

REVIEW

Open Access



Validity and reliability of wearable inertial sensors in healthy adult walking: a systematic review and meta-analysis

Dylan Kobsar^{1,2}, Jesse M. Charlton^{2,3}, Calvin T.F. Tse^{2,3}, Jean-Francois Esculier^{2,4,5}, Angelo Graffos^{2,3}, Natasha M. Krowchuk², Daniel Thatcher² and Michael A. Hunt^{2,4*}

Abstract

Background: Inertial measurement units (IMUs) offer the ability to measure walking gait through a variety of biomechanical outcomes (e.g., spatiotemporal, kinematics, other). Although many studies have assessed their validity and reliability, there remains no quantitative summary of this vast body of literature. Therefore, we aimed to conduct a systematic review and meta-analysis to determine the i) concurrent validity and ii) test-retest reliability of IMUs for measuring biomechanical gait outcomes during level walking in healthy adults.

Methods: Five electronic databases were searched for journal articles assessing the validity or reliability of IMUs during healthy adult walking. Two reviewers screened titles, abstracts, and full texts for studies to be included, before two reviewers examined the methodological quality of all included studies. When sufficient data were present for a given biomechanical outcome, data were meta-analyzed on Pearson correlation coefficients (r) or intraclass correlation coefficients (ICC) for validity and reliability, respectively. Alternatively, qualitative summaries of outcomes were conducted on those that could not be meta-analyzed.

Results: A total of 82 articles, assessing the validity or reliability of over 100 outcomes, were included in this review. Seventeen biomechanical outcomes, primarily spatiotemporal parameters, were meta-analyzed. The validity and reliability of step and stride times were found to be excellent. Similarly, the validity and reliability of step and stride length, as well as swing and stance time, were found to be good to excellent. Alternatively, spatiotemporal parameter variability and symmetry displayed poor to moderate validity and reliability. IMUs were also found to display moderate reliability for the assessment of local dynamic stability during walking. The remaining biomechanical outcomes were qualitatively summarized to provide a variety of recommendations for future IMU research.

(Continued on next page)

* Correspondence: michael.hunt@ubc.ca

²Motion Analysis and Biofeedback Laboratory, University of British Columbia, Vancouver, BC, Canada

⁴Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusions: The findings of this review demonstrate the excellent validity and reliability of IMUs for mean spatiotemporal parameters during walking, but caution the use of spatiotemporal variability and symmetry metrics without strict protocol. Further, this work tentatively supports the use of IMUs for joint angle measurement and other biomechanical outcomes such as stability, regularity, and segmental accelerations. Unfortunately, the strength of these recommendations are limited based on the lack of high-quality studies for each outcome, with underpowered and/or unjustified sample sizes (sample size median 12; range: 2–95) being the primary limitation.

Keywords: Inertial sensors, Inertial measurement units, Gait, Biomechanics, Validity, Reliability, Review

Introduction

Gait analyses are important for evaluating movement in healthy and pathological populations by assessing a range of biomechanical outcomes from simple spatiotemporal parameters to complex three-dimensional (3D) joint angles [1, 2]. While laboratory-based, optical motion analysis systems remain the gold standard for gait analysis, they are expensive, resource intensive, and largely immobile, which limits their accessibility in both research and clinical settings [3]. Alternatively, recent technological advancements have led to the growing popularity of more affordable, easy-to-use, and accessible wearable sensors for the analysis of gait patterns [4].

Wearable technology refers to any electronic device that can be worn, but inertial sensors are the most common type of wearable sensor for measuring gait [5]. These sensors apply the principle of inertia to measure linear accelerations (i.e., accelerometers) or angular velocities (i.e., gyroscopes). Independently, inertial sensors can provide information on the motion of segments, or timing of gait events. Further, inertial sensors can be integrated into what is called an inertial measurement unit (IMU), which contains a 3-axis accelerometer and a 3-axis gyroscope, as well as, in some cases, a 3-axis magnetometer to assess heading direction [6]. The fusion of data from these sensors facilitates the assessment of segment orientations and joint angles [6, 7]. Therefore, inertial sensors, either on their own or combined in an IMU, provide an excellent opportunity to collect a variety of valuable and objective outcomes related to gait.

With the increasing popularity of wearable sensors, there have been an increasing number of studies examining their validity and reliability for gait analysis. Similarly, while there are many reviews of wearable sensor literature available, most have taken a descriptive approach to outline potential applications [5, 8] or methods [4, 9–11]. Therefore, there remains a lack of systematic reviews and meta-analyses which synthesize the results of the many validity and reliability studies which have examined inertial sensor outcomes for gait analysis. Recently, two systematic reviews examined 3D joint kinematics from inertial sensors across a variety of movements and populations [12, 13]. While they were

unable to quantitatively pool data due to study heterogeneity, they were able to qualitatively suggest sagittal, and to a lesser extent frontal, plane lower limb joint kinematics displayed acceptable validity. Nevertheless, these findings remain confounded across a variety of human movements and populations. Therefore, addressing kinematic outcomes in only healthy adult walking may help to homogenize findings and recommendations. Further, there remains a growing body of literature that addresses a variety of spatiotemporal and other biomechanical outcomes assessed across a variety of locations (e.g., back, shank, foot, etc.) in walking which have yet to be addressed in a systematic and quantitative manner. Addressing this gap in the literature will help future researchers to identify not only the most valid and reliable of these variables, but the optimal placement of sensors to measure them. Therefore, our aim was to conduct a systematic review and meta-analysis to determine the i) concurrent validity and ii) test-retest reliability of IMUs for measuring biomechanical gait outcomes (e.g., spatiotemporal, kinematic, or other) during level over-ground or treadmill walking in healthy adults.

Methods

Eligibility criteria

We included journal articles that assessed the validity or reliability of IMUs measuring biomechanical outcomes during walking in healthy adults. For a validity study to be included, it must have assessed the concurrent validity (i.e., simultaneous collection) of inertial sensor measured biomechanical gait outcomes as compared to what we defined to be gold standard devices (See Additional file 1) in healthy adults. Similarly, for a reliability study to be included, it must have assessed the test-retest reliability (i.e., between-day, within-day, or between-tester; involving the same measure/device/placement with removal between sessions) of IMU-measured biomechanical gait outcomes in healthy adult walking. Biomechanical gait outcomes included spatiotemporal parameters (e.g., step time, step length, stance time, etc.), segment or joint kinematics/kinetics, or other biomechanical outcomes (e.g., accelerations, stability, regularity, etc.). However, we did not include per count measures such as gait speed or cadence as these require

two components (e.g., time and distance) and can often be measured as an average over the entire dataset. Additional details on our inclusion and exclusion criteria can be found in Additional file 1.

Study identification and screening

A systematic literature search was conducted with the help of a librarian to identify all relevant journal articles in the following databases: MEDLINE, Embase, CINAHL, Web of Science, and Compendex. Our search criteria were based on the combination of four broad topics: inertial sensors, gait biomechanics, healthy adults, and validity/reliability. Each topic included an expanded set of terms, keywords, and syntax specific to each database to maximize the breadth of our search. A detailed list of our search strategy for each database can be found in Additional file 2. This search was conducted on May 7th, 2019.

Following the removal of duplicate items, titles and abstracts were screened by two independent reviewers (CTFT and DT) to determine their eligibility based on the aforementioned criteria. Studies that were deemed potentially eligible were passed to full-text screening where two independent reviewers (CTFT and DK) conducted a thorough examination of each article to determine if it would be included in our review. Moreover, the reviewers also identified eligible components of the study for future analysis; for example, a study may pass in reliability criteria, but fail validity criteria (or vice versa). Disagreements between reviewers were resolved by consensus, with a third reviewer (MAH) available for arbitration. Most studies defined a clear purpose of assessing the validity and/or reliability of a given IMU outcome in healthy adults, however a number of studies addressed more advanced problems (e.g., clinical populations or new techniques) but still presented results that met our criteria.

Methodological quality

Study quality was assessed by two independent reviewers (JFE and AG) using a modified version of the Critical Appraisal of Study Design for Psychometric Articles [14], which we adapted to studies evaluating the psychometric properties of wearable sensors (Additional file 3). This modified evaluation form contains 12 items evaluating study quality in 5 categories: study question, study design, measurements, analyses, and recommendations. Each item is scored as 2 (satisfactory), 1 (partially satisfactory), or 0 (unsatisfactory), with a total possible score out of 24 converted to a percentage. Raters were blinded to any identifiable information (e.g., author names, study title, publication year, journal) to avoid bias in their quality assessment. Initially, both raters evaluated two articles, after which they met to discuss each item to

clarify their meaning and interpretation. The same process was repeated for each subsequent block of 20 articles. An intraclass correlation coefficient [ICC (3,1)] was calculated to evaluate pre-consensus inter-rater reliability of the total score. Disagreements were discussed and resolved through face-to-face meetings. If a consensus could not be reached, a third rater (DK) served as the tiebreaker. Studies obtaining a quality score between 85 and 100% were classified as high quality (HQ), those scoring between 70 and 85% were classified as moderate quality (MQ) and studies obtaining between 50 and 70% were classified as low quality (LQ). Studies rating below 50% were considered very low quality (VLQ) and were excluded from the quantitative synthesis. However, all studies were still included in the qualitative synthesis. Quality assessment scoring was then used to determine the strength of recommendations [15].

Data extraction

Data were extracted from the included studies by one reviewer (NMK) and checked for accuracy by a second (JMC). Extracted data consisted of study design, sample demographics, inertial sensor specifications and placements, as well as each biomechanical outcome of interest and their reported statistical outcomes. While all statistical outcomes were extracted for the qualitative assessments, data pooling was a priori set to assess only the Pearson correlation coefficients (r) and ICCs for validity and reliability, respectively.

Data pooling

Data pooling was facilitated with a multistage grouping of outcomes. First, all extracted outcomes were dichotomized as assessing either validity or reliability. Outcomes were then separated into overarching outcome groups (e.g., spatiotemporal, kinematic, other), before being grouped by specific outcome names (e.g., step time, stride time, step length, etc.) and finally sensor locations (e.g., foot, shank, thigh, back, etc.). For example, all assessments of “step time” would be grouped together, but further separated based on the placement of the inertial sensor. Data were not further pooled by type of sensor (e.g., accelerometer vs. gyroscope) or algorithm used. Therefore, a single study may contribute to multiple independent data poolings based on validity or reliability, outcome measure, and sensor placements. Biomechanical outcomes with three or more independent study samples using the same sensor location and reporting the desired statistical outcomes (i.e., r , ICC) were quantitatively synthesized. Agreement metrics (i.e., ICC and r) were interpreted as poor (< 0.500), moderate ($0.500–0.749$), good ($0.750–0.899$), and excellent (≥ 0.900).

Data for validity and reliability outcomes were meta-analyzed based on the r and ICC, respectively, and 95%

confidence intervals were generated using a random-effects model (R version 3.6.0 using the meta package with the metacor function [16]). Weighting of individual point estimates was based on study sample sizes. Given the non-normality of Pearson correlation coefficients and ICCs, point estimates were variance-stabilized using Fisher's z-transform [17]. In all cases where an ICC was reported, and as far as we could determine given the information available, the number of measures or comparators was $m = 2$; therefore, Fisher's z-transform applied similarly to both r and ICC. However, for ICCs the standard error was adjusted to $1/\sqrt{(N-3/2)}$ following previous recommendations [18]. Data were then transformed back to their respective original outcome measures for reporting. Heterogeneity was examined using τ^2 , I^2 and Cochran's Q statistic where $\tau^2 = 0$ suggests no heterogeneity, I^2 values < 25 , 26–50%, and $> 75\%$ suggest low, moderate and high heterogeneity [19], and a significant Q statistic indicated that the studies do not share similar effects. Results of the meta-analysis were interpreted using the same agreement metric definitions as outlined above.

Alternatively, qualitative interpretation was conducted on outcomes that were unable to be quantitatively pooled. Additional error metrics (i.e., root-mean-square error (RMSE), standard error of measurement (SEM), minimum detectable change (MDC), limits of agreement (LoA)) were included in this qualitative synthesis to support our interpretations [15]:

- Strong evidence: multiple HQ or MQ studies with consistent results.
- Moderate evidence: multiple studies, including at least one HQ study or multiple MQ studies, presenting consistent results.
- Limited evidence: multiple LQ studies with inconsistent results, or one HQ/MQ study.
- Conflicting evidence: multiple studies providing inconsistent results, regardless of the methodological quality.
- Very limited evidence: only one LQ or MQ study or multiple VLQ

Results

Search results

Our search strategy identified a total of 2804 articles. Following the removal of duplicates, screening of titles/abstracts, and full-text screening, 82 articles [20–101] were included in the current review (Fig. 1). We did not set a date range on the search; however, the number of papers in this area was found to increase heavily from 2008 to 2014, with $> 50\%$ of the included papers published within approximately 5 years, and $> 85\%$ within 10 years (Fig. 2).

Methodological quality

Only 1 article was rated as HQ, 13 as MQ, 50 as LQ and 18 as VLQ (Table 1). Agreement between both raters reached a single-measures ICC (3,1) of 0.83 [95% C.I. 0.75, 0.89]. The items for which articles generally scored higher were “1- Background and research question” and “9- Organization and completeness of study results”. In contrast, 81 papers (95%) did not provide any justification about their sample size and/or appeared to be underpowered.

Study characteristics

The 82 studies included in this review assessed biomechanical outcomes in walking using a variety of IMUs. The most common IMU system used was Xsens Technologies ($n = 9$), followed by Opal ($n = 7$), and finally Dynaport ($n = 5$) and Shimmer ($n = 5$). The most common sampling frequency used to assess walking was 100 Hz (range: 25–2000 Hz). Lastly, data from 1510 healthy adults were included across these studies (mean (sd) sample size: 18 (17) participants; median sample size: 12 participants; range: 2–95 participants). See Table 2 and Table 3 for breakdown of study characteristics separated based on validity and reliability, respectively.

Validity

Overall, a total of 23 spatiotemporal outcomes, 3D lower limb kinematics and kinetics, plus 7 other biomechanical outcomes were assessed across the 63 studies that examined IMU validity. From these outcomes, 12 spatiotemporal parameters presented sufficient study quality and statistical outcomes to allow for data pooling (Fig. 3 and Fig. 4). We were unable to meta-analyze kinematic/kinetic outcomes or other biomechanical outcomes, due to either a limited number of studies or, in many cases, a lack of consistency in data reporting, as many studies reported only RMSE or even a simple mean difference. Studies that were unable to be meta-analyzed were qualitatively summarized by outcomes and placements in Supplementary Table 1 for spatiotemporal outcomes, Supplementary Table 2 for kinematic/kinetic outcomes, and Supplementary Table 3 for other biomechanical outcomes. Therefore, the results presented in the following section represent only outcomes and placements which allowed for quantitative data pooling.

Quantitative pooling of spatiotemporal outcomes for validity

Step time

Data from five low to moderate quality studies (contributing six independent study samples) suggests that the validity for step time measured with IMUs placed on the back was excellent (total $n = 257$; $r = 0.99$, 95% CI [0.97, 1.00], $I^2 = 93\%$, $p < 0.001$) [34, 41, 44, 77, 86]. An additional 10 studies that could not be pooled provided

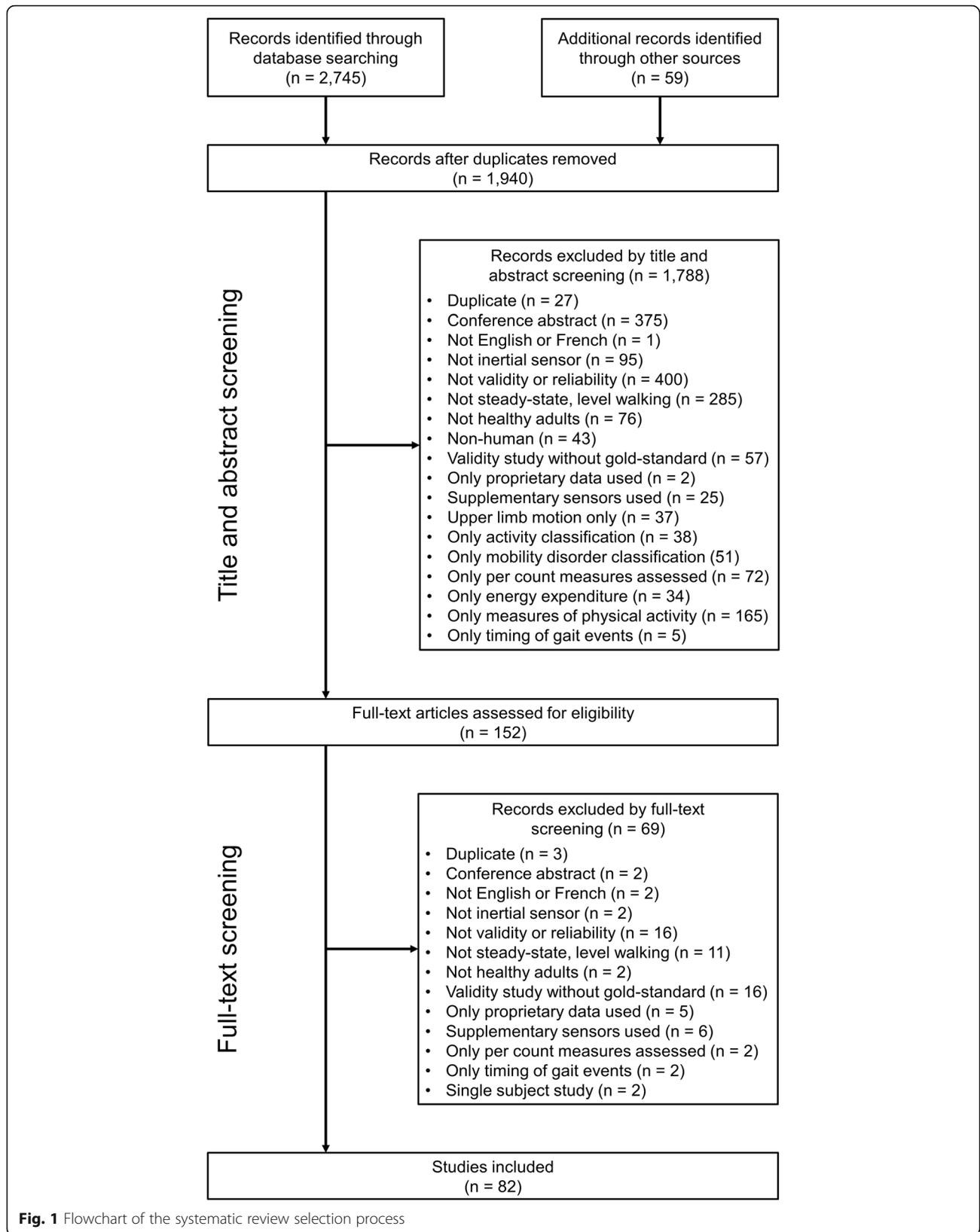


Fig. 1 Flowchart of the systematic review selection process

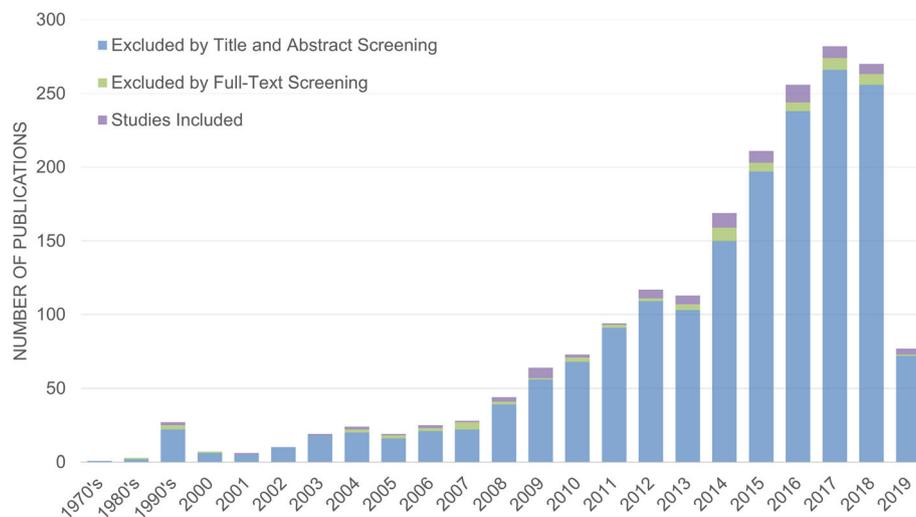


Fig. 2 Number of studies identified, excluded, and included by years

limited evidence for moderate to excellent validity of step times measured at the back or shank/ankle [28, 51, 61, 88, 91, 93].

Step length

Data from five low to moderate quality studies (contributing six independent study samples) suggests that the validity for step length measured with IMUs placed on the back was good (total $n = 234$; $r = 0.88$, 95% CI [0.83, 0.92]; $I^2 = 32\%$; $p < 0.001$) [34, 41, 44, 77, 86]. An additional study that could not be pooled provided limited evidence for excellent validity of step length measured at the back [51].

Stance time

Data from two low quality studies (contributing three independent study samples) suggests that the validity for stance time measured with IMUs placed on the back was excellent (total $n = 107$; $r = 0.91$, 95% CI [0.87, 0.94]; $I^2 = 0\%$; $p < 0.001$) [41, 44]. An additional 5 studies that could not be pooled provided limited evidence for moderate validity of stance times measured at the back [28, 82, 88, 91, 93].

Swing time

Data from two low quality studies (contributing three independent study samples) suggests that the validity of swing time measured with IMUs placed on the back was moderate (total $n = 107$, $r = 0.68$, 95% CI [0.56, 0.77]; $I^2 = 0\%$; $p < 0.001$) [41, 44]. An additional 3 studies that could not be pooled provided very limited evidence for moderate validity of swing times measured at the back [28, 91, 93].

Step time variability

Data from three low to moderate quality studies suggests that the validity of step time variability measured with IMUs placed on the back was poor (total $n = 189$, $r = 0.35$, 95% CI [0.18, 0.50]; $I^2 = 31\%$, $p < 0.001$) [34, 41, 44]. An additional 2 studies that could not be pooled provided limited evidence for excellent validity of step time variability measured at the back [51, 88].

Step length variability

Data from two low quality studies (contributing three independent study samples) suggests that the validity of step length variability measured with IMUs placed on the back was poor (total $n = 107$; $r = 0.06$, 95% CI [-0.14, 0.25]; $I^2 = 0\%$, $p = 0.543$) [41, 44]. An additional study that could not be pooled provided limited evidence for poor validity of step length variability measured at the back [51].

Stance time variability

Data from two low quality two studies (contributing three independent study samples) suggests that the validity of stance time variability measured by IMUs placed at the back was moderate (total $n = 107$; $r = 0.58$, 95% CI [0.35, 0.74]; $I^2 = 0.53\%$; $p < 0.001$) [41, 44]. An additional study that could not be pooled provided very limited evidence for moderate validity of stance time variability measured at the back [88].

Swing time variability

Data from two low quality studies (contributing three independent study samples) suggests that the validity of swing time variability measured by IMUs placed at the back was poor (total $n = 107$; $r = 0.34$, 95% CI [0.11, 0.53]; $I^2 = 30\%$; $p = 0.004$) [41, 44].

Table 1 Quality assessment scoring of 82 included studies

Study Information			Study Ques.	Study Design					Measurement		Analyses			Rec	Total	
Author	Year	Ref	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	/24	%
Abhayasinghe	2019	[20]	1	0	1	2	0	N/A	0	0	2	1	0	0	7	31.8%*
Al-Amri	2018	[21]	2	1	1	2	2	2	2	2	2	1	1	2	20	83.3%
Allseits	2018	[22]	2	1	1	2	0	N/A	0	1	2	2	0	2	13	59.1%*
Allseits	2017	[23]	2	1	1	2	0	N/A	1	2	2	2	0	2	15	68.2%*
Aminian	2004	[24]	1	0	0	2	0	N/A	1	2	2	1	0	1	10	45.5%*
Atallah	2014	[25]	2	1	2	2	0	N/A	0	1	2	2	1	0	13	59.1%*
Backhouse	2013	[26]	2	1	1	2	0	N/A	1	1	2	2	0	1	13	59.1%*
Bautmans	2011	[27]	2	2	0	2	0	2	1	1	2	2	2	2	18	75.0%
Ben Mansour	2015	[28]	2	0	0	2	0	N/A	1	1	2	0	0	1	9	40.9%*
Benoussaad	2016	[29]	2	1	0	2	0	N/A	1	1	1	0	0	1	9	40.9%*
Bertoli	2018	[30]	2	1	0	2	1	N/A	2	2	2	2	0	2	16	72.7%*
Bolink	2016	[31]	2	1	1	2	0	N/A	2	1	2	2	1	2	16	72.7%*
Bruijn	2010	[32]	2	0	1	2	0	N/A	1	2	1	1	1	1	12	54.5%*
Buganè	2012	[33]	2	1	1	2	0	N/A	1	1	2	0	0	1	11	50.0%*
Byun	2016	[34]	2	2	1	2	0	N/A	1	2	2	1	2	1	16	72.7%*
Chalmers	2014	[35]	2	0	0	2	0	N/A	1	0	1	0	0	0	6	27.3%*
Chapman	2019	[36]	1	0	0	2	0	N/A	0	2	2	2	0	1	10	45.5%*
Charlton	2019	[37]	2	1	2	2	2	2	2	0	2	2	2	2	21	87.5%
Cole	2014	[38]	2	1	1	2	0	N/A	1	1	2	2	1	2	15	68.2%*
Cooper	2009	[39]	2	1	0	2	0	N/A	0	2	1	0	0	0	8	36.4%*
Dalton	2013	[40]	2	1	0	2	0	N/A	1	1	2	1	1	2	13	59.1%*
Del Din	2016	[41]	2	1	1	2	0	N/A	1	2	2	1	1	2	15	68.2%*
Esser	2009	[42]	0	0	1	2	0	N/A	1	0	1	1	1	1	8	36.4%*
Furrer	2015	[43]	2	1	1	2	0	N/A	2	1	2	2	1	1	15	68.2%*
Godfrey	2015	[44]	2	1	1	2	0	N/A	1	1	2	1	1	2	14	63.6%*
Gonzalez	2016	[45]	2	1	1	2	1	N/A	1	1	2	0	0	1	12	54.5%*
Gorelick	2009	[46]	2	1	1	2	0	2	1	2	2	1	2	2	18	75.0%
Greene	2012	[47]	2	1	2	2	0	N/A	2	2	2	1	0	0	14	63.6%*
Greene	2010	[48]	2	0	1	2	0	N/A	2	2	2	1	0	1	13	59.1%*
Hamacher	2014	[49]	1	0	1	1	0	2	1	0	2	2	1	1	12	50.0%
Hamacher	2015	[50]	2	0	1	1	0	2	0	0	2	2	2	2	14	58.3%
Hartmann	2009	[51]	2	2	1	2	0	N/A	1	2	2	2	1	2	17	77.3%*
Hartmann	2009	[52]	2	2	1	2	0	2	1	1	2	2	1	2	18	75.0%
Henriksen	2004	[53]	2	1	1	1	0	2	0	2	2	1	2	2	16	66.7%
Huang	2016	[54]	2	1	1	2	0	N/A	1	2	2	0	0	2	13	59.1%*
Hundza	2014	[55]	2	0	0	2	0	N/A	1	0	2	0	0	2	9	40.9%*
Jarchi	2014	[56]	2	0	1	2	0	N/A	1	1	2	2	1	2	14	63.6%*
Karatsidis	2019	[57]	2	1	1	2	0	N/A	2	0	2	1	1	2	14	63.6%*
Kavanagh	2006	[58]	2	1	2	2	0	2	2	2	2	1	1	2	19	79.2%
Kitagawa	2016	[59]	2	1	0	1	0	N/A	0	0	1	0	0	2	7	31.8%*
Kluge	2017	[60]	2	1	1	2	0	0	1	1	2	1	1	1	13	54.2%
Köse	2012	[61]	2	0	1	1	0	N/A	0	0	2	0	0	1	7	31.8%*

Table 1 Quality assessment scoring of 82 included studies (*Continued*)

Study Information			Study Ques.	Study Design						Measurement		Analyses			Rec	Total	
Author	Year	Ref	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	/24	%	
Lebel	2017	[62]	2	1	0	2	0	N/A	1	1	2	1	1	2	13	59.1%*	
L'Hermette	2008	[63]	1	1	0	1	0	N/A	1	0	1	1	0	1	7	31.8%*	
Liikavainio	2007	[64]	2	1	1	1	0	2	2	2	2	0	1	2	16	66.7%	
Liu	2009	[65]	2	1	0	1	0	N/A	0	1	2	1	0	2	10	45.5%*	
Lord	2008	[66]	0	0	1	2	0	N/A	0	0	2	1	2	1	9	40.9%*	
Lyytinen	2016	[67]	2	1	0	2	0	2	1	2	2	1	1	2	16	66.7%	
Maffioletti	2008	[68]	2	1	1	2	0	N/A	2	1	2	1	0	2	14	63.6%*	
Manor	2018	[69]	2	2	1	2	0	2	0	0	1	1	1	2	14	58.3%	
Mariani	2012	[70]	2	1	1	2	0	N/A	1	1	2	2	0	2	14	63.6%*	
Mariani	2013	[71]	2	1	1	2	0	N/A	1	1	2	1	0	2	13	59.1%*	
McGrath	2012	[72]	2	0	1	2	0	N/A	1	2	2	1	1	1	13	59.1%*	
Moe-Nilssen	1998	[73]	2	0	0	1	0	2	1	1	2	2	1	2	14	58.3%	
Nishiguchi	2012	[74]	2	1	0	0	0	2	1	1	2	1	1	1	12	50.0%	
Ohtako	2001	[75]	2	0	0	2	0	N/A	0	1	1	1	0	2	9	40.9%*	
Orlowski	2017	[76]	2	1	1	2	0	2	1	1	2	2	2	2	18	75.0%	
Pepa	2017	[77]	2	0	1	2	0	N/A	0	2	2	2	0	1	12	54.5%*	
Reynard	2014	[78]	2	2	1	1	0	2	1	0	2	2	2	2	17	70.8%	
Sabatini	2015	[79]	2	1	1	2	0	N/A	2	2	1	2	0	2	15	68.2%*	
Saremi	2006	[80]	2	0	1	1	0	2	2	1	2	0	0	2	13	54.2%	
Schmitz-Hübsch	2016	[81]	2	1	1	2	0	N/A	0	1	2	2	0	1	12	54.5%*	
Sejdic	2015	[82]	2	0	1	2	0	N/A	1	2	2	0	0	2	12	54.5%*	
Selles	2005	[83]	2	1	0	2	0	N/A	1	2	2	1	1	1	13	59.1%*	
Senden	2009	[84]	2	0	1	1	0	2	1	1	2	1	1	1	13	54.2%	
Sijobert	2015	[85]	2	1	1	2	0	N/A	1	2	2	0	0	0	11	50.0%*	
Silsupadol	2017	[86]	2	1	1	2	0	N/A	1	2	2	2	1	2	16	72.7%*	
Steins	2014	[87]	1	0	1	1	0	N/A	1	0	2	2	2	1	11	50.0%*	
Storm	2016	[88]	2	0	1	2	0	N/A	1	2	2	0	0	2	12	54.5%*	
Teufl	2019	[89]	2	0	0	1	0	2	2	1	2	1	1	2	14	58.3%	
Teufl	2018	[90]	2	1	1	1	0	2	2	0	2	1	2	2	16	66.7%	
Trojaniello	2014	[91]	2	1	1	2	0	N/A	1	0	2	0	0	1	10	45.5%*	
Trojaniello	2014	[92]	2	1	1	2	0	N/A	2	2	1	2	0	2	15	68.2%*	
Trojaniello	2015	[93]	2	1	1	2	0	N/A	2	2	2	0	0	2	14	63.6%*	
van der Straaten	2018	[94]	2	2	1	2	0	2	1	0	2	2	2	2	18	75.0%	
van Schooten	2013	[95]	2	0	1	2	0	2	1	0	2	2	1	2	15	62.5%	
Washabaugh	2017	[96]	1	1	1	2	0	1	0	1	2	2	1	2	14	58.3%	
Wundersitz	2015	[97]	2	1	1	2	0	N/A	0	2	2	2	1	1	14	63.6%*	
Xia	2017	[98]	2	1	0	2	0	N/A	2	1	2	1	1	1	13	59.1%*	
Zhang	2013	[99]	1	1	1	2	0	N/A	2	0	1	0	1	0	9	40.9%*	
Zijlstra	2013	[100]	2	1	1	1	0	2	1	1	2	1	1	2	15	62.5%	
Zijlstra	2003	[101]	2	1	0	1	0	N/A	0	1	1	0	0	1	7	31.8%*	

*Percentage calculated out of 22 as studies did not qualify for reliability and were not assessed on Q6 for sample retention

Table 2 Details of studies assessing validity for spatiotemporal (ST), kinematic (KIN), and other biomechanical outcomes (OTHER)

Author	Year	Ref	n	Age	ST	KIN	OTHER	Sensor	Hz	Gold Standard System
Abhayasinghe	2019	[20]	19	*	X	X		MPU-9150 (InvenSense)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Al-Amiri	2018	[21]	25	35	X	X		Xsens MTw (Xsens Technologies BV, Netherlands)	60	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Allseits	2017	[23]	11	32	X	X		3D gyroscope (no manufacturer listed)	50	Instrumented Walkway (Matscan, Tekscan, Inc., Boston, MA)
Allseits	2018	[22]	11	32	X	X		3D gyroscope (no manufacturer listed)	*	Instrumented Walkway (Matscan, Tekscan, Inc., Boston, MA)
Aminian	2004	[24]	9	63	X	X		Physilog (BioAGM, CH)	200	Motion Capture Camera System and Force Plates (ELITE System, BTS, Milan, Italy)
Atallah	2014	[25]	14	40	X	X		3D accelerometer (no manufacturer listed)	130	Instrumented Treadmill (hp/cosmos, Munich, Germany)
Backhouse	2013	[26]	12	42	X	X		IDEA (MiniSun LLC., Fresno, CA)	*	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Ben Mansour	2015	[28]	10	29	X	X		MMA8453Q (Freescale Semiconductor); L3G4200D (STMicroelectronics)	200	Instrumented Treadmill (ADAL 3D, Medical Development, Tecmachine Hef, France)
Benoussaad	2016	[29]	10	27	X	X		HikoB Fox (HikoB Villeurbanne, France)	200	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Bertoli	2018	[30]	80	74	X	X		Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Instrumented Walkway (Zeno Walkway, Prokinetics LLC., Havertown, PA)
Bolink	2016	[31]	17	26	X	X		Microstrain Inertia-Link	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Bruijn	2010	[32]	9	*	X	X		PI-node (Philips, The Netherlands)	50	Motion Capture Camera System (Optotrak 3020, NDI, Waterloo, ON)
Buganè	2012	[33]	22	24/27	X	X		Free4Act (F4A)	50	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Byun	2016	[34]	82	69	X	X		FITMETER (FitLifeInc, Suwon, Korea)	32	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Chalmers	2014	[35]	11	22	X	X		ADXL345 (Analog Devices); LPR450 (STMicroelectronics)	60	Motion Capture Camera System (No Brand Reported)
Chapman	2019	[36]	2	50	X	X		APDM	128	Motion Capture Camera System (OptiTrack, Natural Point, Inc., Corvallis, OR)
Cole	2014	[38]	24	71/23	X	X		Inertial Cube3 (InterSense Inc., Bedford, MA, USA)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Cooper	2009	[39]	7	30	X	X		IMU (no manufacture/model reported)	100	Motion Capture Camera System (Qualisys, Sweden)
Dalton	2013	[40]	10	57	X	X		AD_BRC	250	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Del Din	2016	[41]	30	67	X	X		Axivity AX3, UK	50	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Esser	2009	[42]	5	23	X	X		Xsens MTx (Xsens Technologies BV, Netherlands)	100	Motion Capture Camera System (Qualisys, Göteborg, Sweden)
Furrer	2015	[43]	22	27	X	X		Smartphone: Desire HD, HTC Corp, Taiwan	50	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK); Force Plates (OR 6, AMTI, Watertown, MA)
Godfrey	2015	[44]	40/370	29/64	X	X		Axivity AX3, UK	100	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Gonzalez	2016	[45]	5/50	24/68	X	X		LPMS-B (LP Research, Japan)	*	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Greene	2012	[47]	7	*	X	X		Shimmer (Shimmer Sensing, Dublin, IR)	102	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Greene	2010	[48]	9	30	X	X		Shimmer (Shimmer Sensing, Dublin, IR)	102	Motion Capture Camera System (CODA Motion Analysis Leicestershire, UK)
Hartmann	2009	[51]	23	77	X	X		DynaPort (McRoberts BV, Hague, Netherlands)	100	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)

Table 2 Details of studies assessing validity for spatiotemporal (ST), kinematic (KIN), and other biomechanical outcomes (OTHER) (Continued)

Author	Year	Ref	n	Age	ST	KIN	OTHER	Sensor	Hz	Gold Standard System
Huang	2016	[54]	13	30	X			Invensense MPU6050 (San Jose, CA)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Hundza	2014	[55]	7	30	X			Microelectromechanical systems (no manufacture listed)	40	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Jarchi	2014	[56]	10	*	X			E-AR Sensor (Sensixa Ltd., London, UK)	130	Instrumented Treadmill (Gaitway, Kistler Instrument Corp, Amherst, US)
Karatsidis	2019	[57]	11	31	X			Xsens MVm (Xsens Technologies BV, Netherlands)	240	Motion Capture Camera System (Qualisys, Göteborg, Sweden); Force Plates (OR 6, AMTI, Watertown, MA)
Kitagawa	2016	[59]	10	23	X			TSND121 (ATR-Promotions, Japan)	200	Motion Capture Camera System (Motion Analysis Corp, Rohnert Park, CA)
Köse	2012	[61]	9	31	X			Sensorize (FreeSense ApS, Denmark)	100	Motion Capture Camera System (BTS, Milan, Italy)
Lebel	2017	[62]	20	50		X		IGS-180 (Synertial; Inertial Labs)	60	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
L'Hermette	2008	[63]	15	23	X		X	ADXL105-EM3 (Analog Devices)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Liu	2009	[65]	8	25			X	MM-2860 (Sunhayato, Japan)	*	Motion Capture Camera System (NAC Image Technology, Tokyo, Japan)
Lord	2008	[66]	11	73	X			Vitaport Activity Monitor (TEMEC Instruments Inc., Netherlands)	25	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Maffioletti	2008	[68]	10	34	X			IDEEA (MiniSun LLC, Fresno, CA)	32	Force Plates (Kistler Instrumente AG, Winterthur, Switzerland)
Manor	2018	[69]	14	30	X			iPhone 4 s (Apple, CA)	100	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Mariani	2012	[70]	12	32	X			Physilog (BioAGM, CH)	200	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Mariani	2013	[71]	10	*	X			Physilog (BioAGM, CH)	200	Pressure Insoles (Pedar-X, Novel, DE)
McGrath	2012	[72]	5	*	X			Shimmer (Shimmer Sensing, Dublin, IR)	102	Motion Capture Camera System (CODA Motion Analysis, Leicestershire, UK)
Ohtako	2001	[75]	6	*		X		ADXL05 (Analog Devices); ENC03J (Murata)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK); Force Plates (OR 6, AMTI, Watertown, MA)
Pepa	2017	[77]	11	*	X			iPhone 4 s (Apple, CA)	100	Motion Capture Camera System (ELITE System, BTS, Milan, Italy)
Sabatini	2015	[79]	9	30		X		BMA180 (Bosch); ITG-3200 (InvenSense); HMC5843 (Honeywell)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Saremi	2006	[80]	8	*	X			IDEEA (MiniSun LLC, Fresno, CA)	32	Pressure Insoles (B&L Engineering, Los Angeles, CA)
Schmitz-Hübsch	2016	[81]	9	50	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Sejdic	2015	[82]	14	74	X			MMA7260Q (Freescale Semiconductor)	100	Motion Capture Camera System (OptiTrack, Natural Point, Inc., Corvallis, OR)
Selles	2005	[83]	10	29	X			ICSensors 3021 (ICSensors Inc., Fremont, CA)	500	Force Plates (Kistler Instrumente AG, Winterthur, Switzerland)
Sjoberg	2015	[85]	10	*	X			HikoB Fox (HikoB Villeurbanne, France)	200	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Sisupadol	2017	[86]	34	23/74	X			Vivo X5 (Android 4.4.4)	100	Instrumented Walkway (GAITRite; CIR Systems Inc., Franklin, NJ)
Steins	2014	[87]	10	26		X		iPod Touch 4th generation (Apple, CA)	100	Motion Capture Camera System (Qualisys, Sweden)
Storm	2016	[88]	10	28	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Pressure Insoles (F-Scan 3000E, Tekscan, Inc., Boston, MA)
Teufel	2019	[89]	24	*	X			Xsens MTw Awinda (Xsens Technologies BV,	60	Motion Capture Camera System (OptiTrack, Natural Point, Inc., Corvallis,

Table 2 Details of studies assessing validity for spatiotemporal (ST), kinematic (KIN), and other biomechanical outcomes (OTHER) (Continued)

Author	Year	Ref	n	Age	ST	KIN	OTHER	Sensor	Hz	Gold Standard System
Teufel	2018	[90]	28	24		X		Netherlands Xsens MTw Awinda (Xsens Technologies BV, Netherlands)	60	Motion Capture Camera System (OptiTrack, Natural Point, Inc., Corvallis, OR)
Trojaniello	2014	[91]	14	32	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK); Force Plates (OR 6, AMTI, Watertown, MA)
Trojaniello	2014	[92]	10	70	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Trojaniello	2015	[93]	10	70	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	128	Instrumented Walkway (GAITRite, CIR Systems Inc., Franklin, NJ)
Washabaugh	2017	[96]	11	24	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	*	Instrumented Treadmill (Bertec, Columbus, OH)
Wundersitz	2015	[97]	39	24		X		Minimax S4 (Catapult Innovations, Australia)	100	Motion Capture Camera System (Motion Analysis Corp, Rohnert Park, CA)
Xia	2017	[98]	14	25		X		MPU9150 (Invensense, USA)	100	Motion Capture Camera System (Vicon Motion Systems, Oxford, UK)
Zhang	2013	[99]	10	24		X		Xsens MVm (Xsens Technologies BV, Netherlands)	100	Motion Capture Camera System (Optotrak 3020, NDI, Waterloo, ON)
Zijlstra	2003	[101]	15	23	X			Kistler accelerometer	100	Instrumented Treadmill (No Brand Reported)

*Information not reported

Table 3 Details of studies assessing reliability for spatiotemporal (ST), kinematic (KIN), and other biomechanical outcomes (OTHER)

Author	Year	Ref	n	Age	ST	KIN	OTHER	Sensor	Hz
Al-Amri	2018	[21]	24	35		X		Xsens MTw Awinda (Xsens Technologies BV, Netherlands)	60
Bautmans	2011	[27]	20O/20Y	79/22	X		X	DynaPort (McRoberts BV, The Hague, The Netherlands)	100
Charlton	2019	[37]	20	28.3		X		Invensense MPU6050 (San Jose, CA)	100
Gorelick	2009	[46]	8F/10M	25/31	X			IDEEA (MiniSun LLC., Fresno, CA)	*
Hamacher	2014	[49]	19	71	X			Xsens MTw (Xsens Technologies BV, Netherlands)	75
Hamacher	2015	[50]	17O/12Y	71/26			X	Xsens MTw (Xsens Technologies BV, Netherlands)	75
Hartmann	2009	[52]	23	73	X			DynaPort (McRoberts BV, The Hague, The Netherlands)	100
Henriksen	2004	[53]	20	35	X		X	Meac-x (Mega Electronics Ltd., Kuopio, Finland)	250
Kavanagh	2006	[58]	8	23		X		ADXL202 (Analog Devices)	250
Kluge	2017	[60]	11	34	X			Shimmer (Shimmer Sensing, Dublin, IR)	102
Liikavainio	2007	[64]	10	29	X		X	Meac-x (Mega Electronics Ltd., Kuopio, Finland)	2000
Lyytinen	2016	[67]	9	23			X	Meac-x (Mega Electronics Ltd., Kuopio, Finland)	1000
Manor	2018	[69]	14	30	X			iPhone 4 s (Apple, CA)	100
Moe-Nilssen	1998	[73]	19	23			X	Logger Technologi HB (Ostragardsvagen, Sweden)	128
Nishiguchi	2012	[74]	30	21			X	Sony Ericsson, Xperia SO-01B	32
Orlowski	2017	[76]	25	26	X			Shimmer (Shimmer Sensing, Dublin, IR)	102
Reynard	2014	[78]	95	44			X	Physilog (BioAGM, CH)	200
Saremi	2006	[80]	12	31	X			IDEEA (MiniSun LLC., Fresno CA)	32
Senden	2009	[84]	24	21–60			X	McRoberts BV (Hague, Netherlands)	100
Teufl	2019	[89]	24	*	X			Xsens MTw Awinda (Xsens Technologies BV, Netherlands)	60
Teufl	2018	[90]	28	24		X		Xsens MVN (Xsens Technologies BV, Netherlands)	60
van der Straaten	2018	[94]	20	63		X		Xsens MVN (Xsens Technologies BV, Netherlands)	*
van Schooten	2013	[95]	20	29			X	DynaPort (McRoberts BV, Hague, Netherlands)	100
Washabaugh	2017	[96]	19	24	X			Opal (Mobility Lab, APDM Inc., Portland, OR)	*
Zijlstra	2013	[100]	20	74	X			DynaPort (McRoberts BV, Hague, Netherlands)	100

*Information not reported

Step time symmetry

Data from three low to moderate quality studies suggests that the validity of step time symmetry measured by IMUs placed at the back was poor (total $n = 189$; $r = 0.06$, 95% CI $[-0.17, 0.28]$; $I^2 = 55\%$; $p = 0.618$) [34, 41, 44].

Step length symmetry

Data from two low quality studies (contributing three independent study samples) suggests that the validity of step length symmetry measured by IMUs placed at the back was poor (total $n = 107$; $r = 0.06$, 95% IC $[-0.14, 0.25]$; $I^2 = 0\%$; $p = 0.571$) [41, 44].

Stance time symmetry

Data from two low quality studies (contributing three independent study samples) suggests that the validity of stance time symmetry measured by IMUs placed at the back was poor (total $n = 107$; $r = 0.19$, 95% CI $[-0.01, 0.37]$; $I^2 = 0\%$; $p = 0.058$) [41, 44].

Swing time symmetry

Data from two low quality studies (contributing three independent study samples) suggests that the validity of swing time symmetry measured by IMUs placed at the back was poor (total $n = 107$; $r = 0.13$, 95% CI $[-0.17, 0.41]$; $I^2 = 56\%$; $p = 0.395$) [41, 44].

Reliability

Overall, a total of 15 spatiotemporal outcomes, 3D lower limb kinematics, and 8 other biomechanical outcomes were assessed across the 25 studies that examined IMU reliability (See Table 3). From this group, 4 spatiotemporal outcomes and 1 other biomechanical outcome presented sufficient study quality and statistical outcomes for meta-analysis (Fig. 5), but no kinematic outcomes were able to be pooled. Similar to validity, the inability to pool many outcomes was due to either a limited number of studies or, in many cases, a lack of consistency in data reporting. Studies that were unable to be pooled were qualitatively summarized by outcomes and placements in Supplementary

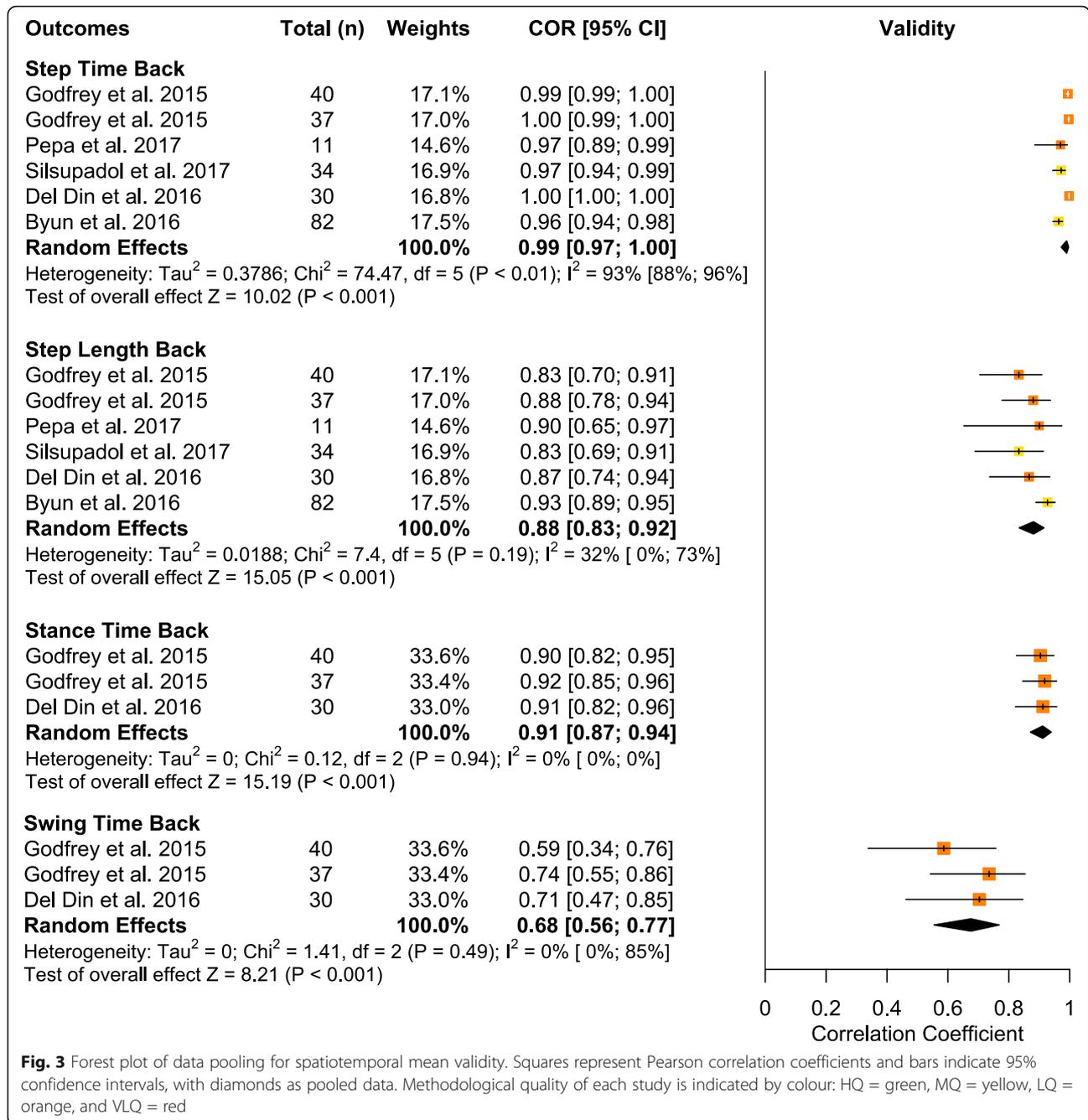


Table 4 for spatiotemporal outcomes, Supplementary Table 5 for kinematic outcomes, and Supplementary Table 6 for other biomechanical outcomes.

Quantitative pooling of spatiotemporal outcomes for reliability

Stride time

Data from three low quality studies suggests that the reliability of stride time measured by IMUs placed at the

foot was excellent (total $n = 38$; ICC = 0.92, 95% CI [0.86, 0.96]; $I^2 = 0\%$; $p < 0.001$) [49, 60, 96].

Stride length

Data from three low quality studies suggests that the reliability of stride length measured by IMUs placed at the foot was excellent (total $n = 38$; ICC = 0.94, 95% CI [0.89, 0.97]; $I^2 = 0\%$; $p < 0.001$) [49, 60, 96].

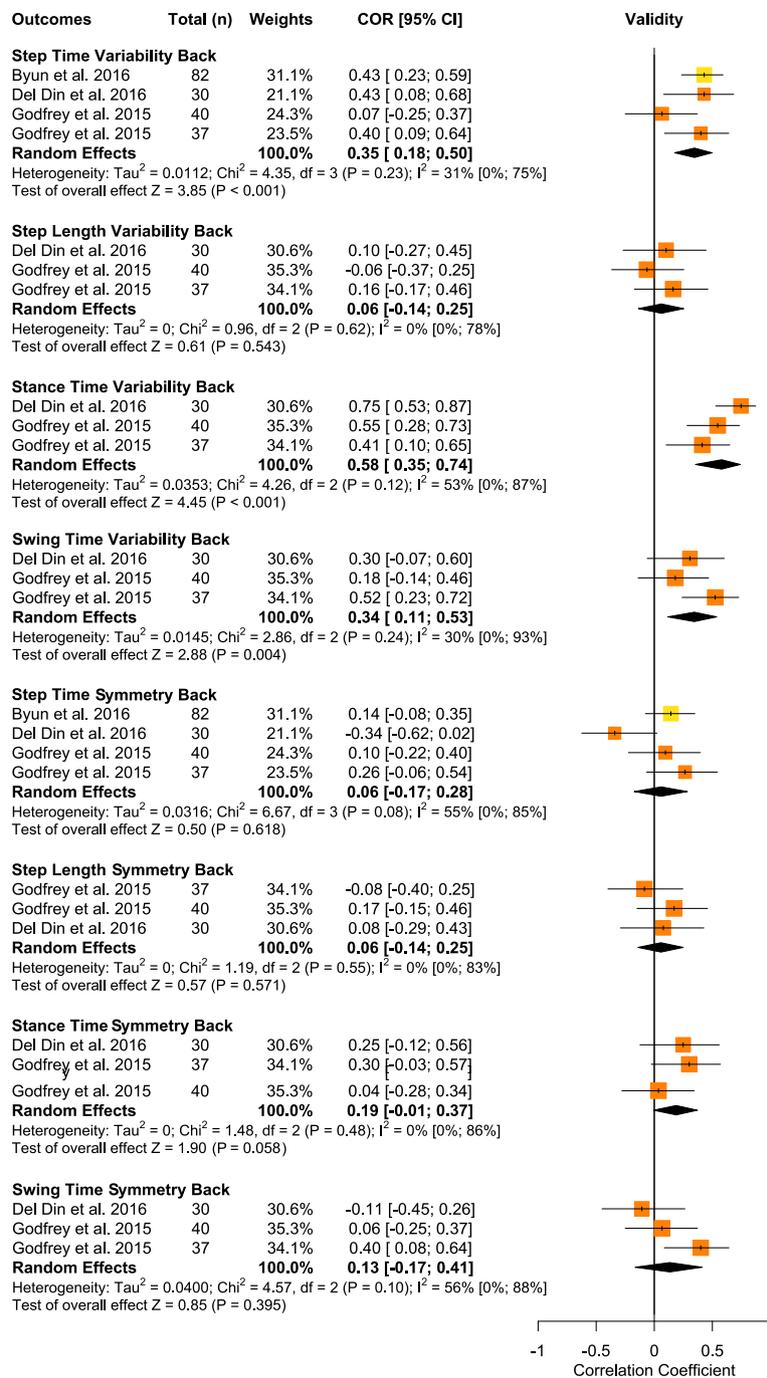


Fig. 4 Forest plot of data pooling for spatiotemporal variability and symmetry validity. Squares represent Pearson correlation coefficients and bars indicate 95% confidence intervals, with diamonds as pooled data. Methodological quality of each study is indicated by colour: HQ = green, MQ = yellow, LQ = orange, and VLQ = red

Stance time

Data from three low quality studies suggests that the reliability of stance time measured by IMUs placed at the foot was good (total $n = 38$; ICC = 0.85, 95% CI [0.72, 0.92]; $I^2 = 0\%$, $p < 0.001$) [49, 60, 96].

Swing time

Data from three low quality studies suggests that the reliability of swing time measured by IMUs placed at the foot was good (total $n = 38$; ICC = 0.89, 95% CI [0.78, 0.95]; $I^2 = 4\%$; $p < 0.001$) [49, 60, 96].

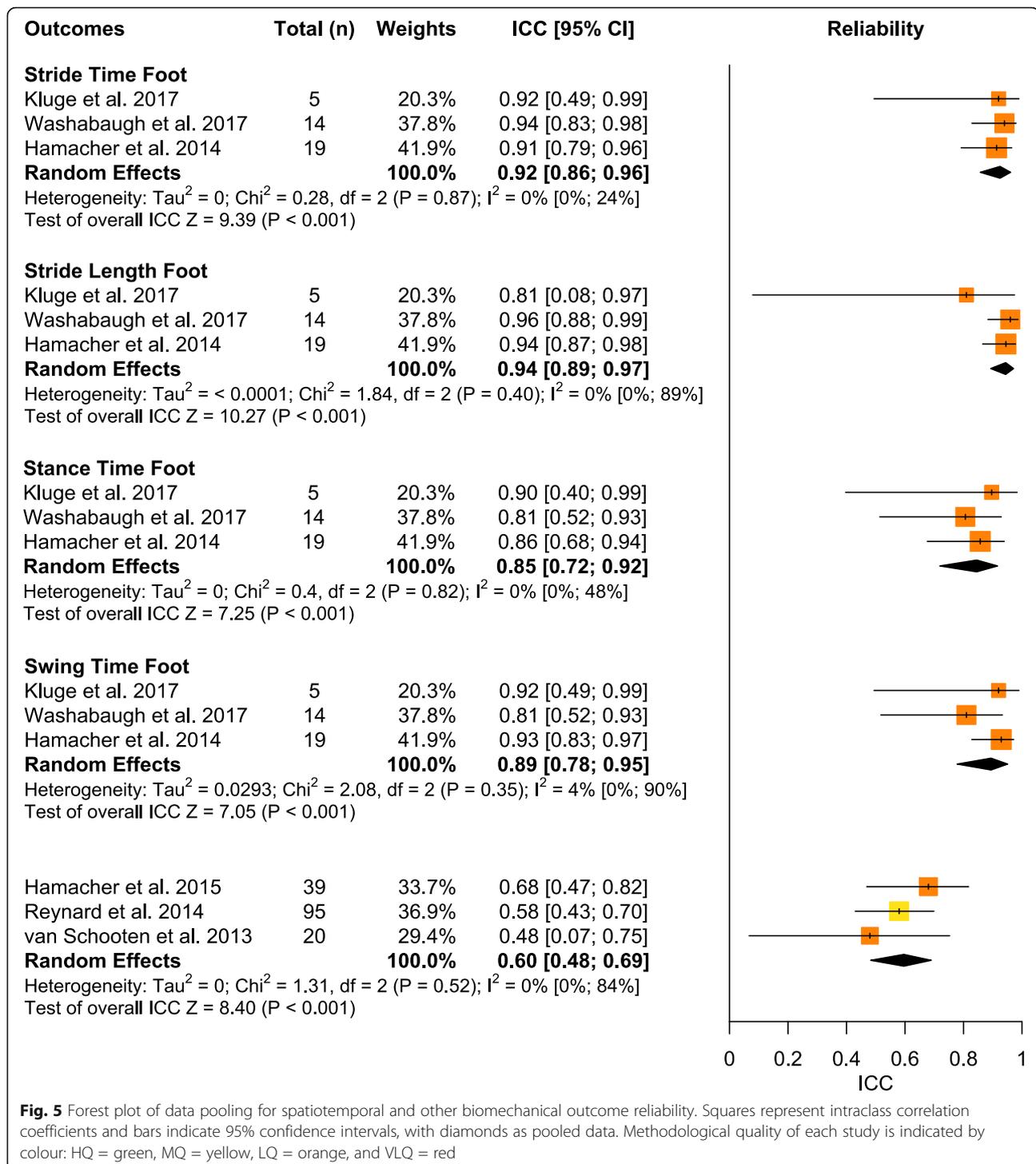


Fig. 5 Forest plot of data pooling for spatiotemporal and other biomechanical outcome reliability. Squares represent intraclass correlation coefficients and bars indicate 95% confidence intervals, with diamonds as pooled data. Methodological quality of each study is indicated by colour: HQ = green, MQ = yellow, LQ = orange, and VLQ = red

Quantitative pooling of other biomechanical outcomes for reliability

Local dynamic stability

Data from three low to moderate quality studies suggests that the reliability of a local dynamic stability outcome, namely short-term, maximum Lyapunov exponent in the mediolateral axis, measured by IMUs placed at the back

was moderate (total $n = 154$; ICC = 0.60, 95% CI [0.48, 0.69]; $I^2 = 0\%$; $p < 0.001$) [50, 78, 95].

Discussion

The aim of this review was to determine the validity and reliability of biomechanical outcomes derived from IMUs during healthy adult walking, with the hope that

we could pool results to provide valuable recommendations based on this immense body of literature. While 82 studies, examining over 100 outcomes, were included in this review, we were able to conduct meta-analysis for only 17 outcomes. Moreover, most data pooling occurred from a limited number of studies (e.g., 3–5). Nevertheless, these findings were able to provide a much-needed synthesis of the validity and reliability data for spatiotemporal, kinematic/kinetic, and other biomechanical outcomes from IMUs, as well as important recommendations for future studies in this growing field of research.

Spatiotemporal parameters presented the most fertile ground to pool results and make recommendations. Most notably, step time and stride time presented the strongest body of evidence for excellent validity and reliability. Although pooling was only possible for step time validity (back) and stride time reliability (foot), the qualitative pooling of results across the back, foot, and other placements also provide relatively consistent, but limited, evidence (based on study quality) for excellent validity and reliability. This limited, but generally consistent evidence was similarly found for good to excellent validity and reliability of step length and stride length across a variety of placements (e.g., back, shank, foot). Lastly, stance time and swing time were examined in fewer studies but were still found to present good to excellent validity and reliability in all pooled data, except swing time validity (moderate validity). Qualitative pooling of these spatiotemporal parameters across a variety of placements generally supported this conclusion with good to excellent validity and reliability. Overall, these findings are supportive of the assessment of mean spatiotemporal outcomes using IMUs, but do not clearly identify any IMU placement to be superior to another. It was only the validity of mean stride length which demonstrated a potential advantage of an IMU at the foot (e.g., excellent validity) compared to the back (e.g., good validity), with reliability metrics remaining excellent at both placements. This provides evidence for improved results of length parameters measured at the foot compared to the back, as one might expect. However, there was only a single study assessing the validity of mean stride length at the back [51] and as such this should be interpreted with caution. To this point, many of the above recommendations were defined as “limited evidence”, but we would argue that this statement of “limited evidence” is primarily based on the limited quality of studies, rather than a limitation of the sensors and outcomes themselves.

Contrary to spatiotemporal mean outcomes, the validity and reliability of spatiotemporal variability and symmetry outcomes were less favourable. Specifically, the validity of pooled variability and symmetry outcomes

(step time, step length, stance time, swing time) measured at the back were poor to moderate, with the qualitative pooling of results providing similar findings on a variety of variability outcomes and placements. The limited studies assessing reliability of these variability and symmetry outcomes fared slightly better, demonstrating poor to good reliability. In contrast to these findings, one study found excellent validity for step time variability [51]. Notably, this study also displayed the highest quality of any in this outcome category at 77.3%. Moreover, step time variability was calculated based from 4 separate walking trials, which may have improved their findings. Nevertheless, these results suggest that unlike mean spatiotemporal outcomes which may mask random error from step to step, variability measures (e.g., standard deviation of individual step or stride-based outcomes) are, by definition, more susceptible to these errors and also require strict and standardized protocols. In general, these findings are similar to a previous review of gait variability across a variety of measurement devices [102], further suggesting that it is more likely the protocol than the IMU itself that limits the validity and reliability of these variability measures. Further, while Lord et al. [102] provided some recommendations (e.g., minimum 12 steps, piloting reliability, etc.), there remains a need for better defined protocols and processing standards for spatiotemporal variability outcomes. For example, variability outcomes computed from, ideally, at least 30 continuous steps [103, 104], or to a lesser extent, multiple walking trials to reach this number [51, 105], may serve to improve the validity and reliability of these important outcomes.

Similar to recent reviews examining the validity and reliability of IMU-derived lower limb joint kinematics [12, 13], we were unable to pool any of these results. This inability to pool data remained even though we had a more homogenous cohort of studies (i.e., healthy adults during walking). Nevertheless, this improved homogeneity did allow us to draw more consistent qualitative interpretations for IMUs in healthy adult walking. For example, while our results support previous conclusions that IMUs provided better estimates of lower limb sagittal joint angles as compared to frontal or transverse angles [12, 13], we also found more consistent levels of good to excellent validity and reliability in the sagittal plane. Further, this translated to RMSEs (Supplementary Tables 2 and 5) approximately half that of previous reviews based on a variety of movements [12, 13]. Similarly, although frontal and transverse plane joint angles displayed less validity and reliability than sagittal joint angles, they were generally found to be moderate to excellent. While this supports the use of IMUs for the measurement of 3D lower limb joint angles, it should be noted that much of this evidence remains limited for the

sagittal plane, and very limited for other planes. Therefore, future research should not only focus on improving these results by examining potential sources of error (e.g., orientation estimates, anatomical calibrations, soft-tissue artifacts, etc.), but doing so in more rigorous research designs. Lastly, in addition to joint angles, we found IMUs displayed excellent validity for obtaining segment angles at the foot, shank, and thigh. Although these findings are also drawn from very limited evidence, this more simplistic approach of measuring segment orientations does not lead to compounding levels of error from multiple sensors across a joint, and as such, may be a better use of IMUs if the information of interest can be derived from a single segment [62].

While IMUs offer the unique opportunity to collect a variety of other biomechanical outcomes, only the reliability results for measures of stability, regularity, and acceleration RMSE were found to have stronger than very limited evidence. Short-term local dynamic stability (mediolateral axis), assessing complex non-linear aspects of gait variability and control [78], was the only outcome to be meta-analyzed and demonstrated moderate reliability. Stride regularity and step symmetry outcomes, assessing the consistency of acceleration waveforms using an autocorrelation procedure [106], demonstrated good and moderate reliability, respectively, but only from qualitative pooling. Further, similar to measures of gait variability, there remains limited information on the best practices for collecting these data. Lastly, acceleration RMS outcomes reported by five studies demonstrated limited evidence for good to excellent reliability in individual axes but could not be meta-analyzed due to incompatibilities in statistical parameters. Together, these results are promising for the reliability of other biomechanical measures that track human motion, but require more high-quality studies to establish better standards for the reliability of these outcomes. While the lack of validity data on these biomechanical outcomes may also be limiting, the unique nature of these outcomes may make establishing a true gold standard validity to optical systems less necessary if more high-quality reliability evidence was present.

One of the most important findings from this review is the lack of high-quality evidence and appropriate statistical outcomes utilized in much of the research in this field. The methodological quality assessment was adapted to best rate IMU validity and reliability studies, and yet many scored poorly. Underpowered and/or unjustified sample sizes were the most glaring issue, with a lack of appropriate statistical outcomes being a common problem as well. For instance, many studies simply reported mean differences as a measure of validity or reliability, which only addresses the bias of the system and not the agreement. Alternatively, reporting only Pearson's r does not

describe any potential systematic bias between measures. Therefore, we strongly advocate for all future work in this area to not only include adequate and/or justified sample sizes [107], but more appropriate statistical outcomes. Specifically, we would advise future work to include both relative (e.g., r , ICC) and absolute (e.g., LOA, SEM) statistical metrics [108, 109]. Further, Bland and Altman plots provide an excellent method to visualize the distribution of scores, but they should always be accompanied with the bias (i.e., mean difference) and an estimate of precision (i.e., standard deviation or 95% confidence interval of mean difference), as well as the limits of agreement with an estimate of precision (95% confidence interval of limits of agreement [110];). While there may be additional metrics that can support the interpretation of results (e.g., RMSE, MDC, etc.), including the aforementioned relative and absolute statistical outcomes as a minimum will provide the reader with an excellent impression of the validity and/or reliability that can be expected on biomechanical outcomes derived from IMUs.

In addition to providing recommendations, we must also acknowledge the limitations in our study. First, we chose not to include per unit measures (counts, cadence, gait speed, etc.) as these can be determined based on post collection estimates (e.g., distance travelled over a given time period = gait speed) which would confound results. Similarly, we chose not to include the direct timing of gait events (e.g., initial contact, toe-off, etc.) as these define the precursors to spatiotemporal outcomes, but not the actual outcomes themselves. Also, due to the already large scope of this review, we did not include within-session reliability or between-session reliability where the device was not removed. For example, Moenilssen [111] examined a variety of outcomes relevant to the current review, but data from that study were not included as the researchers did not remove the device between sessions, and was therefore assessing a different level of IMU reliability. Lastly, we attempted to separate outcomes by walking speed in our synthesis of data and whenever possible used normal or preferred speeds to best represent healthy adult gait. Nevertheless, there were several instances where this was not possible and, as such, some data has mixed speed results.

Future directions

The findings from this comprehensive review and meta-analysis illustrate the vast and continually growing body of literature in this field. Nevertheless, even with this large body of literature, it remains difficult to synthesize findings due to a lack of study quality and standardized protocols. Therefore, we urge the IMU community to focus on quality over quantity in research, as more poor quality, limited sample size studies will not advance the field but only convolute the results. In addition to this

general recommendation, we present four specific recommendations for future directions.

- IMUs consistently demonstrate at least moderate validity and reliability in assessing all mean spatiotemporal parameters. Further, excellent validity and reliability can be expected on measures of step and stride time and length measured at the back and lower limbs. Therefore, we do not recommend the need for future studies to address the validity and/or reliability of mean step and stride time and length during walking as a primary outcome.
- Measures of spatiotemporal parameter variability from IMUs demonstrate inconsistent levels of validity and reliability. However, these inconsistencies are more likely due to variable protocols (i.e., number of steps/trials) and processing techniques, rather than a flaw in the devices themselves. Therefore, future research should seek to identify optimal and standardized protocols and processing techniques best suited to assess measures of gait variability with IMUs.
- While joint kinematics generally demonstrate good to excellent validity and reliability in the frontal and sagittal plane, this information is often drawn from small studies with poor statistical measures. Future research in this area must improve study designs (e.g., justified sample sizes, appropriate statistical outcomes) in order to provide more high-quality evidence and recommendations on these important outcomes.
- Additional biomechanical outcomes such as a stability, regularity, and acceleration RMS demonstrate promising reliability. Unfortunately, much like gait variability, there is a lack information on optimal and standardized protocols. Moreover, similar to joint kinematics, there is a need for more high-quality study designs. Therefore, future research should seek to address the best practices for IMU measures such as stability, regularity, and acceleration RMS using appropriate sample sizes and statistical outcomes.

Conclusion

The findings of this review demonstrate the excellent validity and reliability of IMUs for measuring mean step/stride time and length during walking, but caution the use of spatiotemporal variability and symmetry metrics without strict protocol. Further, this work tentatively supports the use of IMUs for joint angle measurement, especially in the sagittal plane, and other biomechanical outcomes such as stability, regularity, and segmental accelerations. Unfortunately, the strength of these

recommendations are limited based on the paucity of high-quality studies for each outcome. Future work should seek to address these gaps by undertaking more rigorous study designs and statistical considerations for testing the validity and reliability of IMU-derived biomechanical outcomes in walking. We have provided several recommendations for future studies that will strengthen the quality of the results and provide better insights into the validity and reliability of IMUs for gait analysis.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12984-020-00685-3>.

Additional file 1. Complete Inclusion/Exclusion Criteria.

Additional file 2. Complete Search Strategy.

Additional file 3. Critical Appraisal of Study Design for Psychometric Articles.

Additional file 4 Supplementary Table 1. Qualitative summary of validity for spatiotemporal outcomes: *r*/ICC is presented as a weight average and range of reported values, while RMSE, Bias, and LOA widths are provided as the range of reported values. Gray shading identifies outcomes that have been quantitatively pooled in the results section.

Supplementary Table 2. Qualitative summary of validity for other kinematic (and joint moment) outcomes: *r*/ICC is presented as a weight average and range of reported values, while RMSE, Bias, and LOA widths are provided as the range of reported values. **Supplementary Table 3.** Qualitative summary of validity for other biomechanical outcomes: *r*/ICC is presented as a weight average and range of reported values, while RMSE, Bias, and LOA widths are provided as the range of reported values.

Supplementary Table 4. Qualitative summary of reliability for spatiotemporal outcomes: *r*/ICC is presented as a weight average and range of reported values, while SEM, MDC, Bias, and LOA widths are provided as the range of reported values. **Supplementary Table 5.** Qualitative summary of reliability for other kinematic outcomes: *r*/ICC is presented as a weight average and range of reported values, while SEM, MDC, Bias, and LOA widths are provided as the range of reported values.

Supplementary Table 6. Qualitative summary of reliability for other biomechanical outcomes: *r*/ICC is presented as a weight average and range of reported values, while SEM, MDC, Bias, and LOA widths are provided as the range of reported values.

Abbreviations

3D: Three Dimensional; IMU: Inertial Measurement Unit; ICC: Intraclass Correlation Coefficient; *r*: Pearson Correlation Coefficient; RMSE: Root-Mean-Square Error; SEM: Standard Error of Measurement; LoA: Limits of Agreement; HQ: High Quality; MQ: Moderate Quality; LQ: Low Quality; VLQ: Very Low Quality

Acknowledgements

Not applicable.

Authors' contributions

The concept of the review was contributed by DK and MAH. The search was conducted by DK. Abstract screening was conducted by DT and CT, with full-text screening conducted by CT and DK. JFE and AG conducted the quality assessment. NMK extracted data, with support from JMC. JMC and DK analyzed data. DK wrote the majority of the manuscript, with support, edits, revisions, and final approval by all other authors.

Funding

Operational funding was provided, in part, by the Natural Sciences and Engineering Research Council of Canada (MAH). Salary support for this study was provided by the Canadian Institutes of Health Research (DK, JMC, JFE,

MAH) and the Michael Smith Foundation for Health Research (MAH). The funding sources had no role in the design, analysis, or presentation of any findings in this work.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

We have no competing interests to declare.

Author details

¹Department of Kinesiology, McMaster University, Hamilton, ON, Canada. ²Motion Analysis and Biofeedback Laboratory, University of British Columbia, Vancouver, BC, Canada. ³Graduate Programs in Rehabilitation Sciences, University of British Columbia, Vancouver, BC, Canada. ⁴Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada. ⁵The Running Clinic, Lac Beauport, QC, Canada.

Received: 21 November 2019 Accepted: 7 April 2020

Published online: 11 May 2020

References

- Baker R. Gait analysis methods in rehabilitation. *J Neuroeng Rehabil*. 2006;3:1–10.
- Wagenaar RC, Beek WJ. Hemiplegic gait: a kinematic analysis using walking speed as a basis. *J Biomech*. 1992;25:1007–15.
- Simon SR. Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. *J Biomech*. 2004;37:1869–80.
- Tao W, Liu T, Zheng R, Feng H. Gait analysis using wearable sensors. *Sensors*. 2012;12:2255–83.
- Shull PB, Jirattigalachote W, Hunt MA, Cutkosky MR, Delp SL. Quantified self and human movement: a review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. *Gait Posture*. 2014;40:11–9.
- Seel T, Raisch J, Schauer T. IMU-based joint angle measurement for gait analysis. *Sensors*. 2014;14:6891–909.
- Mayagoitia RE, Nene AV, Veltink PH. Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *J Biomech*. 2002;35:537–42.
- Iosa M, Picerno P, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis. *Expert Rev Med Devices*. 2016;4440:1–19.
- Fong DT-P, Chan Y-Y. The use of wearable inertial motion sensors in human lower limb biomechanics studies: a systematic review. *Sensors*. 2010;10:11556–65.
- Chen S, Lach J, Lo B, Yang GZ. Toward pervasive gait analysis with wearable sensors: a systematic review. *IEEE J Biomed Heal Informatics*. 2016;20:1521–37.
- Caldas R, Mundt M, Potthast W, Buarque de Lima Neto F, Markert B. A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. *Gait Posture*. 2017;57:204–10.
- van der Straaten R, De Baets L, Jonkers I, Timmermans A. Mobile assessment of the lower limb kinematics in healthy persons and in persons with degenerative knee disorders: a systematic review. *Gait Posture*. 2018;59:229–41.
- Poitras I, Dupuis F, Biemann M, Campeau-Lecours A, Mercier C, Bouyer L, et al. Validity and reliability of wearable sensors for joint angle estimation: a systematic review. *Sensors*. 2019;19:1555.
- Law M, MacDermid J. Evidence-based rehabilitation: a guide to practice. 2nd ed. Thorofare: Slack Inc; 2008.
- van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane collaboration Back review group. *Spine*. 2003;28:1290–9.
- Schwarzer G. Meta: an R package for meta-analysis. *R News*. 2007;7:40–5.
- Cooper H, Hedges L, Valentine J. The handbook of research synthesis and meta-analysis. 2nd ed. New York: Russell Sage Foundation; 2009.
- Fisher R. Statistical methods for the research worker. 12th ed. New York: Hafner Publishing Company Inc; 1954.
- Higgins J, Thompson S, Deeks J, Altman D. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557–60.
- Abhayasinghe N, Murray I, Sharif BS. Validation of thigh angle estimation using inertial measurement unit data against optical motion capture systems. *Sensors*. 2019;19:596.
- Al-Amri M, Nicholas K, Button K, Sparkes V, Sheeran L, Davies JL. Inertial measurement units for clinical movement analysis: reliability and concurrent validity. *Sensors*. 2018;18:1–29.
- Allseits E, Agrawal V, Lučarević J, Gailey R, Gaunaud I, Bennett C. A practical step length algorithm using lower limb angular velocities. *J Biomech*. 2018;66:137–44.
- Allseits E, Lučarević J, Gailey R, Agrawal V, Gaunaud I, Bennett C. The development and concurrent validity of a real-time algorithm for temporal gait analysis using inertial measurement units. *J Biomech*. 2017;55:27–33.
- Aminian K, Trevisan C, Najafi B, Dejnabadi H, Frigo C, Pavan E, et al. Evaluation of an ambulatory system for gait analysis in hip osteoarthritis and after total hip replacement. *Gait Posture*. 2004;20:102–7.
- Atallah L, Wiik A, Lo B, Cobb JP, Amis AA, Yang GZ. Gait asymmetry detection in older adults using a light ear-worn sensor. *Physiol Meas*. 2014;35.
- Backhouse MR, Hensor EMA, White D, Keenan A-M, Helliwell PS, Redmond AC. Concurrent validation of activity monitors in patients with rheumatoid arthritis. *Clin Biomech*. 2013;28:473–9.
- Bautmans I, Jansen B, Van Keymolen B, Mets T. Reliability and clinical correlates of 3D-accelerometry based gait analysis outcomes according to age and fall-risk. *Gait Posture*. 2011;33:366–72.
- Ben Mansour K, Rezzoug N, Gorce P. Analysis of several methods and inertial sensors locations to assess gait parameters in able-bodied subjects. *Gait Posture*. 2015;42:409–14.
- Benoussaad M, Sijbert B, Mombaur K, Coste CA. Robust foot clearance estimation based on the integration of foot-mounted IMU acceleration data. *Sensors*. 2016;16:1–13.
- Bertoli M, Cereatti A, Trojaniello D, Avanzino L, Pelosin E, Del Din S, et al. Estimation of spatio-temporal parameters of gait from magneto-inertial measurement units: multicenter validation among Parkinson, mildly cognitively impaired and healthy older adults. *Biomed Eng Online*. 2018;17:1–14.
- Bolink SAAN, Naisas H, Senden R, Essers H, Heyligers IC, Meijer K, et al. Validity of an inertial measurement unit to assess pelvic orientation angles during gait, sit-stand transfers and step-up transfers: comparison with an optoelectronic motion capture system. *Med Eng Phys*. 2016;38:225–31.
- Bruijn SM, Ten Kate WRT, Faber GS, Meijer OG, Beek PJ, Dieën JHV. Estimating dynamic gait stability using data from non-aligned inertial sensors. *Ann Biomed Eng*. 2010;38:2588–93.
- Bugané F, Benedetti MG, Casadio G, Attala S, Biagi F, Manca M, et al. Estimation of spatial-temporal gait parameters in level walking based on a single accelerometer: validation on normal subjects by standard gait analysis. *Comput Methods Prog Biomed*. 2012;108:129–37.
- Byun S, Han JW, Kim TH, Kim KW. Test-retest reliability and concurrent validity of a single tri-axial accelerometer-based gait analysis in older adults with normal cognition. *PLoS One*. 2016;11:1–12.
- Chalmers E, Le J, Sukhdeep D, Watt J, Andersen J, Lou E. Inertial sensing algorithms for long-term foot angle monitoring for assessment of idiopathic toe-walking. *Gait Posture*. 2014;39:485–9.
- Chapman RM, Moschetti WE, Van Citters DW. Stance and swing phase knee flexion recover at different rates following total knee arthroplasty: an inertial measurement unit study. *J Biomech*. 2019;84:129–37.
- Charlton JM, Xia H, Shull PB, Hunt MA. Validity and reliability of a shoe-embedded sensor module for measuring foot progression angle during over-ground walking. *J Biomech*. 2019;89:123–7.
- Cole MH, van den Hoorn W, Kavanagh JK, Morrison S, Hodges PW, Smeathers JE, et al. Concurrent validity of accelerations measured using a tri-axial inertial measurement unit while walking on firm, compliant and uneven surfaces. *PLoS One*. 2014;9:e98395.
- Cooper G, Sheret I, McMillian L, Silverdis K, Sha N, Hodgins D, et al. Inertial sensor-based knee flexion/extension angle estimation. *J Biomech*. 2009;42:2678–85.
- Dalton A, Khalil H, Busse M, Rosser A, van Deursen R, ÓLaighin G. Analysis of gait and balance through a single triaxial accelerometer in presymptomatic and symptomatic Huntington's disease. *Gait Posture*. 2013;37:49–54.
- Del Din S, Godfrey A, Rochester L. 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 Heal Informatics*. 2016;20:838–47.
42. Esser P, Dawes H, Collett J, Howells K. IMU: inertial sensing of vertical CoM movement. *J Biomech*. 2009;42:1578–81.
 43. Furrer M, Bichsel L, Niederer M, Baur H, Schmid S. Validation of a smartphone-based measurement tool for the quantification of level walking. *Gait Posture*. 2015;42:289–94.
 44. Godfrey A, Del Din S, Barry G, Mathers JC, Rochester L. Instrumenting gait with an accelerometer: a system and algorithm examination. *Med Eng Phys*. 2015;37:400–7.
 45. González I, López-Nava IH, Fontecha J, Muñoz-Meléndez A, Pérez-SanPablo AI, Quiñones-Urióstegui I. Comparison between passive vision-based system and a wearable inertial-based system for estimating temporal gait parameters related to the GAITrite electronic walkway. *J Biomed Inform*. 2016;62:210–23.
 46. Gorelick ML, Bizzini M, Maffioletti NA, Munzinger JP, Munzinger U. Test-retest reliability of the IDEEA system in the quantification of step parameters during walking and stair climbing. *Clin Physiol Funct Imaging*. 2009;29:271–6.
 47. Greene BR, Foran TG, McGrath D, Doheny EP, Burns A, Caulfield B. A comparison of algorithms for body-worn sensor-based spatiotemporal gait parameters to the gaitrite electronic walkway. *J Appl Biomech*. 2012;28:349–55.
 48. Greene BR, McGrath D, O'Neill R, O'Donovan KJ, Burns A, Caulfield B. An adaptive gyroscope-based algorithm for temporal gait analysis. *Med Biol Eng Comput*. 2010;48:1251–60.
 49. Hamacher D, Hamacher D, Taylor WR, Singh NB, Schega L. Towards clinical application: repetitive sensor position re-calibration for improved reliability of gait parameters. *Gait Posture*. 2014;39:1146–8.
 50. Hamacher D, Hamacher D, Singh NB, Taylor WR, Schega L. Towards the assessment of local dynamic stability of level-grounded walking in an older population. *Med Eng Phys*. 2015;37:1152–5.
 51. Hartmann A, Luzzi S, Murer K, de Bie RA, de Bruin ED. Concurrent validity of a trunk tri-axial accelerometer system for gait analysis in older adults. *Gait Posture*. 2009;29:444–8.
 52. Hartmann A, Murer K, de Bie RA, de Bruin ED. Reproducibility of spatio-temporal gait parameters under different conditions in older adults using a trunk tri-axial accelerometer system. *Gait Posture*. 2009;30:351–5.
 53. Henriksen M, Lund H, Moe-Nilssen R, Bliddal H, Danneskiold-Samsøe B. Test-retest reliability of trunk accelerometric gait analysis. *Gait Posture*. 2004;19: 288–97.
 54. Huang Y, Jirattigalachote W, Cutkosky MR, Zhu X, Shull PB. Novel foot progression angle algorithm estimation via foot-worn, magneto-inertial sensing. *IEEE Trans Biomed Eng*. 2016;63:2278–85.
 55. Hundza SR, Hook WR, Harris CR, Mahajan SV, Leslie PA, Spani CA, et al. Accurate and reliable gait cycle detection in parkinson's disease. *IEEE Trans Neural Syst Rehabil Eng*. 2014;22:127–37.
 56. Jarchi D, Wong C, Kwasnicki RM, Heller B, Tew GA, Yang GZ. Gait parameter estimation from a miniaturized ear-worn sensor using singular spectrum analysis and longest common subsequence. *IEEE Trans Biomed Eng*. 2014; 61:1261–73.
 57. Karatsidis A, Jung M, Schepers HM, Bellusci G, de Zee M, Veltink PH, et al. Musculoskeletal model-based inverse dynamic analysis under ambulatory conditions using inertial motion capture. *Med Eng Phys*. 2019;65:68–77.
 58. Kavanagh JJ, Morrison S, James DA, Barrett R. Reliability of segmental accelerations measured using a new wireless gait analysis system. *J Biomech*. 2006;39:2863–72.
 59. Kitagawa N, Ogihara N. Estimation of foot trajectory during human walking by a wearable inertial measurement unit mounted to the foot. *Gait Posture*. 2016;45:110–4.
 60. Kluge F, Gaßner H, Hannink J, Pasluosta C, Klucken J, Eskofier B. Towards mobile gait analysis: concurrent validity and test-retest reliability of an inertial measurement system for the assessment of spatio-temporal gait parameters. *Sensors*. 2017;17:1522.
 61. Kose A, Cereatti A, Della CU. Bilateral step length estimation using a single inertial measurement unit attached to the pelvis. *J Neuroeng Rehabil*. 2012;9:9.
 62. Lebel K, Boissy P, Nguyen H, Duval C. Inertial measurement systems for segments and joints kinematics assessment: towards an understanding of the variations in sensors accuracy. *Biomed Eng Online*. *BioMed Central*. 2017;16:1–16.
 63. L'Hermette M, Savatier X, Baudry L, Tourny-Chollet C, Dujardin F. A new portable device for assessing locomotor performance. *Int J Sports Med*. 2008;29:322–6.
 64. Liikavainio T, Bragge T, Hakkarainen M, Jurvelin JS, Karjalainen PA, Arokoski JP. Reproducibility of loading measurements with skin-mounted accelerometers during walking. *Arch Phys Med Rehabil*. 2007;88:907–15.
 65. Liu K, Liu T, Shibata K, Inoue Y, Zheng R. Novel approach to ambulatory assessment of human segmental orientation on a wearable sensor system. *J Biomech*. 2009;42:2747–52.
 66. Lord S, Rochester L, Baker K, Nieuwboer A. Concurrent validity of accelerometry to measure gait in Parkinsons disease. *Gait Posture*. 2008;27:357–9.
 67. Lyytinen T, Bragge T, Hakkarainen M, Liikavainio T, Karjalainen PA, Arokoski JP. Repeatability of knee impulsive loading measurements with skin-mounted accelerometers and lower limb surface electromyographic recordings during gait in knee osteoarthritic and asymptomatic individuals. *J Musculoskelet Neuronal Interact*. 2016;16:63–74.
 68. Maffioletti NA, Gorelick M, Kramers-de Quervain I, Bizzini M, Munzinger JP, Tomasetti S, et al. Concurrent validity and intrasession reliability of the IDEEA accelerometry system for the quantification of spatiotemporal gait parameters. *Gait Posture*. 2008;27:160–3.
 69. Manor B, Yu W, Zhu H, Harrison R, Lo O-Y, Lipsitz L, et al. Smartphone app-based assessment of gait during normal and dual-task walking: demonstration of validity and reliability. *JMIR mHealth uHealth*. 2018;6:e36.
 70. Mariani B, Rochat S, Büla CJ, Aminian K. Heel and toe clearance estimation for gait analysis using wireless inertial sensors. *IEEE Trans Biomed Eng*. 2012; 59:3162–8.
 71. Mariani B, Rouhani H, Crevoisier X, Aminian K. Quantitative estimation of foot-flat and stance phase of gait using foot-worn inertial sensors. *Gait Posture*. 2013;37:229–34.
 72. McGrath D, Greene BR, O'Donovan KJ, Caulfield B. Gyroscope-based assessment of temporal gait parameters during treadmill walking and running. *Sport Eng*. 2012;15:207–13.
 73. Moe-Nilssen R. Test-retest reliability of trunk accelerometry during standing and walking. *Arch Phys Med Rehabil*. 1998;79:1377–85.
 74. Nishiguchi S, Yamada M, Nagai K, Mori S, Kajiwara Y, Sonoda T, et al. Reliability and validity of gait analysis by android-based smartphone. *Telemed e-Health*. 2012;18:292–6.
 75. Ohtako Y, Sagawa K, Inooka H. A method for gait analysis in a daily living environment by body-mounted instruments. *JSME Int J Ser C*. 2001;44: 1125–32.
 76. Orlovski K, Eckardt F, Herold F, Aye N, Edelman-Nusser J, Witte K. Examination of the reliability of an inertial sensor-based gait analysis system. *Biomed Eng / Biomed Tech*. 2017;62:615–22.
 77. Pepa L, Verdini F, Spalazzi L. Gait parameter and event estimation using smartphones. *Gait Posture*. 2017;57:217–23.
 78. Reynard F, Terrier P. Local dynamic stability of treadmill walking: Intrasession and week-to-week repeatability. *J Biomech*. 2014;47:74–80.
 79. Sabatini AM, Ligorio G, Mannini A. Fourier-based integration of quasi-periodic gait accelerations for drift-free displacement estimation using inertial sensors. *Biomed Eng Online*. 2015;14:1–18.
 80. Saremi K, Marehbian J, Yan X, Regnaud JP, Elashoff R, Bussell B, et al. Reliability and validity of bilateral thigh and foot accelerometry measures of walking in healthy and hemiparetic subjects. *Neurorehabil Neural Repair*. 2006;20:297–305.
 81. Schmitz-Hübsch T, Brandt AU, Pfueller C, Zange L, Seidel A, Kühn AA, et al. Accuracy and repeatability of two methods of gait analysis – GaitRite™ and mobility lab™ – in subjects with cerebellar ataxia. *Gait Posture*. 2016;48:194–201.
 82. Sejdic E, Lowry KA, Bellanca J, Perera S, Redfern MS, Brach JS. Extraction of stride events from gait accelerometry during treadmill walking. *IEEE J Transl Eng Heal Med*. 2015;4:1–11.
 83. Selles RW, Formanoy MAG, Bussmann JBJ, Janssens PJ, Stam HJ. Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls. *IEEE Trans Neural Syst Rehabil Eng*. 2005;13:81–8.
 84. Senden R, Grimm B, Heyligers IC, Savelberg HHCM, Meijer K. Acceleration-based gait test for healthy subjects: reliability and reference data. *Gait Posture*. 2009;30:192–6.
 85. Sijobert B, Benoussaad M, Denys J, Pissard-Gibollet R, Geny C, Coste CA. Implementation and validation of a stride length estimation algorithm,

- using a single basic inertial sensor on healthy subjects and patients suffering from Parkinson's disease. *Health*. 2015;07:704–14.
86. Silsupadol P, Teja K, Lugade V. Reliability and validity of a smartphone-based assessment of gait parameters across walking speed and smartphone locations: body, bag, belt, hand, and pocket. *Gait Posture*. 2017;58:516–22.
 87. Steins D, Sheret I, Dawes H, Esser P, Collett J. A smart device inertial-sensing method for gait analysis. *J Biomech*. 2014;47:3780–5.
 88. Storm FA, Buckley CJ, Mazzà C. Gait event detection in laboratory and real life settings: accuracy of ankle and waist sensor based methods. *Gait Posture*. 2016;50:42–6.
 89. Teufl W, Lorenz M, Miezal M, Taetz B, Fröhlich M, Bleser G. Towards inertial sensor based mobile gait analysis: event-detection and spatio-temporal parameters. *Sensors*. 2019;19:38.
 90. Teufl W, Miezal M, Taetz B, Fröhlich M, Bleser G. Validity, test-retest reliability and long-term stability of magnetometer free inertial sensor based 3D joint kinematics. *Sensors*. 2018;18.
 91. Trojaniello D, Cereatti A, Della CU. Accuracy, sensitivity and robustness of five different methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the lower trunk. *Gait Posture*. 2014;40:487–92.
 92. Trojaniello D, Cereatti A, Pelosin E, Avanzino L, Mirelman A, Hausdorff JM, et al. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. *J Neuroeng Rehabil*. 2014;11:152.
 93. Trojaniello D, Ravaschio A, Hausdorff JM, Cereatti A. Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, Parkinson's disease and Huntington's disease subjects. *Gait Posture*. 2015;42:310–6.
 94. Van Der Straaten R, Timmermans A, Bruijnes AKBD, Vanwanseele B, Jonkers I, De Baets L. Reliability of 3D lower extremity movement analysis by means of inertial sensor technology during transitional tasks. *Sensors*. 2018;18.
 95. van Schooten KS, Rispens SM, Pijnappels M, Daffertshofer A, van Dieen JH. Assessing gait stability: the influence of state space reconstruction on inter- and intra-day reliability of local dynamic stability during over-ground walking. *J Biomech*. 2013;46:137–41.
 96. Washabaugh EP, Kalyanaraman T, Adamczyk PG, Claffin ES, Krishnan C. Validity and repeatability of inertial measurement units for measuring gait parameters. *Gait Posture*. 2017;55:87–93.
 97. Wundersitz DWT, Gastin PB, Richter C, Robertson SJ, Netto KJ. Validity of a trunk-mounted accelerometer to assess peak accelerations during walking, jogging and running. *Eur J Sport Sci*. 2015;15:382–90.
 98. Xia H, Xu J, Wang J, Hunt MA, Shull PB. Validation of a smart shoe for estimating foot progression angle during walking gait. *J Biomech*. 2017;61:193–8.
 99. Zhang JT, Novak AC, Brouwer B, Li Q. Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. *Physiol Meas*. 2013;34.
 100. Zijlstra A, Zijlstra W. Trunk-acceleration based assessment of gait parameters in older persons: a comparison of reliability and validity of four inverted pendulum based estimations. *Gait Posture*. 2013;38:940–4.
 101. Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture*. 2003;18:1–10.
 102. Lord S, Howe T, Greenland J, Simpson L, Rochester L. Gait variability in older adults: a structured review of testing protocol and clinimetric properties. *Gait Posture*. 2011;34:443–50.
 103. Galna B, Lord S, Rochester L. Is gait variability reliable in older adults and Parkinson's disease? Towards an optimal testing protocol. *Gait Posture*. 2013;37(4):580–5.
 104. Paterson KL, Lythgo ND, Hill KD. Gait variability in younger and older adult women is altered by overground walking protocol. *Age Ageing*. 2009;38(6):745–8.
 105. Brach JS, Perera S, Studenski S, Newman AB. The reliability and validity of measures of gait variability in community-dwelling older adults. *Arch Phys Med Rehabil*. 2008;89(12):2293–6.
 106. Moe-Nilssen R, Helbostad JL. Estimation of gait cycle characteristics by trunk accelerometry. *J Biomech*. 2004;37:121–6.
 107. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res*. 2004;13:251–71.
 108. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64:96–106.
 109. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput Biol Med*. 1990;20:337–40.
 110. Abu-Arafeh A, Jordan H, Drummond G. Reporting of method comparison studies: a review of advice, an assessment of current practice, and specific suggestions for future reports. *Br J Anaesth*. 2016;117:569–75.
 111. Moe-Nilssen R, Aaslund MK, Hodt-Billington C, Helbostad JL. Gait variability measures may represent different constructs. *Gait Posture*. 2010;32:98–101.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

