The Spring Scale Test: A Reliable and Valid Tool for Explaining Fall History

Louis DePasquale, PT, MA;[1] Lisa Toscano, BS, MS, EdD, ATC[2]

[1] Visiting Nurse Service of New York
[2] Physical Education and Human Performance, Manhattan College, Riverdale, NY

ABSTRACT

Background & Purpose: Although perturbations can result in falls, objective practical tests of balance that incorporate perturbations have not received as much attention as other balance measures. The purposes of this study were to: (1) describe a new clinical waist pull perturbation balance test protocol, the Spring Scale Test (SST), (2) determine the test-retest reliability of the SST, and (3) examine the convergent and known groups validity of the SST relative to 4 other balance related measures. **Methods:** Fifty-eight community-dwelling older adults underwent same-day test-retest of the SST, with predictable, one-pound incremental, horizontal sagittal plane manual waist pulls**.** Twenty-nine subjects (mean age 83.5) had at least one fall in the prior 2 years. Twenty-nine subjects (mean age 78.0) were nonfallers. **Results:** The mean (SD) SST performance of 7.5% (1.4) of Total Body Weight (TBW) for fallers and 12.3% (1.7) TBW for nonfallers differed significantly ($p < .001$). High test-retest reliability was supported by ICC of 0.94 [95% CI = 0.89 - 0.96], a method error (ME) of 0.74 and coefficient of variation of the ME of 7.25%. There was moderate to good convergent validity of the SST with all 4 relevant measures of gait and balance: gait speed ($r = 0.53$, $p = .001$), single limb stance ($r = - 0.54$, $p = .001$), tandem stance ($r = 0.55$, $p = .001$), Timed Up and Go ($r = - 0.67$, $p=.001$). The SST score at 10% TBW had high known groups validity with the area under the receiver operating characteristic curve of 0.99, sensitivity of .93 and specificity of .97. Logistic regression revealed that no other variable added to the explanation of fall status once the variance attributable to the SST 10% TBW (91.1%) was included in the predictive equation. **Conclusions:** The SST is a quantitative, reliable, and valid clinical measure of standing balance. The SST 10% TBW cutoff score was the most discriminating indicator of fall history.

Key Words: balance, falls, measurement

INTRODUCTION

Falls are the leading cause of nonfatal injuries in the United States, accounting for approximately 48% to 75% of all unintentional injuries reported for adults 65 years and older.[1] Falls are also the most common cause of death by injury for Americans 65 years and older.[1] Each year 35% to 40 % of adults 65 years and older fall at least once. [1] The total number of hip fractures is predicted to increase because of growth in the older population.[2]

Most falls occur in or just outside the home.[3] Falls in the aging population result in persistent strength and mobility deficits contributing to declines in balance, limiting capacity for independent function, and further increasing risk of recurrent injury.[4]

Fall risk depends on an interaction of many factors, and falls are particularly difficult to explain in older adults who live active, independent lives.[5] Current tests are less successful in explaining falls in active aging adults than in those who are described as frail.[5] Timed mobility performance measures may not adequately challenge systems to detect problems essential in fall risk situations in individuals not exhibiting existing balance deficits or outwardly displaying observable limitations.[6,7] The inability of certain balance and mobility performance measures to explain falls, as well as their lack of responsiveness to falls in active community-dwelling aging adults, has been reported.[5,8] New assessment tools are needed for an active aging population. Such tools should contain more challenging performance-based measures[5] including situations requiring reactive balance, such as recovering from a push to the body or a tripping paradigm.[9] As most balance measures assess static and anticipatory rather than reactionary postural control, researchers need to test individuals as they react to externally imposed perturbations.[9] Following a perturbation that might cause a fall, one must recover balance using a feet-in-place postural strategy response, or a protective compensatory step.[10]

Studies report perturbation forces as percent of total body weight.[10-15] Researchers have examined alternative dominating roles of proactive feedforward, and reactive feedback postural responses in maintaining dynamic standing equilibrium through effective stepping strategies in the aging population.[15-18] The use of a predictable perturbation assessment paradigm is supported by studies concluding that predictability and prior knowledge of magnitude and direction of perturbation forces do not alter EMG latency of lower extremity motor responses and have no effect on automatic postural responses.[19,20] Nevertheless, existing perturbation studies do not provide objective or quantitative stepping performance measures or procedures of safe or feasible practical use in multiple clinical settings.

A new perturbation-based, clinical fall assessment tool has been developed, the Spring Scale Test (SST). The SST consists of repeated incremental rounds of predictable loading and unloading waist pull perturbations. In conjunction with specific performance criteria, the SST protocol attempts to assess and quantify effective limits of anterior-posterior stepping for the purposes of fall risk assessment in the active healthy community dwelling older adult.

The purposes of this study were to: (1) describe the SST and results obtained with it, (2) quantify test-retest reliability of

Address correspondence to: Louis DePasquale, 67 Westchester View Lane, White Plains, NY 10607 (depasqualept@aol.com).

the SST, and (3) examine the convergent construct and known groups validity of the SST% as a measure of fall status.

METHODS

This descriptive study was reviewed and approved by the Visiting Nurse Service of New York Center for Home Care Policy and Research Institutional Review Board (IRB). It involved: (1) screening subjects for eligibility; (2) determining fall history; (3) obtaining information on health, demographics (age, gender, height, and weight), and physical functioning; and 4) administering the SST and 4 other measures potentially explanatory of falls. Test-retest-reliability of the SST was examined by same-day repeated testing. Convergent construct validity of the SST was explored by comparing SST measurements to the other potentially explanatory variables. Known groups validity was assessed by comparing the balance measures of fallers and nonfallers.

Participants

Inclusion criteria required that participants were: (1) willing to provide informed written consent, (2) living independently in the community without assistance, (3) aged 65 years and older, (4) capable of independent unrestricted community ambulation with or without a cane one or more blocks, (5) able to complete the Timed Get Up and Go Test (TUG) in less than 14 seconds, (6) medically stable, not requiring frequent medication adjustments or medical intervention, (7) without hospitalization, spinal or lower extremity fracture within 3 months of participation or not experiencing lower extremity pain, (8) able to stand unsupported without any device, 9() weighing 200 pounds or less, and (10) able to speak and understand English and follow simple instructions. Participants were excluded if they had less than 10° active ankle dorsiflexion/plantarflexion range of motion or less than fair plus (3+) ankle dorsiflexion/plantar flexion strength.

A convenience sample of 61 participants were recruited by the primary investigator, 56 via flyers, formal presentations at local senior centers, one on one presentations, and through word of mouth from the local community; 5 were recruited from home care patients treated by the primary investigator. Following screening by the primary investigator, 3 recruits were excluded due to acute medical or personal time constraint reasons. The remaining 58 participants (39 females, 19 males) provided written informed consent. An a priori power analysis was performed based on pilot data and the assumption that a 2% point difference in SST pull force at failure would be meaningful and likely between fallers and nonfallers. With an alpha of 0.05, power would exceed 80% with 9 in each group. Over sampling of both groups (n=29) was employed to adequately address the convergent and known groups validity aspects of this study and in case assumptions were too optimistic since pilot data provided information on SST performance, not between group differences.

Fall history was obtained from all prospective participants by the primary investigator during recruitment, 3 to 4 months prior to testing. Fall history information was noted solely to tabulate some semblance of equal numbers in both groups. This initial inquiry about fall status was not documented on the participant data record. Formal judgment about fall history was ascertained at the conclusion of the formal testing session. Although not technically blinded to fall history, several months transpired between recruitment and data collection, with many intervening patients and activities. This minimized the primary investigator's recall of fall history of individual subjects. All novice examiners were blinded to fall history.

Subjects were categorized as fallers if they reported at least one fall in the past 2 years. Nonfallers reported no falls in the past 2 years.[21] A fall was defined as "any disturbance of balance during routine activities that resulted in a person's trunk, knee, or hand unintentionally coming to rest on the ground, wall, table, chair, or other surface,"[5] which otherwise, in the estimation of the participant, would have resulted in contact with the ground." For this study a 2 year fall incident recall period was chosen in an attempt to include individuals beyond acute and subacute stages of fall recovery. Fall exclusions included overwhelming environmental hazards (eg, violence, slipping on ice) or acute medical conditions (eg, syncopal collapse, stroke).[22,23] Participants provided updated, self- reported fall history at the conclusion of same-day, in-home SST testing.

Table 1 displays demographic data including age, gender, height and weight. Table 2 provides a comparison of fallers and nonfallers by age decade.

Table 1. Descriptive Characteristics of Participants

FALLERS (n=29)	Minimum	Maximum	Mean	Std. Dev.
Age	75.0	94.1	85.6	5.55
Weight (lbs)	106	200	145.8	26.4
Weight (kg)	48.1	90.7	66.22	11.97
Height (meters)	1.35	1.85	1.64	0.11
BMI*	17.8	34.4	24.5	3.27
SF-36** score	25.0	95.0	63.3	18.58
NONFALLERS (n=29)	**Minimum**	**Maximum**	**Mean**	**Std. Dev.**
Age	65	90	78.0	7.75
Weight (lbs)	103	200	153.0	28.18
Weight (kg)	46.7	90.7	69.4	12.78
Height (meters)	1.48	1.83	1.67	0.089
BMI	19.4	32.5	24.9	3.12
SF-36 score	55.0	100.0	86.2	12.15

* BMI = Body mass index
**SF-36 = Health status survey- Physical functioning subscale

Table 2. Frequency Distribution of Fallers and Nonfaller by Decade

Age groups	Fallers N (%)		Nonfallers N (%)	
60 - 69	0	(0)	6	(100)
70 - 70	7	(44)	9	(66)
80 - 89	18	(60)	12	(40)
90 +	4	(67)	2	(33

The Spring Scale Test

The spring scale test (SST) is a standing test of reactive and proactive balance that assesses the maximum waist-pull perturbation force against which a person can maintain postural

stability against both anterior and posterior direction perturbations. The test is completed in 2 steps. The first step consists of rounds of anterior direction waist-pull perturbation loading and unloading force delivered until the anterior direction maximum performance criteria is achieved. This is followed by rounds of posterior direction waist-pull perturbation loading and unloading delivered until the maximum posterior direction performance criteria is achieved. The waist-pull perturbation, delivered with a spring scale device attached to a belt around the subject's waist (Figure 1), progressively increases waist-pull force in one pound increments for each round of testing. Each round of testing assesses balance reactions in response to (1) a spring scale loading force and (2) an unloading of the spring scale loading force.

Figure 1. SST Anterior directional waist pull testing-(rear stepping): set up with spring scale and subject's feet in clear view.

The SST performance criteria for the maximum spring scale loading force is the greatest amount of waist-pull force the person can accommodate by maintaining both feet flat (heel-sole) on the ground. The SST balance performance criteria for the maximum spring scale unloading force is the greatest amount of waist pull force against which, when unloaded, the person can regain balance in less than 4 steps without any contact with an outside support. Unloading of anterior direction loading force seeks to elicit rear stepping responses; unloading of posterior direction loading forces seeks to elicit forward stepping responses.

The maximum waist-pull force against which the participant can maintain postural stability under all 4 conditions (anterior direction loading, unloading of anterior direction loading, posterior direction loading, and unloading of posterior direction loading), reported as a percentage of the person's total body weight (TBW %), defines the SST directional limit score. The SST TBW % performance measure is the lower of two directional limit scores.

Instrumentation

The spring scale has a 26-pound (12-kilogram) capacity. The pocket sized 8 inch linear spring scale is capable of quantifying manual waist-pull forces in one pound increments with a 0 set point turn dial calibration capacity (Pelouze/Pelstar LLC, Product of Pelstar; Bridgeview, IL). Calibration accuracy of the linear spring scale-measuring instrument is achieved through suspension of a 5-pound weight prior to and at mid-point of each test day with a range of 4 to 6 participants tested per test day. At no time was re-calibration of the linear spring scale required during testing. Spring scale replacement was required on 4 occasions over the course of 6 months of testing due to fatigue of the metal spring scale-sliding indicator. Replacement linear spring scale instruments were identical to the original and calibrated in the same way prior to use.

Procedures for the SST

All testing and data collection were performed in the home of each participant during one same-day visit. Volunteers were screened for inclusion/exclusion criteria and provided informed written consent. Participants then underwent 2 complete SST trials 15 minutes apart. Between the SST trials, gait speed, timed get up and go test (TUG), single limb stance (SLS), tandem stance (TS), and completion of the SF-36 physical functioning subscale questionnaire were assessed. These measures are described later in this section.

The SST protocol employs a padded 5 inch wide belt secured around the subject's waist. The spring scale is attached to the belt (Figure 1) on one end and held by the examiner at the other end. For safety purposes and IRB requirements, a 4 foot tether strap was secured via belt hardware at waist level to both the examiner and participant enabling unrestricted responses. The direction of the waist-pull is in the sagittal plane with the scale held parallel to the floor. There is only one round of testing at each spring-scale weight. Participants are tested wearing normal footwear and standing on a firm unpadded surface in a comfortable normal foot- flat position (heel-sole floor contact) in an area that allows unrestricted stepping responses without any support surfaces to grab hold of. Grasping the available pull strap, the examiner assumes a stable stance in close proximity (2 feet) of the participant, ensuring clear sight of the participant's feet and linear spring scale measurement instrument while maintaining anterior-posterior orientation line of pull (Figure 1). During unloading the examiner steps toward the participant to avoid tether strap tightening, which could interfere with the participant's movements.

Prior to each SST directional limit testing, participants are provided a practice trial of continuous spring scale loading waist-pull forces to the maximum limits of foot-flat accommodation, the postural response SST performance criteria for the loading response. The purpose of a pretesting practice trial is to familiarize the participant with the spring scale device and the spring scale loading forces delivered. Subjects do not have a practice trial of sudden 'unloading.' Following the practice trial, the test begins with rounds of anterior direction loading and unloading of spring scale waist-pull forces, beginning at one-pound of force. Each successive round increases the loading force by one-pound. During each round, the loading force is gradually increased to allow the participant time to accommodate. The loading phase is followed by the unloading phase, administered in a sudden, semi-random manner at the discretion of the examiner within a subjective 5 second window from the point of successful foot flat accommodation. Each successive round of increasing loading and unloading force continues until the

subject fails either the loading or unloading SST performance criteria. Once the anterior direction waist-pull spring scale force limit is established, the test is repeated in the posterior direction.

Prior to and during SST testing, examiners instruct participants to maximally resist waist pull loading forces while maintaining foot-flat (heel-sole) floor contact and to use the fewest steps possible, but no more than 3 steps, in response to the sudden unloading/release of the loading waist pull forces. Participants are informed that waist pull loading forces will suddenly be unloaded without warning at each additional one pound incremental force at the discretion of the examiner within a subjective 5-second window time frame from the point of successful foot-flat loading accommodation. No other specific postural strategy cues are provided.

Reliability Testing of SST

Repeated measures SST protocol trials were administered to each subject 15 minutes apart to examine test-retest reliability and to eliminate the effect of time in an attempt to consider only the effects of using multiple measures.[24] Precedent exists for same day reliability testing of physical performance.[25-28] Trial 1 SST testing was performed by the primary investigator who has more than 10 years experience using the SST.SST trial 2 testing was performed by 1 of 3 participating novice examiners (0 to 3 months SST use). The second author performed 47 of the retests. Novice retest examiners observed trial 1, were blinded to fall status and did not have access to trial 1 numeric data. The primary investigator recorded data for both SST trials. Developed as a first attempt, single-trial, "no practice" test, practice is intrinsic due to the SST procedure of repeated incremental rounds of loading/unloading waist-pull forces. For purposes of this study, in an attempt to minimize examiner bias or variations in technique, a second chance attempt (1 per direction) was provided only when incomplete foot contact with the floor was observed during the loading phase of directional limit testing for only loading waist-pull forces up to and including 10% TBW. Spring Scale Test directional limit testing proceeded until a maximal direction limit SST score was reached.

Additional Measures

Prior to SST testing, demographic data including age, sex, height and body weight (using a calibrated digital floor scale), were obtained (Table 1). Vital signs included blood pressure, resting heart rate and respirations, which were recorded prior to, between, and at the conclusion of all SST trials.

Between the 2 SST trials, the first author, who has more than 25 years clinical experience, administered 4 balance and mobility performance measures in the following order: (1) gait speed,[29] (2) Timed Up and Go (TUG),[30] (3) single limb stance (SLS) with eyes open,[31] and (4) tandem stance (TS).[32,33] Time was measured with a digital stop watch. Novice examiners recorded the balance and mobility performance scores and administered the SF-36 physical functioning subscale.[34] This subscale of the SF-36 assesses a person's self-perception of health related limits in physical functioning.

Gait speed is a good predictor of incident disability.[32] Participants completed one gait speed trial. Participants were instructed to walk at their comfortable "normal" speed, along a 3 meter (10 feet) measured length walkway with 5 foot acceleration/deceleration zones at each end. Gait speed data was expressed as meters/sec (m/sec.). One trial gait speed testing has been addressed[29] as well as the legitimacy of obtaining gait speed measurements for distances less than 20 feet.[35] Gait speed has been shown to possess high multiple rater[36] and test-retest reliability.[37]

The TUG is a measure of balance and anticipatory postural control. The TUG measures the time required for an individual to stand from an armchair, walk a distance of 3 meters, turn and walk back to the chair, and sit down. For this study, timing began upon the verbal command "go." Participants stood up from the chair (using their upper limbs as needed), walked a distance of 10 feet (3 meters) to a mark on the floor,[30] turned and walked back to the chair, turned again and sat down. The stop watch timing ended as soon as the person sat down. Participants were instructed to walk in a quick but safe speed.[30] The best time of 2 trials was recorded.[3] Reliability is supported by an intraclass correlation coefficient (ICC) of 0.98.[30] Using a 13.5 second cutoff value, the TUG has a sensitivity of 87% and a specificity of 87% for identifying older adults with a fall history.[30] One TUG trial may adequately represent performance.[29]

The SLS test is a timed test of static balance that requires the participant to stand without support on one lower limb.[8,31] For this study, participants were tested wearing shoes, eyes open, and arms at their side. Subjects were supported by the tester while assuming a comfortable position standing on preferred lower limb. Once stable, support was withdrawn. The tester timed performance until the unsupported foot touched the floor[8] or until 30 seconds had elapsed.[31] Participants were not allowed to hop on, or change foot position of the weight bearing extremity. Trunk inclination was not accounted for. The best time of 2 trials was recorded. Excellent inter-rater reliability for the SLS test has been described (ICC= 0.99).[38] Sensitivity of 95% and specificity of 58% at a 30 second cutoff value has been reported[39] in correctly classifying community-living fallers and nonfallers.

Similar to the SLS test, the TS test is an assessment of static balance included in the Established Populations for Epidemiologic Research in the Elderly (EPESE), Short Performance Battery.[32,33] It measures the time, in seconds, that an individual can maintain balance with a narrow (heel-toe) base of support, without assistance or taking a step. For this study, manual support was provided until the subject assumed a comfortable position (dominant foot behind). Timing began upon release of manual support until a step occurred or 30 seconds had elapsed. The best time of 2 trials was recorded. Test-retest reliability of ICC (0.70) for TS with eyes open has been reported[40] in adults with neurological deficits. Sensitivity of 55% and specificity of 94 % at a 10 second cutoff value[41] differentiates fall risk in community dwelling elderly adults.

Data Analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS, version 11.0 SPSS Inc, Chicago, Il) and Med Calc (Mariakerke, Belgium). Conventional descriptive statistics (means and standard deviations) were calculated separately for fallers and nonfallers for all variables.

The single measure, 'two-way mixed model' ICC was used to evaluate test-retest reliability of the SST. Response stability

across repeated trials was further described using: (1) method error (ME) to determine absolute error (discrepancy between two sets of repeated scores) and (2) coefficient of variation of ME_{cv} to determine the percentage error between the same repeated scores.

Two aspects of validity were examined, convergent validity and known groups validity. Measures that conceptually converge should be strongly correlated whereas those measures with less in common should display weaker correlations. Pearson product correlation coefficients were calculated between all balance-related measures to establish convergent validity. The ability to distinguish between subjects with and without a history of falling (known groups) first involved t-test analysis of data for the SST, TUG, gait speed, SLS test, and the TS test. Thereafter, receiver operating characteristic (ROC) curve analysis was performed including development of criterion-referenced cut scores for all balance and fall related measures examined in this study. Of particular interest of this analysis was the area under the curve (AUC), which allows for the determination of an optimal cutoff value that maximizes sensitivity and specificity and provides positive predictive value (PV+) and negative predictive value (PV-). Each AUC was compared statistically to every other AUC using the MedCalc program. Finally, logistic linear regression, forward stepwise (Wald) method, was used to assess the ability of the SST 10% TBW performance measure, relative to other balance related measures, to explain group (faller versus nonfaller) membership.

RESULTS

The SST was administered to 58 active independent and community-living older adults who either had or had not fallen in the previous 2 years. Expressed as a percent of total body weight, the mean SST was 7.5% (1.4) for fallers, 12.3% (1.7) for nonfallers, which represented a significant difference ($p < .001$) between groups. All participants completed the SST protocol without any falls, uncontrolled responses, reported or observed adverse outcomes, injury or discomforts. Fatigue was not reported by any participant and therefore was not considered a factor. Fallers and nonfallers differed significantly on all balance-related measures (Table 3). Single trial SST administration required less than 15 minutes, with a typical clustering between 8 to 12 minutes depending upon body weight and skill level of the participant. Twenty- eight nonfallers (96.5% of the nonfallers) displayed effective balance stepping strategies within 1 to3 steps at 10% TBW waist pull forces. One nonfaller was unable to achieve SST 10% TBW initially, requiring a second chance attempt. This participant reported a fear of falling when outdoors despite recorded performance measures well within age referenced norms for the other 4 balance related tests examined in this study. The mean number of steps required to recover from directional limit perturbations (anterior-rear stepping / posterior-forward stepping respectively) was 1.9/1.4 for nonfallers and 2.4/2.2 for fallers. Directional failure was slightly biased towards the anterior (rear stepping) direction, in both fallers (55%) and nonfallers (51%). Ceiling effects were recorded for 2 nonfaller participants (3.5%) in the 65 to 69 age group, each weighing 200 pounds. These resulted in a SST 13% TBW limit score (26 pound scale capacity divided by 200 pound body weight), exceeding the 12.3% TBW mean SST % TBW limit value for nonfallers. Two fallers (7%) achieved SST10%TBW, both requiring a second chance attempt. In addition no participant indicated a change in self-reported fall status from recruitment to test day.

The 2 groups studied were of equal sizes determined by self-reported 2-year fall history. The 80 to 89 year old age group was the largest sample subgroup. The percentage of fallers increased with increasing decade of groups (Table 2).

Reliability of the SST

The SST repeated measures demonstrated excellent same day test-retest reliability (ICC = 0. 94 [95 % CI = 0.89-0.96]). Reliability of the SST was further supported by a ME of 0.74 and by a ME_{cv} of 7.25%. The ME, a measure of response stability quantifies absolute instrument error, and indicates that measurements will need to differ by more than 0.74 lb. to be considered truly different.

Validity of the SST

The SST% TBW score demonstrated moderate to good convergent validity with the other balance related measures (Table 4). The correlation of the SST%TBW score with each of the other measures was significant ($p < .001$) as were correlations of the alternative balance related measures with one another.

Table 3 reports analysis for known groups validity. The t-tests revealed significant differences between fallers and nonfallers for

Table 3. Comparison of Measurements Obtained from 58 Older Adults Classified as Fallers (n=29) or Nonfallers (n=29)

Variable	Fallers Mean (SD)	Nonfallers Mean (SD)	T-test (p)	ROC Results** Area Under Curve	Cut Point	Sensitivity/ Specificity (%)	PV+/PV- %)***
SST (% weight)*	7.5 (1.4)	12.3 (1.7)	11.63 (.001)	.992	10.0	93.1 / 96.6	96.4 / 93.3
Gait speed (m/s)	1.0 (0.2)	1.3 (0.2)	4.623 (.001)	.820	1.2	65.5 / 75.9	73.1 / 68.8
Timed Up & Go (s)	9.2 (1.3)	7.0 (0.9)	-7.22 (.001)	.916	7.4	79.3 / 93.1	92.0 / 81.8
Single limb stance (s)	3.2 (3.3)	10.3 (9.6)	3.78 (.001)	.766	6.5	48.3 / 89.7	82.4 / 63.4
Tandem stance (s)	12. 7 (10.8)	23.9 (9.9)	4.12 (.001)	.765	22.0	72.4 / 75.9	75.0 / 73.3

* SST%= Spring Scale Test percentage
** ROC = Receiver operating characteristic curve
*** PV+ = Positive predictive value, PV- = Negative predictive value

Table 4. Pearson Correlations Between Variables Predicting Falls

Variable	SST%	GS	TUG	SLS
SST%*				
Gait speed (GS)	0.53 (.001)			
Timed Up & Go (TUG)	-0.67 (.001)	-0.51 (.001)		
Single limb stance (SLS)	0.54 (.001)	0.34 (.009)	-0.48 (.001)	
Tandem stance (TS)	0.55 (.001)	0.36 (.005)	-0.46 (.001)	0.50 (.001)

*SST% = Spring Scale test percentage

all balance related measures. The SST 10% TBW performance measure was judged to represent the best combination of sensitivity and specificity, possessing the highest level of discrimination to fall status compared to available alternative fall related variables examined in this study. The statistical comparison of the ROC curves (Figure 2) showed the area under the curve for the SST at a cut score of 10% TBW was statistically larger than for other measures of mobility and balance (AUC = 0. 992), sensitivity 93.1% and specificity 96.6%, positive predictive value, (PV=96.4%), negative predictive value, (PV-93.3%). Logistic regression revealed that once the variance in fall status attributable to the SST 10% TBW (91.1%) was included in the predictive equation (Wald 8.78, p= .003), no other variable added to the explanation (p = .097 [TUG] to .994 [TS]).

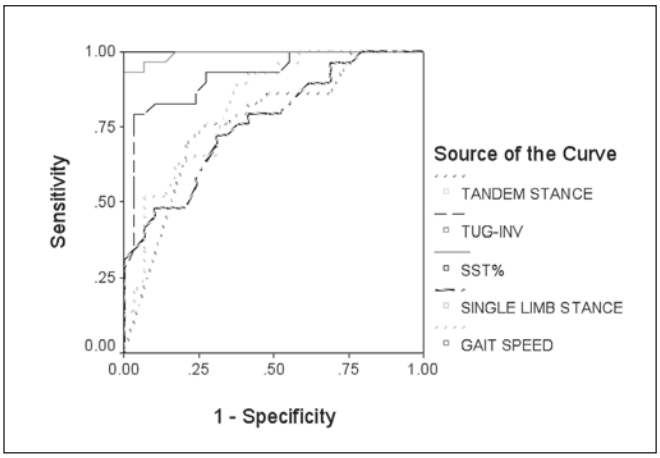

Figure 2. ROC curves illustrating sensitivity and specificity of balance related measures (TUG –INV = inverse of timed-up-and go time, SST% = Spring Scale Test measure as a % of body weight).

DISCUSSION

The SST protocol reported in this study differs significantly from previous reported waist pull methods[10-15] which employ multiple examiners, nonportable equipment, either one-directional, nonstepping, or random, unpredictable methods. The SST has yielded an innovative, reliable, safe, practical, simple, and clinically feasible single examiner waist-pull, fall risk assessment method requiring less than 15 minutes for proficient completion. The SST loading foot-flat nonstepping accommodation, and the SST unloading 3 step limit balance response performance criteria are components of other described waist-pull procedures.[11,13,14] Nonfallers typically displayed fewer, well-defined steps compared to fallers during SST trials. Posterior SST directional limit testing yielded a 0.8 mean step difference between fallers and nonfallers compared to 0.5 anterior SST directional limit testing. Similar findings by Shultz et al[10] were obtained using a random, unpredictable, multidirectional laboratory based method within a 5% TBW maximum load limit, demonstrating a 0.5 step difference in both anterior and posterior directions between normal and balance impaired older adults, considered functionally significant.[15] In distinct contrast to random, unpredictable methods, the ability of the predictable loading and semi-random unloading principles of the SST protocol to obtain similar directional stepping limit data reported by others[10] supports the clinical relevance and feasibility of this new approach.

Percent of total body weight (%TBW) is a conventional measure used to quantify perturbation forces regardless of method.[10-15] These studies[10-15] describe waist-pull perturbation methods entailing issues of safety and/or logistical practicalities. Concern regarding ordinal scoring validity and low sensitivity to subtle age related deteriorations in postural control associated with one method has been reported.[11] Collectively, these factors have perhaps hindered the development of a practical, safe, and reliable clinical perturbation method.

Reliability

The test-retest reliability of the SST is supported by its high ICC and low ME, and $ME_{cv.}$ The high ICC value (0.94) compares favorably to ICC values reported by other investigators for the alternative balance related measures examined in this study: gait speed (0.90),[37] TUG (0.98),[30] (SLS (0.99),[38] and TS (0.70).[40] The ME analysis indicates that measurements will need to differ by more than 0.74 lb to be considered truly different. Since the SST protocol employs 1 lb units of measurement, the ME value represents a fractional portion of this 1 lb incremental unit of measure. A limitation is that these measures were repeated within a single session, making it uncertain as to what kind of difference between sessions exceeds measurement error and can be considered real. The ME was relatively minor and acceptable based on the nature of the measurements and the possible ranges of response variability. In addition, provision of the ME_{cv} (7.25%) provides a minimal % change value between 2 SST % TBW scores indicating that the differences between two SST % TBW scores needs to be at least 7.25% to be considered real change in performance which would be on the magnitude of 1 lb approximately for body weight <135 lbs and 2 pounds for body weight values > 135 lbs. Information on the response stability of perturbation and other tests is largely lacking. Test-retest reliability was determined through same-day testing. The high test-retest ICC scores suggest that the test protocol provides predictable and repeatable performance with no apparent short-term learning. The 15 minute time interval provided adequate time for completion of the additional measures and SF-36 subscale survey data collection, and rest for participants resulting in a practical and manageable method facilitating timely completion of study data collection tasks and enhancing retention of volunteer participants. The

purpose for selecting the 15-minute interval was to avoid genuine changes in the measured variable. Use of the 15 minute interval in this study was appropriate based upon the test-retest ICC (0.94), suggesting the effect of memory and learning was apparently negligible, yielding data which could be interpreted and applied with greater confidence, thus strengthening the generalizability of the research outcomes. Same day test-retest is a conventional method previously reported.[24-28] The nature of the measurements prevents simultaneous testing by 2 examiners.

Validity

Our results uphold the convergent validity of the SST. Although the SST was the only test in this study which involved externally applied postural perturbations for the purposes of quantifying stepping and nonstepping responses, the SST demonstrated moderate to good convergent validity with the other balance related measures. In our study correlations between the SST and the other balance related measures ($r \geq 0.53$) were higher than those between the other balance related measures themselves ($r \leq -0.51$). In addition, correlations between the SST and the other balance related measures in our study were comparable to or higher than correlations reported elsewhere between balance related measures other than the SST. For example, correlations of $r = 0.32$ to -0.50 have been reported between SLS performance and gait speed. Correlations of $r = 0.40$ to -0.56 and $r = 0.56$ have been reported between SLS time and TUG and TS times respectively. This reinforces that the SST displays higher correlations with all the other balance related measures examined than any of the other measures have with each other. This information would support justification for not performing multiple tests in this population.

The SST 10% TBW value was the best performance measure for discriminating fall status and for identifying fall group membership. Compared to published values for the other balance related performance measures used: gait speed (80%/89%; descriptive same-day, at 34m/min cutoff), [47] TUG (87%/87%; retrospective 1 year, 13.5 sec. cutoff), [30] SLS (95%/58%; retrospective 1 year), [39] and TS (55%/94%; prospective 14 months) [41] the SST possessed the highest sensitivity (93.1%) and specificity (96.6%), as well as excellent positive and negative predictive values. Reported values will depend upon criterion score and fall definition. Although all of the balance related performance measures examined were able to discriminate with regards to fall status, our analysis confirmed that the SST 10 % TBW value was the best performance measure for discriminating fall status and for identifying fall group membership. Good diagnostic measures are expected to have sensitivity greater than 80% and specificity greater than 60%. [48] Measures used for screening, as opposed to diagnosis, may need to be more sensitive (90%), if it is important that no one with the problem is missed.[48] The excellent discriminant ability of the SST 10% TBW cut off is further supported by the ROC (AUC = 0.99) the highest value obtained for all comparative balance related measures (Table 3), which further supports the SST 10% TBW as a quantitative fall risk screening performance measure in the active community living aging adult. In addition to high levels of combined sensitivity and specificity, the excellent positive and negative predictive values (PV+ 96%, PV- 93%), demonstrate the efficiency of the SST protocol in terms of time and resources. The time required for proficient administration of the SST protocol (8 – 12 minutes), would seem justified, yielding a sufficient number of accurate responses to be clinically useful. Results of logistic regression confirmed the high discriminant ability of the SST to fall status indicating that SST 10% TBW was the variable most likely to identify fall group membership in the sample tested. Because this was not a prospective study, conclusions about the cut score to use for predicting future fall risk are not possible.

The ROC generated cut points for the balance related measures examined in this study are comparable to, or exceed, published mean performance values for community living adults, (gait speed, 0 .71 –1.24 m/sec),[49] (TUG, 8 –11 sec.),[29] (TS, 10 sec.),[33] with the exception of SLS (6.5 sec.) compared to (SLS, 8.5-17 sec.).[50] Nevertheless, the results of this study suggest that use of ROC determined cut point values for other balance related performance tests examined in this study failed to detect many active community living older adults with a moderate to high risk of falling. These tests may not have provided adequate stress for balance deficits to emerge The 10% TBW value may represent a minimal performance measure indicative of low fall risk useful as a screening tool in the active community living aging adult. Intuitively, the SST protocol would appear useful for identifying fall status in a less active aging sub group with more pronounced deficits, an issue which needs further examination.

Of note, 30 of the participants were aged 80 to 89 years, comprising the largest subgroup of the sample (52%) containing the most fallers. This is consistent with national statistics indicating the incidence of falls is higher in this age group, accounting for nearly 75% of all reported injuries in this group living in the United States.[1] In addition, history of a prior fall is a significant, if not the most important predictor of a future fall.[51] In this study the mean percent of total body weight (% TBW) load that fallers could manage was 7.5% compared to 12.3% for nonfallers. The SST 10% TBW performance measure effectively discriminated and identified fall group membership between known groups of active and independent community living aging adults. Individuals unable to achieve SST 10% TBW limit scores would share group membership with individuals reporting a fall history within the past 2 years.

Clinical applications of the SST % TBW performance measure could include screening for fall risk, clinical decision making regarding community integration, quantifying responses to balance improvement interventions, designing an induced stepping treatment paradigm, and determining the need for specific assistive devices. Further study is warranted to establish the responsiveness to change in fall status and predictive validity of the SST 10 % TBW performance value through prospective design. In addition further study is warranted to duplicate the results of this study, determine the usefulness of the SST in different subgroups of the aging population, examine the usefulness of the SST protocol as an induced stepping treatment paradigm, describe and examine the reliability and validity of a predictable lateral testing spring scale waist pull protocol and compare to other available performance measures including the protocol examined in this study.

Limitations

The SST protocol is predicated on prior knowledge of both direction and magnitude of all loading waist pull perturbation

forces in combination with sudden controlled, semi-random timed unloading (5 sec. window) of successfully accommodated loading spring scale forces. The SST method of semi-random unloading of spring scale forces attempts to avoid predictable release patterns minimizing anticipatory postural responses. The 26-pound linear spring scale capacity (13% TBW) determined the 200-pound body weight limit inclusion criteria. The linear spring scale features calibrations in 1-pound units thus rendering fractional readings unfeasible. Fifteen minutes was the test-retest time interval used.

Limitations included use of volunteers as a sample of convenience. Two participants were former patients of the experienced examiner, several participants were either friends or family members of former patients of the experienced examiner, and other participants were members of community senior centers who are highly motivated to perform optimally. Women constituted the majority of the sample (67%). Multiple versus single fall status was not differentiated over the 2-year period. Vision acuity and sensation were not evaluated as potential contributors to performance. Medical and fall history were dependent upon accurate recall by the subject. Inclusion criteria did account for stable well-managed chronicities to represent a range of older adults who are fairly active and in fairly good health in spite of the presence of some pathology. This judgment is further supported by the reported mean SF-36 physical functioning subscale scores (63.3 for fallers) and (86.3 for nonfallers), which exceeded the published mean subscale score of 57.1 reported for 368 community dwelling older adults 65years and older. [52] Despite these limitations, it appears that fallers may perform quantitatively different in clinically meaningful ways based on the results on this study.

CONCLUSION

The SST protocol described in this study is a quantitative, reliable and valid clinical assessment method of standing balance appropriate for the active community living older adult population. Compared to other balance related performance measures (gait speed, TUG, SLS, and TS), known groups and convergent analysis revealed that the SST 10% TBW performance measure best identified fall group membership.

ACKNOWLEDGEMENT

The authors extend special thankful gratitude to Richard Bohannon, DPT, EdD, NCS, FAPTA, FAHA for invaluable mentoring and loyal commitment to this project providing guidance and assistance with study design, statistical analysis, manuscript preparation, and revisions. Special thanks to VNSNY executive and regional management for ongoing support, to C. Scott, PT; D. Gulle, PT; and G. Octagon, PT for assistance with study implementation, K. DePasquale for technical assistance and to the gracious spirit of our volunteer participants.

REFERENCES

1. Falls Among Older Adults: An Overview. Centers for Disease Control Web site. Available at: http://www.cdc.gov/ncipc/factsheets/adultfalls.htm. Accessed July 30, 2008.
2. Cumming RG, Nevitt MC, Cummings SR. Epidemiology of hip fractures. *Epidemiol Rev.* 1997;19:244-257.
3. Arnold CM, Faulkner RA. The history of fall and the association of the timed up and go test to fall and near-falls in older people with osteoarthritis. *BMC Geriatr.* 2007;7:17-26.
4. Binder EF, Brown M, Sinacore DR, Stega-May K, et al. Effects of extended outpatient rehabilitation after hip fracture: a randomized controlled trial. *JAMA.* 2004;18:837-846.
5. 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.
6. Thrane G, Joakimsen RM, Thornquist E. The association between the timed up and go test and history of falls: The Tromso study. *BMC Geriatr.* 2007;7:1-7.
7. Nordin E, Lidelof N, Rosendahl E, Jensen J, Lundin-Olsson L. Prognostic validity of the timed up-and go test, a modified get-up-and-go test, staff's global judgment and fall history in evaluating fall risk in residential care facilities. *Age Aging.* 2008;37:442-448.
8. Lin MR, Hwang HF, Hu MH, Wu HD, Wang YW, Huang FC. Psychometric comparisons of the timed "up and go", one-leg stand, functional reach and Tinetti balance measures in community-dwelling older people. *J Am Geriatr Soc*. 2004;52:1343-1348.
9. Harris JE, Eng JJ, Marigold DS, Tokuno CD, Louis CL. Relationship of balance and mobility to fall incidence in people with chronic stroke. *Phys Ther.* 2005;85:150-158.
10. Schultz BW, Ashton-Miller JA, Alexander NB. Compensatory stepping in response to waist pulls in balance-impaired and unimpaired women. *Gait Posture.* 2005;22:198-209.
11. Wolfson LI, Whipple R, Amerman P, Kleinberg A. Stressing the posture response: a quantitative method for testing balance. *J Am Geriatr Soc.* 1986;34:845-850.
12. Duncan PW, Studenski S, Chandler J, Bloomfield R, Lapointe LK. Electromyographic analysis of postural adjustments in two methods of balance testing. *Phys Ther.* 1990;70:88-96.
13. Chandler JM, Duncan PW, Studenski SA. Balance performance on the postural stress test: comparison of young adults; healthy, elderly, and fallers. *Phys Ther.*1990;70:410-415.
14. Lee WA, Deming L, Sahgal V. Quantitative and clinical measures of static standing balance in hemiparetic and normal subjects. *Phys Ther.*1988;68:970-976.
15. Mansfield A, Peters AL, Liu BA, Maki BE. A perturbation-based balance training program for older adults: study protocol for a randomized controlled trial. *BMC Geriatr.* 2007;7:12-29.
16. Pai YC, Wening JD, Runtz EF, Iqbal K, Pavol MJ. Role of feedforward control of movement stability in reducing slip-related balance loss and falls among older adults. *J Neurophysiol*. 2003;90:755-762.
17. Pavol MJ, Pai YC. Feedforward adaptations are used to compensate for a potential loss of balance. *Exp Brain Res*. 2002;145:528-538.
18. Pai YC, Bhatt TS. Repeated-slip training: an emerging paradigm for prevention of slip-related falls among older adults. *Phys Ther.*2007;87:1-13.
19. Badke MB, Duncan PW, DiFabio RP. Influence of prior knowledge on automatic and voluntary postural adjust-

ments in healthy and hemiplegic subjects. *Phys Ther.* 1987;67:1495-1500.
20. Diener HC, Horak F, Stelmach G, et al. Direction and amplitude precuing has no effect on automatic postural responses. *Exp Brain Res.* 1991;89:219-223.
21. Richardson JK. Factors associated with fallers in older patients with diffuse polyneuropathy. *J Am Geriatr Soc.* 2002;50:1767-1773.
22. Tiedemann A, Shimada H, Sherrington C, Murray S, Lord S. The comparative ability of eight functional mobility tests predicting falls in community-dwelling older people. *Age Aging.* 2008;37:430-435.
23. Dite W, Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. *Arch Phys Med Rehabil.* 2002;83:1566-1571.
24. Rothstein JM. *Measurement and Clinical Practice*. 1st ed. New York, NY: Churchill Livingstone; 1985.
25. Giorgetti MM, Harris BA, Jette A. Reliability of clinical balance outcome measures in the elderly. *Physiother Res Int.* 1998;3:274-283.
26. Melzer I, Shtilman I, Rosenblatt N, Oddsson L. Reliability of voluntary step execution behavior under single and dual task conditions. *J Neuroengineering Rehabil.* 2007;4:16-25.
27. Morris S, Morris ME, Iansek R. Reliability of measurements obtained with the timed "up & go" test in people with Parkinson disease. *Phys Ther.*2001;81:10-818.
28. Rome K, Hanchard NC. Within-day reliability of temporal-spatial gait parameters associated with rheumatoid arthritic feet. *Musculoskeletal Care.* 2005;3:17-23.
29. Steffen TM, Hacker TA, Mollinger L. Age and gender-related test performance in community-dwelling elderly people; six-minute walk test, berg balance scale, timed up and go test and gait speed. *Phys Ther.* 2002;82:28-137.
30. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the timed up and go test. *Phys Ther*. 2000;80:846-853.
31. Bohannon RW, Larkin PA, Cook AC, Gear J, Singer J. Decrease in timed balance test scores with aging. *Phys Ther.* 1984;64:1067-1070.
32. Guralnik JM, Ferricci L, Pieper CF, et al. Lower extremity function and subsequent disability: Consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci.* 2000; 55:221-M231.
33. Shubert TE, Schrodt LA, Mercer VS, Busby-Whitehead J, Giuliani CA. Are scores on balance screening tests associated with mobility in older adults? *J Geriatr Phys Ther.* 2006;29:33-39.
34. Ware JE. *SF-36 Health Survey: Manual and Interpretation Guide.* Boston, MA: The Health Institute, New England Center; 1997.
35. Bohannon RW. Population representative gait speed and its determinants. *J Geriatr Phys Ther.* 2008;31:49-52.
36. Wolf SL, Catlin PA, Gage K, et al. Establishing the reliability and validity of measurements of walking time using the Emory Functional Ambulation Profile. *Phys Ther.* 1999;79:1122-1133.
37. Bohannon RW. Comfortable walking speeds of adults aged 20-79 years: reference values and determinants. *Age Aging*. 1997;26:15-19.
38. Springer BA, Marin R, Cyhan T, et al. Normative values for the unipedal stance test with eyes open and closed. *J Geriatr Phys Ther.* 2007;30:8-15.
39. Hurvitz EA, Richardson JK, Werner RA, Ruhl AM, Dixon MR. Unipedal stance testing as an indicator of fall risk among older outpatients. *Arch Phys Med Rehabil.* 2000;81:587-591.
40. Steffen T, Seney M. Test-retest and minimal detectable change on balance and ambulation tests, the 36-item short form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther.* 2008;88:733-746.
41. Murphy MA, Olson SL, Protas EJ, Overby AR. Screening for falls in community-dwelling elderly. *J Aging Phys Act.* 2003;11:66-80.
42. Ringsberg K, Gerdhem P, Johansson J, Obrant KJ. Is there a relationship between balance,gait performance and muscular strength in 75-year-old women? *Age Aging.* 1999;28:289-293.
43. Kinugasa T, Nagasaki H, Furuna T, Itoh H. Physical performance measures for characterizing high functioning older persons. *J Aging Phys Act.* 1996;4:338-248.
44. Valliant J, Vuillerme N, Martigne P, et al. Balance, aging, and osteoporosis: effects of cognitive exercises combined with physiotherapy. *J Bone Spine.* 2006;73:414-418.
45. Cho B, Scarpace D, Alexander N. Tests of stepping as indicators of mobility, balance, and fall risk in balance-impaired older adults. *J Am Geriatr Soc.* 2004;52:1168-1173.
46. Franchignoni F, Tesio L, Martino MT, Ricupero C. Reliability of four simple quantitative tests of balance and mobility in healthy elderly females. *Aging Clin Exp Res.* 1998;10:26-31.
47. Harada N, Chiu V, Damron-Rodriquez J, et al. Screening for balance and mobility impairment in elderly individuals living in residential care facilities. *Phys Ther.* 1995;75:462-469.
48. Lincoln NB, Nicholl CR, Flannaghan T, Leonnard M, Van der Gucht E. The validity of questionnaire measures for assessing depression after stroke. *Clin Rehabil.* 2003;17:840-846.
49. Lusardi MM, Pellecchia GL, Schulman M. Functional performance in community-living older adults. *J Geiatr Phys Ther*. 2003; 26:14-22.
50. Bohannon RW. Single limb stance times: a descriptive meta-analysis of data from individuals at least 60 years of age. *Top Geriatr Rehab.* 2006;22:70-77.
51. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. *Can J Public Health.* 1992; 83:S7-S11.
52. Sherman SE, Reuben D. Measures of functional status in community-dwelling elders. *J Gen Intern Med.*1998;13:817-823.