

TEST-RETEST RELIABILITY OF THE SWAY BALANCE MOBILE APPLICATION

Ryan Z. Amick, PhD¹, Alex Chaparro, PhD², Jeremy A. Patterson, PhD³, Michael J. Jorgensen, PhD⁴

¹Department of Human Performance Studies, College of Education, Wichita State University, Wichita, KS; ²Department of Psychology- Human Factors, Aging, Perception & Performance Laboratory, College of Liberal Arts and Sciences, Wichita State University, Wichita, KS; ³Department of Human Performance Studies, Human Performance Laboratory, Center for Physical Activity and Aging, College of Education, Wichita State University, Wichita, KS; ⁴Department of Biomedical Engineering, College of Engineering, Wichita State University, Wichita, KS

Corresponding Author: ryan.amick@wichita.edu

Background: The SWAY Balance Mobile Application is an FDA-cleared balance testing system which uses the built-in tri-axial accelerometers of a mobile electronic device to objectively assess postural movement. The system was designed to provide a means of quantitative balance assessment in clinical and on-field environments. The purpose of this study was to determine the intrasession and intersession reliability, as well as the minimum difference to be considered real, of the SWAY Balance Mobile Application.

Methods: 24 individuals (15 male, 9 female; aged 25.96 (± 5.78 years)) performed the SWAY Balance protocol twice per testing session over a period of three testing sessions. Each testing session was separated by a minimum of seven days. Interclass Correlation Coefficients were calculated as an indication of the test-retest reliability. The minimum difference to be considered real was calculated to determine the minimum score change necessary to indicate an actual change in balance performance.

Results: Mean SWAY Balance scores ranged from 86.90 (± 14.37) to 89.90 (± 11.19). Repeated measures ANOVA revealed no significant mean differences between SWAY balance scores of the experimental trials ($F_{(5,115)} = 0.673$; $p = 0.65$). Excellent reliability was found (ICC(3,1) = 0.76; SEM = 5.39) with a minimum difference to be considered real of approximately 15.

Conclusions: Results indicate that SWAY provides excellent overall reliability. However, it may be appropriate to have subjects perform a familiarization trial at the beginning of each testing session. Additionally, SWAY may demonstrate a ceiling effect when assessing balance improvements in those who already demonstrate good balance.

Introduction

Recent advances in mobile consumer electronics technology has significantly changed the way business is conducted in many different industries, including healthcare. This is due to the capabilities of devices such as smartphones, tablet computers, and

multimedia devices, which incorporate various sensors including accelerometers, gyroscopes, cameras, global positioning systems (GPS), magnetometers, and microphones with increasingly efficient operating systems, microprocessors, and batteries.^{1,2} Within the healthcare field, these advancements have allowed for

an increasingly prominent role of consumer electronic devices in patient monitoring, diagnostics, communication, and medical education.¹ The benefit of these devices is that they are relatively affordable, portable, and require only the addition of a software application to access the data outputs of the installed sensors.

An area of growing interest is the application of mobile consumer electronic devices in the role of patient care and monitoring. For this purpose, a number of different smartphone based mobile device applications have been developed utilizing the various sensors contained in these devices. This includes, among others, applications which monitor the travel patterns of patients with dementia and/or Alzheimer disease, the activity levels of patients in cardiac and stroke rehabilitation programs, as well as at home monitoring of those with sleep apnea, diabetes, and mental disorders.³⁻⁸

A number of mobile device applications have also been developed for the biomechanical assessment of human movement. These applications tend to utilize data gathered from the micro-electromechanical systems (MEMS) tri-axial accelerometers and/or gyroscopes to provide information utilized for gait analysis, falls detection, and activity recognition.^{2,9-16} However, despite the increasing potential, availability, and access of mobile device applications which provide patient monitoring functions, caution should be exercised when utilizing these systems. This is due to a lack of oversight during the development of these software applications by governing and standards agencies such as the U.S. Food and Drug Administration (FDA), with developers being unverified sources of clinical diagnostic information.¹⁷ Additionally, the degree and quality of reliability and validity testing varies greatly among these systems.¹

One recently developed method for assessing standing balance is the SWAY Balance Mobile Application (SWAY) (SWAY Medical, Tulsa, OK, USA). SWAY is an FDA approved mobile device software application which, when installed on a mobile consumer electronic device, accesses the MEMS tri-axial accelerometer output to assess balance through a series of balance tests. This assessment method is intended to provide professionals in various healthcare fields the ability to perform quantitative functional limitations assessments and fall risk assessments.¹⁸ Additionally, it can potentially be utilized by practitioners in sports medicine,

such as athletic trainers, to provide supporting information to be utilized when making return-to-play decisions after an athlete has suffered an injury.

The SWAY protocol consists of five stances (Figure 1) including bipedal (feet together), tandem stance (left foot forward), tandem stance (right foot forward), single leg stance (right), and single leg stance (left). Each stance is performed on a firm surface with eyes closed for a period of 10 seconds. For the duration of each stance, the subject holds the measuring device upright against the mid-point of their sternum. Deflections of the tri-axial accelerometer are then recorded throughout each of the balance test stances. Upon completion of the five stances, these deflections are utilized to calculate a final balance score ranging from 0–100, with higher scores indicating better balance. The final balance score is unit-less. It is an interpretation of the acceleration of deflections within the accelerometer and is derived by undisclosed calculations from SWAY Medical.

Preliminary testing has indicated that SWAY yields consistent and reliable measures of human standing balance. Pilot testing was performed comparing the consistency of SWAY balance scores to those measured concurrently with the BIODEX Balance System SD (BBS).¹⁵ Postural stability was recorded in the anterior-posterior direction as subjects performed a static Athlete's Single Leg Test protocol for a period of 10 seconds. Overall results showed no significant difference between mean SWAY scores and Anterior-Posterior stability scores recorded with the BBS. Additionally, subject feedback was recorded with regard to the usability of SWAY. Subjects reported that the SWAY software application was easy to navigate and that testing instructions were clear and easy to follow. However, this study was limited to measuring posture only in the anterior-posterior directions. The rationale for this is that an early iteration of the SWAY platform was used which had not yet incorporated accelerations detected in the medial-lateral direction.

The current version of the SWAY software (version 1.6) does now incorporate accelerations in all axes for the calculation of a balance score. However, the reliability of the fully developed SWAY software has yet to be established. Therefore, the purpose of this study was to determine the test-retest reliability of the SWAY Balance Mobile Application balance scores.

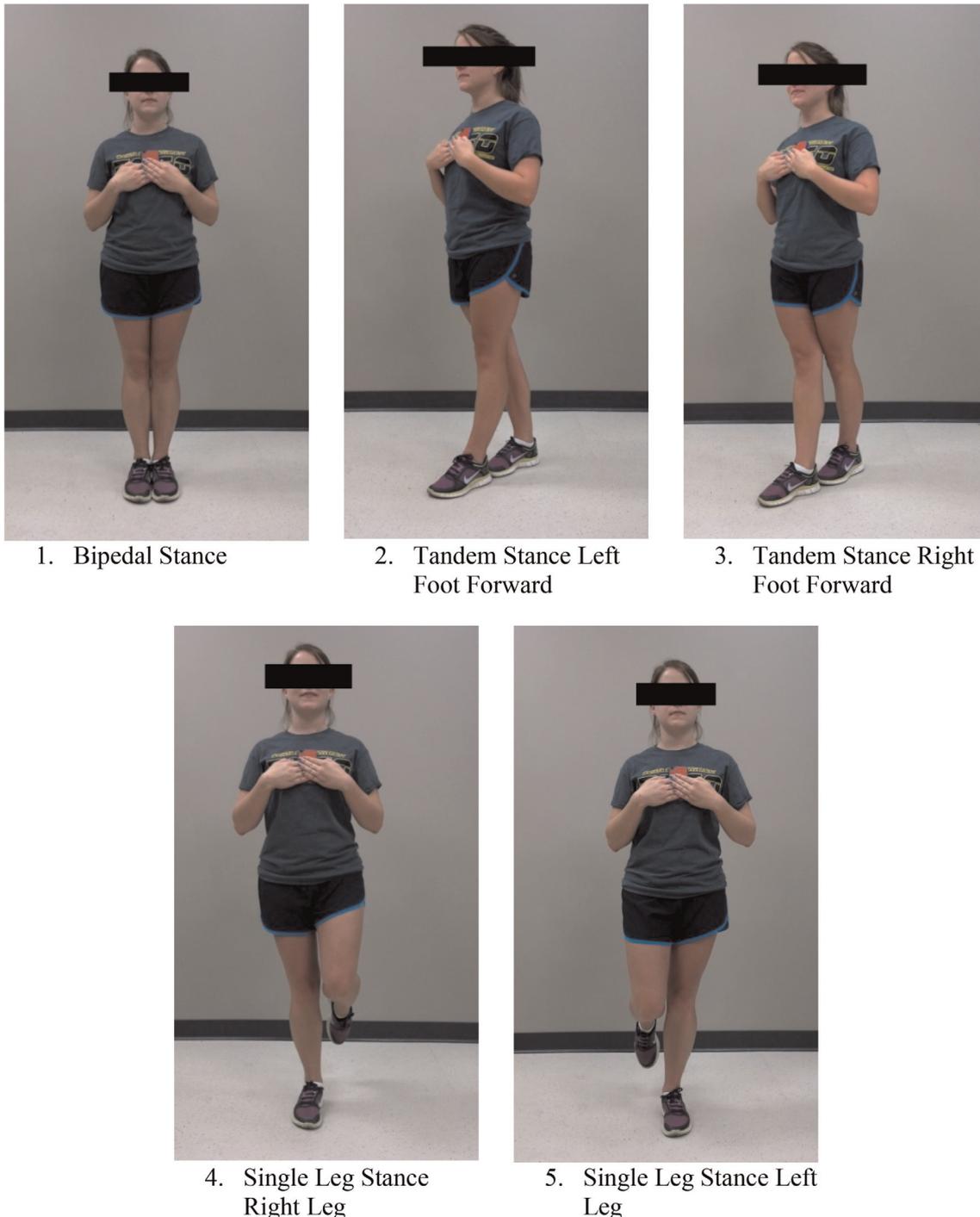


Figure 1: SWAY Balance Mobile Application Balance Stances While Holding Measuring Device Against Chest

Methods

Participants

A sample of convenience totaling 24 individuals (15 male, 9 female; aged 25.96 (± 5.78 years)) volunteered to participate in this study. All participants were university graduate and undergraduate students free from any condition or injury which may limit their ability to balance. All methods and procedures were approved by the Wichita State

University Institutional Review Board for Human Subjects. An Informed Consent form describing the nature of the testing to be completed, as well as exclusion criteria, was provided to all participants upon arrival to the testing facility. Testing procedures were then explained to all participants and exclusion criteria confirmed verbally. Participants were excluded if they reported any pre-existing condition that may alter their ability to balance normally. Upon receiving approved informed

consent, demographic and anthropometric measures were recorded.

SWAY Balance Mobile Application

SWAY Balance (SWAY Medical, Tulsa, OK, USA) is a mobile device software application which accesses the MEMS tri-axial accelerometer output to measure balance during a series of balance tests. The SWAY Balance testing protocol developed by SWAY Medical, LLC consists of five stances each performed for 10 seconds. Stances include bipedal standing (feet together), tandem standing (heel-to-toe with right foot behind left), tandem standing (heel-to-toe with left foot behind right), single leg standing (right foot), and single leg standing (left foot). Each stance is performed on a firm surface with eyes closed (Figure 1).

The SWAY balance test was administered utilizing an Apple iPod Touch (5th Generation) (Apple Computer Inc., Cupertino, CA, USA) loaded with the SWAY software (version 1.6). For each balance stance, subjects were instructed to hold the device upright, using both hands to press the face of the device against the mid-point of their sternum, so that the top of the device was below a line horizontal with the clavicles. Instructions for each balance stance were presented on the iPod screen sequentially so that upon the completion of a balance stance, instructions for the next stance were automatically displayed. Once all balance stances were completed, a final balance score ranging from 0-100 was produced and recorded, with a higher score indicating better balance.

SWAY Balance tests were administered during three testing sessions. Test-retest reliability was determined within sessions (intrasession reliability), and across sessions (intersession reliability). To minimize the potential for introducing a training effect, each testing session was separated by a minimum of seven days.¹⁹ During the first testing session, a familiarization trial was administered, per

SWAY recommendations. The familiarization trial was then followed by two experimental SWAY trials. Here, an experimental trial is defined as the full completion of the SWAY protocol (i.e., testing of the five stances). During the second and third testing sessions, only two experimental trials were administered. All intrasession trials were separated by a minimum of two minutes. The testing protocol is illustrated in Figure 2. Subjects performed all evaluations without shoes.

Data Analysis

Statistical analysis for this study was completed with the use of Statistical Packages for the Social Science (SPSS) version 21.0 (Chicago, Ill.) with a level of significance set at $\alpha < 0.05$. A Kolmogorov-Smirnov test was performed to evaluate all balance scores for normality of distribution. A 2x6 (SEX x TRIAL) mixed factorial ANOVA with repeated measures was used to analyze changes in SWAY scores between groups and across trials. An intraclass correlation coefficient (ICC) was calculated, which represents a ratio of actual score variance to overall variance. Here, an ICC(3,1) model was utilized as this model assesses only the reliability of the measurement by considering subjects as random effects and the measurement tool as a fixed effect.^{20,21} Additionally, for interpreting the ICC values, “excellent” reliability is indicated by an ICC > 0.75, ICC ranging from 0.40 – 0.75 indicates “fair to good” reliability, and an ICC < 0.40 indicates “poor” reliability.²²

The ICC was then used to determine the standard error of the measure (SEM), as well as the minimum difference to be considered real (MD). Here the SEM represents an absolute estimate of the reliability of the test by providing an indication of the expected variation in observed scores that occur due to measurement error. This allows for the determination of a range of scores within which a true score is likely to fall based upon an observed score.²⁰ A low SEM would produce a smaller range around an



Figure 2: Experimental Protocol

observed score, indicating better reliability of the test. The MD represents the change in score on a repeated evaluation necessary to reflect an actual change in performance. Similarly, the percent coefficient of variation (%CV) was calculated which represents the percent score change necessary to be considered an actual change in performance.²⁰

Results

Subject demographic information is presented in Table 1. All balance scores were found to be normally distributed. Mean SWAY scores ranged from 86.90 (± 14.37) to 89.90 (± 11.19). Descriptive statistics for each of the six experimental trials are reported in Table 2 and illustrated in Figure 3. A 2x6 (SEX x TRIAL) mixed factorial ANOVA with repeated measures revealed that male and female subjects did not differ significantly with regard to SWAY scores ($F(1,22) = 0.075, p = 0.79$). Repeated measures ANOVA revealed no significant mean differences between SWAY balance scores of the experimental trials ($F(5,115) = 0.673; p = 0.65$).

SWAY balance scores from trials 1, 3, and 5, and then 2, 4, and 6 were used to calculate the intersession reliability. For trials 1, 3, and 5 (Table 3), a good

degree of intersession reliability was found ($ICC(3,1) = 0.61; SEM = 7.51$). The MD and %CV were also determined to be 20.81 and 10.47% respectively. For trials 2, 4, and 6 (Table 3), an excellent degree of intersession reliability was found ($ICC(3,1) = 0.76; SEM = 5.39$). The MD and %CV were also determined to be 14.95 and 5.95% respectively. Intrasection reliability of SWAY Balance scores are summarized in Table 4. Here, the degree of reliability ranged from good in week one ($ICC(3,1) = 0.47$), to excellent in weeks two ($ICC(3,1) = 0.78$) and three ($ICC(3,1) = 0.75$).

Discussion

This study is the first to examine the test-retest reliability of the SWAY Balance Mobile Application. To determine the reliability of the SWAY protocol, subjects were asked to perform two balance trials per testing session with each session separated by a minimum of seven days. To investigate intersession reliability, ICC and SEM values were calculated utilizing balance scores from the 1st, 3rd, and 5th trials (trials 1-3-5), and again calculated utilizing the 2nd, 4th, and 6th trials (trials 2-4-6). When comparing the two data sets, slight differences can be seen. The ICC of trials 1-3-5 indicate

	Mean (SD)
Age (years)	25.96 (5.78)
Weight (kg)	78.20 (16.52)
Stature (cm)	173.22 (11.09)
BMI (kg/m ²)	25.93 (4.26)
3 rd Lumbar Vertebrae (cm)	108.45 (7.64)
Sternal Mid-Point (cm)	131.10 (9.05)

BMI = Body Mass Index, cm = Centimeters, kg = Kilograms, m² = Meters Squared

Table 1: Subject Demographic Information

	Mean (SD)
SWAY Trial 1 (Week 1)	87.93 (9.61)
SWAY Trial 2 (Week 1)	88.46 (11.12)
SWAY Trial 3 (Week 2)	86.90 (14.37)
SWAY Trial 4 (Week 2)	89.57 (10.59)
SWAY Trial 5 (Week 3)	88.49 (11.71)
SWAY Trial 6 (Week 3)	89.90 (11.19)

Table 2: Sway Score Descriptive Statistics

	Trials 1-3-5	Trials 2-4-6
ICC (3,1)	0.61	0.76
SEM	7.51	5.39
MD	20.81	14.95
%CV	10.47%	5.95%

ICC = Intraclass Correlation Coefficient, SEM = Standard Error of the Measure, MD = Minimum Difference to be Considered Real, %CV = Percent Coefficient of Variation

Table 3: Sway Intersession Reliability

	Trials 1-2 (Week1)	Trials 3-4 (Week 2)	Trials 5-6 (Week 3)
ICC (3,1)	0.47	0.78	0.75
SEM	7.56	5.82	5.77
MD	20.96	16.13	15.99
%CV	8.41%	6.8%	6.43%

ICC = Intraclass Correlation Coefficient, SEM = Standard Error of the Measure, MD = Minimum Difference to be Considered Real, %CV = Percent Coefficient of Variation

Table 4: Sway Intrasection Reliability

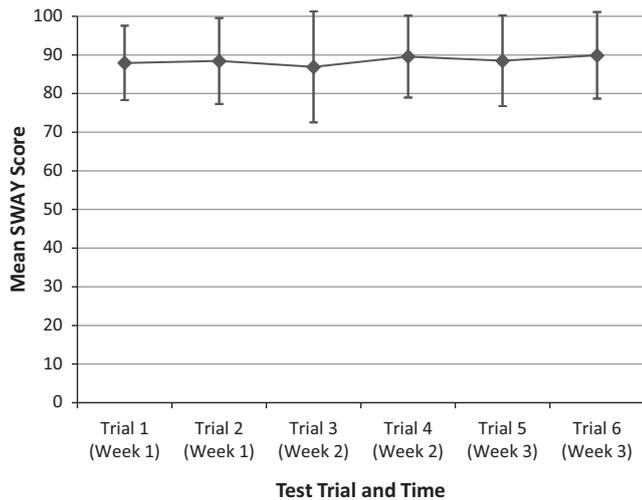


Figure 3: Mean SWAY Balance Scores by Trial and Time. Error bars equal to one standard deviation.

good intersession reliability ($ICC = 0.61$), compared to an excellent interrater reliability for trials 2-4-6 ($ICC = 0.76$). Additionally, the SEM was slightly higher for trials 1-3-5 ($SEM = 7.51$) when compared to trials 2-4-6 ($SEM = 5.39$). This indicates that for each testing session, subjects tended to perform slightly better on the second trial than they did on the first. This may indicate that a practice effect occurs after the first assessment.

Further evidence supporting poorer performance on the first trial of each testing session when compared to the second, can be seen when comparing the MD and % CV of each data set. For trials 1-3-5, a relatively large score change of approximately 21 must occur before an actual change in balance can be assumed to occur. However, for trials 2-4-6, a score change of approximately 15 must occur. Additionally, for trials 1-3-5, the percent change in score to be considered real is nearly 10.5%, whereas it is approximately 6% for trials 2-4-6.

Taking into account the differences between the two data sets of trials 1-3-5 and trials 2-4-6, it may be possible to observe an increased degree of reliability of SWAY. One method to improve the overall reliability of SWAY may be to have subjects perform a familiarization trial at the beginning of each testing session, especially when testing sessions are separated by seven days or more. A comparison between the trials 1-3-5 and trials 2-4-6 data sets revealed that subjects tended to perform better on the second within day trial, which may indicate that a practice effect occurs after the first assessment. Therefore, administering a familiarization trial at the beginning of each testing session would retain

the excellent intersession reliability values realized utilizing the trials 2-4-6 dataset. Additionally, it has been suggested that administering a familiarization trial may be an effective method for reducing random error.²³

Similar to intersession reliability, overall intrasession reliability was found to range from good to excellent. Intrasession reliability values for week 2 and week 3 were found to be excellent, with ICC's of 0.78 ($SEM = 5.82$) and 0.75 ($SEM = 5.77$) respectively, while week 1 was found to be good ($ICC = 0.47$; $SEM = 7.56$). The improved reliability values may be an indication that a learning effect occurred after week 1. However, the repeated measures ANOVA revealed this not to be the case as no significant differences between mean SWAY scores were observed across all six experimental trials. Despite a lower intrasession ICC for week 1 compared to week 2 and 3, we are confident in stating that the intrasession reliability of SWAY is excellent as multiple testing sessions indicated an ICC of 0.75 or greater. Additionally, when comparing the SEM across weeks, we see that while the week 1 SEM is higher than those for week 2 and 3, it is by less than 2. Furthermore, while the % CV for week 1 compared to weeks 2 and 3 was higher, the difference was less than 2%. Therefore, while the week 1 ICC differs from weeks 2 and 3, the error within the measurements are relatively similar.

Despite findings of excellent intersession and intrasession reliability, the results of this study may indicate that SWAY demonstrates a ceiling effect. A ceiling effect refers to when a measurement or assessment tool exhibits a specific upper limit for possible scores, and a majority of those evaluated score near that limit.²⁴ For SWAY, the balance scores produced fall within a possible range of 0-100, with higher scores indicating better balance. Here, for all experimental trials, the overall mean SWAY balance scores ranged from 86.90 to 89.90. Additionally, the SEM values for trials 1-3-5 and trials 2-4-6 are 7.51 and 5.39 respectively, and MD for trials 1-3-5 and trials 2-4-6 were 20.81 and 14.95 respectively. If we utilize the SEM and MD values from trials 2-4-6, using the SEM we would be able to determine if a subject's true score demonstrates a positive change, so long as their observed score is below approximately 94. However, we would not be able to determine if that change was due to a real change in balance. This is due to a balance score change of approximately 15 needing to occur before

we can say that an actual change in balance has occurred, as indicated by the MD.

In these healthy subjects, SWAY may be limited in its ability to detect a positive change in balance because the average score is within 15 of the maximum possible score. This could ultimately limit the usability of SWAY, especially to detect subtle balance changes in those who already demonstrate high SWAY balance scores. However, to assess those who already have good balance, it may be possible to modify the SWAY protocol. One potential method to do this would be to have subjects perform SWAY while standing on a compliant floor surface, such as a foam pad. This would potentially alter somatosensory feedback, thus increasing the difficulty of maintaining balance. This in turn may result in reduced balance scores generated by SWAY, allowing for balance assessments to be conducted on those with good balance. However, if the protocol is to be modified in this manner, baseline and familiarization trials would likely need to be completed with the compliant floor surface as well.

While the primary finding of this study was that the SWAY Balance Mobile Application demonstrated overall excellent intersession and intrasession reliability, caution should be utilized when generalizing these findings to different population samples. First, the data from this study were obtained from a sample of convenience, where subjects were healthy young adults. Therefore, both intersession and intrasession reliability may be significantly different in other populations such as older adults and those with physical or physiological conditions affecting balance performance. Second, this study did not control for leisure time physical activity.

Conclusion

In conclusion, this study determined the test-retest reliability of the SWAY Balance Mobile Application in a sample of young healthy adults. Results indicate that SWAY provides excellent overall reliability. However, results indicated that within a testing session, SWAY balance scores tended to be slightly higher on the second trial when compared to the first. This may indicate that subjects balance performance improved after the first trial. Therefore, to improve reliability, it may be appropriate to have subjects perform a familiarization trial at the beginning of each testing session, especially when testing sessions are separated by seven days or more. Additionally, SWAY may demonstrate a ceiling

effect when assessing balance improvements in those who already demonstrate good balance.

Competing Interests

The author's declare that they have no competing interests.

Authors Contributions

RZA conceived the study, participated in its design and coordination, data analysis and interpretation, and helped to draft the manuscript. AC participated in the study design, data analysis, and helped to draft the manuscript. JAP participated in the study design and coordination, and helped to draft the manuscript. MJJ participated in the study design and coordination, data analysis and interpretation, and helped to draft the manuscript. All authors read and approved the final manuscript.

References

1. Ozdalga E, Ozdalga A, Ahuja, N. The smartphone in medicine: a review of current and potential use among physicians and students. *Journal of medical Internet research* 2012;**14**(5):e128.
2. Mellone S, et al. Smartphone-based solutions for fall detection and prevention: the FARSEEING approach. *Zeitschrift Für Gerontologie Und Geriatrie* 2012;**45**(8):722–7.
3. Sposaro F, Danielson J, Tyson, G. iWander: an Android application for dementia patients. In: Engineering in Medicine and Biology Society (EMBC). 2010 Annual International Conference of the IEEE, IEEE 2010.
4. Worringham C, Rojek A, Stewart I. Development and feasibility of a smartphone, ECG and GPS based system for remotely monitoring exercise in cardiac rehabilitation. *PloS one* 2011;**6**(2):e14669.
5. Edgar SR, et al. Wearable shoe-based device for rehabilitation of stroke patients. In: Engineering in Medicine and Biology Society (EMBC). 2010 Annual International Conference of the IEEE, IEEE 2010.
6. Bsoul M, Minn H, Tamil L. Apnea MedAssist: real-time sleep apnea monitor using single-lead ECG. *Information Technology in Biomedicine. IEEE Transactions on* 2011;**15**(3):416–27.
7. Harvey P, et al. Data acquisition in a wireless diabetic and cardiac monitoring system. In: Engineering in Medicine and Biology Society, EMBC. 2011 Annual International Conference of the IEEE, IEEE 2011.

8. Puiatti A, et al. Smartphone-centred wearable sensors network for monitoring patients with bipolar disorder. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, IEEE 2011.
9. Galán-Mercant A, Cuesta-Vargas AI. Differences in trunk accelerometry between frail and non-frail elderly persons in functional tasks. *BMC research notes* 2014;7(1):100.
10. Nishiguchi S, et al. Reliability and validity of gait analysis by android-based smartphone. *Telemedicine and e-Health* 2012;18(4):292–6.
11. Yamada M, et al. Objective assessment of abnormal gait in patients with rheumatoid arthritis using a smartphone. *Rheumatology international* 2012;32(12):3869–74.
12. Guimarães V, et al. Phone based fall risk prediction, in Wireless Mobile Communication and Healthcare. Springer 2012:135–42.
13. Lee RY, Carlisle AJ. Detection of falls using accelerometers and mobile phone technology. *Age and ageing* 2011;40(6):690–6.
14. Rigoberto, MM, Toshiyo, T, Masaki, S. Smart phone as a tool for measuring anticipatory postural adjustments in healthy subjects, a step toward more personalized healthcare. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, IEEE 2010.
15. Patterson JA, et al. Validation of measures from the smartphone SWAY Balance application: A pilot study. *International Journal of Sports Physical Therapy* 2014;9(3):135–9.
16. Patterson JA, et al. Comparison of the SWAY Balance mobile application to a commonly used balance assessment method. *International Journal of Athletic Therapy & Training* 2014;19(3):4–7.
17. Franko O. Validate an App: How to Design Your Study and Get Published. *Journal of Mobile Technology in Medicine* 2012;1(2):1–4.
18. SWAY Medical. [cited 2014 July, 22]; Available from: <http://swaymedical.com/>.
19. Steffen T, Seney M. Test-retest reliability 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. *Physical Therapy* 2008;88(6):733–46.
20. Vincent, WJ, Weir, JP. Statistics in Kinesiology. 4th edn. Champaign, IL: Human Kinetics 2012.
21. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research* 2005;19(1):231–40.
22. Fleiss, J. Reliability of Measurement. The design and analysis of clinical experiments. New York, NY: Wiley 1986.
23. Evans MD, Goldie PA, Hill KD. Systematic and random error in repeated measurements of temporal and distance parameters of gait after stroke. *Archives of physical medicine and rehabilitation* 1997;78(7):725–9.
24. Lewis-Beck, MS., Bryman, A, Liao, TF. Encyclopedia of Social Science Research Methods. Thousand Oaks, CA: SAGE Publications, Inc. 2004.