Evaluation of Dynamic Balance Among Community-Dwelling Older Adult Fallers: A Generalizability Study of the Limits of Stability Test

Sean Clark, PhD, Debra J. Rose, PhD

ABSTRACT. Clark S, Rose DJ. Evaluation of dynamic balance among community-dwelling older adult fallers: a generalizability study of the Limits of Stability Test. Arch Phys Med Rehabil 2001;82:468-74.

Objective: To establish reliability estimates of the 75% Limits of Stability® Test (75% LOS test) when administered to community-dwelling older adults with a history of falls.

Design: Generalizability theory was used to estimate both the relative contribution of identified error sources to the total measurement error and generalizability coefficients. A random effects repeated-measures analysis of variance (ANOVA) was used to assess consistency of LOS test movement variables across both days and targets.

Setting: A motor control research laboratory in a university setting.

Participants: Fifty community-dwelling older adults with 2 or more falls in the previous year.

Main Outcome Measures: Spatial and temporal measures of dynamic balance derived from the 75% LOS test included average movement velocity, maximum center of gravity (COG) excursion, end-point COG excursion, and directional control.

Results: Estimated generalizability coefficients for 2 testing days ranged from .58 to .87. Total variance in LOS test measures attributable to inconsistencies in day-to-day test performance (Day and Subject × Day facets) ranged from 2.5% to 8.4%. The ANOVA results indicated that no significant differences were observed in the LOS test variables across the 2 testing days.

Conclusions: The 75% LOS test administered to older adult fallers on 2 consecutive days provides consistent and reliable measures of dynamic balance.

Key Words: Accidental falls; Balance; Elderly; Rehabilitation.

THE ABILITY TO CONTROL intentional movements of the center of gravity (COG) when leaning or performing weight-shifting activities is critical to the successful performance of various functional tasks associated with activities of daily living (ADLs).[1] Many older adults, however, experience difficulties and are often at increased risk for falls when performing ADLs that require dynamic postural control. Decrements in dynamic postural control have been attributed to both age and pathology-associated changes in spatial and temporal parameters associated with movements of the COG within the stability region. Compared with younger adults, healthy older adults exhibit smaller voluntary COG excursions, reach maximal lean more slowly, and exhibit less postural control once they have reached maximum lean.[2,3] Dynamic postural control is even further compromised as a result of underlying pathology and/or physical deconditioning in older adult populations.[4] The ability to quantify reliably age- and/or pathology-associated declines in dynamic postural control has implications for both the early identification of individuals at risk for falls and for the evaluation of treatment interventions.

Recent advances in computerized forceplate technology have provided researchers and clinicians a way to quantify objectively an individual's performance during various dynamic balance tasks. One dynamic balance assessment test increasingly reported in the clinical and research literature is the Limits of Stability® Test (LOS test). The LOS test provides spatial and temporal measures (eg, movement velocity, maximum excursion, directional control) of COG movements as a person volitionally leans to various positions in space. Previous investigators[5-7] have used these temporal and spatial measures from the LOS test to elucidate dynamic balance capabilities in both healthy and patient populations. Although sophisticated measures of dynamic postural control can be derived from performance on the LOS test, the clinical value of these movement-related variables depends on their reliability.[8]

The reliability of the LOS test has been studied both in young populations and in healthy older adult populations.[7,9-11] However, previous investigators,[9-11] with the exception of Clark et al,[7] based their reliability estimates on performance variables that are no longer available on current versions of the LOS test software.[a] Potential problems associated with the calculations of the original LOS test movement variables (ie, movement time, path sway, target sway, distance error) may have produced biases in previous reliability estimates of the LOS test. For example, earlier test versions required that subjects actually reach each of the 8 test targets to receive a performance score. Failure to reach the target resulted in a default score of 8 seconds in the case of the movement time variable, and subsequently an inaccurate estimate of the performance variables. The current LOS test movement variables no longer require that subjects actually reach the target, providing a more accurate assessment of dynamic postural control.[12]

Although the study by Clark[7] indicates that the LOS test, performed at either 75% or 100% of maximum limits of stability, is reliable when administered to healthy older adults on 2 separate occasions, clinical measurements and the treatment of balance-related disorders are almost exclusively performed with patient populations or with individuals at risk for falls. Differences in movement strategies between older adults with

---

From the Department of Movement Science, Gordon College, Wenham, MA (Clark); and Center for Successful Aging, California State University, Fullerton, Fullerton, CA (Rose).

Accepted in revised form May 30, 2000.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the author(s) or upon any organization with which the author(s) is/are associated.

Reprint requests to Sean Clark, PhD, Dept of Movement Science, Gordon College, 255 Grapevine Rd, Wenham, MA 01984, e-mail: *sclark@faith.gordon.edu*.

0003-9993/01/8204-5839$35.00/0

doi:10.1053/apmr.2001.21859

and without a history of falls during dynamic postural tasks may have implications for the consistency and/or stability (ie, reliability) of dynamic balance measures. Consequently, investigators must be cautious when generalizing reliability estimates of the LOS test from older adults with no prior history of falls to those who have experienced repeated falls.

The present study sought to estimate the reliability of the LOS test when administered to older adults experiencing disturbances in balance and gait. The 75% LOS test was selected because it is likely to more than adequately challenge the postural control system in a group of older adults experiencing disorders of balance and gait.

METHODS

Subjects

Fifty older adults (35 women, 15 men; age range, 62–90yr; mean age ± standard deviation, 77.5 ± 6.6yr) volunteered to participate in the present study. These subjects were a subgroup selected from a larger sample of community-dwelling older adults ($n = 75$) recruited to participate in a balance intervention program. Participants for the intervention program were solicited through newspaper advertisements and presentations to physician groups within the community. Once enrolled in the balance intervention program, individuals completed a comprehensive background and medical history questionnaire. The primary investigators reviewed the questionnaires and 50 older adults were identified as meeting the specific selection criteria for the present study. These inclusion criteria included: having had 2 or more falls within the previous year; living independently in the community (ie, noninstitutionalized setting); having no known medical diagnosis that might account for balance difficulties (eg, Parkinson's disease, stroke, multiple sclerosis); having no known cognitive impairments; not currently taking any medications known to adversely affect balance or to compensate for balance-related problems (eg, Antivert [meclizine hydrochloride], certain classes of psychotropic drugs); and normal or corrected vision (eg, glasses, contact lens). Additionally, participants had to be able to ambulate without an assistive device and to maintain an upright stance independently for a minimum of 2 minutes. Before participating in the investigation, each participant signed an informed consent document approved by the university's institutional review board.

Instrumentation

Spatial and temporal measures of dynamic postural control were obtained from each subject's LOS test performance on the PRO Balance Master® system,[a] version 6.11. The PRO Balance Master system has 4 symmetrically positioned force transducers that measure vertical pressures applied by a standing person to the support surface. These vertical pressure data were used to derive anteroposterior and mediolateral coordinates of the center of pressure, which were subsequently used to calculate the spatial and temporal characteristics of the projected COG movements. The forceplate system was also interfaced with a model 486 PC computer to acquire and store test data.

Procedures

A standard 75% LOS test was administered on the PRO Balance Master on 2 consecutive days. During each testing session, subjects were assisted onto the force platform and asked to maintain an upright stance with their arms resting by their sides and their feet in a standardized foot position as recommended by the equipment manufacturer.[12] A video screen was positioned at eye level, directly in front of the

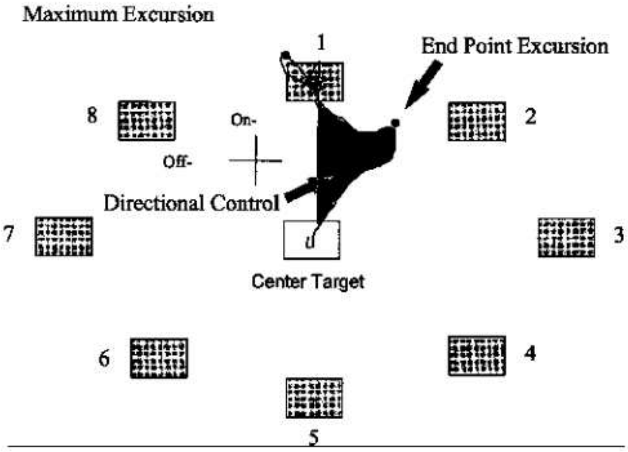

Fig 1. Target set-up and dynamic balance measures for the LOS test.

individual being tested. The video screen provided an on-screen visual display of the test set-up as well as concurrent visual biofeedback of the subject's COG position. The on-screen test set-up (fig 1) consisted of 8 visual targets (ie, small squares) displayed in a circular fashion positioned at 75% of the subject's theoretic limits of stability. The appropriate 75% LOS test target locations for each subject were derived by using the PRO Balance Master software.[12] Calculations of the limits of stability target locations were based on the subject's predicted COG height (ie, derived from standing height) and previously determined maximum COG sway angles.[12]

Before testing, subjects were informed that the on-screen COG cursor (ie, visual biofeedback) moved in response to the movements of their body COG. They were then encouraged to produce movements of the cursor by leaning the body away from the midline. This 3- to 5-minute familiarization period provided subjects with an opportunity to explore movements of the on-screen COG cursor to promote an understanding of the relationship between movements of the cursor and the actual movements of the body's COG.

Testing procedures as described in the PRO Balance Master operators manual were initiated by having subjects position the COG cursor in the center target.[12] Subjects were subsequently instructed to move the COG cursor as quickly and directly as possible in the direction of the highlighted target as soon as the start signal appeared on the screen.[12] The start signal was in the form of a small blue circle that moved from the center target to the highlighted test target. Once subjects moved the position of the COG cursor either within the test target or as close to it as possible (ie, reached maximal lean), they were instructed to hold their position as still as possible until the blue circle and start signal disappeared from the screen. The subject then repositioned the COG cursor in the center target and waited for the start signal for the next highlighted target. The standard LOS test protocol required participants to lean out toward each of the 8 targets in a sequential clockwise direction.

During the test, subjects were required to stand with their arms by their sides and to maintain their feet in the standardized foot position. A reference grid superimposed on the force plate allowed for careful monitoring of the feet during the testing procedures. If the subject lost balance while leaning (eg, took a step) or shifted foot position during testing, his/her feet were repositioned and the trial was repeated. Also, as a protective measure against potential falls, subjects wore a properly fitted safety harness during all testing procedures.

Dependent Variables

The following spatial and temporal measures of dynamic balance were derived for each of the 8 test targets comprising the LOS test: average movement velocity (MV), directional control (DC), end point excursion (EE), and maximum excursion (ME). MV, expressed in degrees per second, quantified the speed at which a subject was able to displace the COG (ie, lean) during the first sustained movement excursion toward the test target. The degree to which the COG was controlled during the first sustained movement excursion was expressed as DC. The DC value was derived from the amount of on-axis movement of the COG relative to off-axis COG movement and was expressed as a percentage of the total on-axis movement. EE and ME provided measures of the distance each participant was able to lean through his/her theoretic limits of stability. EE indicated the on-axis distance the COG was displaced from the center target during the first sustained movement excursion toward the test target. The ME data quantified the maximal distance that COG was displaced from the center target in the on-axis direction of the test target. Both EE and ME were expressed as a percentage of the test target distance (ie, 75% limits of stability). Each limits of stability variable provided specific information regarding the subject's ability to control intentional movements of his/her COG to the 8 predetermined positions in space (ie, to the 8 on-screen visual targets). See figure 1 for a graphic illustration of DC, ME, and EE.

Data Analysis

Reliability estimates across the 2 test days and 8 test targets were determined for each of the 75% LOS test movement variables. Analyses of both measurement consistency and generalizability were conducted by using a fully crossed 50 × 2 × 8 (Subjects × Day × Target) random effects repeated-measures design. The GENOVA computer program, version 2.2, was used to analyze all data.[13]

Generalizability analysis. When performing clinical assessments, a patient's test score may be viewed as a sample score from the universe (ie, infinite distribution) of possible scores under the specific measurement protocol used.[14-17] Consequently, a test score may be influenced by multiple sources of measurement error. Differences between the observed score and the expected or universe score (ie, true score) may be attributed to measurement errors associated with the specific conditions or "facets"—ie, raters, days, trials—under which the testing was performed.[14,16,18] Generalizability analysis enables the researcher to identify and estimate the relative contribution of various sources of measurement error within a single model (ie, generalizability study [G study]); and to determine the optimal measurement schedule for controlling measurement error and for increasing reliability (ie, decision study [D study]).[14-18] A more detailed review of generalizability theory is provided elsewhere.[7,14,19]

Generalizability study. The first step of the G study required the identification of each source of error or "facet" that may have contributed to the variability in the subjects' scores. Each facet was then identified as either a random or fixed measurement effect. In the present investigation, days and targets were identified as random facets. That is, these facets were identified as being a random representative sample of all possible observations of that facet.[18,19] The 2 testing days in the present measurement design were considered to be representative of a random selection of all possible test days (ie, universe of days) from which balance measures could have been obtained. Similarly, the 8 test targets were identified as a random representative sample of all possible target or spatial positions located at 75% of the subject's theoretic limits of stability.

After identification of each of the facets in the measurement design, a fully crossed 50 × 2 × 8 (Subjects × Day × Targets) random effects repeated-measures analysis of variance (ANOVA) was performed. This analysis technique provided calculated mean square values for each source of measurement variation in the design (ie, each facet and their interactions). Variance components for the object of measurement (ie, subjects), Day (D), Target (T), Subject by Day (S × D), Subject by Target (S × T), Day by Target (D × T), and the Subject by Day by Target interaction combined with the residual random error (S × D × T-E) were then estimated based on the expected mean squares and calculated mean squares for each source of variance. When negative variance components were obtained, a 0 value was substituted for the negative value and the 0 value was used for any further calculations involving these variance components.[16,18,20]

After identification of the various facets and calculations of the variance components, a G study was performed. In the G study, the relative contribution of each variance component to the total measurement error was determined.[14,16,19] These estimates of the percentage of variance attributed to subjects, D, T, S × D, S × T, D × T, and S × D × T-E indicated which measurement condition(s) were contributing to the variability in the subjects' scores.

Decision study. A D study was performed after completion of the G study. The D study enables the investigator to determine the optimal measurement design. Specifically, the D study yields generalizability coefficients (G coefficients) that reflect the reliability or generalizability of the measures for a specified measurement design.[15-17] The calculated G coefficient serves as a reliability index and can be interpreted as a reliability coefficient across the universe(s) of the various facets included in the study.[14-17] In the present investigation, G coefficients were calculated across the universe of days and targets where the Day facet was varied across the 2 days and the Target facet was generalized across the 8 targets.

Measurement consistency. Although investigators generally recognize that some variability in test scores occurs when conducting repeated evaluations, the magnitude of the observed differences in these scores should not be statistically or clinically significant.[17,21,22] Because a reliable measurement system, by definition, provides consistent test scores that are free from error across multiple evaluations, investigators interested in establishing reliability estimates must evaluate the consistency of test results obtained from repeated assessments. In the present study, measurement consistency (ie, differences in mean scores) of the 4 LOS test movement variables across the 2 days of testing and the 8 targets was assessed by performing tests of statistical significance for the calculated quasi F ratios based on the mean squares from the random effects ANOVA output.[23] To prevent the inflation of type I error, the alpha level of significance was adjusted to $p < .01$. Tukey's post hoc comparisons of means were conducted when significant differences were observed in either the Day and/or Target main effects. Post hoc comparisons were also conducted at an adjusted alpha level of $p < .01$.

Absolute reliability. To provide an indication of the absolute reliability of the measures, the standard error of the measurement (SEM) was calculated for each LOS test movement variable. Each SEM was derived as the positive square root of the absolute error variance for each of the respective LOS test movement variables.[14,18,19] The calculated SEM val-

Table 1: Mean Values for Days Collapsed Across 8 Targets

Day	MV	ME	EE	DC
1	2.11 ± 0.30	87.67 ± 5.42	69.05 ± 9.65	0.33 ± 0.15
2	2.25 ± 0.30	88.15 ± 5.68	71.25 ± 9.13	0.38 ± 0.15
SEM	0.30	5.61	7.63	0.13

NOTE. Data presented as means ± standard deviation.

ues reflect the amount of error that can be expected in the subject's performance scores.

RESULTS

Measurement Consistency

Table 1 contains mean values and standard deviations for each LOS test performance variable for the 2 days of testing collapsed across the 8 test targets. Nonsignificant F ratios for the Day main effect in each ANOVA result indicated that LOS test performance as measured by each of its 4 movement variables was consistent across the 2 days of testing. In contrast to the findings for the Day effect, variability in LOS test performance across the 8 test targets was determined to be statistically significant for all 4 LOS test movement variables examined. The ANOVA results for the Target main effect indicated significant differences in MV ($F_{7,18} = 32.98, p < .001$), ME ($F_{7,19} = 7, p < .001$), EE ($F_{7,25} = 29.52, p < .001$), and DC ($F_{7,19} = 21.26, p < .001$).

Follow-up Tukey post hoc comparisons were conducted independently for each LOS test movement variable to identify which target differences contributed to the significant Target main effect. Post hoc analysis for MV indicated that the COG excursions toward the forward and rear targets (targets 1 and 5, respectively) were significantly slower than the COG excursions toward all other targets. Results from post hoc comparisons for EE indicated that initial COG excursions within the 75% theoretic limits of stability were also significantly smaller for targets 1 and 5 when compared with both the lateral targets (targets 3, 7) and the forward diagonal targets (targets 2, 8). Additionally, EE values for the rear diagonal targets (targets 4, 6) were significantly smaller than values for the right forward target (target 2). Post hoc analyses further revealed that ME values were significantly larger for target 2 than for all other test targets, except the right lateral target (target 3). Also, ME values for target 3 were significantly larger than ME values for the rear target (target 5). Finally, post hoc comparisons for DC indicated that COG movement control when leaning toward targets 4 and 6 was poorer (ie, significantly larger) than that observed for all other targets, except for target 5. DC values for target 5 were also significantly larger than the DC values for targets 3 and 7 and target 2.

Table 3: Variance Components and Percentage of Variation for ME

Source of Variation	Variance Component	% of Variation
Subject (S)	186.48	37.13
Day (D)	.69	.14
Target (T)	23.41	4.66
X × D	11.49	2.29
X × T	79.08	15.75
D × T	.63	.13
S × D × T-E	200.46	39.91
Total	502.24	100.00

G Study

G study results, including the estimated variance components and the percentages of variation for each facet, are presented in tables 2–5. As indicated in these tables, the total variation in LOS test performance attributed to the Day facet was less than 1% for each of the LOS test movement variables examined. These findings indicate that the contributions of the Day variance to the total measurement error for each LOS test variable were negligible. Moreover, a summation of the Day facet with both the S × D and D × T interactions yielded percentage variance values that ranged from only 2.55% to 8.39% across the 4 LOS test variables. Collectively, the G study findings indicate that the total variance in LOS test performance associated with administering the 75% LOS test on 2 separate days was minimal (<9%).

In comparison to variance estimates for the Day facet, variability in LOS test performance attributed to differences across the 8 test targets accounted for a larger proportion of the total measurement error in each LOS test movement variable (tables 2–5). Approximately 5% (ME) to 14% (DC) of the total variation in the LOS test measures was attributable to the Target facet. Additionally, the S × T interaction yielded estimated variance values that ranged from approximately 8% (MV) to 16% (ME). The larger variance estimates associated with the S × T interaction indicated that subjects varied in their abilities to control COG movements to the different test targets.

The largest proportion of measurement variability in each of the LOS test movement variables was attributed to the residual error variance (S × D × T-E). The S × D × T-E interaction contributed between 39.91% (ME) and 53.26% (MV) to the total variation in the dependent variables examined (see tables 2–5). These results indicated that a large percentage of the variability in the LOS test was associated with (1) the highest order interaction term (ie, S × D × T), (2) sources of mea-

Table 2: Variance Components and Percentage of Variation for MV

Source of Variation	Variance Component	% of Variation
Subject (S)	.299	26.32
Day (D)	.009	.79
Target (T)	.085	7.48
S × D	.050	4.40
S × T	.088	7.75
D × T	.000	.00
S × D × T-E	.605	53.26
Total	1.136	100.00

Table 4: Variance Components and Percentage of Variation for EE

Source of Variation	Variance Component	% of Variation
Subject (S)	183.47	24.84
Day (D)	1.53	.21
Target (T)	84.05	11.38
S × D	28.23	3.82
S × T	84.29	11.41
D × T	0.00	.00
S × D × T-E	357.14	48.35
Total	738.71	100.00

Table 5: Variance Components and Percentage of Variation for DC

Source of Variation	Variance Component	% of Variation
Subject (S)	.018	12.59
Day (D)	.000	.00
Target (T)	.020	13.99
S × D	.012	8.39
S × T	.021	14.69
D × T	.000	.00
S × D × T-E	.072	50.35
Total	.143	100.00

surement error or facets not identified in the present investigation, or (3) random measurement error.

D Study

D Study results for each of the 4 LOS test movement variables are presented in table 6. As indicated in this table, a single administration of the 75% LOS test (ie, 8 targets) yielded estimated G coefficients ranging from .44 (DC) to .80 (ME), whereas, G coefficients derived for the present measurement protocol (8 targets, 2 test days) ranged from .58 (DC) to .87 (ME). The calculated G coefficient for DC indicated a moderate reliability estimate[24]; whereas the G coefficients for the MV, ME, and EE measures yielded high reliability estimates when generalized across the complete LOS test and 2 testing days.[24]

Standard Error of Measurement

Calculations of the SEM values for each of the 4 LOS test movement variables were based on the estimated variance components derived from the G study implementing the full measurement protocol (8 test targets, 2 days of testing). The calculated mean score for each of the 2 test days and the respective SEM values for each movement variable are presented in table 1. Comparison of the SEM values with the calculated mean scores for test days indicated that the SEM values were relatively small for each of the reported LOS test movement variables.

DISCUSSION

The present investigation was prompted by the need to establish reliability estimates of the 75% LOS test when conducted with independent community-dwelling older adults with a history of falls. Although the reliability of this test has previously been established when conducted with healthy community-dwelling older adults,[6] no attempt has been made to determine the reliability of the 75% LOS test when conducted with older adults who experience disorders of balance and gait. Results of the present analyses indicate that the spatial and temporal measures of COG movement for the LOS test conducted at 75% of the subject's theoretic limits of stability provide consistent and reliable measures of dynamic balance when performed by independent community-dwelling older adults with a history of falls. The reported G coefficients for the 4 LOS test movement variables when generalized across 2 days of testing and 8 limits of stability targets ranged from moderate to high. Additionally, results of the ANOVA indicated that the measures of dynamic balance derived from the LOS test were consistent across the 2 test days.

Caution is often advised when interpreting or generalizing reliability estimates because issues may exist regarding both the size and homogeneity of the subject sample.[25] The use of small samples or samples that fail to represent the population adequately may be a concern in generalizability analyses because they may include potential inaccuracies or instabilities in the variance estimates.[26] Although formulas for sample size estimates are not readily found in the generalizability theory literature, previous investigators[25,27,28] have reported that samples of 30 to 50 participants are appropriate when using intraclass correlation analyses to establish test reliability. Given that generalizability theory is an extension of the intraclass reliability model, the inclusion of 50 subjects in the present study is consistent with both suggested intraclass correlation sample size estimates and sample sizes previously reported in G studies.[7,19]

Generalizability Analysis

Similar to the work of Clark et al,[7] reliability in the present investigation was estimated by using generalizability analysis. Unlike reliability estimates from classical test theory, generalizability analysis provides researchers and clinicians with estimates of both the magnitude and the relative contribution of identified sources of measurement error.[14,15,18] This information helps investigators determine a measurement protocol that provides optimal, adequate, and/or cost-effective reliability estimates.[15]

Day facet. Generalizability analysis in the present investigation provided estimates of the total variance in LOS test movement scores attributable to differences or inconsistencies in day-to-day test performance. Estimates of the Day variance are valuable for researchers and clinicians because variation in day-to-day performance contributes to measurement error and consequently may have negative implications on the reliability of measures. Variance estimates derived for the Day facet in the present investigation indicated, however, that when the 75% LOS test is administered to older adult fallers, little variation is evident in performance scores across days. Our findings ranged from 2.5% to 8.4%. Clark[7] reported similar findings. The investigators reported that the Day facet was a relatively small source of measurement error (2%–12%) when administering the 75% LOS test to a sample of healthy community-dwelling older adults. Additionally, findings from both investigations indicate that the LOS test movement variables are reliable across repeated evaluations. The implications for practitioners are that, though variation in scores during repeated evaluations of the 75% LOS test is expected, the extent of differences in movement variables is statistically and presumably clinically nonsignificant.

Target and Subject by Target facets. The variance estimates attributed to the subjects by targets interaction indicated that subjects differed in their LOS test performance scores as a function of the 8 LOS test targets. These differences or inconsistencies in the subjects' performance may be attributed to the inability of some subjects to move the COG to various positions in space located at 75% of their theoretical limits of stability. Age-related declines in the voluntary excursions of the COG to various regions within the limits of stability have been previously identified.[2,3] Consequently, the 8 target positions of the 75% LOS test derived from the subject's theoretic

Table 6: Coefficients for Days and 8 Targets

Day	MV	ME	EE	DC
1	.69	.80	.69	.44
2	.80	.87	.80	.58

maximum stability limits may have exceeded the actual limits of stability of some older adult subjects.

Variability in the LOS test measures associated with the Target facet and the Subject by Target interaction may also be attributed to differences in the selection of postural strategies for producing displacements of the COG. Although subjects in the present study were encouraged to produce movements of the COG cursor by leaning or rotating about the ankle joints (ie, use an ankle strategy), some subjects may have explored the effectiveness of different postural strategies for producing displacements of the COG cursor. For example, a subject may have adopted an ankle strategy to produce COG movements to the mediolateral targets, but may have selected a hip strategy for COG excursions to the anteroposterior targets. Several possible explanations to account for the exploration of postural strategies during the LOS test could be forwarded, including: adopting a biomechanically "safer" strategy (ie, hip strategy) for situations of perceived instability or fear of falling; compensating for self-perceived cognitive and physical demands associated with implementing only an ankle strategy; and/or limitations in movement strategies because of undiagnosed pathologic conditions.

Unexplained Variance

In the present investigation, the unexplained variance component (S × D × T-E) accounted for the largest percentage of variability in each of the LOS test movement variables. A portion of this measurement variability may be attributable to random measurement error. Possible sources of random measurement error in the present investigation include inherent electrical noise in the PRO Balance Master system, disturbances in the testing environment, subject's motivation level, and misinterpretations of the COG visual biofeedback.

Variability attributed to the S × D × T-E interaction may also be attributable to sources of measurement error not identified in the present measurement design. That is, the present design only calculated variance estimates for the object of measurement (ie, subjects), the Day facet, the Target facet, and the Subject, Day, and Target interaction effects. Additional sources of measurement error in the present investigation may have included the subject's age and biomechanic factors (eg, muscular strength, joint range of motion), which can limit postural movements.

Absolute Reliability

Results from G study provide practitioners and researchers with information regarding the relative variance contributions attributed to each of the various sources of measurement error included in the design. Although this information is valuable, especially when optimizing a measurement protocol, absolute differences in the measures are unknown. For the practitioner, the SEM or absolute differences in measures is an important and practical component of measurement reliability.[17] That is, the practitioner is often concerned with how closely the obtained score on a test reflects the true score for that test. The SEM value provides the expected range about the observed score in which the true score lies. The calculated SEM values reported in the present and previous investigations were relatively small compared with the mean scores. Thus, when administering the 75% LOS test to older adult fallers and nonfallers, the practitioner can expect the true score to lie within a limited range of the observed score. Additionally, the small SEM values may also be beneficial when evaluating the effectiveness of a balance intervention program. Specifically, a criterion for an effective program is that postintervention assessment scores do not overlap (±SEM) with scores obtained during the preintervention evaluations.

Clinical Implications

The present investigation provides clinicians with estimates of the relative contribution of several error sources associated with LOS test performance. A clinician's knowledge of these relative variance contributions affords the opportunity to modify a measurement protocol to minimize measurement error and obtain acceptable levels of reliability when administering the 75% LOS test to independent community-dwelling older adults with a previous history of falls. For example, a clinician can conclude from the present findings that the residual error variance is a significant source of measurement error. It is possible, therefore, to reduce this residual error variance by standardizing both testing instructions and procedures and by providing sufficient practice time for patients to understand the relationship between their movements and the movement of the COG cursor. Also, by recognizing the relatively low variance estimates associated with the Day facet, a clinician may determine that 3 days of testing is not more cost effective than 2 days, according to generalizability estimates.

CONCLUSIONS

The 75% LOS test administered to older adult fallers on 2 consecutive days is a reliable test of dynamic balance. The G coefficients for the MV, ME, and EE measures indicated high reliability estimates when generalizing across the 2 days of testing. Performance scores for the test's 4 LOS test movement variables were consistent across the 2 test days. A minimum of 2 testing days (or 2 administrations of the test on the same day) is recommended to obtain reliable and consistent measures of dynamic balance when administering the 75% LOS test to independent community-dwelling older adults with a history of falls.

References

1. Topp R, Mikesky A, Thompson K. Determinants of four functional tasks among older adults: an exploratory regression analysis. J Orthop Sport Phys Ther 1998;27:144-53.
2. Schieppati M, Hugon M, Grasso M, Nardone A, Galante M. The limits of equilibrium in young and elderly normal subjects and in parkinsonians. Electroencephalogr Clin Neurophysiol 1994;93:286-98.
3. Blaszczyk JW, Lowe DL, Hanseb PD. Ranges of postural stability and their changes in the elderly. Gait Posture 1994;2:11-7.
4. Rose DJ, Dickin C. Postural instability and age-related changes in dynamic limits of stability. J Sport Exerc Psychol 1999;21:S93.
5. Hamman RG, Mekjavic I, Longridge NS. Training effects during repeated therapy sessions of balance training using visual feedback. Arch Phys Med Rehabil 1992;73:738-44.
6. Rose DJ, Clark S. Can the control of bodily orientation be significantly improved in a group of older adults with a history of falling? J Am Geriatr Soc 2000;48:275-82.
7. Clark S, Rose DJ, Fujimoto K. Generalizability of the limits of stability test in the evaluation of dynamic balance among older adults. Arch Phys Med Rehabil 1997;78:1078-84.
8. Shephard NT, Telian SA, Smith-Wheelock M, Raj A. Vestibular and balance rehabilitation therapy. Ann Otol Rhinol Laryngol 1993;102:198-205.
9. Brouwer B, Culham EG, Liston RA, Grant T. Normal variability of postural measures: implications for the reliability of relative balance performance outcomes. Scand J Rehabil Med 1998;30:131-7.
10. Henderson NE, Overby AS, Panzer VP. Internal consistency and stability of balance measures among different age groups [abstract]. In: Proceedings of the 12th International Congress of the World Federation for Physical Therapy; 1995 June 25-30; Washington (DC). Alexandria (VA): American Physical Therapy Assoc; 1995. p 751.

11. Hageman PA, Leibowitz M, Blanke D. Age and gender effects on postural control measures. Arch Phys Med Rehabil 1995;76:961-5.
12. NeuroCom International Inc. Balance Masters operators manual. Clackamas (OR): NeuroCom International; 1993.
13. Crick JE, Brennan RL. GENOVA: a general purpose analysis of variance system [computer program]. Dorchester (MA): Univ Massachusetts at Boston, Computer Facilities; 1984.
14. Morrow JR Jr. Generalizability theory. In: Safrit MJ, Wood TM, editors. Measurement concepts in physical education and exercise science. Champaign (IL): Human Kinetics; 1989. p 73-96.
15. Shavelson RJ, Webb NM, Rowley GL. Generalizability theory. Am Psychol 1989;44:922-32.
16. Shavelson RJ, Webb NM. Generalizability theory: a primer. Newbury Park (CA): Sage; 1991.
17. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Norwalk (CT): Appleton & Lange; 1993.
18. Brennan RL. Elements of generalizability theory. Iowa City (IA): American College Testing Program; 1983.
19. Roebroeck ME, Hariaar J, Lankhorst GJ. The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. Phys Ther 1993;73:386-401.
20. Cardinet J, Tourneur Y, Allel L. Extension of generalizability theory and its applications in educational measurement. J Educ Meas 1981;13:183-204.
21. Mitchell SK. Interobserver agreement, reliability and generalizability of data collected in observational studies. Psychol Bull 1979;86:376-90.
22. Thomas JR, Nelson JK. Introduction to research in health, physical education, recreation and dance. Champaign (IL): Human Kinetics; 1985. p 253-68.
23. Myers JL, Wells AD. Research design and statistical analysis. New York: Harper-Collins; 1991.
24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:240-8.
25. Morrow JR Jr, Jackson AW. How significant is your reliability? Res Q Exerc Sport 1993;64:352-5.
26. Smith P. A confidence interval approach for variance component estimates in the context of generalizability theory. Educ Psychol Meas 1982;42:459-66.
27. Donner A, Eliasziw M. Sample size requirements for reliability studies. Stat Med 1987;6:441-8.
28. Baumgartner TA. Norm-referenced measurement: reliability. In: Safrit MJ, Wood TM, editors. Measurement concepts in physical education and exercise science. Champaign (IL): Human Kinetics; 1989. p 45-72.

Supplier

a. NeuroCom International, Inc, 9570 SE Lawnfield Rd, Clackamas, OR 97015-9611.