Gait & Posture 31 (2010) 307–310

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance

Ross A. Clark [a,*], Adam L. Bryant [a], Yonghao Pua [b], Paul McCrory [a], Kim Bennell [a], Michael Hunt [a]

[a] Centre for Health, Exercise and Sports Medicine, Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Carlton, Victoria 3010, Australia
[b] Department of Physiotherapy, Singapore General Hospital, Singapore

ARTICLE INFO

Article history:
Received 9 July 2009
Received in revised form 10 November 2009
Accepted 15 November 2009

Keywords:
Balance
Motor control
Movement disorder
Rehabilitation
Force plate
Biomechanics
Gait
Posture

ABSTRACT

Impaired standing balance has a detrimental effect on a person's functional ability and increases their risk of falling. There is currently no validated system which can precisely quantify center of pressure (COP), an important component of standing balance, while being inexpensive, portable and widely available. The Wii Balance Board (WBB) fits these criteria, and we examined its validity in comparison with the 'gold standard'—a laboratory-grade force platform (FP). Thirty subjects without lower limb pathology performed a combination of single and double leg standing balance tests with eyes open or closed on two separate occasions. Data from the WBB were acquired using a laptop computer. The test–retest reliability for COP path length for each of the testing devices, including a comparison of the WBB and FP data, was examined using intraclass correlation coefficients (ICC), Bland–Altman plots (BAP) and minimum detectable change (MDC). Both devices exhibited good to excellent COP path length test–retest reliability within-device (ICC = 0.66–0.94) and between-device (ICC = 0.77–0.89) on all testing protocols. Examination of the BAP revealed no relationship between the difference and the mean in any test, however the MDC values for the WBB did exceed those of the FP in three of the four tests. These findings suggest that the WBB is a valid tool for assessing standing balance. Given that the WBB is portable, widely available and a fraction of the cost of a FP, it could provide the average clinician with a standing balance assessment tool suitable for the clinical setting.

1. Introduction

Impaired standing balance associated with many health conditions manifests itself in reduced functional ability [22]. Assessment of standing balance has been shown to provide important information in a variety of situations, ranging from prediction of falls in the elderly [14] through to examining technique during surgery [19]. Consequently, a number of assessment protocols have been devised [3,22]. Laboratory-based assessment using measures of center of pressure (COP) recorded from a force platform (FP) – considered the gold standard measure of balance [12] – have identified important outcome measures which are too subtle to detect using a subjective scale [16]. Using a FP to assess standing balance provides useful information, however they are often expensive, difficult to setup and cumbersome to transport and therefore this form of balance assessment is often not feasible in a clinical setting. Consequently, subjective assessment tools which do not require specialized equipment, such as the Berg Balance Scale, are commonly used and have also been shown to provide valuable information [3,5]. While these protocols are more clinically applicable, they suffer from limitations including ceiling effects and a limited precision to detect small changes in performance [5,11]. In addition, previous research indicates that the relationship between scores on subjective tests and measures of COP displacement is only moderate [10], and that a combination of the two protocols may provide important information which cannot be obtained by either subjective or quantitative assessment alone [1,5].

This highlights the need to create a portable, inexpensive balance assessment system that has widespread availability. The Wii Balance Board (WBB) (Nintendo, Kyoto, Japan), part of the popular video game WiiFit, satisfies all of these criteria. The WBB possesses similar characteristics to a laboratory-grade FP in that it contains four transducers which are used to assess force distribution and the resultant movements in COP. Originally designed as a video game controller, the WBB is predominantly used in combination with a video game console and its associated software. Given the capacity for providing instant feedback and the potential for enhanced motivation levels [17], this system has already been integrated into the rehabilitation programs of neurological patients with balance defects [6]. In addition to its use as a biofeedback and gaming tool, the WBB could potentially be

* Corresponding author. Tel.: +61 4 3173 7609; fax: +61 3 8344 4188.
E-mail address: raclark@unimelb.edu.au (R.A. Clark).

0966-6362/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.gaitpost.2009.11.012

used by clinicians to collect and analyse laboratory-grade balance data using the techniques and outcome measures most specific to the patient population of interest. The WBB is a small fraction of the cost of a laboratory-grade FP, is mass-marketed and is portable, and consequently it has the potential to become a key component of a clinician's testing battery if it can be shown to produce reliable and valid results. Therefore, the aim of our study was to compare COP data collected on a WBB with that of a laboratory-grade FP during a variety of balance tests.

2. Methods

2.1. Participants

Thirty young (age = 23.7 ± 5.6 years), injury free individuals (gender = 10 male, 20 female; height = 1.68 ± 0.09 m; body mass = 63.80 ± 15.20 kg) were tested on two occasions, completed within 2 weeks and at least 24 h apart. No participant reported a major back or lower limb pathology, use of medication, or a history of neurologic disease that may influence standing balance. The study was approved by the institution's Human Research Ethics Committee and all participants provided informed consent.

2.2. Procedures

On each of the two test occasions, participants performed a series of four different standing balance tasks on a laboratory-grade FP (AMTI Model OR6-5, Watertown, MA, U.S.A.), which measured 50 cm × 46 cm in size and was mounted flush with the laboratory floor, and a WBB, which has a useable surface of 45 cm × 26.5 cm and was located on the laboratory floor beside the FP. The FP was calibrated in accordance with the manufacturer's recommendations. The WBB was interfaced with a laptop computer using custom-written software (Labview 8.5 National Instruments, Austin, TX, U.S.A.), and was calibrated by placing a variety of known loads at different positions on the WBB, a protocol discussed in detail in Appendix A.

Four standing balance tasks were chosen based on their varying difficulty and common use in previous literature [2,21]. These balance tests were: single limb standing (on the dominant limb) with eyes closed, single limb standing with eyes open, double limb standing with eyes closed and feet together and double limb standing with eyes open and feet a comfortable distance apart (measured and kept consistent for both testing sessions). The order of tasks and testing device was randomly assigned for each participant, but remained consistent between testing sessions. During each trial the participants were instructed to keep their hands placed on their hips and to remain as still as possible for the duration of the trial. Data were collected for 10 s during single limb trials and for 30 s during double limb trials. A total of three successful trials (maximum of three unsuccessful attempts) were conducted for each task and device with 15 s of rest between trials and a minimum of 60 s between-device or task.

2.3. Data analysis

A preliminary inspection of the time-frequency characteristics of the data revealed high frequency noise contamination of both the FP and WBB signal. Consequently, data for both devices were sampled at 40 Hz and filtered using an eighth order Butterworth filter with a lowpass cut-off frequency of 12 Hz. This cut-off frequency is slightly higher than the 10 Hz recommended by Salavati et al. [18] due to the higher frequency content of the signal observed in the single limb trials performed in the present study, and provided the optimal compromise between noise attenuation and maintenance of signal power.

The outcome measure used in this study was total COP path length. Given that the trials were for a fixed time interval, these COP path length results in this study are analogous to a measure of average COP velocity (path length per time interval tested). Therefore total COP path length was chosen as the primary outcome measure because it is known to be a reliable and valid measure of standing balance [18]. Based on the median of the three repetitions, which was performed to remove the potential for outlying data to influence the results, a single value for each of the outcome measures was obtained for each task (single limb eyes open or closed, double limb eyes open or closed), device (FP or WBB), and test occasion (Day 1 or Day 2).

2.4. Statistical analysis

The first step was to examine agreement between the two devices by creating a Bland–Altman plot for the COP path lengths of each testing protocol. Specifically, this was performed by plotting the difference in COP measures between the two methods against the mean results [4]. A two-way, random-effects, single measure (median of the three trials) intraclass correlation coefficients ($ICC_{(2,1)}$) model was used to assess reliability. In conjunction with the ICC values, standard error of measurement (SEM) and minimum detectable change (MDC) values were calculated to assess the concurrent validity between the WBB and the FP as well as the within-device test–retest reliability and measurement error over the two testing sessions for all outcome measures [13,20]. Point estimates of the ICCs were interpreted as follows: excellent (0.75–1), modest (0.4–0.74), or poor (0–0.39) [9]. All statistical analyses were conducted using the Statistical Package for the Social Sciences (SPSS Inc. Version 15.0, Chicago, IL, U.S.A.). The MDC, which is otherwise known as the reliable change index score, was calculated using the equations reported previously by Jacobson and Truax [13]. It is expressed as the percentage test–retest change in COP path length required to find a significant difference at an alpha level of 0.05 based on the Day 1 mean value.

3. Results

The results for the (1) single limb with eyes open, (2) single limb with eyes closed, (3) double limb with eyes open and feet apart and (4) double limb with eyes closed and feet together tests are provided in Table 1. Two participants were unable to successfully complete three trials of single limb balance with eyes closed on Day 1, and therefore did not undergo testing for this task on Day 2. Consequently, test–retest statistical analysis was performed on the data for 30 participants in all tests except the single limb eyes closed trial, which was limited to the data of 28 participants. The Bland–Altman plots for the COP path length balance test are provided in Fig. 1. While no obvious relationship between the difference and the mean was observed for any of the balance tests, three of the four balance trials showed a bias towards higher mean COP path length values in the tests performed on the WBB.

In general, both devices showed excellent COP path length test–retest reliability (Table 1), with only the performance on the WBB during double limb standing with eyes open and feet apart (ICC = 0.66) failing to reach an ICC value of 0.75. Additionally, concurrent validity was shown to be consistently excellent across balance tasks and testing sessions (ICC = 0.77–0.89). The SEM (FP range = 5.3–13.2%, WBB range = 8.7–13.1%) and MDC (FP range = 14.5–34.7%, WBB range = 24.5–29.4%) values were reasonably high for both devices, with the WBB MDC values higher than the FP values in three of the four trials.

4. Discussion

The ability for a clinician to objectively assess standing balance using a portable, inexpensive and valid system could provide numerous benefits in a wide range of patient populations. In this respect, we have shown that the WBB exhibits excellent test–retest reliability for COP path length assessment and possesses concurrent validity with a laboratory-grade FP. This provides the impetus for further research into the clinical applications of this video game equipment and the creation of software to facilitate uptake of WBB assessment of balance in the clinical setting.

This study did not attempt to examine the reliability of different balance testing protocols in specific patient populations, which has been the focus of previous research [2,18], as our between-device results suggest that this should be similar whether a FP or the WBB is incorporated into the testing battery. With regards to concurrent validity, examination of the Bland–Altman plots reveals a small difference in COP path length values between the two devices. These disparate values were consistent between days, and were possibly due to device specific factors such as the precision and sensitivity of the sensors or the difference in surface texture and hardness [7]. The consistency of the within-device results indicates that this would not have an effect on comparisons performed on patients using the same device. The results for the MDC revealed that reasonably large variations in COP path length are required in a test–retest study to detect a significant change in performance. These values were commonly in excess of 20%, and would therefore imply that low magnitude changes in standing balance performance would not be statistically detected by either the FP or WBB systems. However, in patient populations such as those with severe movement disorders or those recovering from lower limb

Table 1
Reliability and concurrent validity analysis of COP path length (cm) measures during each of the four standing balance trials.

	FP	WBB	Mean diff (95%CI)	ICC (95%CI)
Single limb, eyes open				
Day 1	42.2 (10.6)	48.3 (13.7)	−6.1 (−9.2, −3.0)	0.81 (0.39, 0.92)
Day 2	40.3 (8.8)	47.6 (12.9)	−7.2 (−9.9, −4.6)	0.80 (0.02, 0.93)
Mean Diff (95% CI)	1.9 (−.48, 4.2)	0.7 (−2.8, 4.3)		
ICC (95%CI)	0.89 (0.76, 0.95)	0.86 (0.70, 0.93)		
SEM	3.5	5.1		
MDC (%)	23.0	29.4		
Single limb, eyes closed				
Day 1	80.2 (21.9)	87.0 (20.4)	−6.6 (−11.5,−1.8)	0.88 (0.69, 0.95)
Day 2	75.8 (18.5)	90.9 (26.1)	−15.2 (−21.3, −9.0)	0.77 (0.03, 0.92)
Mean Diff (95% CI)	4.4 (−2.2, 11.0)	−4.0 (−11.2, 3.3)		
ICC (95%CI)	0.79 (0.54, 0.90)	0.81 (0.59, 0.91)		
SEM	10.0	11.4		
MDC (%)	34.7	28.3		
Double limb, eyes open				
Day 1	41.3 (5.8)	38.7 (6.7)	2.6 (0.8, 4.5)	0.77 (0.46, 0.90)
Day 2	41.8 (4.8)	38.6 (6.8)	3.2 (1.2, 5.2)	0.78 (0.54, 0.90)
Mean Diff (95% CI)	0.3 (−1.7, 1.1)	0.1 (−1.9, 2.1)		
ICC (95%CI)	0.86 (0.71, 0.93)	0.66 (0.20, 0.85)		
SEM	2.2	4.0		
MDC (%)	14.5	27.9		
Double limb, eyes closed				
Day 1	68.2 (16.2)	74.3 (21.9)	−6.1 (−10.1, −2.1)	0.89 (0.71, 0.95)
Day 2	69.4 (14.8)	75.7 (21.6)	−6.3 (−10.4, −2.3)	0.88 (0.67, 0.95)
Mean Diff (95% CI)	0.7 (−3.5, 2.1)	1.2 (−6.0, 3.6)		
ICC (95%CI)	0.94 (0.87, 0.97)	0.91 (0.80, 0.96)		
SEM	4.0	6.6		
MDC (%)	16.1	24.5		

FP: force plate; WBB: Wii Balance Board; COP: center of pressure; AP: anteroposterior; ML: mediolateral; SD: standard deviation; CI: confidence interval; ICC: intraclass correlation coefficient; Diff: difference; SEM: standard error of the measurement; MDC: minimum detectable change, expressed as a percentage of the Day 1 mean value.

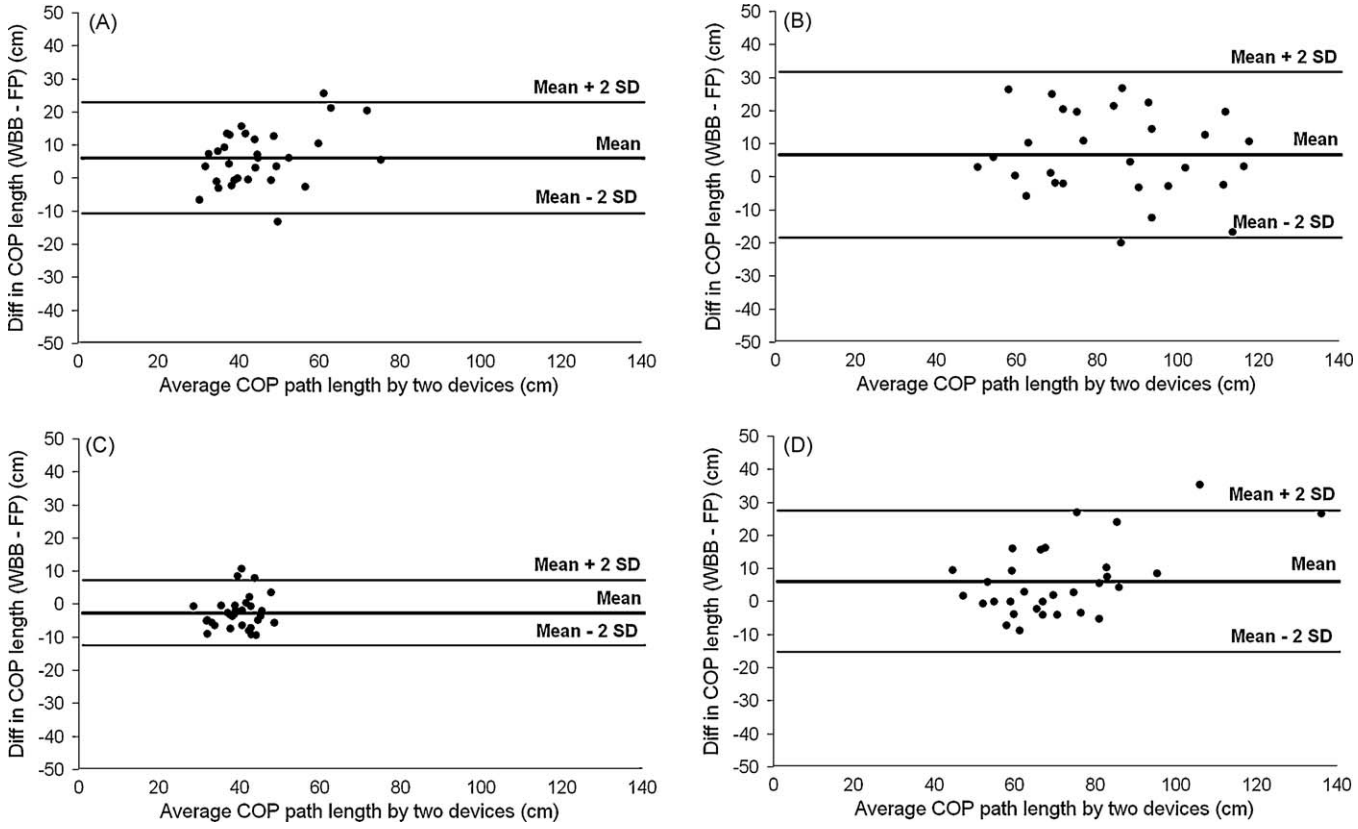

Fig. 1. Bland–Altman plots representing comparisons between the laboratory-grade force platform (FP) and the Wii Balance Board (WBB) for the four testing conditions: (A) single limb, eyes open; (B) single limb, eyes closed; (C) double limb, eyes open; (D) double limb, eyes closed. The mean line represents the mean difference between the devices, with the upper and lower lines representing the limits of agreement (2SD).

surgery this form of balance assessment may be feasible. Additionally, these high MDC values appear to be consistent with previous studies which have reported similar ICC and standard deviation values, the two test-specific components of the MDC equation, during standing balance trials [8,18]. In regard to the specific systems, the WBB possessed higher MDC values in three of the four trials, most notably in the two double limb tasks. This should be considered if the WBB is to be implemented for balance assessment, however it does not disqualify it from usage in further research.

Although the results of the present study are promising, the creation of the WBB as primarily a video gaming controller, and the subsequent requirement for it to be inexpensive while remaining profitable, does result in a number of limitations. This prevents the WBB from being a direct replacement for a FP in activities that require rapid, high force movements such as jumping and running. A limitation of more importance in the assessment of balance is the inability to assess force in the horizontal axes, which are important components of the standard COP equations. This limitation would have a significant impact on the results of balance tests where sizeable momentum along these axes is produced, with a previous study suggesting that the *X* and *Y* axis coordinates derived from vertical plane forces in the absence of correction for horizontal plane force should be referred to as center of balance instead of COP [15]. However, as explained in more detail in the calibration notes (refer to Appendix A), the force levels in these two axes only rarely exceeded ±5 N. This was expected due to the low movement velocity inherent in the balance tasks, particularly the double leg trials. Therefore, although the lack of correction for *X* and *Y* axes force is an inherent limitation when deriving COP values from the WBB, its excellent concurrent validity when compared with the gold standard suggests that it is a satisfactory device for assessing standing balance.

With regard to the custom-written software that was created to interface the WBB with the laptop, we acknowledge that the majority of clinicians are not equipped with the necessary technical skills to create these programs. However, because this study examined the data acquired and analysed through a computer using common calibration techniques, and found it to be valid, it is conceivable that the results achieved using the Nintendo Wii based software programs could also provide valid results. While the computer based data collection and analysis technique allows for the creation of custom programs more specifically designed for each patient population, the balance tests contained in the standard Nintendo Wii programs (such as WiiFit) may provide sufficient data.

In conclusion, the WBB provides comparable data to a FP when assessing COP path length during standing balance trials. Consequently, the WBB has the potential to 'bridge the gap' between laboratory testing and clinical assessment of standing balance. Instead of replacing subjective-based balance protocols, the WBB could provide practitioners from a range of medical specialties and disciplines with supplementary balance information that is not discernible using visual assessment alone. Further research should examine the diagnostic and prognostic accuracy of balance measures collected using the WBB in patient populations. If the WBB data is found to provide important information, it could allow for more sensitive monitoring of change in balance over time and a better evaluation of the effectiveness of treatment for an individual, thereby improving evidence-based clinical practice.

Source of funding

No external funding was received specifically for this project. No support was received from Nintendo. The first author's research fellow position is funded by ASICS Oceania.

Conflict of interest

There is no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gaitpost.2009.11.012.

References

[1] Adkin AL, Frank JS, Jog MS. Fear of falling and postural control in Parkinson's disease. Mov Disord 2003;18(5):496–502.
[2] Bauer C, Gröger I, Rupprecht R, Gabmann KG. Intrasession reliability of force platform parameters in community-dwelling older adults. Arch Phys Med Rehabil 2008;89(10):1977–82.
[3] Berg K, Norman KE. Functional assessment of balance and gait. Clin Geriatr Med 1996;12(4):705–23.
[4] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307–10.
[5] Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. Phys Ther 2008;88(5):559–66.
[6] Deutsch JE, Borbely M, Filler J, Huhn K, Guarrera-Bowlby P. Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy. Phys Ther 2008;88(10):1196–207.
[7] Dickinson JI, Shroyer JL, Elias JW. The influence of commercial-grade carpet on postural sway and balance strategy among older adults. Gerontologist 2002;42(4):552–9.
[8] Doyle TL, Newton RU, Burnett AF. Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people. Arch Phys Med Rehabil 2005;86(10):2034–40.
[9] Fleiss J. The Design and Analysis of Clinical Experiments. New York, USA: John Wiley & Sons; 1986.
[10] Frykberg GE, Lindmark B, Lanshammar H, Borg J. Correlation between clinical assessment and force plate measurement of postural control after stroke. J Rehab Med 2007;39(6):448–53.
[11] Gustavsen M, Aamodt G, Mengshoel AM. Measuring balance in sub-acute stroke rehabilitation. Adv Physiother 2006;8(1):15–22.
[12] Haas BM, Burden AM. Validity of weight distribution and sway measurements of the Balance Performance Monitor. Physiother Res Int 2000;5(1):19–32.
[13] Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59(1):12–9.
[14] Jarnlo GB. Functional balance tests related to falls among elderly people living in the community. Eur J Geriatr 2003;5(1):7–14.
[15] Nichols DS, Glenn TM, Hutchinson KJ. Changes in the mean center of balance during balance testing in young adults. Phys Ther 1995;75(8):699–706.
[16] Piirtola M, Era P. Force platform measurements as predictors of falls among older people—a review. Gerontology 2006;52(1):1–16.
[17] Ramchandani A, Carroll K, Buenaventura R, Douglas J, Liu J. Wii-habilitation increases participation in therapy. In: Proceedings of 2008 Virtual Rehabilitation, IWVR; 2008. p. 69.
[18] Salavati M, Hadian MR, Mazaheri M, Negahban H, Ebrahimi I, Talebian S, et al. Test–retest reliability of center of pressure measures of postural stability during quiet standing in a group with musculoskeletal disorders consisting of low back pain, anterior cruciate ligament injury and functional ankle instability. Gait Posture 2009;29(3):460–4.
[19] Savoie S, Tanguay S, Centomo H, Beauchamp G, Anidjar M, Prince F. Postural control during laparoscopic surgical tasks. Am J Surg 2007;193(4):498–501.
[20] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420–8.
[21] Springer BA, Marin R, Cyhan T, Roberts H, Gill NW. Normative values for the unipedal stance test with eyes open and closed. J Geriatr Phys Ther 2007;30(1):8–15.
[22] Winter DA, Patla AE, Frank JS. Assessment of balance control in humans. Med Prog Technol 1990;16(1–2):31–51.