-hand corner of the graph, determined from the ROC curve.

For likelihood ratios, raw BBS scores were aggregated into 5 levels each 5 points wide, with scores of less than 40 being placed into one category because of small cell sizes. The formation of the quintiles was chosen to reflect potential clinical utility and the format used by Riddle and Stratford.[22] The percentage of fallers in each quintile was calculated as the number of fallers in a quintile divided by the total number of subjects in the quintile to quantify the risk gradient across the quintiles for each fall outcome. The positive likelihood ratio [sensitivity/(1−specificity)] and the 95% confidence interval were calculated for each quintile of the scale for the 3 fall outcomes. The first set of likelihood ratios was calculated for the outcome any fall to compare fallers with nonfallers. The second set of likelihood ratios was calculated for the outcome multiple falls to compare people who sustained ≥2 falls with a combined group of nonfallers and people who sustained single falls. The third set of likelihood ratios was calculated for the outcome injurious falls to compare people who sustained at least 1 injurious fall with a combined group of nonfallers and fallers who did not sustain any injury. A likelihood ratio of 1.0 indicates no additional risk prediction from the use of a test and is equivalent to a test in which sensitivity and specificity are both equal to 50%.

* SAS Institute Inc, PO Box 8000, Cary, NC 27511.

Results

Dichotomizing the BBS at a score of 45 resulted in groups with different 1-year probabilities for falling, in the expected direction: 58% of people (20/34) with BBS scores at or below 45 fell, and 39% of people (60/153) with scores above 45 fell. As determined from ROC analysis, the optimal single cutoff value for any fall was 54 (AUC=0.59), that for multiple falls was 53 (AUC=0.68), and that for injurious falls was 54 (AUC=0.60) (Fig. 2). Sensitivity and specificity values for fall outcomes at both the cutoff value of ≤45 and the optimal value obtained from the ROC curve are shown in Table 2. People in all 3 fall outcome groups had BBS

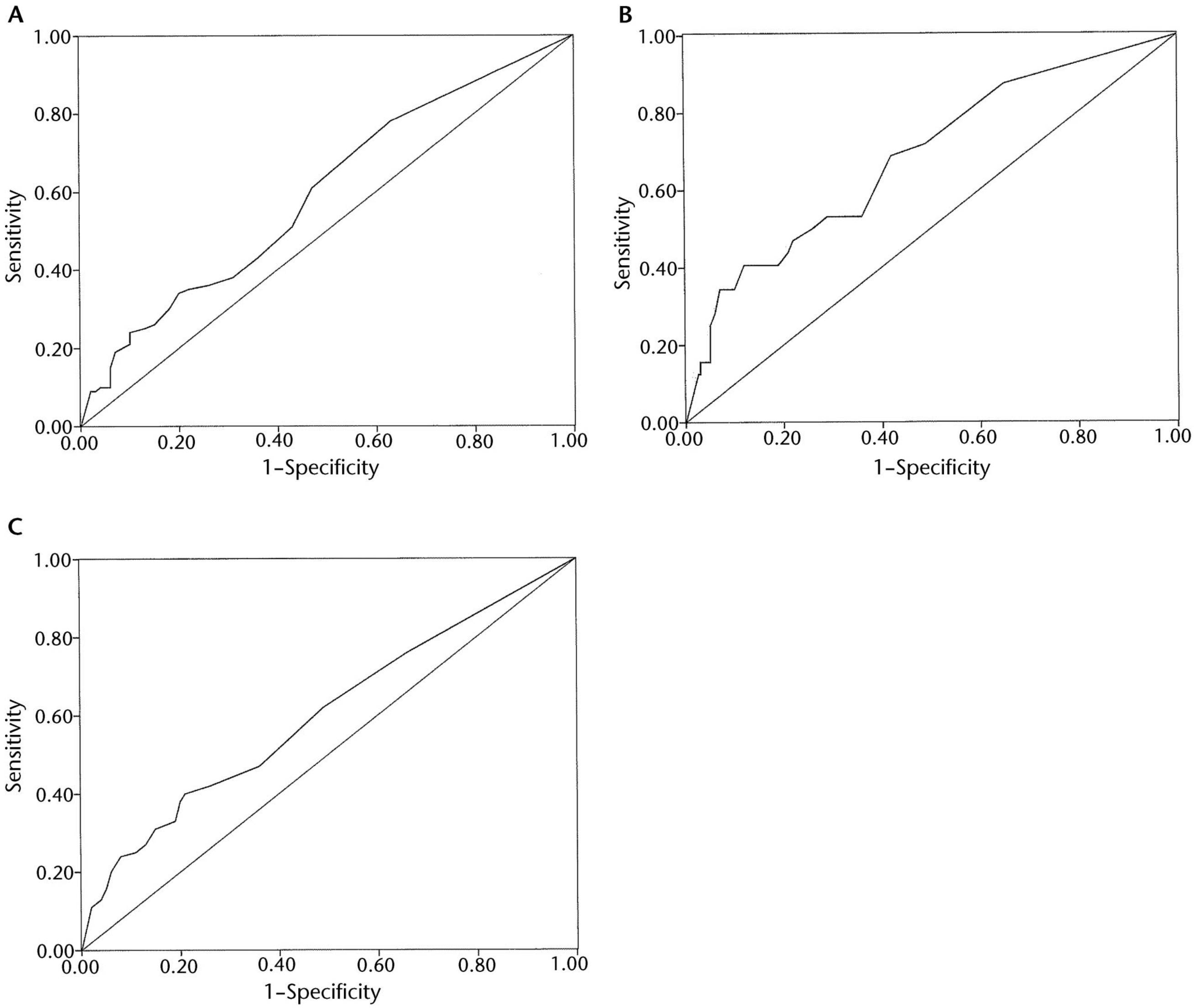

Figure 2.
Receiver operating characteristic (ROC) curves and area under the curve (AUC) for the outcomes any fall (all fallers compared with nonfallers, AUC=0.59), multiple falls (people with ≥2 falls compared with nonfallers and people with single falls, AUC=0.68), and injurious falls (people sustaining fall-related injuries compared with nonfallers and fallers without injuries, AUC=0.60) (n=187).

Table 2.
Sensitivity and Specificity Values for Fall Outcomes at the Cutoff Value of 45 and the Optimal Cutoff Value Determined From the Receiver Operating Characteristic Curve (n=187)

Outcome (No. of Participants)	Cutoff Value (No. of Participants Below Threshold)	No. of Participants Who Fell and Were Identified as at Increased Risk for Falling	Sensitivity (95% Confidence Interval)	Specificity (95% Confidence Interval)
Any fall (80)	≤45 (57)	20	0.25 (0.16–0.36)	0.87 (0.79–0.92)
	≤54 (122)	49	0.61 (0.50–0.72)	0.53 (0.43–0.63)
Multiple falls (33)	≤45 (57)	14	0.42 (0.26–0.61)	0.87 (0.79–0.92)
	≤53 (110)	22	0.69 (0.50–0.83)	0.57 (0.47–0.66)
Injurious falls (55)	≤45 (57)	16	0.29 (0.18–0.43)	0.86 (0.79–0.92)
	≤54 (122)	30	0.62 (0.48–0.74)	0.51 (0.42–0.60)

Table 3.
Positive Likelihood Ratios Across 5 Intervals of Berg Balance Scale (BBS) Scores for the Outcome Any Fall (n=187)[a]

BBS Scores (No. of Participants)	Fallers (n=80)		Nonfallers (n=107)		Positive Likelihood Ratio (95% Confidence Interval)
	No.	Proportion	No.	Proportion	
<40 (19)	12	12/80=0.15	7	7/107=0.07	2.29 (0.95–5.56)
40–44 (13)	7	7/80=0.09	6	6/107=0.06	1.56 (0.55–4.47)
45–49 (18)	9	9/80=0.11	9	9/107=0.08	1.34 (0.56–3.22)
50–54 (49)	21	21/80=0.26	28	28/107=0.26	1.003 (0.62–1.63)
≥55 (88)	31	31/80=0.39	57	57/107=0.53	0.73 (0.52–1.01)

[a] "Fallers" were people who sustained a fall over the study period, and "nonfallers" were people who did not fall over the study period.

scores ranging from 30 to 56, as did those who did not fall. In this sample of community-dwelling older adults, the cutoff value of 45 recommended from previous studies produced sensitivities of 25% to 42% and specificities of 86% to 87% for the 3 fall outcomes. The high cutoff values required to optimize sensitivity in each fall outcome category indicate that balance impairment alone does not define increased fall risk and that falls are frequent among people with scores above 45.

Dividing BBS scores into quintiles and using this multilevel scale produced the positive likelihood ratios for each fall outcome; the results are shown in Tables 3, 4, and 5. Across the quintiles for the outcome any fall (Tab. 3), the risk for falling increased as the quintile score decreased: fall risk started at 35% (31/88) for BBS scores of ≥55 and increased for each consecutively lower band at 43% (21/49), 50% (9/18), 53% (7/13), and 63% (12/19) for BBS scores of <40. These data corresponded to likelihood ratio estimates demonstrating an increased risk for falling from baseline risk with BBS scores of 45 to 49 (likelihood ratio=1.34) and increasing progressively as the scores decreased.

For the outcome multiple falls (Tab. 4), the risk for falling also demonstrated an increasing gradient from the highest score to the lowest score: fall risk was 10% (9/88) for BBS scores of ≥55 and increased successively for each lower band at 16% (8/49), 11% (2/18), 31% (4/13), and 54% (10/19) for BBS scores of <40. The transition point for increased risk according to likelihood ratios started at BBS scores of 40 to 44 (likelihood ratio=2.07), and there was a marked increase for scores of less than 40 (likelihood ratio=5.19). The outcome injurious falls demonstrated a gradient of increasing risk from the highest score to the lowest score (Tab. 5): fall risk started at 24% (21/88) for BBS scores of ≥55 and increased successively for each lower band at 24% (12/49), 39% (4/13), and 58% (11/19) for BBS scores of <40. The likelihood ratios started to increase at BBS scores of 45 to 49 (likelihood ratio=1.53) and continued to increase as BBS scores decreased. A clear gradient of increasing risk with decreasing BBS score was demonstrated for each fall outcome, although the confidence intervals included 1.0 for all levels except scores of less than 40 for multiple falls and injurious falls, which were both statistically significant. The likelihood ratios for multiple falls at BBS scores of less than 40 produced a moderate shift from the baseline probability of falling, indi-

Table 4.
Positive Likelihood Ratios Across 5 Intervals of Berg Balance Scale (BBS) Scores for the Outcome Multiple Falls (n=187)[a]

BBS Scores (No. of Participants)	Fallers (n=33)		Nonfallers and Single Fallers (n=154)		Positive Likelihood Ratio (95% Confidence Interval)
	No.	Proportion	No.	Proportion	
<40 (19)	10	10/33=0.30	9	9/154=0.06	5.19 (2.29–11.75)
40–44 (13)	4	4/33=0.12	9	9/154=0.06	2.07 (0.68–6.33)
45–49 (18)	2	2/33=0.06	16	16/154=0.10	0.58 (0.14–2.42)
50–54 (49)	8	8/33=0.24	41	41/154=0.27	0.91 (0.47–1.76)
≥55 (88)	9	9/33=0.27	79	79/154=0.51	0.53 (0.30–0.95)

[a] "Fallers" were people who sustained a fall over the study period, "nonfallers" were people who did not fall over the study period, and "single fallers" were people who sustained single falls.

Table 5.
Positive Likelihood Ratios Across 5 Intervals of Berg Balance Scale (BBS) Scores for the Outcome Injurious Falls (n=87)[a]

BBS Scores (No. of Participants)	Fallers Sustaining Injuries (n=55)		Nonfallers and Fallers Not Sustaining Injuries (n=132)		Positive Likelihood Ratio (95% Confidence Interval)
	No.	Proportion	No.	Proportion	
<40 (19)	11	11/55=0.2	8	8/132=0.06	3.3 (1.40–7.76)
40–44 (13)	4	4/55=0.07	9	9/132=0.07	1.07 (0.34–3.32)
45–49 (18)	7	7/55=0.13	11	11/132=0.08	1.53 (0.63–3.73)
50–54 (49)	12	12/55=0.22	37	37/132=0.28	0.78 (0.44–1.38)
≥55 (88)	21	21/55=0.38	67	67/132=0.51	0.75 (0.52–1.1)

[a] "Fallers" were people who sustained a fall over the study period, and "nonfallers" were people who did not fall over the study period.

cating a value that may be clinically useful in predicting risk.[23] The gradient of likelihood ratios for injurious falls was not continuous across the BBS scores and was nearly equal to the baseline for BBS scores of 40 to 44. This result may have been attributable to the small number of fallers in that band of scores resulting in an unstable estimate.

Discussion

When conventional statistical techniques of sensitivity and specificity were used to evaluate the performance of the BBS in predicting falls, the previously recommended cutoff value of 45 was inadequate for identifying the majority of future fallers in this community sample. Dichotomizing the score did produce 2 groups with different risks of falling, but many falls occurred in the low-risk group. The ROC curve results demonstrated that to optimize sensitivity, higher cutoff scores would be needed to predict each outcome. The use of these higher cutoff scores would change the operational definition of balance impairment, so that some people defined as having balance impairment would actually have an independent and functional range of balance abilities. Even with the higher cutoff scores, the sensitivities would be suboptimal for a clinical setting. For example, the higher cutoff score for the outcome any fall only provided a sensitivity of 61%. Rather than revising the operational definition of balance impairment to make it better fit the outcome, it makes more sense to emphasize that falls in older people are multifactorial, so that any dichotomous view of balance impairment alone does not explain or adequately quantify all future risk. When the BBS was used as a multilevel tool, rather than a dichotomous tool, however, likelihood ratios demonstrated a clear gradient of risk for each fall outcome. The best discriminative ability was found for multiple falls, for which the likelihood ratios indicated that risk increased below a score of 45, with a significant increase below 40. Likelihood ratios have 3 important properties that overcome the limitations of other measures of diagnostic test accuracy, such as sensitivity and specificity, for evaluating screening or diagnostic tests: (1) they do not change with the pretest probability of disease, (2) they can be calculated for multiple levels of test results and preserve useful information that is lost when a multiple-level test is dichotomized, and (3) they allow for the assessment of the impact of a test result on the odds that a person will have a particular outcome.[24,25] Information from prospective epidemiological studies has demonstrated that the risk for falling increases as the number of risk factors increases.[1,26,27] Balance is but one domain that is evaluated in screening for falls; therefore, the next step in gaining an understanding of fall risk is the integration of information from each of the individual components of a fall assessment to obtain an overall risk. The use of likelihood ratios facilitates this integration of information by allowing for the modification of pretest probability upon testing of balance.

Methodological limitations have affected previous studies assessing the predictive validity of the BBS.[9,11,12,14] In the original study by Berg et al,[2] the value of 45 was used to allow the calculation of relative risk estimates as a demonstration of predictive validity, but it was not recommended that the scale be dichotomized at this value for use. This aspect of the original validation study has been misunderstood and has affected studies that have used this value for research purposes. Additionally, the majority of validity studies have used a case-control design to evaluate whether the cutoff value of 45 can distinguish people with a past history of falls from those without such a history.[9,11,12,14] The most significant limitation of this design is that the assessment of balance at unspecified and variable times after the outcome of falling does not allow establishment of the temporal order of balance impairment and falling. Balance impairment may be the result of fall-related injuries or subsequent selfimposed limitations in activity rather than having preceded falling. Also, case-control studies cannot be used to estimate probabilities, which are required to evaluate predictive validity. In the present study, we used data from a cohort with fall information collected prospectively over a 1-year time frame and avoided the limitations associated with case-control studies.

No previous study of prospective falls calculated the ROC curve in order to determine the AUC in evaluating the predictive validity of the BBS.[9,10] The AUC evaluates the diagnostic accuracy of the test because the area is equal to the probability of accurately discriminating between a randomly chosen person with the outcome and a randomly chosen person without the outcome.[28] An uninformative test will have an AUC of 0.5, in which case determining a cutoff value that maximizes sensitivity and specificity is moot. Three case-control studies[9,11,14] evaluated the optimal value for dichotomizing the BBS, and although there was variability in the values (≤45 for any fall, ≤49 and ≤38 for multiple falls, and ≤47 for single falls), the limitations of the case-control study design greatly affected the utility of this information. In one of these studies,[11] the ROC curve and the AUC were calculated, but the value of the information was limited by the retrospective nature of distinguishing fallers from nonfallers. In their original validation study, Berg et al[2] demonstrated, through multiple logistic regression, that the BBS score as a continuous value, a history of falls in the preceding 3 months, and visual problems predicted future falls.

The only prospective study[10] evaluating the BBS as a dichotomous scale did not evaluate an optimal threshold; instead, the threshold was set at ≤45 and demonstrated a sensitivity of 53%. The results of the present study are consistent with those of Bogle Thorbahn and Newton[10] in that the magnitude of the sensitivity indicated that the use of the BBS as a dichotomous scale and its use alone do not predict future fall risk well. The present study is the first to use data on prospective falls collected over a 1-year period to evaluate the predictive validity of the BBS and to use a comprehensive methodological analysis of sensitivity, specificity, ROC curves, AUC, and likelihood ratios. Our study results reaffirm the finding of the original validation study of the BBS by Berg et al[2] that falls are multifactorial. Additionally, the BBS in the original validation study was used to evaluate people for multiple falls, and in the present study, we found that the power of the measurement scale was better for evaluating multiple falls than for evaluating any falls or injurious falls.

The ROC curves demonstrated that the AUCs for the outcomes any fall and injurious falls were close to the uninformative test value of 0.50, at 0.59 and 0.60, respectively. An AUC of 0.8 has been stated to represent a reasonably powerful model.[24] The AUC for the outcome multiple falls was 0.68; this value indicated that the test provides moderate discriminative value for this outcome and that this outcome should be the outcome of interest when the BBS is used to evaluate future risk in samples that are similar to ours.

Riddle and Stratford[22] calculated likelihood ratios for a multilevel BBS using previously published data and were able to demonstrate a gradient of risk. Unfortunately, these values will be influenced by the limitations of the study methodologies used to derive the data on falls in case-control studies and the use of study designs with different outcomes, a past history of falls and future falls.[22] A gradient of scores on the BBS was previously reported for the categories visual impairment and ambulatory ability, with the BBS score decreasing as impairment increased.[2,29] Shumway-Cook et al[14] found that a 1-point change in the BBS score led to different probabilities of falling, a finding that emphasizes the presence of a gradient of risk across the scale. The present research findings further support the gradient of information across the range of scores that can be lost when the BBS is used as a dichotomous scale. Additionally, the findings demonstrate that the type of fall outcome has an influence on the transition point in the range of scores indicating an increased fall risk above the baseline level of risk.

The present study has several limitations. One is that the data on prospective falls were derived from one arm of an intervention study in which the BBS formed part of the baseline risk factor assessment. Because the participants were exposed to information on the prevention of falls, they may have received recommendations related to an identified balance impairment. If these recommendations were followed and were successful in improving balance, then the results presented here may underestimate the predictive validity of falls in community-dwelling older people. We believe that this probability is unlikely because there was no difference in fall risk between the 2 study arms (Speechley and colleagues, unpublished data). The self-report of injurious falls was not validated with medical records or documentation of a clinic visit for medical attention; therefore, there may have been a risk of underreporting of this fall outcome, resulting in the findings being conservative and underestimating the risk of sustaining injurious falls.

Another limitation is the potential lack of generalizability of the results to the general population of older adults. Our sample participants may not be representative of the general population of older adults because they were volunteers for a study on the preventions of falls. Our sample had more men than most other studies of older adults, and the men were military veterans. As such, they may have been different from their contemporaries who were not veterans as a result of their military experience. Finally, the overall risk for falling in our sample was 42.8%, a value that was higher than the values of 29% to 35% typically cited for community-dwelling older people.[26,30,31] This increased risk could be partially explained by the loss of follow-up information on healthier people. As determined by a sensitivity analysis assuming that the losses to follow-up remained in the study under the conditions "all did not have the outcome" and "all did have the outcome," the overall risk for falling would range from 38% to 55% and still would represent a higher-risk group.

A second set of sensitivity analyses was performed to evaluate the impact of the composition of the comparison group on the calculation of the likelihood ratios for the outcomes multiple falls and injurious falls. Likelihood ratios were recalculated with only nonfallers as the comparison group. The new calculations, not presented here, produced likelihood ratios of the same magnitude, and the trend of increasing risk with decreasing score was consistent with the original calculations presented in Tables 3, 4, and 5.

The sample size available for data analysis was based on a study that was not designed specifically to answer our research question. The sample was large enough to find statistically significant differences with the BBS used as a dichotomous scale. The sample size had limitations with the BBS applied as a multilevel scale, as demonstrated by the wider confidence intervals and small cell sizes in some bands. The numbers of multiple falls and injurious falls were small, and this factor may have affected the stability of the measures calculated. Larger studies with more events are required for a full evaluation of the predictive ability of the BBS across its full scale range. A total of 23 subjects with incomplete follow-up data were excluded from the analysis. As discussed in the "Method" section, the subjects with missing data were not statistically different from those with complete follow-up information, except for "decreased physical activity in the last 3 days." Excluding the subjects with incomplete follow-up data could have led to an increased observed association between the BBS scores and the fall outcomes. Although they represented 11% of the original sample, their effect on the association was probably minimal because no association was found when dichotomized BBS scores were used and the ranges of scores were the same for fallers and nonfallers.

The strengths of the present study are the prospective design and the method used to ascertain the fall outcomes in a large sample of community-dwelling older adults. A comprehensive statistical methodology was used to evaluate the predictive validity of the BBS, examining the measurement tool as both a dichotomous scale and a continuous scale. The present study represents the most comprehensive use of statistical analysis techniques (sensitivity, specificity, ROC curves, AUC, and likelihood ratios) to evaluate the predictive validity of the BBS with the ideal format of data on the occurrence of prospective falls that were collected over a 1-year follow-up period. Likelihood ratios are another method of describing the performance of a diagnostic test and one that is more common in the medical literature than the physical therapy literature; therefore, this research also demonstrates the utility of the use of likelihood ratios in physical therapy research and practice. Because falls have consistently been found to be multifactorial, future research should now be directed at prospective methods for evaluating risk with multiple domains. Although derivations of multifactor predictive scales have been published in the literature, there is a dearth of research establishing the predictive validity of these scales.[4,32–37]

Conclusion

The goal of screening is to identify people at risk for falling in the imminent future for further in-depth evaluation. Conventional data analysis techniques with ROC curves indicate that the BBS has better discriminatory ability for identifying people sustaining recurrent or multiple falls than for identifying those who fall once or sustain an injury among community-dwelling older people. The effectiveness of the BBS as a dichotomous scale in identifying people at risk for falling is lower than that of a multilevel form of the scale with likelihood ratios. This new information, derived from a prospective study, and the fact that the BBS was not originally recommended as a dichotomous scale justify the recommendation that the use of a cutoff value of 45 should cease. We evaluated a method for integrating balance assessment through the BBS with other fall risk information into a prediction of future fall risk. Likelihood ratios preserve the risk gradient that is present over the whole scale, overcome the problem that fall risk is substantial above a cutoff value of 45, and permit calculation of the probability of falling for a given individual.

Ms Muir and Dr Speechley provided concept/idea/research design. All authors provided writing. Dr Speechley provided data collection, fund procurement, and subjects. Ms Muir provided data analysis.

This project was approved by the University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects.

This article was submitted August 29, 2007, and was accepted November 28, 2007.

DOI: 10.2522/ptj.20070251

References

1 Berg KO, Wood-Dauphinée SL, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can.* 1989;41:304–311.

2 Berg KO, Wood-Dauphinée SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. *Can J Publ Health.* 1992;S2:S7–S11.

3 Berg KO, Wood-Dauphinée SL, Williams JI. The balance scale: reliability assessment with elderly residents and patients with an acute stoke. *Scand J Rehabil.* 1995;27:27–36.

4 American Geriatrics Society, British Geriatrics Society, and American Academy of Orthopaedic Surgeons Panel of Falls Prevention. Guidelines for the prevention of falls in older persons. *J Am Geriatr Soc.* 2001;49:664–672.

5 National Institute of Clinical Excellence. Clinical guideline 21: the assessment and prevention of falls in older people, 2004. Available at: http://www.nice.org.uk. Accessed July 11, 2005.

6 Gillespie LD, Gillespie WJ, Robertson MC, et al. Interventions for preventing falls in elderly people. *Cochrane Database Syst Rev.* 2003;(4):CD000340.

7 Perell KL, Nelson A, Goldman RL, et al. Fall risk assessment measures: an analytic review. *J Gerontol A Biol Sci Med Sci.* 2001;56:M761–M766.

8 Scott V, Votova K, Scanlan A, Close J. Multifactorial and functional mobility assessment tool for fall risk among older adults in community, home-support, long-term and acute care settings. *Age Ageing.* 2007;36:130–139.

9 Lajoie Y, Gallagher SP. Predicting falls within the elderly community: comparison of postural sway, reaction time, the Berg Balance scale and the Activities-specific Balance Confidence (ABC) scale for comparing fallers and non-fallers. *Arch Gerontol Geriatr.* 2004;38:11–26.

10 Bogle Thorbahn LD, Newton RA. Use of the Berg Balance Test to predict falls in elderly persons. *Phys Ther.* 1996;76:576–585.

11 Chiu AYY, Au-Yeung SSY, Lo SK. A comparison of four functional tests in discriminating fallers from non-fallers in older people. *Disabil Rehabil.* 2003;25:45–50.

12 O'Brien K, Pickles B, Culham E. Clinical measures of balance in community-dwelling elderly female fallers and non-fallers. *Physiother Can.* 1998;3:212–217.

13 Boulgarides LK, McGinty SM, Willett JA, Barnes CW. Use of clinical and impairment-based tests to predict falls by community-dwelling older adults. *Phys Ther.* 2003;83:328–339.

14 Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. *Phys Ther.* 1997;77:812–819.

15 Speechley M, Belfry S, Borrie MJ, et al. Risk factors for falling among community-dwelling veterans and their caregivers. *Can J Aging.* 2005;24:261–274.

16 Alessi CA, Josephson KR, Harker JO, et al. The yield, reliability, and validity of a postal survey for screening community-dwelling older people. *J Am Geriatr Soc.* 2003;51:194–202.

17 Institute of Gerontology, University of Michigan. interRAI. Available at: http://www.interrai.org. Accessed December 11, 2007.

18 Morris JN, Bernabei R, Ikegami N, et al. *RAI-HOME CARE (RAI-HC) Assessment Manual for Version 2.0: Primer on Use of the Minimum Data Set for Home Care (MDS-HC) Version 2.0 and the Client Assessment Protocols (CAPS)—for use in Canada and the United States.* Washington, DC: interRAI Corp; 2002.

19 Landi F, Tua E, Onder G, et al. Minimum data set for home care: a valid instrument to assess frail older people living in the community. *Med Care.* 2000;38:1184–1190.

20 Morris JN, Fries BE, Steel K, et al. Comprehensive clinical assessment in community setting: applicability of the MDS-HC. *J Am Geriatr Soc.* 1997;45:1017–1024.

21 Flahault A, Cadilhac M, Thomas G. Sample size calculation should be performed for design accuracy in diagnostic test studies. *J Clin Epidemiol.* 2005;58:859–862.

22 Riddle DL, Stratford PW. Interpreting validity indexes for diagnostic tests: an illustration using the Berg Balance Test. *Phys Ther.* 1999;79:939–948.

23 Jaeschke R, Guyatt G, Lijmer J. Diagnostic tests. In: Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: a Manual for Evidence-based Clinical Practice.* Chicago, Ill: AMA Press; 2002:121–140.

24 Haynes RB, Sackett DL, Guyatt GH, Tugwell P. *Clinical Epidemiology: How To Do Clinical Practice Research.* 3rd ed. Philadelphia, Pa: Lippincott Williams & Wilkins; 2006.

25 Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol.* 1991;44:763–770.

26 Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls: a prospective study. *JAMA.* 1989;261:2663–2668.

27 Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *N Engl J Med.* 1988;319:1701–1707.

28 Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall Health; 2000.

29 Lee HKM, Scudds RJ. Comparison of balance in older people with and without visual impairment. *Age Ageing.* 2003;32:643–649.

30 Campbell JA, Borrie MJ, Spears GF. Risk factors for falls in a community-based prospective study of people 70 years and older. *J Gerontol.* 1989;44:M112–M117.

31 O'Loughlin JL, Robitaille Y, Boivin JF, Suissa S. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. *Am J Epidemiol.* 1993;137:342–354.

32 Stalenhoef PA, Diederiks JPM, Knottnerus JA, et al. A risk model for the prediction of recurrent falls in community-dwelling elderly: a prospective cohort study. *J Clin Epidemiol.* 2002;55:1088–1093.

33 Tromp AM, Smit JH, Deeg DJH, et al. Predictors for falls and fractures in the Longitudinal Aging Study Amsterdam. *J Bone Miner Res.* 1998;13:1932–1939.

34 Graafmans WC, Ooms ME, Hofstee HMA, et al. Falls in the elderly: a prospective study of risk factors and risk profiles. *Am J Epidemiol.* 1996;143:1129–1136.

35 Covinsky KE, Kahana E, Kahana B, et al. History and mobility exam index to identify community-dwelling elderly persons at risk of falling. *J Gerontol A Biol Sci. Med Sci.* 2001;56:M253–M259.

36 Pluijm SMF, Smit JH, Tromp EAM, Stel VS, Deeg DJH, Bouter LM, Lips P. A risk profile for identifying community-dwelling elderly with a high risk of recurrent falling: results of a 3-year prospective study. *Osteoporosis International.* 2006;17:417–425.

37 Stel VS, Smit JH, Pluijm SMF, Lips P. Balance and mobility performance as treatable risk factors for recurrent falling in older persons. *J Clin Epidemiol.* 2003;56:659–668.