

Smartphone-based gait and balance accelerometry is sensitive to age and correlates with clinical and kinematic data

Sharon Olsen^{a,*}, Usman Rashid^{a,b}, Celia Allerby^a, Eliza Brown^a, Michaela Leyser^a, Gabrielle McDonnell^a, Gemma Alder^a, David Barbado^{c,d}, Nusratnaaz Shaikh^a, Sue Lord^a, Imran Khan Niazi^{a,b,e}, Denise Taylor^a

^a Rehabilitation Innovation Centre, Health & Rehabilitation Research Institute, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand

^b Centre for Chiropractic Research, New Zealand College of Chiropractic, PO Box 113–044, Newmarket, Auckland 1149, New Zealand

^c Department of Sport Science, Sports Research Centre, Miguel Hernandez University of Elche, Avda. de la Universidad s/n, Elche 03202, Spain

^d Institute for Health and Biomedical Research (ISABIAL Foundation), Avda. Pintor Baeza, 12 HGUA, Alicante 03550, Spain

^e Centre for Sensory-Motor Interaction (SMI), Department of Health Science and Technology, Aalborg University, 9220 Aalborg, Denmark

ARTICLE INFO

Keywords:

Balance
Gait
Smartphone
Accelerometer
Validity
Gait&Balance App

ABSTRACT

Background: The Gait&Balance (G&B) App has produced valid and reliable measures of gait and balance in young healthy adults but has not been tested in older adults.

Research question: In healthy middle-to-older aged adults, are G&B App measurements sensitive to age, valid against clinical and kinematic measures, and reliable?

Method: Healthy participants ($n = 34$, 14 male, 42–94 years) completed the G&B App protocol three times within a single session. 3D kinematics were collected concurrently. Clinical balance measures were collected (Modified Clinical Test of Sensory Interaction in Balance (mCTSIB), Mini Balance Evaluation Systems Test (MBT), and Functional Gait Assessment (FGA)). Sensitivity to age was assessed with Pearson's correlations. Validity tests included Pearson's correlations and Bland-Altman limits of agreement. Reliability tests included intra-class correlation coefficients and standard error of the measure.

Results: During quiet stance on a compliant surface, the G&B App was sensitive to age-related differences not detectable with the mCTSIB. During walking tasks, there was adequate convergent validity between the MBT and G&B App measures of step length, and between the FGA and G&B App measures of walking speed, step length, and periodicity. The G&B App had moderate-to-excellent validity against 3D kinematics for postural stability during quiet stance (r 0.98 [0.98, 0.99]), step time (r 0.97 [0.96, 0.98]), walking speed (r 0.79 [0.7, 0.86]), and step length (r 0.73 [0.61, 0.81]). Test-retest reliability was moderate-to-excellent for G&B App measures of postural stability, walking speed, periodicity, step length, and step time. G&B App measures of step length asymmetry, step length variability, step time asymmetry, and step time variability had poor validity and reliability.

Significance: The G&B App was sensitive to age-related differences in balance not detectable with clinical measurement. It provides valid and reliable measures of postural stability, step length, step time, and periodicity, which are not currently available in standard practice.

1. Introduction

Maintaining balance requires a highly responsive central nervous system which rapidly integrates information from the visual, vestibular, and/or somatosensory systems, and transforms this into motor actions that maintain the body's orientation [1]. Balance impairments

commonly arise with older age and with clinical conditions such as vestibular disorders, stroke, and Parkinson's Disease. Balance impairments affect the ability to maintain an upright position during quiet stance and/or during dynamic gait tasks, and can severely limit mobility and increase the risk of falling [2]. Physiotherapists assess balance using standard clinical measures, however these tests are prone to ceiling

* Correspondence to: Rehabilitation Innovation Centre, Health and Rehabilitation Research Institute, Private Bag 92006, Auckland 1142, New Zealand.

E-mail address: sharon.olsen@aut.ac.nz (S. Olsen).

<https://doi.org/10.1016/j.gaitpost.2022.11.014>

Received 7 August 2022; Received in revised form 16 November 2022; Accepted 24 November 2022

Available online 28 November 2022

0966-6362/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

effects [3,4]; that is, they are not sensitive to detecting balance dysfunction associated with ageing or in those with mild impairments [5]. Many standard clinical tests are vulnerable to tester bias as they require physiotherapists to use their judgement to classify balance performance under broad categories [6] and they lack responsiveness to small changes over time [4]. Gold-standard technology for measuring balance such as force plates, 3D motion capture, computerised dynamic posturography, and instrumented walkways, can provide more sensitive and responsive measures of balance [7–10]. However, these are expensive and rarely accessible in standard clinical practice.

A potential solution is the development of mobile applications (apps) that measure balance using the inertial data from accelerometers and gyroscopes embedded within standard smartphones [11–14]. Our interdisciplinary team has developed a smartphone app for gait and balance assessment called the Gait&Balance App (G&B App) [15]. The feature which distinguishes this app from others, such as Sensor Kinetics Pro [12], APP-Coo-Balance-Test [13], and those developed for ankle instability [11] and stroke [14], is the inclusion of tasks that provide progressive balance challenges in both quiet stance and dynamic gait. The G&B App includes a well-defined assessment protocol that can be executed within limited clinic space. To date, the G&B App has been validated against gold-standard measures of balance and gait in young, healthy adults [15], but has not been tested in older adults or compared with clinical balance measures. As balance declines with age, evaluating the sensitivity of the G&B App to age-related differences in balance is important for establishing its validity as a measurement tool [16]. Therefore, this study investigated the validity of the G&B App through determining its sensitivity to age, correlation with clinical balance measures, and correlation with gold-standard 3D motion capture, in healthy middle-to-older aged adults. Test-retest reliability was also investigated.

2. Method

2.1. Study design

In this cross-sectional study, participants attended a single session where their balance and gait were assessed using the G&B App, 3D motion capture, and clinical measures.

2.2. Participants

Participants were healthy community-dwelling adults over 40 years of age. To ensure participants were healthy, the following exclusion criteria were applied: falling in the past year, feeling unsteady when standing or walking or worried about falling, walking aids, limited ability to balance and walk around the home, diagnosed vestibular disorders, neurological conditions that impaired movement, cognitive impairment, history of major lower limb orthopaedic surgery, and acute illness. All participants provided written informed consent.

2.3. Sample size

The sample size calculation accounted for both validity and test-retest reliability and required 27 participants to estimate ICC scores ≥ 0.8 (based on 3 observations per participant, power 0.8, α 0.05, and lower bound 95 % CI > 0.6) but was increased to 34 participants to account for dropouts and ensure representation of a range of ages. The study received ethical approval from the Auckland University of Technology Ethics Committee (21/51).

2.4. Procedures

The study flow is shown in Fig. 1. Four retro-reflective 18 mm markers were secured to the participant's bare feet at the posterior calcaneus and head of the fifth metatarsal bilaterally. A smartphone (iPhone SE, Apple Inc, Cupertino, CA, USA) was secured to the lower lumbar spine (centre of phone approximately L5/S1) using an elasticated core stability belt (Whiteley Allcare, Auckland, New Zealand) that had been customised by attaching a phone casing (Sports armband, Tech.Inc, Auckland, New Zealand) (see Fig. 1). A retro-reflective marker was attached to the centre of the phone screen. Participants completed six standardised balance tasks [15] while accelerometry data was collected with the G&B App: standing, firm surface, eyes open (1) and eyes closed (2); standing, compliant surface, eyes open (3) and eyes closed (4); walking looking straight ahead (5); walking with head turning (6). For standing tasks, participants stood for 30 s with feet hip-width apart and arms by sides [17], and tasks were terminated if participants took a step, opened their eyes, or required physical assistance. The compliant surface was medium-density foam (Airex Balance Pad, Airex AG, Sins, Switzerland). All six tasks were completed three times with 5-minute rests between each set of six tasks. The first four

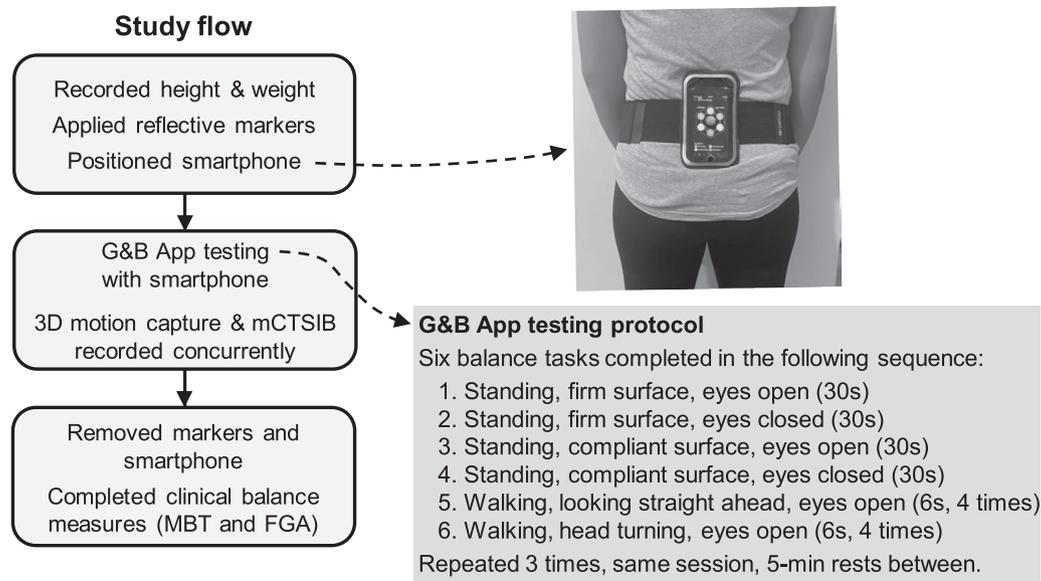


Fig. 1. Study flow and G&B App testing protocol.

G&B App tasks also make up the Modified Clinical Test of Sensory Interaction in Balance (mCTSIB) [18], a clinical test which quantifies postural stability by timing how long an individual can maintain stability in quiet stance under different sensory conditions for up to 30 s. To complete the mCTSIB, each participant's ability to complete the 30-s standing tasks was recorded with a stop watch and averaged across the three trials. During the six tasks, kinematic data were collected using 12-camera 3D motion capture (Qualisys, Gothenburg). Then the markers and phone were removed and participants donned walking shoes to perform two further clinical balance tests, the Mini Balance Evaluation Systems Test (MBT) [19] and the Functional Gait Assessment (FGA) [20], in a carpeted corridor. The MBT tests balance on a 3-point scale during 14 standing and walking tasks; there are subscores for anticipatory balance, reactive balance, sensory orientation, and balance during gait, and a total score out of 28 [19]. The FGA tests balance on a 4-point scale during 10 dynamic gait tasks and is scored out of 30 [20]. For all tests, stand-by supervision was provided.

2.5. Data processing

The processing of accelerometry data from the G&B App and 3D kinematic data has been described previously [15]. From the standing tasks, outcomes of postural stability (PS), anteroposterior PS (PS_{AP}) and mediolateral PS (PS_{ML}) were determined. The postural stability measures were obtained as the negative natural log of the mean absolute acceleration referenced to the laboratory co-ordinates; with the negative natural log, higher accelerations (higher postural sway) result in lower postural stability scores, and vice versa. These outcomes were also obtained from the iPhone retro-reflective marker by double numerical differentiation of displacement. From the walking tasks, the following spatiotemporal parameters were determined: walking speed, periodicity index (a measure between zero and 100 % that encompasses step symmetry between the right and left step within a stride and the consistency across strides, where perfect symmetry = 100 % [21]), mean step length, mean step time, step length variability, step time variability, step length asymmetry, and step time asymmetry. Step length and step time were calculated from the left and right separately and averaged together for mean step length and mean step time. The difference between the two limbs was represented with step length asymmetry and step time asymmetry. The difference across strides was represented with step length variability and step time variability which were calculated from the mean of the SD of step length and step time from the left and right. The same spatiotemporal parameters were obtained from the four retro-reflective foot markers using foot velocity algorithm to identify heel-strike and toe-off events.

2.6. Data analysis

The sensitivity of G&B App data and clinical data to age was evaluated with partial Pearson's product-moment correlation (*Pearson's r_p*) which accounted for repeated measures, height, and BMI. The analysis of G&B App measures against clinical measures was carried out on clinical measures that correlated with age (with lower bound of 95 % CI of *r_p* >0.3), to ensure the clinical outcomes were valid measures of age-related changes in our healthy sample. These clinical measures were evaluated against the G&B App outcomes using Pearson's product-moment correlation (*Pearson's r*); data from trial 3 was used to minimise confounding caused by potential repetition effects. Interpretation was based on the lower bound of the 95 % CI of *r* as follows: excellent (0.6–1), adequate (0.3–0.599), and poor (< 0.3) [22].

Agreement between the G&B App and 3D kinematic data was evaluated for consistency, using *Pearson's r*, and for absolute agreement, using Bland-Altman limits of agreement percentage (LoA %). Interpretation was based on the lower bound of the 95 % CI of *Pearson's r* and LoA % as follows: excellent (>0.90, 0.0–4.9 %), good (0.75–0.89, 5.0–9.9 %), moderate (0.50–0.74, 10.0–49.9 %) and poor (<0.50,

>50.0 %) [23,24]. For the standing tasks, all three trials were combined in one model as the postural stability scores obtained from kinematic data were affected by high noise caused by double numerical differentiation of the displacement time series. The presence of between-task differences in the same model expanded the outcome variance. Periodicity index was not included in the analysis between the G&B App and kinematic data as it is an acceleration-based measure and double differentiation of the kinematic data for computing acceleration results in high noise.

Test-retest reliability across the three sets of G&B App outcomes was assessed with a 2-way random effects model to estimate intra-class correlation coefficients (ICC) for absolute agreement between single measures [25]. From the ICC model, the standard deviation of the residuals was taken as the standard error of the measure (SEM). SEM was also expressed as a percentage of the trial 3 mean (SEM %). Interpretation was based on the lower bound of the 95 % CI of ICCs as follows: excellent (0.90–1), high (0.7–0.89), moderate (0.50–0.69), and poor (0–0.49) [26]. Statistical analysis utilised R software [27].

3. Results

3.1. Participants

A total of 34 participants were enrolled in and completed the study (42–94 years of age, 20 females, median BMI 25.3 (interquartile range 5.9)). See Table 1 for clinical balance measures.

3.2. Are Gait&Balance App measures sensitive to age-related differences in balance?

The most sensitive G&B App measures to age were the periodicity index when walking looking straight ahead (*r_p* −0.62 [−0.79, −0.35]) and PS_{AP} on a compliant surface with eyes open (*r_p* −0.64 [−0.81, −0.38]), which both had adequate correlations (see Fig. 2A). For the clinical measures, there were adequate correlations between age and the MBT total (*r_p* −0.66 [−0.82, −0.42]), MBT dynamic gait (*r_p* −0.62 [−0.79, −0.35]), and FGA (*r_p* −0.65 [−0.81, −0.41]). Other clinical measures showed correlations with age but had the lower bound of the 95 % CIs in the poor range: MBT anticipatory (*r_p* −0.57 [−0.76, −0.29]), MBT reactive (*r_p* −0.54 [−0.74, −0.25]), MBT sensory (*r_p* −0.25 [−0.55, 0.09]) (see full results in supplementary file). There was a strong ceiling effect in the mCTSIB, the clinical measure of postural stability during quiet stance (see Fig. 2B).

3.3. Do Gait&Balance App measures correlate with clinical measures?

The MBT total, MBT dynamic gait, and FGA correlated with age, and

Table 1
Participant characteristics.

Age bands	<i>n</i>	Male: Female	mCTSIB (score/120)	MBT (score/28)	FGA (score/30)
40–49	2	1:1	120 (120–120)	26.5 (26–27)	30 (30–30)
50–59	6	3:3	120 (120–120)	25 (24–27)	30 (25–30)
60–69	9	4:5	120 (120–120)	23 (18–27)	29 (24–30)
70–79	10	5:5	120 (120–120)	22 (14–25)	27 (21–29)
80–89	5	1:4	120 (116–120)	18 (15–21)	24 (19–27)
90 +	2	0:2	113.5 (107–120)	10 (8–12)	9.5 (8–11)

Median (range) provided. Higher scores indicate more stable balance. G&B App measures of postural stability, walking speed, and periodicity index, by age band, can be found in the supplementary file.

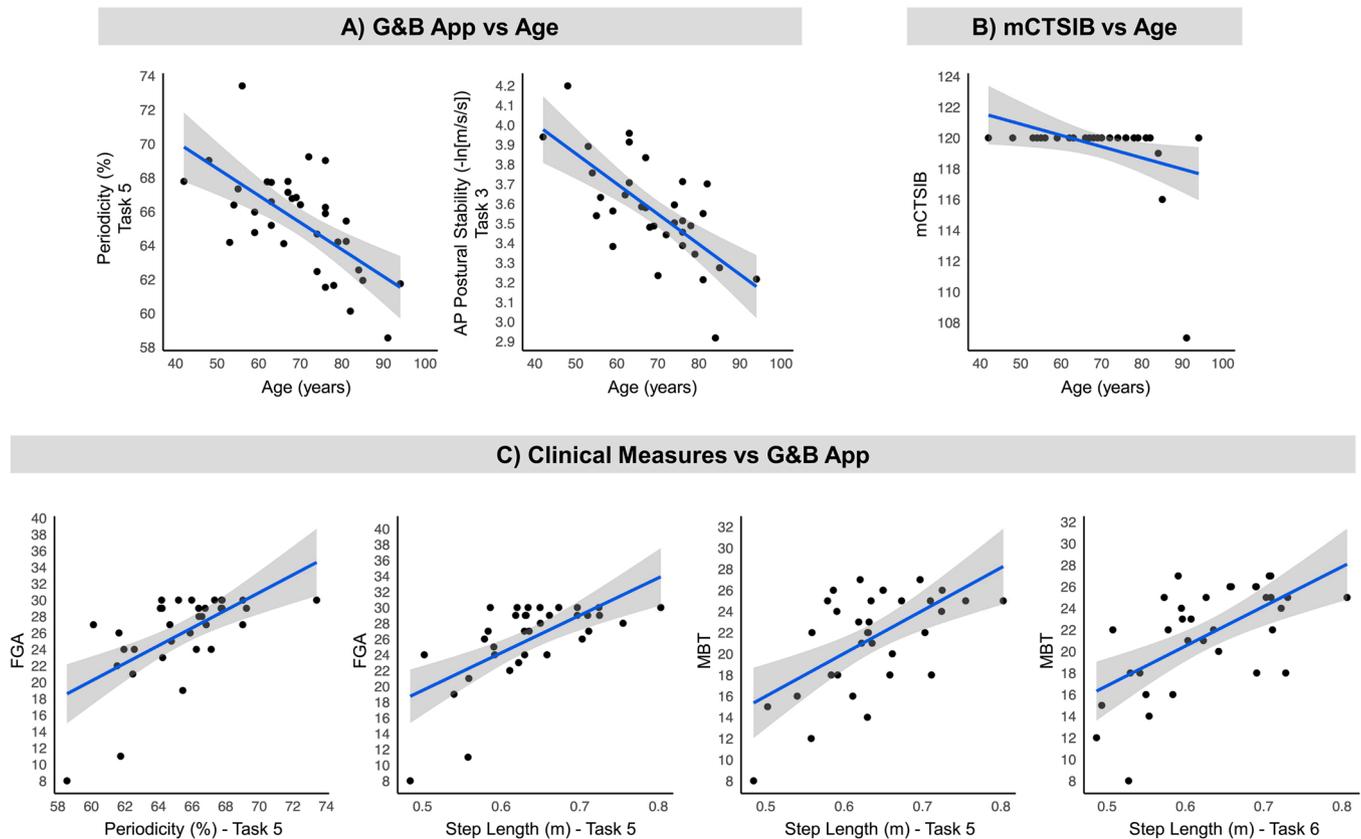


Fig. 2. Plots demonstrating correlations between: A) G&B App parameters (periodicity index during walking looking straight ahead and AP postural stability on a compliant surface with eyes open) and age, B) mCTSIB (the clinical measure of postural stability during quiet stance) and age, and C) clinical measures (FGA and MBT) and G&B App parameters.

therefore were deemed valid clinical measures in our healthy sample; these measures were used to determine the validity between the G&B App and clinical measures. There were adequate correlations between the MBT and FGA and several G&B App measures (see Table 2 and Fig. 2C).

3.4. Do Gait&Balance App measures correlate with gold-standard kinematic data?

For balance tasks in quiet stance, there was excellent validity between G&B App measures of postural stability and gold-standard kinematic data. For walking tasks, there was moderate-to-excellent validity for G&B App measures of walking speed, step length and step time, but step length variability, step length asymmetry, step time variability and step time asymmetry had poor validity. See Table 3. (See supplementary file for Bland-Altman plots for all comparisons).

3.5. Are Gait&Balance App measures reliable?

G&B App test-retest reliability is shown in Table 4. For balance tasks in quiet stance, most postural stability measures demonstrated moderate reliability. For walking tasks, G&B App parameters of walking speed, periodicity index, step length and step time, had moderate-to-excellent reliability. However, reliability for G&B App measures of step length/step time variability and asymmetry was poor, except for Task 5 step length asymmetry which was moderately reliable. The reliability of the equivalent gait parameter collected with 3D motion capture shows comparable reliability of the two systems (Table 4).

4. Discussion

This study demonstrates the validity and reliability of several gait and balance parameters derived from the G&B App in healthy middle-to-older aged adults.

4.1. Are Gait&Balance App measures sensitive to age-related differences in balance?

The G&B App measure of PS_{AP} on a compliant surface with eyes open was correlated with age ($r_p -0.64 [-0.81, -0.38]$) (Fig. 2A). This study enrolled a healthy sample, and the clinical measure of postural stability (mCTSIB) had a strong ceiling effect (Fig. 2B); that is, almost all participants scored maximally during tests on firm and compliant surfaces with eyes open and closed. Interestingly, when participants performed those same standing tasks with the G&B App, the App could differentiate between younger and older participants; this is promising and suggests the G&B App can detect age-related balance changes that some clinical measures cannot detect.

The G&B App periodicity index measure correlated with age ($r_p -0.62 [-0.79, -0.35]$) (Fig. 2A). This measure encompasses gait variability, through determining the consistency of gait cycles over time, and gait symmetry, through determining the consistency of half gait cycles within whole gait cycles [21]. The correlation between the periodicity index and age aligns with literature demonstrating the relationship between increased gait variability or gait asymmetry and balance impairment [2,28]. It is promising that the G&B App could detect age-related changes in this measure, and this supports the App's validity. These sensitivity findings also suggest the G&B App might be responsive to small changes that occur during rehabilitation; this should be investigated in a future study.

Table 2
Validity analysis between G&B App and clinical measures.

Task	G&B App parameter	Clinical measure	r [95 % CI]	Correlation	
Task 5 Walking looking straight ahead	Walking speed	MBT total	0.53 [0.23, 0.74]	Poor	
	Walking speed	MBT Dynamic Gait	0.42 [0.09, 0.66]	Poor	
	Walking speed	FGA	0.74 [0.54, 0.86]	Adequate	
	Periodicity index	MBT total	0.53 [0.23, 0.73]	Poor	
	Periodicity index	MBT Dynamic Gait	0.49 [0.19, 0.71]	Poor	
	Periodicity index	FGA	0.63 [0.37, 0.8]	Adequate	
	Step length	MBT total	0.59 [0.32, 0.78]	Adequate	
	Step length	MBT Dynamic Gait	0.51 [0.2, 0.72]	Poor	
	Task 6 Walking head turning	Walking speed	MBT total	0.48 [0.16, 0.7]	Poor
		Walking speed	MBT Dynamic Gait	0.37 [0.04, 0.63]	Poor
Walking speed		FGA	0.66 [0.41, 0.81]	Adequate	
Periodicity index		FGA	0.41 [0.08, 0.66]	Poor	
Step length		MBT total	0.62 [0.35, 0.79]	Adequate	
Step length		MBT Dynamic Gait	0.56 [0.28, 0.76]	Poor	
Step length		FGA	0.61 [0.34, 0.78]	Adequate	

Comparisons not shown were deemed to have no correlation (lower bound 95 % CI of r below zero).

4.2. Do Gait&Balance App measures correlate with clinical measures?

We anticipated a correlation between G&B App Tasks 1–4 and the mCTSIB as these assess similar balance domains [19,29]. However, the strong ceiling effect in the mCTSIB data meant this could not be correlated with G&B App measures. It is anticipated that if this study was repeated in a population with balance or sensory deficits, the mCTSIB would correlate with G&B App measures of postural stability. Two other studies in people with stroke [14] and ataxia [13] have shown correlations between other smartphone-based balance measures in quiet stance, and clinical measures (Berg Balance Scale, Scale for Assessment and Rating of Ataxia) with r values of 0.7–0.91 [13,14]. Thus, further research is needed to test the convergent validity of the G&B App during balance in quiet stance in clinical populations.

Several G&B App parameters from walking tasks were correlated with clinical measures. G&B App measures of walking speed (looking straight ahead and with head turning) correlated with the total FGA score, but not the total MBT score. This is not surprising as the FGA focuses on dynamic gait tasks [20], whereas the MBT covers a range of balance domains (i.e. postural stability, anticipatory adjustments, reactive responses, sensory orientation, and gait stability) [19]. G&B

Table 3
Validity analysis between G&B App data and kinematic data.

	G&B App parameter	r [95 % CI]	Consistency	LOA %	Agreement
Standing tasks 1 – 4	PS	0.98 [0.98, 0.99]	Excellent		
	PS _{ML}	0.96 [0.95, 0.97]	Excellent		
	PS _{AP}	0.97 [0.96, 0.97]	Excellent		
Task 5 Walking looking straight ahead	Walking speed	0.79 [0.7, 0.86]	Moderate	18.2	Moderate
	Mean step length	0.73 [0.61, 0.81]	Moderate	16.3	Moderate
	Mean step time	0.97 [0.96, 0.98]	Excellent	5.2	Good
	Step length variability	0.29 [0.09, 0.47]	Poor	147.1	Poor
	Step time variability	0.49 [0.31, 0.63]	Poor	156.5	Poor
	Step length asymmetry	0.14 [– 0.06, 0.34]	Poor	215.6	Poor
	Step time asymmetry	0.2 [– 0.01, 0.39]	Poor	282.6	Poor
	Task 6 Walking head turning	Walking speed	0.87 [0.8, 0.91]	Good	17.4
Mean step length		0.8 [0.71, 0.87]	Moderate	17.3	Moderate
Mean step time		0.98 [0.97, 0.99]	Excellent	4.8	Excellent
Step length variability		0.28 [0.06, 0.47]	Poor	137.8	Poor
Step time variability		0.36 [0.16, 0.54]	Poor	120.4	Poor
Step length asymmetry		0.06 [– 0.16, 0.28]	Poor	234.1	Poor
Step time asymmetry		0.21 [– 0.01, 0.41]	Poor	215.5	Poor

App measures of step length correlated with both the total MBT and FGA. This correlation is understandable given step length decreases with increasing age and is a strategy used to prevent falling [2,30]. Periodicity index correlated with the FGA which aligns with the relationship between gait variability and falls described above [2]. We anticipated there might also be correlations between G&B App measures of step length/step time variability/asymmetry and the FGA, but this was not the case. This may be because these G&B App measures were similar across our sample, as demonstrated by their lack of sensitivity to age. Further research should explore the validity of these measures in clinical populations where there is likely to be more marked differences between each side and each step.

4.3. Do Gait&Balance App measures correlate with gold-standard kinematic data?

During balance tasks in quiet stance (Tasks 1–4), G&B App measures of postural stability had excellent validity against gold-standard 3D motion kinematic data. This suggests that for balance measures during quiet stance, the G&B App could be a low-cost alternative to 3D motion

Table 4
Reliability of G&B App parameters.

	G&B App Parameter	Trials Mean (SD)			SEM	SEM %	ICC [95 % CI]	Reliability with G&B App	Reliability 3D system
		Trial 1	Trial 2	Trial 3					
Task 1 Firm Eyes Open	PS	3.49 (0.22)	3.43 (0.24)	3.45 (0.23)	0.1	3	0.82 [0.7, 0.9]	High	
	PS _{ML}	4.26 (0.29)	4.19 (0.31)	4.20 (0.30)	0.1	2	0.88 [0.79, 0.93]	High	
	PS _{AP}	4.11 (0.20)	4.03 (0.21)	4.08 (0.18)	0.1	3	0.71 [0.54, 0.84]	Moderate	
Task 2 Firm Eyes Closed	PS	3.29 (0.34)	3.32 (0.26)	3.37 (0.25)	0.1	4	0.75 [0.61, 0.85]	Moderate	
	PS _{ML}	4.08 (0.40)	4.10 (0.31)	4.16 (0.32)	0.2	4	0.79 [0.66, 0.88]	Moderate	
	PS _{AP}	3.86 (0.34)	3.86 (0.40)	3.93 (0.34)	0.2	4	0.65 [0.48, 0.79]	Poor	
Task 3 Compliant Eyes Open	PS	2.46 (0.30)	2.58 (0.34)	2.64 (0.33)	0.1	6	0.78 [0.6, 0.89]	Moderate	
	PS _{ML}	3.61 (0.32)	3.73 (0.33)	3.78 (0.32)	0.1	3	0.82 [0.62, 0.91]	Moderate	
	PS _{AP}	3.46 (0.28)	3.53 (0.26)	3.56 (0.26)	0.1	3	0.76 [0.61, 0.87]	Moderate	
Task 4 Compliant Eyes Closed	PS	1.86 (0.50)	1.97 (0.47)	2.10 (0.42)	0.2	9	0.8 [0.6, 0.9]	Moderate	
	PS _{ML}	3.15 (0.45)	3.28 (0.41)	3.36 (0.39)	0.2	5	0.78 [0.58, 0.89]	Moderate	
	PS _{AP}	2.87 (0.41)	2.95 (0.34)	3.05 (0.33)	0.2	5	0.78 [0.59, 0.89]	Moderate	
Task 5 Walking looking straight ahead	Walking speed (m/s)	1.15 (0.15)	1.22 (0.15)	1.23 (0.15)	0.05	4	0.82 [0.6, 0.91]	Moderate	Moderate
	Periodicity index (%)	64.01 (4.11)	64.91 (3.90)	65.42 (2.99)	2	3	0.74 [0.58, 0.85]	Moderate	
	Mean step length (m)	0.62 (0.07)	0.63 (0.07)	0.64 (0.07)	0.01	2	0.94 [0.84, 0.97]	High	High
	Mean step time (s)	0.54 (0.05)	0.52 (0.04)	0.52 (0.04)	0.02	3	0.76 [0.55, 0.88]	Moderate	Moderate
	Step length variability (%)	5.40 (1.55)	5.10 (1.75)	4.78 (1.64)	1	23	0.53 [0.33, 0.7]	Poor	Poor
	Step time variability (%)	5.79 (2.56)	4.70 (1.99)	4.70 (1.51)	2	34	0.38 [0.18, 0.59]	Poor	Poor
	Step length asymmetry (%)	4.14 (3.35)	4.23 (3.63)	3.77 (3.25)	2	45	0.75 [0.61, 0.86]	Moderate	Poor
	Step time asymmetry (%)	4.85 (3.73)	4.73 (4.05)	4.16 (3.72)	2	56	0.64 [0.46, 0.78]	Poor	Poor
Task 6 Walking head turning	Walking speed (m/s)	1.09 (0.18)	1.13 (0.19)	1.13 (0.19)	0.05	5	0.91 [0.83, 0.95]	High	High
	Periodicity index (%)	58.48 (5.99)	59.62 (5.08)	60.13 (6.02)	3	5	0.7 [0.54, 0.82]	Moderate	
	Mean step length (m)	0.63 (0.08)	0.63 (0.08)	0.63 (0.08)	0.02	2	0.96 [0.94, 0.98]	Excellent	High
	Mean step time (s)	0.59 (0.07)	0.57 (0.06)	0.56 (0.06)	0.02	4	0.86 [0.72, 0.93]	High	High
	Step length variability (%)	5.80 (1.88)	6.36 (2.48)	5.79 (1.97)	2	27	0.45 [0.24, 0.64]	Poor	Moderate
	Step time variability (%)	5.64 (2.34)	5.81 (2.25)	5.35 (1.94)	2	29	0.51 [0.31, 0.69]	Poor	Poor
	Step length asymmetry (%)	3.87 (3.36)	3.39 (3.09)	3.58 (2.47)	2	50	0.65 [0.47, 0.79]	Poor	Poor
	Step time asymmetry (%)	3.95 (3.74)	3.34 (3.46)	3.43 (3.18)	3	78	0.41 [0.19, 0.61]	Poor	Poor

The number of decimal places for SEM was chosen such that the SEM had one significant digit [1]. For ICCs for 3D system please refer to the [supplementary file](#). [1] T.J. Cole, Too many digits: the presentation of numerical data, Arch. Dis. Child. 100(7) (2015) 608–9.

capture. These findings support previous findings of good-to-excellent validity for the G&B App postural stability measures in quiet stance in younger adults [15]. The slightly higher validity reported in our middle-to-older age sample may relate to greater unsteadiness in this population, which may have been more easily detected by the G&B App, producing a larger correlation with the gold-standard.

During walking tasks, the validity of G&B App parameters against kinematic data was excellent for step time, and moderate-to-good for walking speed and step length. Currently, clinicians do not have access to quick and accurate measures of step time and step length, and therefore the G&B App offers a new tool to assess these parameters. This

may be valuable in assessing gait changes in older adults and those at risk of falling. The standard clinical measure for walking speed, the 10 m-walk test, has excellent validity against instrumented systems [31], however gait speed alone is unlikely to be sufficient to assess dynamic balance [32]. With the G&B App, walking speed can be assessed along with other important balance and gait parameters. G&B App measures of step length asymmetry/variability and step time asymmetry/variability had poor validity against the 3D motion system. Other studies have reported poor agreement between trunk accelerometry and an instrumented walkway for measures of gait variability and asymmetry in healthy young and older adults [23,33]. The lack of agreement

between the two systems for step length/step time asymmetry and variability is most likely a manifestation of these measures being unreliable; this is discussed further below. Other authors have explained the lack of agreement in terms of differences in the way heel-strike and toe-off events are detected [34], errors in left-right classification [23], influences of gravity and skin motion [33], or signal noise caused by the integration procedure applied within the inverted pendulum model [35].

4.4. Are Gait&Balance App measures reliable?

Most G&B App postural stability measures in quiet stance demonstrated moderate test-retest reliability, which was lower than previously-reported for G&B App reliability in young adults [15] but comparable with other accelerometry research in a mixed-age sample [36]. G&B App measures of walking speed, step length, and step time, had higher reliability when walking with head turns than looking straight ahead. This may reflect tighter constraints on neuromuscular control required under more challenging balance conditions [37]. Reliability for G&B App measures of step length/time variability and asymmetry was generally poor; however, these measures also had poor reliability with the gold-standard 3D system. Low ICCs for both accelerometry and gold-standard systems may have resulted from low between-participant variance in our healthy sample or may relate to the use of short walking bouts, in which spatiotemporal parameters have less time to become established [38]. Gait variability between sessions is likely to be more reliable [39] and is more relevant to clinical practice where the G&B App could be used to track changes during rehabilitation; this should be a focus of future research.

4.5. Study limitations

Despite attempting to recruit equal numbers of males and females, fewer males were enrolled. The MBT and FGA were not randomised resulting in a learning effect which may have raised the FGA scores. The study demonstrates the performance of the G&B App in a healthy ageing cohort, and therefore, the results may not be generalisable to those with more substantial balance impairments.

5. Conclusion

This study established the validity and reliability of several gait and balance parameters measured with the G&B App. The analysis against age and clinical measures supported the validity of the G&B App. The analysis against gold-standard kinematic measures demonstrated that the G&B App offers valid measures of postural stability, walking speed, step length and step time. G&B App measures of postural stability, walking speed, periodicity index, step length and step time had moderate-to-excellent reliability. The G&B App has potential to be used in older adults to detect balance deficits that are not detectable with standard clinical measures. Further research is needed to determine if smartphone measures are responsive to small changes that occur during rehabilitation.

Funding

This study was supported by PBRF funding from the School of Clinical Sciences, Auckland University of Technology.

Author contributions

SO, UR, GA, and DT designed the study. SO managed ethical approval. SO, CA, EB, ML, and GM completed data collection. UR managed data processing and completed the statistical analysis. All authors contributed to data interpretation. SO drafted the manuscript and all authors contributed to the final version.

Acknowledgements

Priyadharshini Suresh and Shobika Ravindran contributed to data collection and processing.

Conflict of Interest

The authors declare no conflict of interest.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.gaitpost.2022.11.014.

References

- [1] R.J. Peterka, Sensory integration for human balance control, *Handb. Clin. Neurol.* 159 (2018) 27–42.
- [2] M.Y. Osoba, A.K. Rao, S.K. Agrawal, A.K. Lalwani, Balance and gait in the elderly: a contemporary review, *Laryngoscope Invest. Otolaryngol.* 4 (1) (2019) 143–153.
- [3] J. Bernhardt, P. Ellis, S. Denisenko, K. Hill, Changes in balance and locomotion measures during rehabilitation following stroke, *Physiother. Res. Int.* 32 (2) (1998) 109–122.
- [4] M. Godi, F. Franchignoni, M. Caligari, A. Giordano, A.M. Turcato, A. Nardone, Comparison of reliability, validity, and responsiveness of the MiniBEST and berg balance scale in patients with balance disorders, *Phys. Ther.* 93 (2) (2013) 158–167.
- [5] P.K. Pardasany, N.K. Latham, A.M. Jette, R.C. Wagenaar, P. Ni, M.D. Slavin, et al., Sensitivity to change and responsiveness of four balance measures for community-dwelling older adults, *Phys. Ther.* 92 (3) (2012) 388–397.
- [6] M. Mancini, F.B. Horak, The relevance of clinical balance assessment tools to differentiate balance deficits, *Eur. J. Phys. Rehabil. Med.* 46 (2) (2010) 239–248.
- [7] H. El-Kashlan, N. Shepard, A. Asher, M. Smith-Wheelock, S. Telian, Evaluation of clinical measures of equilibrium, *Laryngoscope* 108 (3) (1998) 311–319.
- [8] R. Sabchuk, P. Bento, A. Rodacki, Comparison between field balance tests and force platform, *Rev. Bras. Med. Esport.* 18 (6) (2012) 404–408.
- [9] L. Prosperini, C. Pozzilli, The clinical relevance of force platform measures in multiple sclerosis: a review, *Mult. Scler. Int.* 2013 (2013), 756564.
- [10] A.I. Perez-Sanpablo, J. Quinzanos-Fresnedo, R. Loera-Cruz, I. Quinones-Uriostegui, G. Rodriguez-Reyes, R. Perez-Zavala, Validation of the instrumented evaluation of spatio-temporal gait parameters in patients with motor incomplete spinal cord injury, *Spinal Cord.* 55 (7) (2017) 699–704.
- [11] Y.L. Chiu, Y.J. Tsai, C.H. Lin, Y.R. Hou, W.H. Sung, Evaluation of a smartphone-based assessment system in subjects with chronic ankle instability, *Comput. Methods Prog. Biomed.* 139 (2017) 191–195.
- [12] S. Han, D. Lee, S. Lee, A study on the reliability of measuring dynamic balance ability using a smartphone, *J. Phys. Ther. Sci.* 28 (2016) 2515–2518.
- [13] G. Arcuria, C. Marcotulli, R. Amuso, G. Dattilo, C. Galasso, F. Pierelli, et al., Developing a smartphone application, triaxial accelerometer-based, to quantify static and dynamic balance deficits in patients with cerebellar ataxias, *J. Neuro.* 267 (3) (2020) 625–639.
- [14] Y.R. Hou, Y.L. Chiu, S.L. Chiang, H.Y. Chen, W.H. Sung, Development of a smartphone-based balance assessment system for subjects with stroke, *Sensors* 20 (1) (2019).
- [15] U. Rashid, D. Barbado, S. Olsen, G. Alder, J.L.L. Elvira, S. Lord, et al., Validity and reliability of a smartphone app for gait and balance assessment, *Sensors* 22 (1) (2021).
- [16] L.E. Cofre Lizama, M. Pijnappels, G.H. Faber, P.N. Reeves, S.M. Verschueren, J. H. van Dieen, Age effects on mediolateral balance control, *PLoS One* 9 (10) (2014), e110757.
- [17] A. Shumway-Cook, F.B. Horak, Assessing the influence of sensory interaction on balance. suggestion from the field, *Phys. Ther.* 66 (10) (1986) 1548–1550.
- [18] D.M. Wrisley, S.L. Whitney, The effect of foot position on the modified clinical test of sensory interaction and balance, *Arch. Phys. Med. Rehabil.* 85 (2) (2004) 335–338.
- [19] F. Franchignoni, F. Horak, M. Godi, A. Nardone, A. Giordano, Using psychometric techniques to improve the balance evaluation systems test: the mini-BESTest, *J. Rehabil. Med.* 42 (4) (2010) 323–331.
- [20] D.M. Wrisley, G.F. Marchetti, D.K. Kuharsky, S.L. Whitney, Reliability, internal consistency, and validity of data obtained with the functional gait assessment, *Phys. Ther.* 84 (10) (2004) 906–918.
- [21] W. Zhang, M. Smuck, C. Legault, M.A. Ith, A. Muaremi, K. Aminian, Gait symmetry assessment with a low back 3D accelerometer in post-stroke patients, *Sensors* 18 (10) (2018).
- [22] J.H. Kahn, A. Ohlendorf, A. Olsen, K.E. Gordon, Reliability and validity of the functional gait assessment in incomplete spinal cord injury, *Top. Spinal Cord. Inj. Rehabil.* 26 (4) (2020) 268–274.
- [23] A. Godfrey, S. Del Din, G. Barry, J.C. Mathers, L. Rochester, Instrumenting gait with an accelerometer: a system and algorithm examination, *Med Eng. Phys.* 37 (2015) 400–407.

- [24] L.G. Portney, M.P. Watkins, *Foundations of Clinical Research: Applications to Practice*, Pearson/Prentice Hall, Upper Saddle River, NJ, 2009.
- [25] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, *J. Chiropr. Med* 15 (2016) 155–163.
- [26] B. Munro, *Statistical Methods for Health Care Research*, Vol. 1, Lippincott Williams & Wilkins, Philadelphia, 2005.
- [27] R Core Team, *R: a language and environment for statistical computing*, R foundation for statistical computing, Vienna, Austria (2022).
- [28] G. Yogev, M. Plotnik, C. Peretz, N. Giladi, J.M. Hausdorff, Gait asymmetry in patients with Parkinson's disease and elderly fallers: when does the bilateral coordination of gait require attention? *Exp. Brain Res.* 177 (3) (2007) 336–346.
- [29] D.S. Peterson, C. Van Liew, S. Stuart, P. Carlson-Kuhta, F.B. Horak, M. Mancini, Relating Parkinson freezing and balance domains: a structural equation modeling approach, *Park. Relat. Disord.* 79 (2020) 73–78.
- [30] D.D. Espy, F. Yang, T. Bhatt, Y.C. Pai, Independent influence of gait speed and step length on stability and fall risk, *Gait Posture* 32 (3) (2010) 378–382.
- [31] D.M. Peters, S.L. Fritz, D.E. Krotish, Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults, *J. Geriatr. Phys. Ther.* 36 (1) (2013) 24–30.
- [32] S. Lord, B. Galna, L. Rochester, Moving forward on gait measurement: toward a more refined approach, *Mov. Disord.* 28 (11) (2013) 1534–1543.
- [33] A. Hartmann, S. Luzi, K. Murer, R.A. de Bie, E.D. de Bruin, Concurrent validity of a trunk tri-axial accelerometer system for gait analysis in older adults, *Gait Posture* 29 (3) (2009) 444–448.
- [34] S. Del Din, A. Godfrey, L. Rochester, Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: toward clinical and at home use, *IEEE J. Biomed. Health Inf.* 20 (3) (2016) 838–847.
- [35] Q. Zhao, B. Zhang, J. Wang, W. Feng, W. Jia, M. Sun, Improved method of step length estimation based on inverted pendulum model, *Int J. Distrib. Sens Netw.* 13 (4) (2017).
- [36] S.L. Whitney, J.L. Roche, G.F. Marchetti, C.C. Lin, D.P. Steed, G.R. Furman, et al., A comparison of accelerometry and center of pressure measures during computerized dynamic posturography: a measure of balance, *Gait Posture* 33 (4) (2011) 594–599.
- [37] H. Lee, K.P. Granata, Process stationarity and reliability of trunk postural stability, *Clin. Biomech. (Bristol, Avon)* 23 (6) (2008) 735–742.
- [38] K.L. Paterson, N.D. Lythgo, K.D. Hill, Gait variability in younger and older adult women is altered by overground walking protocol, *Age Ageing* 38 (6) (2009) 745–748.
- [39] S. Lord, T. Howe, J. Greenland, L. Simpson, L. Rochester, Gait variability in older adults: a structured review of testing protocol and clinimetric properties, *Gait Posture* 34 (4) (2011) 443–450.