Published in final edited form as:

J Appl Biomech. 2019 February 01; 35(1): 11–18. doi:10.1123/jab.2017-0322.

Reliability of Postural Sway Measures of Standing Balance Tasks

Saud F. Alsubaie¹, Susan L. Whitney², Joseph M. Furman³, Gregory F. Marchetti⁴, Kathleen H. Sienko⁵, Patrick J. Sparto²

¹Department of Physical Therapy and Health Rehabilitation, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

²Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA, United States

³Department of Otolaryngology, University of Pittsburgh, Pittsburgh, PA, United States

⁴Department of Physical Therapy, Duquesne University, Pittsburgh, PA, United States

⁵Departments of Mechanical Engineering and Biomedical Engineering, University of Michigan, Ann Arbor, MI, United States

Abstract

The reliability of balance exercises performance in experimental and clinical studies has typically been confined to a small set of exercises. In order to advance the field of assessing balance exercise intensity, establishing the reliability of performance during a more diverse array of exercises should be undertaken. The purpose of this study was to investigate the test-retest reliability of postural sway produced during performance of 24 different balance tasks, and to evaluate the reliability of different measures of postural sway. Sixty-two healthy subjects between the ages of 18 and 85 years of age (50% female, mean age 55 ± 20 years) participated. Subjects were tested during two visits one week apart and performed two sets of the 24 randomized standing tasks per visit. The tasks consisted of combinations of the following factors: surface (firm and foam), vision (eyes open and eyes closed), stance (feet apart and semi-tandem), and head movement (no movement, yaw, and pitch). Angular position displacement, angular velocity, and linear acceleration postural sway in the pitch and roll planes was recorded via an inertial measurement unit. The postural sway measures demonstrated at fair to good test-retest reliability with few exceptions, and angular velocity measures demonstrated the greatest reliability. The between-visit reliability of two averaged trials was excellent for most tasks. The study indicates that performance of most balance tasks used as part of balance rehabilitation is reliable, and quantitative assessment could be used to document change.

Keywords

rehabilitation; prescription; posturography; intensity

Correspondence Address: Saud F. Alsubaie, PhD, PT, Assistant Professor at Department of Physical Therapy & Health Rehabilitation, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia, Phone: +966540588079, Sfa5@pitt.edu.

Conflict of Interest Disclosure: None

Introduction

Balance training has been found to be useful for all age groups in improving mobility and functionality¹⁻³. Customized balance exercises and vestibular rehabilitation therapy are considered to be effective options to improve balance by facilitating the central nervous system's ability to compensate for balance deficits¹⁻⁵. These treatments have elicited beneficial results in improving balance in older adults and people with vestibular disorders, lessening the symptoms of vestibular disorders, and reducing falls¹⁻⁴.

Prior to beginning a balance exercise intervention, patients have their balance assessed across a variety of domains⁶. For example, the ability to use various sensory systems to maintain upright standing balance can be assessed using the Clinical Test of Sensory Interaction and Balance⁷. The six conditions in the Clinical Test of Sensory Interaction and Balance represent a relatively small subset of conditions that may be used in balance therapy. In addition to examining the effect of reducing sensory input (e. g. standing on foam or closing eyes), assessing a change in the base of support (e. g. standing in semi-tandem stance) and perturbing the vestibular system (e.g. moving the head in yaw or pitch directions) may be undertaken before determining the exercise prescription.

In order to properly assess subjects' performance during balance and vestibular exercises so that decisions can be made about progression of those exercises, the reliability of performance should be established. This study includes a comprehensive set of tasks which are used as a part of balance and vestibular rehabilitation programs, which distinguishes this study from previous studies that considered the reliability of a limited number of tasks^{8,9}. This study also included tasks that incorporate head movements and this is the only study to our knowledge that includes this type of tasks. In addition, we examined the performance using inertial measurement units, which are being used with increasing frequency in clinical research because of their greater portability and lower cost compared with force platforms. Therefore, the purpose of this study was to examine the test-retest reliability of subjects' postural sway during 24 standing balance tasks, within and between two visits occurring one week apart. Secondary to this main purpose, we wanted to examine which kinematic variables of postural sway provided the best reliability, and also report on the standard error of measurement and minimal detectable change of these measures.

Methods

Sixty-two healthy subjects (out of 72 screened) who were independent in performance of daily activities and were between the ages of 18 and 85 years old (31 females and 31 males, mean age 55 ± 20 years) completed the study. The subjects were recruited through local advertisements and from previous balance research studies. Study participants were distributed into four age groups: young (18–44 years old), middle-aged (45–59 years old), old (60–74 years old), and very old (75–85 years old) to ensure a representative distribution of balance ability across the age spans. The University of Pittsburgh Institutional Review Board approved all procedures in this study and all subjects provided written informed consent prior to participation.

Subjects were excluded if they were unable to stand for 3 minutes without rest, were unable to complete the Romberg test for 30 seconds, had distal sensory loss (unable to feel a pressure of a 4.31 rated monofilament applied on the dorsum of the foot and the medial side of the foot below the medial malleolus with eyes closed ¹⁰), had visual acuity worse than 20/40, had a diagnosis of benign paroxysmal positional vertigo (positive Dix–Hallpike test or positive Roll test) or a peripheral vestibular disorder (positive head impulse test), had a history of neurological or orthopedic disorders severe enough to affect standing balance, used an assistive device for ambulation, were pregnant, had excessive weight (body mass index, BMI > 35), had cognitive impairment (< 25 points on the Montreal Cognitive Assessment), or had a history of falling 2 times or more within the last 12 months while performing activities of daily living. A total of 10 subjects were excluded based on the following reasons: eight subjects did not pass the screening tests (4 did not pass the cognitive test; 3 did not pass the monofilament test; 1 did not pass the horizontal roll test for benign paroxysmal positional vertigo), 1 subject was excluded due to a behavioral issue (under the influence of alcohol), and 1 subject did not come back for the follow up visit. The 10 subjects who did not complete the study had a mean age of 64 ± 14 years.

Subjects attended two testing sessions one week apart. During each experimental session, participants performed two sets of 24 standing balance tasks in randomized order. The tasks included a full-factorial combination of the following independent variables: vision (eyes open and eyes closed); surface (firm and foam surfaces); stance (feet apart and semi-tandem); and head movements (head still, yaw, and pitch) as shown in Table 1. Included in the table is a count of how many participants were able to successfully complete at least one trial of each task.

Participants stood without shoes to avoid the confounding effect of variation in footwear. During conditions of the foam surface, subjects stood on a foam pad (AIREX Balance Pad S34–55) and the room's temperature was kept at approximately 72° F during all visits to avoid differences in the foam properties. During the various base of support stances, subjects were instructed to distribute their body weight equally on each foot. For the semi-tandem stance position, subjects stood with the front foot touching the medial side of the other foot by a half of a foot length, with the dominant foot in the back. The dominant foot was determined by asking the subjects about the foot that they would use to kick a ball ¹¹. During the eyes closed conditions, subjects wore opaque goggles. During yaw and pitch conditions, subjects were instructed to move their head at a frequency of 1 Hz by moving their head to the beat of a metronome ¹² within a range of 45 degrees in the yaw direction and 30 degrees in pitch direction. Subjects practiced the head movement in these directions before they started the experiment.

Subjects were instructed to stand as stable as possible with arms at their side during all trials. The trial length was 35 s, which was chosen because it is a common time for standing balance exercise performance during rehabilitation and this duration has also been shown to provide reliable performance ^{13–16}. A trial was considered a failure and data collection was stopped if a subject demonstrated any of the following postural alterations: stepping out of position, changing their feet or arms from the starting position, and/or touching something for support. Subjects repeated failed trials once in each set if they lost their balance before

completing at least 25 seconds of a trial. Subjects were guarded by a physical therapist during all tasks to prevent falling and wore a safety harness attached to an anchor point in the ceiling that did not let subject reach the ground in case of a fall but would allow them to move freely. There was a seated rest break for 1 minute after every 3 tasks to avoid fatigue.

During the performance of the standing tasks, an inertial measurement unit (IMU, Xsens Technologies B.V., Enschede, The Netherlands)^{17,18} was mounted on each subject's posterior lower back at the level of the iliac crest (L4). The inertial measurement unit measured trunk angular displacement and velocity in the pitch and roll planes, and linear acceleration in the antero-posterior (A/P) and medio-lateral (M/L) directions at a sampling rate of 100 Hz. The inertial measurement unit was placed on a static surface to measure the amount of error due to just the instrument; the root-mean-square error in tilt = 0.08 deg, tilt velocity = 0.08 deg/s, and acceleration = 0.004 m/s².

Demographic data including age, gender, weight, height and BMI were summarized by descriptive statistics. Additionally, the mean scores of the Functional Gait Assessment, Activities-specific Balance Confidence Scale questionnaire, and gait speed for all groups were recorded.

Sway measures were recorded during all trials for 35 seconds and the first five seconds of data collection were removed in order to avoid the effect of the subject's initial establishment of balance¹³. The data were low-pass filtered using a second order Butterworth filter with a cut-off frequency of 3 Hz¹⁹. During the analysis, each trial was plotted individually and inspected visually using MATLAB software to ensure that there were no extraneous movements. The Root Mean Square (RMS) of the trunk angular displacement and velocity in the pitch and roll directions, and linear acceleration in the A/P and M/L directions were calculated and used in the analysis to test the hypotheses. The RMS was calculated as follows:

$$RMS = \sqrt{\frac{\sum_{i=0}^n (a_i^2)}{n}}$$

where a is instantaneous sway value with mean value subtracted, and i is an individual data sample, and n is the total number of samples.

Demographic characteristics were compared among groups using a one-way ANOVA for dependent variables that were continuous and normally distributed (gait speed) and post-hoc comparisons were conducted to evaluate pairwise differences among the groups, using the Sidak approach to control for Type 1 error. The Kruskal-Wallis test was used with dependent variables that were continuous but not normally distributed (BMI, Functional Gait Assessment, and Activities-specific Balance Confidence Scale questionnaire) and Dunn's procedure was used for pairwise comparisons with a Bonferroni correction for multiple comparisons.

To explore the test-retest reliability of the healthy subjects' performance during the stance balance tasks, absolute and relative measures of reliability were computed. For

relative reliability, the intra-class correlation coefficient (ICC) was used for variables with continuous characteristics (RMS of the trunk angular displacement velocity, and linear acceleration) ²⁰. Model (3) and form (1) of the ICC was used which indicates that each task was performed by each subject, as the subjects were the only subjects of interest, and that a single measurement was the unit of interest, for both the within- and between-visit analyses. For each task, test-retest reliability was assessed within the two sets of each visit, between the first set of both visits, and between the average of both sets from each visit. Intra-class correlation coefficient (ICC) reliability scores range from 0 to 1.0 where excellent reliability ranges from 0.75 to 1.0, fair to good reliability ranges from 0.4 to 0.74 and poor reliability ranges from 0 to 0.4 ²¹.

To assess the absolute reliability, the standard error of measurement (SEM), and minimal detectable change (MDC) were assessed. The Standard Error of Measurement (SEM) was calculated as follows: $SEM = SD * \sqrt{(1 - r)}$ where SD is the standard deviation, and r equals the reliability coefficient (i.e. ICC value) ²². The Minimal Detectable Change (MDC) was calculated as follows: $MDC = 1.96 * \sqrt{2} * SEM$ ²².

Results

Of 72 people who underwent onsite screening, 62 participants completed the study and comprised the four age groups as follows: young (n=17), middle-aged (n=15), older (n=15), and very old group (n=15) (Table 2). The participants had a mean age of 55 ± 20 years. There was a significant difference between groups on BMI, Activities-specific Balance Confidence scale, gait speed and the Functional Gait Assessment (Table 2). The young adults had significantly lower BMI than the old ($p = 0.001$) and very old participants ($p = 0.009$). The very old group had significantly lower scores on the Activities-specific Balance Confidence Scale questionnaire compared with the young adults ($p = 0.023$). Gait speed was significantly faster in the young ($p = 0.01$) and the middle-aged groups ($p = 0.007$) compared with the very old group. Finally, the very old had significantly lower (worse) scores than all the other groups on the Functional Gait Assessment ($p = 0.003$). Furthermore, referring back to Table 1, as the age of the participants increased, a greater percentage was not able to complete some of the more challenging tasks.

The overall group mean (SD) values of the six postural sway measures varied greatly among the 24 tasks (Table 3). General observations include greater magnitude of the values with tasks performed on foam compared with firm surface, with eyes closed compared with eyes open, and with head movement compared with head still. Furthermore, pitch head movements induced greater postural sway specifically in the pitch and AP directions compared with head still and yaw conditions, whereas yaw head movements induced greater movement in roll and ML directions compared with head still and pitch movements.

The test-retest reliability (ICC) was calculated for the RMS of trunk angular displacement (pitch and roll planes), angular velocity (pitch and roll planes), and linear acceleration (A/P and M/L directions) for each task separately. Average reliability coefficients across all 24 exercises were then computed in order to examine patterns with respect to the different sway measures (Table 4). Across all six measures, the best reliability coefficients were obtained

for the average of 2 trials, between visits, and none of the tasks had poor reliability ($ICC < 0.4$). The next highest reliability coefficients were calculated within visit 2, which were somewhat better than the reliability within visit 1. It is also of note that there were less individual tasks of poor reliability in visit 2 compared with visit 1. Finally, the reliability of the first set between visit one and two was the lowest and had the greatest number of individual tasks with poor reliability.

Considering the six postural sway measures, the average scores of the ICCs of the RMS of trunk angular velocity were greater than the trunk angular displacement and linear acceleration measures in general (Table 4). Based on this finding, we decided to illustrate the ICCs within visit 1 for all 24 balance tasks, for the RMS of trunk angular velocity in the pitch and roll directions. In addition, we computed the associated standard error of measurement and minimal detectable change for these measures (Table 5). There was considerable variation in test-retest reliability for pitch velocity among the 24 tasks within visit one, ranging from 0.18 (task 23, eyes closed in semi-tandem stance on foam with yaw head movement) to 0.87 (task 19, eyes closed in feet apart stance on foam with head still). For pitch velocity, 19 of the tasks had good to fair reliability, two had excellent reliability, and three had poor reliability. The three tasks with poor reliability were tasks 20, 23, and 24, which all were performed on foam surface with eyes closed: task 20 was with feet apart with yaw head movement, and tasks 23 and 24 were with feet semi-tandem with yaw and pitch head movements. It is recognized that these tasks are some of the most challenging that can be performed at any age. For roll velocity, 19 tasks had good to fair reliability and 5 had excellent reliability. The task-specific ICC, SEM, and MDC for the between-visit average of two set is reported in Table 6. The minimum ICC scores were 0.67 in pitch velocity and 0.66 in roll velocity for task 13 (standing on foam, eyes open, feet apart, and the head not moving). For both pitch and roll velocity, 20 tasks or more had excellent reliability between visits (i.e. average of 2 measures), and the remaining tasks had good to fair reliability.

Discussion

The test-retest reliability of performance across a diverse set of 24 balance tasks was good to excellent with few exceptions. The mean scores of the ICC coefficients of RMS of trunk tilt velocity were higher than the average scores of the coefficients of RMS of trunk tilt acceleration and displacement. Consistently, several studies that have examined the reliability of sway measures during different postural control tasks in various age populations found that the mean velocity is the most reliable measure of postural sway^{23–27}. A systematic review that included thirty-two studies revealed that among all other COP measures, the velocity measure was generally a reliable measure²⁷. A possible biomechanical explanation for sway velocity being the most reliable measure might be that body sway displacement is more prone to drift (for example going from a forward lean to a backward lean) and thus have more variation from trial to trial despite a person having stable posture. Conversely, the body sway velocity should have no drift over time, and thus the RMS value of velocity would be more consistent.

The sway measures in the frontal plane (i.e. roll and M/L directions) were more reliable in most of the tasks compared to sway measures in the sagittal plane (i.e. pitch and

A/P directions). Others have noted that sway measures in the M/L direction were more reliable compared to sway measures in the A/P direction, despite different instrumentation and populations utilized ^{25,26,28,29}. Overall, the between-subject variability was greater for the measures in the pitch and AP directions compared with the roll and ML directions, suggesting that the increased reliability values in the roll and ML directions were due, in fact, to improved consistency of performance in this direction, and not because of greater variation in the data.

This study determined that test-retest reliability coefficients of sway measures within visits appear to have higher reliability values compared to a single trial between visits, which is consistent with previous studies ^{24,26,30}. Several studies have proposed that this difference in test-retest reliability coefficients may be attributed to a change in postural control over time which could be the result of a learning effect ^{30,31}, whereas Fisher attributed the disparity in reliability scores to biological reasons such as stress of daily life that cannot be controlled ³².

After averaging the performance within visits, the ICC coefficients increased substantially for all variables compared to the ICC coefficients obtained from a single trial. Averaging sway measures from two trials or more leads to a better estimate of the true value which may explain the improvement of the ICC coefficients after averaging. Assuming that the construct of balance performance can be estimated to the same degree by inertial measurement units and force platforms, studies have been designed to determine the appropriate number of trials to be averaged in order to obtain reliable measures and concluded that averaging sway measures from at least two trials can improve the reliability coefficients, especially the velocity measure ^{23,27,33}. Lafond et al. recommended averaging two trials to obtain a reliable measure (ICC > 0.90) of the COP mean velocity and averaging 4 trials was needed to obtain a reliable measure of the COP range and displacement ²³. In light of our study's results, it is recommended to average sway measures from two trials in order to obtain reliable results especially for the velocity measure, while other measures may need averaging from more than two trials to obtain reliable results.

The standing balance tasks that were studied are commonly prescribed in the clinic, and encompass a wide variety of conditions that are used in balance and vestibular rehabilitation. However, some tasks had poor sway reliability. After reviewing the average value of sway measures and the missing rates of the balance tasks, it became clear that some of the tasks with low reliability coefficients were relatively easy tasks, resulting in limited between-subject variability which may explain the poor reliability ⁸. Other tasks (23 and 24) with low reliability were very difficult tasks in which subjects, especially older subjects, could not maintain their balance throughout these tasks, resulting in a greater proportion of missing data for those tasks. Furthermore, only subjects with good balance, which we expect to have less variability in sway, were able to maintain their balance throughout those tasks.

The experimental visits in this study lasted for one hour and forty-five minutes on average, which may have caused fatigue, especially for older adults who required more time for breaks. However, randomizing the testing conditions within sets and visits was designed to mitigate the order effect due to practice or fatigue.

Performance during standing balance tasks has acceptable reliability for most tasks. The RMS of trunk tilt velocity in the roll direction was the most reliable measure. These data can help researchers select reliable tasks of varying levels of difficulty that can be used to assess balance performance in experimental studies. Furthermore, clinicians can use the absolute reliability measures (SEM, MDC) to assess whether patients are improving during rehabilitation.

Acknowledgments

We would like to acknowledge the assistance of the following people in helping with data collection: Dr. Abdulaziz Alkathiry, Anita Lieb, Susan Strelinski, Dr. Chia-Cheng Lin, Dr. Mohammed Alyabroudi, Bader Alqahtani, Sahar Abdulaziz, Kefah Alshebbier, Brooke Klatt, Carrie Hoppes, David Fear, and Mohammed Almotairi.

We would like to acknowledge the Clinical and Translational Science Institute (CTSI) for helping with subject recruitment. The project described was supported in part by the National Institutes of Health through Grants UL1TR001857 and 5-R21-DC-012410-02.

References

1. Horak FB, Jones-Rycewicz C, Black FO, Shumway-Cook A. Effects of vestibular rehabilitation on dizziness and imbalance. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery*. 1992;106(2):175–180. [PubMed: 1738550]
2. Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. *The Cochrane database of systematic reviews*. 2012(9):CD007146.
3. Howe TE, Rochester L, Neil F, Skelton DA, Ballinger C. Exercise for improving balance in older people. *The Cochrane database of systematic reviews*. 2011(11):CD004963.
4. Hillier SL, McDonnell M. Vestibular rehabilitation for unilateral peripheral vestibular dysfunction. *The Cochrane database of systematic reviews*. 2011(2):CD005397.
5. Shepard NT, Telian SA. Programmatic vestibular rehabilitation. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery*. 1995;112(1):173–182. [PubMed: 7816453]
6. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. *Physical therapy*. 2009;89(5):484–498. [PubMed: 19329772]
7. Shumway-Cook A, Horak FB. Assessing the influence of sensory interaction of balance. Suggestion from the field. *Physical therapy*. 1986;66(10):1548–1550. [PubMed: 3763708]
8. Whitney SL, Roche JL, Marchetti GF, et al. A comparison of accelerometry and center of pressure measures during computerized dynamic posturography: a measure of balance. *Gait & posture*. 2011;33(4):594–599. [PubMed: 21333541]
9. Marchetti GF, Bellanca J, Whitney SL, et al. The development of an accelerometer-based measure of human upright static anterior- posterior postural sway under various sensory conditions: test-retest reliability, scoring and preliminary validity of the Balance Accelerometry Measure (BAM). *J Vestib Res*. 2013;23(4–5):227–235. [PubMed: 24284603]
10. Holewski JJ, Stess RM, Graf PM, Grunfeld C. Aesthesiometry: quantification of cutaneous pressure sensation in diabetic peripheral neuropathy. *Journal of rehabilitation research and development*. 1988;25(2):1–10.
11. Gabbard C, Hart S. A question of foot dominance. *The Journal of general psychology*. 1996;123(4):289–296. [PubMed: 9042740]
12. Hall CD, Herdman SJ. Reliability of clinical measures used to assess patients with peripheral vestibular disorders. *J Neurol Phys Ther*. 2006;30(2):74–81. [PubMed: 16796772]
13. Rine RM, Schubert MC, Whitney SL, et al. Vestibular function assessment using the NIH Toolbox. *Neurology*. 2013;80(11 Suppl 3):S25–31. [PubMed: 23479540]

14. Muehlbauer T, Roth R, Bopp M, Granacher U. An exercise sequence for progression in balance training. *Journal of strength and conditioning research / National Strength & Conditioning Association*. 2012;26(2):568–574.
15. Allum JH, Carpenter MG, Horslen BC, et al. Improving impaired balance function: real-time versus carry-over effects of prosthetic feedback. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference*. 2011;2011:1314–1318.
16. Le Clair K, Riach C. Postural stability measures: what to measure and for how long. *Clinical biomechanics*. 1996;11(3):176–178. [PubMed: 11415618]
17. Blair Stephanie, Duthie Grant, Robertson Sam, Hopkins William, Ball K. Concurrent validation of an inertial measurement system to quantify kicking biomechanics in four football codes. *Journal of biomechanics*. 2018.
18. Al-Amri M, Nicholas K, Button K, Sparkes V, Sheeran L, Davies JL. Inertial Measurement Units for Clinical Movement Analysis: Reliability and Concurrent Validity. *Sensors (Basel)*. 2018;18(3).
19. Dozza M, Chiari L, Horak FB. Audio-biofeedback improves balance in patients with bilateral vestibular loss. *Archives of physical medicine and rehabilitation*. 2005;86(7):1401–1403. [PubMed: 16003671]
20. Fisher RA. *Statistical methods for research workers*. 12th ed. Edinburgh,: Oliver and Boyd; 1954.
21. Fleiss JL. *The design and analysis of clinical experiments*. Wiley classics library ed. New York: Wiley; 1999.
22. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Physical therapy*. 2006;86(5):735–743. [PubMed: 16649896]
23. Lafond D, Corriveau H, Hebert R, Prince F. Intrasection reliability of center of pressure measures of postural steadiness in healthy elderly people. *Archives of physical medicine and rehabilitation*. 2004;85(6):896–901. [PubMed: 15179642]
24. Benvenuti F, Mecacci R, Gineprari I, et al. Kinematic characteristics of standing disequilibrium: reliability and validity of a posturographic protocol. *Archives of physical medicine and rehabilitation*. 1999;80(3):278–287. [PubMed: 10084435]
25. Swanenburg J, de Bruin ED, Favero K, Uebelhart D, Mulder T. The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers. *BMC musculoskeletal disorders*. 2008;9:162. [PubMed: 19068125]
26. Rafal S, Janusz M, Wieslaw O, Robert S. Test-retest reliability of measurements of the center of pressure displacement in quiet standing and during maximal voluntary body leaning among healthy elderly men. *Journal of human kinetics*. 2011;28:15–23. [PubMed: 23487011]
27. Ruhe A, Fejer R, Walker B. The test-retest reliability of centre of pressure measures in bipedal static task conditions--a systematic review of the literature. *Gait & posture*. 2010;32(4):436–445. [PubMed: 20947353]
28. Heebner NR, Akins JS, Lephart SM, Sell TC. Reliability and validity of an accelerometry based measure of static and dynamic postural stability in healthy and active individuals. *Gait & posture*. 2015;41(2):535–539. [PubMed: 25544692]
29. Moe-Nilssen R. Test-retest reliability of trunk accelerometry during standing and walking. *Archives of physical medicine and rehabilitation*. 1998;79(11):1377–1385. [PubMed: 9821897]
30. Lin D, Seol H, Nussbaum MA, Madigan ML. Reliability of COP-based postural sway measures and age-related differences. *Gait & posture*. 2008;28(2):337–342. [PubMed: 18316191]
31. Tjernstrom F, Fransson PA, Hafstrom A, Magnusson M. Adaptation of postural control to perturbations--a process that initiates long-term motor memory. *Gait & posture*. 2002;15(1):75–82. [PubMed: 11809583]
32. Fisher ST. *The intra-session and inter-session reliability of centre-of-pressure based measures of postural sway within a normal population*. UNITEC Institute of Technology. 2010.
33. Corriveau H, Hebert R, Prince F, Raiche M. Intrasection reliability of the “center of pressure minus center of mass” variable of postural control in the healthy elderly. *Archives of physical medicine and rehabilitation*. 2000;81(1):45–48. [PubMed: 10638875]

Table 1:

Balance tasks conditions. The order of the tasks was randomized for each subject, and each set. The number of task completers is the number of participants out of 62 who were able to successfully complete at least one trial of the task.

<u>Task</u>	<u>Surface</u>	<u>Visual input</u>	<u>Base of support</u>	<u>Head movement</u>	<u>Task completers (n)</u>
1	Firm	Eyes open	Feet apart	Head still	62
2	Firm	Eyes open	Feet apart	Yaw	62
3	Firm	Eyes open	Feet apart	Pitch	62
4	Firm	Eyes open	Semi-tandem	Head still	62
5	Firm	Eyes open	Semi-tandem	Yaw	62
6	Firm	Eyes open	Semi-tandem	Pitch	62
7	Firm	Eyes closed	Feet apart	Head still	62
8	Firm	Eyes closed	Feet apart	Yaw	62
9	Firm	Eyes closed	Feet apart	Pitch	62
10	Firm	Eyes closed	Semi-tandem	Head still	62
11	Firm	Eyes closed	Semi-tandem	Yaw	59
12	Firm	Eyes closed	Semi-tandem	Pitch	60
13	Foam	Eyes open	Feet apart	Head still	62
14	Foam	Eyes open	Feet apart	Yaw	62
15	Foam	Eyes open	Feet apart	Pitch	62
16	Foam	Eyes open	Semi-tandem	Head still	62
17	Foam	Eyes open	Semi-tandem	Yaw	57
18	Foam	Eyes open	Semi-tandem	Pitch	57
19	Foam	Eyes closed	Feet apart	Head still	62
20	Foam	Eyes closed	Feet apart	Yaw	62
21	Foam	Eyes closed	Feet apart	Pitch	61
22	Foam	Eyes closed	Semi-tandem	Head still	62
23	Foam	Eyes closed	Semi-tandem	Yaw	37
24	Foam	Eyes closed	Semi-tandem	Pitch	41

Table 2:

Demographic characteristics and clinical balance measures.

	All (18–85)	Young (18–44)	Middle-aged (45–59)	Older (60–74)	Very old (75–85)
	n=62	n=17	n=15	n=15	n=15
Age, years Mean (SD)	55 (20)	28 (8)	53 (4)	67 (4)	79 (3)
Gender, female n (%)	31 (50)	9 (53)	8 (53)	7 (47)	7 (47)
Body Mass Index, kg/m ² Median (Range)	26.3 (15.5–35.8)	21.8 (18.1–33.5)	27.5 (18.1–32.1)	29.9 (15.5–34.8)	27.8 (19.9–35.8)
Semmes-Weinstein Monofilament, log ₁₀ (force)Median (Range)	4.08 (2.83–4.31)	3.84 (2.83–4.08)	4.08 (2.83–4.31)	4.08 (3.61–4.17)	4.17 (3.61–4.31)
Montreal Cognitive AssessmentMedian (Range)	29 (26–30)	29 (26–30)	28 (26–30)	29 (26–30)	28 (26–30)
The Activity-specific Balance Confidence ScaleMedian (Range)	97 (81–100)	99 (89–100)	94 (83–100)	98 (81–100)	91 (88–99)
Gait Speed, m/sMean (SD)	1.30 (0.20)	1.38 (0.20)	1.39 (0.21)	1.26 (0.19)	1.16 (0.12)
Functional Gait AssessmentMedian (Range)	28 (19–30)	29 (27–30)	29 (23–30)	28 (19–30)	24 (19–29)

Table 3:

Mean (SD) of postural sway measures for each of the balance tasks across all four trials (2 sessions x 2 visits).

Task	RMS Pitch Position deg	RMS Roll Position deg	RMS Pitch Velocity deg/s	RMS Roll Velocity deg/s	RMS AP Accel. m/s ²	RMS ML Accel. m/s ²
1	0.39 (0.18)	0.13 (0.08)	0.45 (0.20)	0.17 (0.08)	0.07 (0.03)	0.02 (0.01)
2	0.40 (0.18)	0.16 (0.07)	0.71 (0.27)	0.74 (0.49)	0.07 (0.03)	0.06 (0.03)
3	0.56 (0.22)	0.14 (0.06)	2.03 (1.15)	0.37 (0.20)	0.12 (0.04)	0.03 (0.01)
4	0.51 (0.28)	0.34 (0.13)	0.70 (0.37)	0.46 (0.24)	0.09 (0.05)	0.07 (0.03)
5	0.57 (0.27)	0.45 (0.23)	1.14 (0.52)	1.09 (0.59)	0.11 (0.05)	0.12 (0.05)
6	0.68 (0.30)	0.45 (0.20)	2.16 (1.24)	0.87 (0.43)	0.15 (0.05)	0.10 (0.04)
7	0.41 (0.25)	0.13 (0.09)	0.52 (0.53)	0.19 (0.17)	0.07 (0.05)	0.02 (0.02)
8	0.44 (0.19)	0.21 (0.10)	0.90 (0.39)	1.05 (0.77)	0.08 (0.04)	0.08 (0.05)
9	0.66 (0.27)	0.16 (0.07)	2.50 (1.33)	0.45 (0.21)	0.14 (0.05)	0.03 (0.01)
10	0.57 (0.38)	0.39 (0.20)	0.86 (0.43)	0.58 (0.28)	0.10 (0.06)	0.09 (0.04)
11	0.72 (0.39)	0.56 (0.29)	1.47 (0.67)	1.41 (0.69)	0.14 (0.07)	0.15 (0.06)
12	0.79 (0.36)	0.53 (0.26)	2.61 (1.38)	1.09 (0.57)	0.17 (0.06)	0.13 (0.05)
13	0.53 (0.23)	0.25 (0.10)	0.67 (0.28)	0.37 (0.15)	0.10 (0.04)	0.05 (0.02)
14	0.73 (0.27)	0.34 (0.14)	1.20 (0.51)	0.99 (0.57)	0.14 (0.05)	0.10 (0.04)
15	0.89 (0.36)	0.33 (0.16)	2.46 (1.21)	0.66 (0.26)	0.18 (0.06)	0.07 (0.03)
16	0.69 (0.35)	0.60 (0.29)	1.13 (0.61)	1.10 (0.59)	0.12 (0.06)	0.12 (0.05)
17	0.96 (0.46)	0.95 (0.44)	1.89 (0.90)	2.30 (1.14)	0.18 (0.08)	0.21 (0.08)
18	1.03 (0.48)	0.87 (0.40)	2.82 (1.35)	1.87 (0.83)	0.20 (0.08)	0.19 (0.07)
19	0.68 (0.30)	0.29 (0.15)	0.90 (0.45)	0.45 (0.25)	0.13 (0.05)	0.06 (0.03)
20	1.01 (0.48)	0.41 (0.16)	1.69 (0.79)	1.30 (0.70)	0.20 (0.08)	0.12 (0.05)
21	1.15 (0.49)	0.38 (0.15)	3.04 (1.34)	0.81 (0.31)	0.24 (0.09)	0.09 (0.04)
22	1.01 (0.77)	0.80 (0.44)	1.74 (1.31)	1.65 (1.04)	0.19 (0.12)	0.17 (0.08)
23	1.25 (0.47)	1.42 (0.64)	2.55 (1.14)	3.49 (1.77)	0.24 (0.07)	0.30 (0.11)
24	1.41 (0.68)	1.28 (0.72)	3.47 (1.58)	2.96 (1.78)	0.28 (0.11)	0.27 (0.12)

Table 4:

Average intraclass correlation coefficients (ICC model 3,1) of the Root-Mean-Square (RMS) of trunk tilt displacement, velocity, and acceleration across the 24 tasks, standard deviation (SD), and the number of tasks with poor reliability (ICC < 0.4).

	Within 1 st visit (Set 1 & 2)		Within 2 nd visit (Set 1 & 2)		Between visits (1 st Set)		Between visits (Average of 2 Sets)	
	ICC (SD)	# of <u>tasks</u> with poor reliability	ICC (SD)	# of <u>tasks</u> with poor reliability	ICC (SD)	# of <u>tasks</u> with poor reliability	ICC (SD)	# of <u>tasks</u> with poor reliability
RMS of pitch displacement	0.38 (0.16)	11	0.50 (0.14)	5	0.39 (0.14)	9	0.68 (0.11)	0
RMS of roll displacement	0.51 (0.17)	4	0.56 (0.19)	4	0.44 (0.16)	8	0.84 (0.09)	0
RMS of pitch velocity	0.58 (0.15)	3	0.60 (0.15)	1	0.47 (0.14)	6	0.81 (0.06)	0
RMS of roll velocity	0.64 (0.12)	0	0.63 (0.19)	3	0.47 (0.14)	8	0.84 (0.09)	0
RMS of A/P acceleration	0.46 (0.17)	7	0.53 (0.13)	4	0.40 (0.15)	9	0.73 (0.09)	0
RMS of M/L acceleration	0.57 (0.12)	2	0.59 (0.16)	2	0.49 (0.12)	6	0.87 (0.06)	0

A/P: antero-posterior; M/L: medio-lateral.

Table 5:

Intraclass correlation coefficients model (3, 1), standard deviation (SD), standard error of measurement (SEM), and minimal detectable change (MDC) for RMS of pitch and roll trunk tilt velocity within the 1st visit (Sets 1 and 2).

Task	RMS of pitch velocity				RMS of roll velocity			
	ICC	SD	SEM	MDC	ICC	SD	SEM	MDC
1	0.71	0.17	0.09	0.26	0.79	0.08	0.04	0.10
2	0.62	0.26	0.16	0.44	0.80	0.52	0.23	0.64
3	0.59	1.06	0.68	1.88	0.80	0.21	0.09	0.26
4	0.67	0.25	0.14	0.40	0.70	0.29	0.16	0.44
5	0.63	0.45	0.27	0.76	0.67	0.63	0.36	1.01
6	0.47	1.34	0.97	2.70	0.53	0.47	0.32	0.89
7	0.56	0.22	0.15	0.41	0.81	0.08	0.03	0.10
8	0.64	0.41	0.25	0.69	0.53	0.85	0.58	1.62
9	0.63	1.34	0.81	2.25	0.68	0.22	0.12	0.34
10	0.59	0.49	0.31	0.86	0.35	0.34	0.27	0.76
11	0.45	0.56	0.41	1.15	0.55	0.67	0.45	1.24
12	0.51	1.49	1.04	2.90	0.59	0.61	0.39	1.08
13	0.73	0.26	0.13	0.37	0.63	0.15	0.09	0.25
14	0.61	0.52	0.32	0.90	0.63	0.61	0.37	1.02
15	0.48	1.25	0.90	2.50	0.67	0.30	0.17	0.48
16	0.75	0.46	0.23	0.64	0.70	0.52	0.28	0.78
17	0.57	0.91	0.60	1.65	0.61	1.18	0.73	2.04
18	0.67	1.52	0.87	2.42	0.66	0.86	0.50	1.39
19	0.87	0.49	0.18	0.49	0.66	0.20	0.12	0.33
20	0.32	0.91	0.75	2.07	0.78	0.76	0.36	0.99
21	0.51	1.25	0.87	2.42	0.51	0.33	0.23	0.63
22	0.70	1.22	0.67	1.86	0.47	0.85	0.62	1.71
23	0.18	0.93	0.85	2.35	0.71	1.75	0.94	2.61
24	0.37	2.04	1.62	4.49	0.43	1.73	1.31	3.62

Table 6:

Intraclass correlation coefficients model (3, 1), standard deviation (SD), standard error of measurement (SEM), and minimal detectable change (MDC) for RMS of pitch and roll trunk tilt velocity between the 1st and 2nd visit, using the average of sets 1 and 2.

Task	RMS of pitch velocity				RMS of roll velocity			
	ICC	SD	SEM	MDC	ICC	SD	SEM	MDC
1	0.73	0.20	0.10	0.29	0.90	0.08	0.03	0.07
2	0.86	0.27	0.10	0.28	0.87	0.49	0.18	0.49
3	0.82	1.15	0.49	1.36	0.80	0.20	0.09	0.24
4	0.78	0.37	0.17	0.48	0.80	0.24	0.11	0.29
5	0.83	0.52	0.21	0.59	0.92	0.59	0.17	0.47
6	0.87	1.24	0.45	1.24	0.95	0.43	0.10	0.27
7	0.74	0.53	0.27	0.75	0.76	0.17	0.08	0.23
8	0.85	0.39	0.15	0.42	0.89	0.77	0.26	0.71
9	0.81	1.33	0.58	1.61	0.69	0.21	0.12	0.33
10	0.79	0.43	0.20	0.54	0.75	0.28	0.14	0.38
11	0.76	0.67	0.33	0.92	0.68	0.69	0.39	1.09
12	0.87	1.38	0.50	1.38	0.93	0.57	0.15	0.42
13	0.67	0.28	0.16	0.44	0.66	0.15	0.09	0.24
14	0.85	0.51	0.20	0.54	0.92	0.57	0.16	0.45
15	0.87	1.21	0.44	1.21	0.80	0.26	0.12	0.33
16	0.76	0.61	0.30	0.83	0.88	0.59	0.20	0.57
17	0.72	0.90	0.47	1.32	0.89	1.14	0.38	1.05
18	0.91	1.35	0.41	1.12	0.87	0.83	0.30	0.83
19	0.84	0.45	0.18	0.49	0.91	0.25	0.08	0.21
20	0.76	0.79	0.39	1.07	0.89	0.70	0.23	0.64
21	0.86	1.34	0.50	1.39	0.76	0.31	0.15	0.41
22	0.84	1.31	0.53	1.46	0.89	1.04	0.35	0.96
23	0.82	1.14	0.48	1.34	0.93	1.77	0.47	1.30
24	0.85	1.58	0.61	1.69	0.81	1.78	0.78	2.15