

GENERAL ORTHOPAEDICS

Validity and reliability of inertial measurement units on gait, static balance and functional mobility performance among community-dwelling older adults: a systematic review and meta-analysis

Lulu Yin^{1,*}, Xin Xu^{2,*}, Ruiyan Wang³, Feifei Li⁴, Yushan Wang⁴ and Lin Wang^{5,6}¹Key Laboratory of Exercise and Health Sciences (Shanghai University of Sport), Ministry of Education, Shanghai, China²Shanghai Normal University Tianhua College, Shanghai, China³School of Rehabilitation Science, Shanghai University of Traditional Chinese Medicine, Shanghai, China⁴Guangdong OPPO Mobile Telecommunications Co. Ltd, Shenzhen, Guangdong, China⁵Sports Medicine and Rehabilitation Center, Shanghai University of Sport, Shanghai, China⁶Shanghai Shangti Orthopaedic Hospital, Shanghai, ChinaCorrespondence should be addressed to L Wang: wanglin@sus.edu.cn

*(L Yin and X Xu contributed equally to this work)

- **Purpose:** This systematic review and meta-analysis investigated validity and test-retest reliability of inertial measurement units (IMUs) in gait metrics, static balance and functional mobility performance in community-dwelling older adults.
- **Methods:** Spatiotemporal/biomechanical outcomes were meta-analyzed using intraclass correlation coefficients (ICCs) or Pearson correlation coefficients (r) for validity and reliability, respectively.
- **Results:** In our systematic review of 56 articles and meta-analysis of 38 articles, the included studies varied in quality from low-to-moderate. During validity analysis, IMU-derived metrics, including walking speed, cadence, step/stride time, step time variability, step/stride length and duration of sit-to-stand (STS) test/timed up, and go test (TUGT) exhibited excellent (ICCs) or good-to-excellent (r values) agreement with gold standards. In terms of reliability, excellent test-retest consistency was found for walking speed, cadence, step/stride time, stance/swing time, step/stride length during gait, individual STS duration, TUGT duration and walking speed during the 6-min walk test.
- **Conclusions:** Due to consistently high levels of validity and reliability, the present study supported the use of IMUs for measuring gait spatiotemporal outcomes. However, caution was advised when applying spatiotemporal variability and symmetry metrics. In addition, characterized by moderate-to-good validity and reliability, current review provides evidence of a neutral nature regarding the utilization of IMUs for static balance and functional mobility performance.

Keywords: validity; reliability; inertial measurement units; postural control; older adults

Introduction

Population aging is a growing global concern, with the number of individuals aged 65 and older projected to rise from 761 million in 2021 to 1.6 billion by 2050 (1). This

demographic shift brings increased health challenges, particularly falls, which contribute to injuries, fractures, functional impairments and psychological consequences

such as the fear of falling (2). Falls result from a complex interaction of biological, behavioral, environmental and socioeconomic factors (3). Postural control, involving the intricate coordination among these – encompassing the coordination of sensory, neural and motor systems – emerges as a key biological risk factor. Understanding the mechanisms underlying postural control and employing accurate assessment methods is essential for addressing fall risk and enhancing the quality of life in older adults (4, 5, 6).

Postural control is commonly assessed using observational methods, questionnaire-based assessments and instrumented measurements (7, 8, 9). While observational and questionnaire approaches are convenient for clinical settings, their reliability and validity can be affected by subjective bias (10, 11). In contrast, instrumented measurements offer greater precision but are limited by high costs, complexity and limited portability (12). Given these limitations, there is a need for innovative methods to accurately assess postural control in older adults. Wearable technologies have emerged as promising alternatives, providing both accuracy and portability for daily gait analysis (13, 14).

Inertial measurement units (IMUs), consisting of accelerometers, gyroscopes and magnetometers, are commonly used to measure postural control by capturing real-time motion data from key body segments. Despite their widespread use, the reliability and validity of IMUs in postural control measurement remain uncertain due to technical challenges such as cumulative errors, positional drift and the need for recalibration. Kobsar *et al.* (15) conducted investigation on the validity and reliability of wearable inertial sensors in the context of healthy adult walking. Zeng *et al.* (16) investigated the validity and reliability of IMUs in lower-extremity kinematics during running. In the elderly population, some systematic reviews have focused on the validation and reliability of IMUs for people with diseases, such as stroke (17), multiple sclerosis (18) and Parkinson diseases (19); few studies have summarized the application situations of IMUs for elderly people in communities. Nonetheless, a notable gap persists in the form of the dearth of systematic reviews and meta-analyses aimed at quantifying the reliability and validity of IMUs concerning gait, postural control and functional mobility in community-dwelling elderly populations. Therefore, evidence supporting the recommendation for the use of IMUs in conducting postural control tests for community-dwelling elderly individuals is currently lacking.

Therefore, the purpose of this review was to undertake a systematic review and subsequent meta-analysis to ascertain the levels of test-retest reliability and concurrent validity provided by IMUs in the assessment of gait, postural control and functional mobility performance among community-dwelling older adults.

By integrating qualitative and quantitative analyses, this study would provide sufficient evidence-based evidence and valuable recommendations for real-world use of IMUs during elderly individuals.

Materials and methods

The protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42023459956) and followed the preferred reporting items for systematic reviews and meta-analysis (PRISMA) guidelines (<https://www.prisma-statement.org/>) (20).

Search strategy

PubMed, Embase, Scopus, Cochrane library, Ovid MEDLINE and web of science electronic databases were searched from inception until Nov 30, 2024. The search terms and strategies included: (Wearable Electronic Devices (Mesh) OR Micro-Electrical-Mechanical Systems (Mesh) OR Smartphone (Mesh) OR wearable sensor* OR inertial sensor* OR inertial motion capture OR IMU* OR IMU OR IMUs OR acceleromet* OR gyrosco* OR magnetomet*) AND (spatio-temporal analysis (Mesh) OR postural balance (Mesh) OR gait* OR cadence OR step frequency OR stride frequency OR step time OR stride time OR cycle time OR contact time OR swing time OR step length OR stride length OR sit-to-stand (STS) OR chair-stand OR chair-rise OR timed up and go OR timed-up-and-go OR timed-up and go OR timed up & go OR timed-up-&-go OR Timed up-and-go OR timed up-&-go OR get up and go OR TUG OR TUGT OR six min walk OR six min walk OR six min walk OR 6 min walk OR 6 min walk OR 6 min walk OR 6 min walk OR 6 MW OR 6 MWT (6 min walk test) OR 6-MW OR 6-MWT) AND (aging OR elder* OR older adult*) AND (Reproducibility of Results (Mesh) OR data accuracy (Mesh) OR validity OR reliability OR feasibility OR repeatability OR consistency). Minor adjustments were made to different databases. The full search strategy for each database is described in Supplementary Information 1 (see section on [Supplementary materials](#) given at the end of the article).

Inclusion and exclusion criteria

Articles that met the following criteria were included in this systematic review: i) evaluated the validity or reliability of IMUs; ii) measured specific gait spatiotemporal parameters, static postural balance and functional mobility test outcomes including STS test, timed up and go test (TUGT) and 6 MWT; iii) compared the measurements captured by IMUs with those obtained using reference systems; iv) collected

data among community-dwelling older adults; and v) were published in English. Studies that only measured event detection or energy expenditure were excluded from this review. Additional details on the inclusion and exclusion criteria and definitions of spatiotemporal parameters are provided in Supplementary Information 2.

Study selection

After duplicate articles were removed, two independent reviewers screened titles and abstracts according to the eligibility criteria. Full-text screening of potentially eligible articles was performed by one author and rechecked by a second author. All reference lists and bibliographies of the retrieved studies were reviewed if relevant studies were missing in the electronic search. Disagreements were resolved by a third reviewer. According to the standards for systematic review, qualitative synthesis and quantitative synthesis were performed (21).

Qualitative analysis

Assessment of risk of bias

Risk of bias was assessed using a modified version of the Critical Appraisal of Study Design for Psychometric Articles (22). This checklist contains 12 items that assess the methodological quality of five domains: the study question, study design, measurements, analyses and recommendations. Each item comprises three descriptors. The maximum score was 24. Initially, two assessors independently reviewed two articles. For articles with any disagreements, they discussed the scoring and interpretation together to reach a consensus before proceeding with the evaluation of the remaining articles. The assessors were blinded to any identifiable information related to the studies to avoid bias in the quality assessment. Furthermore, agreement between the two assessors was calculated using the Cohen's kappa coefficient, with a 95% confidence interval (95% CI). Cohen's kappa coefficients of <0.40, 0.40–0.75 or >0.75 were regarded as poor, fair-to-good or excellent, respectively, (23).

Assessment of quality

To grade the quality of the study, a previously described classification scheme was applied, namely, high quality (HQ) with 85–100% scores, moderate quality (MQ) with 70–85% scores, low quality (LQ) with 50–70% scores and very low quality (VLQ) with <50% scores. Quality assessment scoring was used to determine the strength of the recommendations (24).

Assessment of heterogeneity

Heterogeneity was examined using Tau², Chi² and I² statistics, where, Tau² = 0 suggested no heterogeneity;

I² values of <25%, 26%–50% and >75% indicated low, moderate and high heterogeneity, respectively; and a significant Chi² indicates heterogeneity.

Data extraction

Data extraction was completed by two authors using a predefined form. The data consisted of: i) study identification information; ii) participant characteristics: sample size, sex, age, height and weight; iii) IMUs' specifications: name, manufacturer, composition, number used, placement and sample frequency; iv) reference systems used; v) study design: walking speed or distance and different conditions during postural balance; vi) specific parameters; and vii) reported statistical outcomes.

Quantitative analysis

Statistical outcomes for reliability and validity

For both reliability and validity analyses, statistical outcomes included Pearson correlation coefficient (r), coefficient of determination (r²), coefficient of multiple correlation, intraclass correlation coefficient (ICC) with 95% CI, root mean square error (RMSE), bias (mean difference), limits of agreement (LoA), coefficient of variation (CV) and standard error of the mean (SEM) (25). Outcomes shown graphically without specific values were excluded. Data pooling was predefined to focus on ICCs, r and sample size for validity and reliability.

Data pooling and dichotomization

Validity and reliability were dichotomized for data pooling, with further division based on specific outcome parameters (e.g. walking speed and step time). Data were not pooled by sensor type (e.g., accelerometer vs gyroscope) or algorithm. A single study may contribute to multiple independent pools based on validity, statistical outcomes and measured parameters. ICCs were categorized as poor (<0.500), moderate (0.500–0.749), good (0.750–0.899) or excellent (≥0.900) (26), and r as no correlation (<0.250), fair (0.250–0.500), moderate-to-good (0.500–0.750) or good-to-excellent (≥0.750) (27).

Subgroup and sensitivity analysis

Subgroup analyses were performed when at least two studies were available, exploring heterogeneity by movement speed (slow, preferred or fast), visual conditions during static balance (eyes open/closed) and IMU placement (back, shank or ankle/foot). Sensitivity analyses were conducted by sequentially removing one study at a time and recalculating summary correlation coefficients, weighted by sample size. Fisher's z-transformation was applied to stabilize variance due to the non-normality of ICCs and r (28):

$$\text{Fisher's } Z_{ICC} = 0.5 \times \ln \frac{1 + ICC}{1 - ICC}$$

$$\text{Fisher's } Z_r = 0.5 \times \ln \frac{1 + r}{1 - r}$$

$$SE_{ICC} = \frac{1}{\sqrt{n - 3/2}}$$

$$SE_r = \sqrt{\frac{1}{n - 3}}$$

$$\text{Summary } ICC/r = \frac{e^{2z} - 1}{e^{2z} + 1}$$

where, 'n' represents sample sizes, 'SE' represents the standard error and 'Z' signifies the summary Fisher's Z value (29). The data were then transformed back to ICCs or r for reporting.

Random effects model

The review manager (RevMan 5.4) was used for the meta-analysis (<https://training.cochrane.org/revman>). A random effects model was used for the meta-analysis due to the heterogeneity of the experimental conditions and population (30). Statistical significance was $P < 0.05$.

Meta-analysis interpretation and level of evidence

The results of the meta-analysis were interpreted by using the agreement metrics outlined above. Statistical results that were not included in the quantitative analysis were included in the qualitative analysis to support the interpretation. An adapted rating system from the Cochrane collaboration back-review group was used to determine the level of evidence for each parameter (Table 1) (24).

Table 1 Definitions of levels of evidence.

Level of evidence	Criteria
Strong evidence	Consistent results in HQ studies ($n \geq 2$)
Moderate evidence	Consistent results among multiple MQ studies ($n \geq 2$)
Limited evidence	Consistent results among multiple LQ studies ($n \geq 2$)
Conflicting evidence	Inconsistent results among multiple studies
Very limited evidence	Only one LQ or MQ study or multiple VLQ studies

HQ, high quality; MQ, moderate quality; LQ, low quality; VLQ, very low quality.

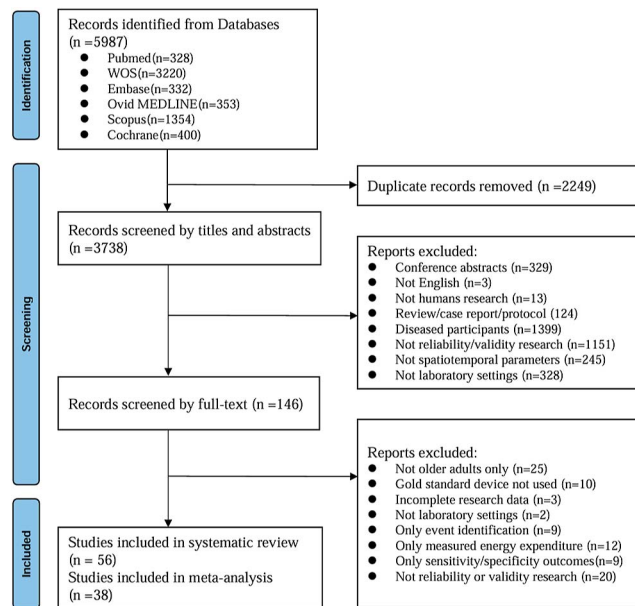


Figure 1

Flowchart of the screen outline of this systematic review and meta-analysis.

Results

Search results

The search strategy outline was recommended by the PRISMA (Fig. 1) A total of 5987 articles were identified. Following the removal of duplicate literature, screening of titles and abstracts and full-text readings, 56 articles were deemed eligible for the scope of this systematic review. 38 articles were included in the meta-analysis due to the availability of sufficient data.

Methodological quality

As shown in Table 2, 36 articles were rated as MQ (31, 32, 34, 35, 36, 37, 38, 40, 41, 42, 43, 45, 47, 49, 50, 55, 56, 57, 59, 60, 62, 64, 65, 66, 68, 72, 73, 74, 75, 76, 77, 78, 80, 81, 83, 85) and 20 articles as LQ (33, 39, 44, 46, 48, 51, 52, 53, 54, 58, 61, 63, 67, 69, 70, 71, 79, 82, 84, 86). Notably, no articles received HQ or VLQ ratings. The agreement between the two raters reached a good/excellent level (Cohen's kappa = 0.914, 95% CI = 0.90, 0.93). The items generally rated higher score were 'Q1 background and research question', 'Q4 scope/design' and 'Q12 conclusions/recommendations'. In contrast, most of the studies performed a sample size calculation process (Q5), and a few articles described participant retention (Q6) and estimates of variance (Q11).

Characteristics of included studies

The detailed characteristics of the 56 articles are presented in Supplementary Information 3.

Table 2 Quality assessment scoring of 56 included studies.

Reference	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Total	%	Quality
Adamowicz <i>et al.</i> (31)	2	2	0	2	1	NA	2	2	2	2	1	2	18/24	75%	MQ
Alqahtani <i>et al.</i> (32)	2	2	0	2	1	2	1	2	2	2	1	2	19/24	79.16%	MQ*
Álvarez <i>et al.</i> (33)	2	2	1	0	1	NA	1	2	1	1	1	2	14/24	58.33%	LQ*
Bäcklund <i>et al.</i> (34)	2	2	1	2	1	2	1	2	2	2	1	2	20/24	83.33%	MQ
Bautmans <i>et al.</i> (35)	2	2	1	2	1	NA	1	2	1	2	2	2	18/24	75%	MQ*
Bochicchio <i>et al.</i> (36)	2	2	2	2	1	NA	2	2	1	2	2	2	20/24	83.33%	MQ*
Burton <i>et al.</i> (37)	2	2	1	2	1	NA	1	2	2	2	1	2	18/24	75%	MQ
Byun <i>et al.</i> (38)	2	2	2	2	1	NA	1	2	2	2	1	2	19/24	79.16%	MQ*
Byun <i>et al.</i> (39)	2	2	1	2	1	NA	1	2	2	1	2	0	16/24	66.67%	LQ*
Cerrito <i>et al.</i> (40)	2	2	1	2	1	NA	1	2	1	2	2	2	18/24	75%	MQ*
Chan <i>et al.</i> (41)	2	2	2	2	1	NA	1	2	1	2	1	1	17/24	70.83%	MQ*
Cole <i>et al.</i> (42)	2	2	1	2	0	NA	1	2	1	2	2	2	17/24	70.83%	MQ
Contreras <i>et al.</i> (43)	2	2	1	2	2	NA	2	2	2	2	1	2	20/24	83.33%	MQ*
De Groote <i>et al.</i> (44)	2	1	0	2	1	NA	1	2	2	1	2	2	16/24	66.67%	LQ*
Digo <i>et al.</i> (45)	2	2	1	2	0	NA	2	2	1	2	1	2	17/24	70.83%	MQ*
Donath <i>et al.</i> (46)	1	1	1	2	1	NA	1	2	1	2	1	2	15/24	62.50%	LQ*
Ensink <i>et al.</i> (47)	2	2	2	2	1	NA	1	2	2	2	2	2	20/24	83.33%	MQ
Ferrari <i>et al.</i> (48)	2	2	2	2	0	NA	1	2	2	1	1	2	17/24	70.83%	LQ
Foster <i>et al.</i> (49)	2	2	1	2	1	NA	1	2	1	2	2	2	18/24	75%	MQ
Fudickar <i>et al.</i> (50)	2	2	1	2	2	NA	2	2	2	2	1	2	20/24	83.33%	MQ*
Greene <i>et al.</i> (51)	2	1	1	2	1	NA	2	2	2	1	0	2	16/24	66.67%	LQ*
Grimpampi <i>et al.</i> (52)	2	2	1	2	2	NA	2	2	2	1	0	1	17/24	70.83%	LQ
Hamacher <i>et al.</i> (53)	1	1	0	2	0	NA	2	2	1	2	0	1	13/24	54.17%	LQ*
Hartmann <i>et al.</i> (54)	2	1	1	2	1	NA	1	2	1	2	1	2	16/24	66.67%	LQ*
Hartmann <i>et al.</i> (55)	2	1	1	2	1	NA	2	2	2	2	2	2	19/24	79.16%	MQ*
Hellmers <i>et al.</i> (56)	2	2	1	2	1	NA	1	2	2	2	1	2	18/24	75%	MQ*
Kobsar <i>et al.</i> (57)	2	2	2	2	1	NA	1	2	1	2	1	2	18/24	75%	MQ
Koose <i>et al.</i> (58)	2	2	0	1	1	NA	1	1	2	1	2	2	15/24	62.50%	LQ*
Kuntapun <i>et al.</i> (59)	2	2	2	2	1	NA	1	2	2	1	0	2	17/24	70.83%	MQ*
Maganja <i>et al.</i> (60)	2	2	1	2	2	NA	1	2	2	2	2	2	20/24	83.33%	MQ
Maggio <i>et al.</i> (61)	2	2	1	2	1	NA	0	2	1	1	0	2	14/24	58.33%	LQ*
Magistro <i>et al.</i> (62)	2	2	2	1	1	2	1	2	1	1	1	2	18/24	75%	MQ*
Mancini <i>et al.</i> (63)	2	1	2	2	0	NA	1	2	2	1	1	2	16/24	66.67%	LQ*
Marques <i>et al.</i> (64)	2	2	2	2	1	NA	1	2	1	2	2	2	19/24	79.16%	MQ*
Matikainen-Tervola <i>et al.</i> (65)	2	2	2	2	0	1	1	2	2	2	2	2	20/24	83.33%	MQ*
Micó-Amigo <i>et al.</i> (66)	2	2	2	2	1	NA	1	2	2	1	1	2	18/24	75%	MQ*
Motti Ader <i>et al.</i> (67)	2	2	1	2	1	NA	2	2	2	1	0	2	17/24	70.83%	LQ*
Orange <i>et al.</i> (68)	2	2	0	2	1	NA	1	2	2	2	2	2	18/24	75%	MQ*
Ozinga <i>et al.</i> (69)	2	2	0	2	0	NA	1	2	2	1	2	0	14/24	58.33%	LQ
Pedrero-Sánchez <i>et al.</i> (70)	2	2	0	1	1	NA	1	2	2	1	0	2	14/24	58.33%	LQ
Peller <i>et al.</i> (71)	2	2	0	1	1	NA	0	1	2	1	0	2	12/24	50%	LQ
Phillips <i>et al.</i> (72)	2	1	1	2	1	NA	1	2	1	2	2	2	17/24	70.83%	MQ
Pooranawatthanakul <i>et al.</i> (73)	2	2	1	2	2	NA	2	2	2	2	1	2	20/24	83.33%	MQ*
Rantalainen <i>et al.</i> (74)	2	2	1	2	1	NA	1	2	1	2	2	2	18/24	75%	MQ*
Rantalainen <i>et al.</i> (75)	2	2	1	2	1	NA	1	2	1	2	2	2	18/24	75%	MQ*
Regterschot <i>et al.</i> (76)	2	2	2	2	2	NA	1	2	2	2	1	2	20/24	83.33%	MQ*
Regterschot <i>et al.</i> (77)	2	2	2	2	1	NA	1	2	2	2	2	2	20/24	83.33%	MQ
Rogan <i>et al.</i> (78)	2	2	1	2	1	NA	1	2	2	2	1	2	18/24	75%	MQ*
Rüdiger <i>et al.</i> (79)	2	2	1	2	1	NA	1	2	2	1	0	2	16/24	66.67%	LQ
Rudisch <i>et al.</i> (80)	2	2	1	2	1	NA	1	2	2	2	2	2	19/24	79.16%	MQ*
Saunders <i>et al.</i> (81)	2	2	2	2	1	NA	1	2	1	2	1	2	18/24	75%	MQ*
Smith <i>et al.</i> (82)	2	2	1	2	0	NA	1	2	1	2	1	2	16/24	66.67%	LQ*
Song <i>et al.</i> (83)	2	2	2	2	1	NA	1	2	1	2	1	2	18/24	75%	MQ
Walgaard <i>et al.</i> (84)	2	1	1	2	1	NA	1	2	2	1	1	2	16/24	66.67%	LQ
Werner <i>et al.</i> (85)	2	2	1	2	2	2	0	2	2	2	2	2	21/24	87.5%	MQ*
Zhang <i>et al.</i> (86)	2	2	1	2	1	NA	1	2	1	2	0	2	16/24	66.67%	LQ*

*Studies included in the meta-analysis.

NA, not mentioned; LQ, low quality; MQ, medium quality.

A total of 2658 community-dwelling adults were included, comprising 1013 males and 1562 females (gender composition was not reported in Mancini *et al.* (63), Pedrero-Sánchez *et al.* (70) and Rogan *et al.* (78)). The sample size of most studies ranged from 12 to 197 participants. Some studies focused specifically on subgroups, such as healthy older adults or elderly individuals with chronic diseases. To minimize confounding factors of diseases, this study only included data from healthy older adults. The studies spanned from 2009 to 2024, reflecting the recent applications of IMU technology in gait and balance assessment. The study designs primarily consisted of cross-sectional studies and cohort studies.

The most common applications of IMU systems are as follows: smartphones ($n = 6$) (41, 44, 59, 68, 73, 87), Dynaport ($n = 5$) (35, 54, 55, 66, 84), followed by Xsens ($n = 4$) (45, 47, 53, 63), Shimmer ($n = 3$) (51, 67, 82), Fitbit ($n = 3$) (37, 60, 72), ADXL ($n = 3$) (32, 57, 62) and Apple devices ($n = 3$) (58, 69, 83).

Of these studies, 21 exclusively utilized the accelerometer (32, 35, 37, 49, 54, 55, 57, 58, 59, 60, 61, 62, 63, 64, 72, 74, 75, 79, 81, 83, 86), while 12 explicitly mentioned employing both accelerometer and gyroscope components in their data analysis (38, 39, 44, 51, 66, 67, 69, 70, 78, 82, 84). Furthermore, one study utilized a gyroscope and magnetometer (36) and another study employed a triaxial accelerometer and a rotation vector sensor (40). In contrast, the remaining 19 studies either did not provide explicit information on the components used (31, 33, 34, 41, 42, 47, 52, 53, 68, 71, 73, 80, 85, 87, 88) or implied they used all accelerometer, gyroscope and magnetometer components for the data analysis (45, 46, 50, 56, 76, 77, 89).

Across these studies, the most common sampling frequency was sequentially 100 Hz ($n = 12$) (35, 42, 47, 50, 54, 55, 56, 57, 66, 69, 70, 84, 87), 50 Hz ($n = 9$) (31, 45, 59, 62, 63, 73, 76, 77, 86) and 500 Hz ($n = 5$) (36, 44, 46, 78, 88). Most of the studies employed either one ($n = 34$) (32, 35, 36, 38, 39, 41, 42, 44, 50, 52, 53, 54, 55, 56, 57, 58, 61, 63, 64, 66, 68, 69, 70, 71, 73, 74, 76, 79, 81, 83, 84, 85, 86, 88) or two ($n = 15$) IMU sensors (34, 37, 40, 46, 51, 59, 62, 66, 67, 75, 77, 78, 80, 82, 87).

The most commonly utilized inertial wearable locations were the back ($n = 30$) (31, 32, 35, 38, 39, 40, 42, 44, 45, 47, 50, 52, 54, 55, 56, 57, 58, 59, 61, 63, 66, 69, 70, 72, 75, 81, 84, 88, 89), predominantly located at the level of the 3rd–5th lumbar vertebra, varying locations of chest (31, 40, 41, 47, 51, 68, 73, 77, 86), wrist (31, 37, 49, 60, 62, 64, 72), arm (62, 79), thigh (36, 49, 51, 59, 60, 76, 77, 83, 85, 87), shank (34, 45, 67, 82, 89), ankle (45, 72, 74, 75, 78), foot (33, 47, 53, 66, 80) and shoes (46).

Qualitative and quantitative synthesis

Validity

As shown in Supplementary Table 1 presented in Supplementary Information 5, during gait-related

articles, a total of 18 spatiotemporal outcomes and one biomechanical outcome were extracted from 21 articles (33, 34, 38, 39, 42, 45, 49, 54, 59, 60, 61, 62, 66, 72, 74, 75, 78, 79, 80, 87, 89), which assessed the validity of IMUs. Ten spatiotemporal outcomes had sufficient statistical qualifications for data pooling (Table 3). Walking speed, cadence, step/stride time, step time variability, stance time and step/stride length during gait exhibited excellent (ICCs) or good-to-excellent (r values) agreement with gold standards. Notably, subgroup analysis results showed significant effect based on walking speed on validity of walking speed ($P < 0.001$), step time ($P < 0.001$) and step length ($P < 0.001$). However, moderate-to-good agreement for stride time SD (standard deviation) and poor agreement for swing time were observed during walking gait. Forest plots of quantitative analysis are presented in Supplementary Information 4 Figs 1, 2, 3, 4, 5, 6, 7, 8.

In static balance-related articles, as shown in Supplementary Table 2 presented as Supplementary Information 5, seven biomechanical outcomes were extracted from three articles (44, 69, 73), which assessed the IMU validity. The validity of the measurements varied, with some showing strong agreement ($r > 0.9$), such as ellipsoid volume and RMS (root mean square) COM (center of mass), while others, such as mean acceleration and RMS acceleration, exhibited moderate validity (r between 0.7 and 0.9). No outcomes could be pooled because of the lack of consistency among the outcomes.

During the STS test, as shown in Supplementary Table 3 during Supplementary Information 5, nine biomechanical outcomes were extracted from nine articles (31, 36, 40, 41, 50, 64, 68, 77, 83) that assessed IMU validity. The validity of measurements varied from moderate-to-excellent. Duration of STS test exhibited excellent (ICCs) or good-to-excellent (r values) agreement with gold standards (Table 3). However, moderate-to-good agreement was observed for mean power and mean velocity of STS (Table 3). Forest plots of quantitative analysis are presented in Supplementary Information 4 Fig. 9.

During TUGT, as shown in Supplementary Table 4 presented in Supplementary Information 5, six biomechanical outcomes were extracted from five articles (41, 50, 56, 62, 84), which assessed the validity. Only TUGT duration exhibited excellent (ICCs) or good-to-excellent (r values) agreement with gold standards (Table 3). Forest plots of quantitative analysis are presented in Supplementary Information 4 Fig. 9.

During the 6 MWT, as shown in Supplementary Table 5 presented in Supplementary Information 5, five spatiotemporal outcomes were extracted from three articles (37, 46, 57) that assessed validity. No outcomes could be pooled because of the lack of consistency among the outcomes.

Table 3 Quantitative pooling results for validity of outcomes derived from IMUs.

Outcomes	Articles			Results		I ²	Subgroup analysis, P [†]		Sensitivity analysis
	QTY	n	Reference	Summary	ICC/r (95%CI)		Locations	Speed	
Gait									
Walking speed	MQ	5	(37, 73, 77, 79, 88)	Excellent	0.954 (0.903,0.978)	98%	0.59	<0.001	Stable
	LQ	2	(38, 53)						
	LQ	1	(60)	Good-to-excellent	0.869 (0.744, 0.937)*	88%	NA	NA	Stable
Cadence	MQ	3	(37, 58, 86)	Excellent	0.992 (0.947, 0.999)	98%	0.57	0.11	Stable
	LQ	1	(53)						
	LQ	1	(37)	Good-to-excellent	0.993 (0.848,0.999)*	98%	NA	NA	Stable
Step time	MQ	2	(58, 86)						
	MQ	4	(37, 65, 77, 88)	Excellent	0.960 (0.854,0.990)	97%	0.75	<0.001	NA
	LQ	1	(53)						
Step time variability	MQ	3	(37, 44, 58)	Good-to-excellent	0.965 (0.903, 0.987)*	93%	0.19	0.90	Stable
	MQ	2	(37, 53)	Excellent	0.998 (0.995, 0.999)	74%	NA	NA	NA
	MQ	4	(73, 74, 79, 88)	Excellent	0.997 (0.990, 0.999)	89%	NA	NA	Stable
Stride time	MQ	4	(73, 74, 79, 88)	Excellent	0.997 (0.990, 0.999)	89%	NA	NA	Stable
Stride time SD	MQ	2	(73, 74)	Moderate	0.565 (−0.327, 0.925)	94%	NA	NA	NA
Stance time	MQ	2	(73, 88)	Excellent	0.973 (−0.808, 0.999)	99%	NA	NA	NA
Swing time	MQ	2	(73, 88)	Poor	0.310 (0.040, 0.537)	100%	NA	NA	NA
Step length	MQ	4	(37, 77, 79, 88)	Excellent	0.960(0.854, 0.990)	97%	0.39	<0.001	Stable
	LQ	1	(53)						
	MQ	3	(37, 58, 86)	Good-to-excellent	0.834 (0.774,0.879)*	0%	NA	NA	Unstable
Stride length	MQ	1	(73)	Excellent	0.992 (0.964, 0.998)	93%	NA	NA	NA
	LQ	1	(53)						
STS test									
Individual STS duration	MQ	2	(35, 39)	Good-to-excellent	0.989 (0.982, 0.993)*	23%	NA	NA	NA
STS mean power	MQ	2	(35, 67)	Moderate-to-good	0.726 (0.523, 0.854)*	55%	NA	NA	NA
STS mean velocity	MQ	2	(35, 67)	Moderate-to-good	0.749 (0.641, 0.831)*	0%	NA	NA	NA
Timed up-and-go test									
TUGT duration	MQ	2	(49, 55)	Good-to-excellent	0.959 (0.793, 0.992)*	99%	NA	NA	NA

IMUs, inertial measurement units; SD, standard deviation; STS, sit to stand test; TUGT, timed up and go test; MQ, moderate quality; LQ, low quality; ICCs, intraclass correlation coefficients; r, Pearson correlation coefficients; NA, no subgroup/sensitivity analysis was carried out due to insufficient data; QTY, quality.

*Indicates r (95% CI) values; †Statistically significant values are in bold.

Meta-analysis was not possible for other outcomes due to the limited quantity or consistency in data reporting; many studies only reported RMSE or even a simple bias. Articles that were not eligible for data pooling are qualitatively summarized in Supplementary Information 5.

Reliability

As shown in Supplementary Table 6 presented in Supplementary Information 5, during gait-related articles, a total of 28 spatiotemporal outcomes were extracted from 12 articles (34, 35, 38, 55, 56, 59, 60, 62, 67, 74, 75, 87) that assessed IMU reliability during different walking conditions. Fourteen spatiotemporal outcomes showed sufficient statistical qualifications for data pooling (Table 4) Excellent test-retest consistency was found for walking speed, cadence, step/stride time, stance/swing time and step/stride length during gait; good consistency was observed for swing time variability; and moderate consistency was noted for step/stride time variability, step time asymmetry and stride time SD during gait. Notably, subgroup analysis results showed significant effect based on walking

speed on validity of walking speed ($P < 0.001$). Forest plots of quantitative analysis are presented in Supplementary Information 4 Figs 10, 11, 12, 13, 14, 15, 16.

In static balance-related articles, as shown in Supplementary Table 7 presented in Supplementary Information 5, a total of ten biomechanical outcomes were extracted from eight articles (32, 44, 58, 63, 70, 71, 73, 81) that assessed IMU reliability during different balance measurement conditions. These biomechanical outcomes provided sufficient statistical qualifications for data pooling (Table 4). Good consistency was observed for RMS acceleration during double-leg stance, while moderate consistency was noted for RMS acceleration during semi-tandem stance. Forest plots of quantitative analysis are presented in Supplementary Information 4 Figs 17 and 18.

During the STS test, as shown in Supplementary Table 8 presented in Supplementary Information 5, twelve biomechanical outcomes were extracted from five articles (40, 51, 64, 76, 86) that assessed IMU reliability. Six biomechanical outcomes showed sufficient statistical

Table 4 Quantitative pooling results for reliability of outcomes derived from IMUs.

Outcomes	Articles			Results		I ²	Subgroup analysis, P [†]	Sensitivity analysis
	QTY	n	Reference	Summary	ICC (95%CI)			
Gait								
Walking speed	MQ	6	(34, 37, 54, 58, 73, 86)	Excellent	0.962 (0.935, 0.977)	90%	LS: 0.001	Stable
Cadence	MQ	4	(37, 54, 58, 86)	Excellent	0.953 (0.917, 0.973)	62%	NA	Unstable
Step time	MQ	3	(37, 54, 58)	Excellent	0.941 (0.872, 0.973)	90%	NA	Stable
	LQ	1	(66)					
Step time variability	MQ	2	(37, 54)	Moderate	0.653 (0.578, 0.716)	78%	NA	Stable
	LQ	1	(66)					
Step time asymmetry	MQ	2	(34, 37)	Moderate	0.647 (0.397, 0.808)	90%	NA	Stable
	LQ	1	(66)					
Stride time	MQ	2	(73, 74)	Excellent	0.971 (0.915, 0.990)	91%	NA	Unstable
	LQ	1	(66)					
Stride time SD	MQ	2	(73, 74)	Moderate	0.515 (0.327, 0.664)	94%	NA	NA
Stride time variability	MQ	1	(74)	Moderate	0.716 (0.438, 0.869)	89%	NA	Stable
	LQ	2	(52, 66)					
Stance time	MQ	1	(73)	Excellent	0.972 (0.917, 0.991)	92%	NA	Stable
	LQ	2	(52, 66)					
Stance time variability	LQ	2	(52, 66)	Good	0.757 (0.501, 0.892)	84%	NA	NA
Swing time	MQ	1	(73)	Excellent	0.978 (0.937, 0.992)	91%	NA	Stable
	LQ	2	(52, 66)					
Swing time variability	LQ	2	(52, 66)	Good	0.818 (0.558, 0.933)	89%	NA	NA
Step length	MQ	4	(37, 54, 58, 86)	Excellent	0.947 (0.881, 0.977)	89%	NA	Unstable
Stride length	MQ	2	(73, 86)	Excellent	0.976 (0.938, 0.991)	91%	NA	Stable
	LQ	2	(52, 66)					
Static balance test								
RMS acceleration during double leg stance	MQ	3	(31, 72, 80)	Good	0.774 (0.711, 0.824)	75%	VS: 0.68	Unstable
	LQ	3	(43, 57, 62)					
RMS acceleration during semi-tandem stance	MQ	1	(31)	Moderate	0.749 (0.611, 0.843)	79%	NA	Stable
	LQ	2	(43, 57)					
Sit to stand test								
Individual STS duration	MQ	2	(39, 75)	Good	0.770 (0.647, 0.856)	82%	MS: 0.44	Stable
	LQ	2	(50, 85)					
Total STS duration	MQ	1	(75)	Good	0.856 (0.814, 0.890)	37%	NA	Stable
	LQ	2	(50, 85)					
Stand-up duration	MQ	1	(63)	Good	0.778 (0.354, 0.938)	97%	NA	NA
	LQ	1	(50)					
STS maximal velocity	MQ	1	(75)	Good	0.869 (0.811, 0.909)	0%	MS: 0.37	NA
	LQ	1	(85)					
STS peak power	MQ	2	(35, 67)	Moderate	0.887 (0.762, 0.948)	79%	MS: 0.07	NA
STS maximal jerk	MQ	1	(75)	Moderate	0.716 (0.572, 0.814)	37%	MS: 0.05	NA
	LQ	1	(85)					
TUGT								
TUGT duration	MQ	1	(75)	Excellent	0.935 (0.793, 0.980)	88%	MS: 0.98	Stable
	LQ	2	(81, 85)					
6 MWT								
Walking speed	MQ	1	(84)	Excellent	0.989 (0.971, 0.996)	89%	MS: 0.26	Stable
	LQ	1	(45)					
Stride length	MQ	1	(84)	Good	0.894 (0.831, 0.935)	63%	MS: 0.82	Stable
	LQ	1	(45)					

IMUs, inertial measurement units; 6 MWT, 6 min walk test; RMS, root mean square; SD, standard deviation; STS, sit to stand test; TUGT, timed up and go test; MQ, moderate quality; LQ, low quality; ICCs, intraclass correlation coefficients; r, Pearson correlation coefficients; NA, no subgroup/sensitivity analysis was carried out due to insufficient data; LS, locations subgroup; VS, visual subgroup; MS, movement speed subgroup.

[†]Statistically significant values (P < 0.05) are in bold.

qualification for data pooling (Table 4). Excellent test-retest consistency was found for individual STS duration; good consistency was observed for total STS duration, stand-up duration and STS maximal velocity; and moderate consistency was noted for STS maximal jerk and STS peak power. Forest plots of quantitative analysis are presented in Supplementary Information 4 Figs 19 and 20.

During TUGT, as shown in Supplementary Table 9 presented in Supplementary Information 5, five biomechanical outcome was extracted from three articles (62, 76, 82, 86) that assessed reliability. The TUGT duration was used for data pooling (Table 4). Excellent test-retest consistency was found for TUGT duration. Forest plots of quantitative analysis are presented in Supplementary Information 4 Fig. 21.

During the 6 MWT, as shown in Supplementary Table 10 presented in Supplementary Information 5, five spatiotemporal outcomes and one biomechanical outcome were extracted from four articles (37, 46, 52, 85), which assessed reliability. Walking speed and stride length had sufficient statistical qualifications for data pooling (Table 4). Excellent test-retest consistency was found for walking speed during the 6 MWT, while good consistency was observed for stride length during the 6 MWT. Forest plots of quantitative analysis are presented in Supplementary Information 4 Figs 22 and 23.

Discussion

This review aimed to determine the concurrent validity and test-retest reliability of gait analysis, static balance and functional mobility performance derived from IMUs in community-dwelling older adults. 56 articles, examining spatiotemporal parameters and biomechanical outcomes during gait, static balance test, STS test, TUGT and 6 MWT, were included in this review. Among them, 38 articles were included in the quantitative data synthesis.

Qualitative analysis

Study quality and reporting limitations

Of the included studies, 20 were rated as LQ, 36 as MQ and none as HQ or VLQ. Only a few studies performed sample size calculations (Q5), described sensor calibration procedures (Q7) or provided variance estimates (Q11). The main factors contributing to LQ were unjustified sample sizes and the lack of multilevel statistical parameters. A few studies (47, 50, 52, 60, 73, 76, 85) provided detailed sample size calculations. Notably, the calibration process, critical for IMU data accuracy, was rarely reported across various tests (gait (45, 53, 55, 67), static balance (51), STS (31, 36, 51), TUGT (50) and 6 MWT (52)). The lack of high-quality studies and detailed reporting on sample size calculations and calibration procedures significantly limits the ability to draw firm conclusions.

Furthermore, many studies reported only relative reliability (ICCs) but lacked absolute reliability measures (SEM or LoA). Most studies presented only confidence intervals for variance estimates, often without comparison to appropriate benchmarks or standards (22). In line with Kobsar *et al.* (57), we recommend including sample size power estimation, calibration procedures and comprehensive statistical analyses. Both relative and absolute reliability should be assessed, and additional metrics such as SEM, *r* and CV could provide a more comprehensive evaluation of IMU validity and reliability.

Strengths and limitations of IMU applications

Most studies have used standardized IMU devices to measure gait parameters, which enhance the comparability of results. However, most studies failed to report detailed IMU calibration processes, which may have affected the accuracy of the measurements. Although there are currently no guidelines in the field of sports science regarding the use of IMUs, several key factors have been recommended for inclusion in the checklist, including sensor mass and dimension, sampling frequency, sensor capacity, sensor placement, etc. (90). In addition, different studies have differences in the selection of parameters. Some studies only focus on walking speed or steps count, while other studies include multiple parameters such as step frequency and step length. This difference has led to inconsistencies between studies, especially in terms of gait variability.

Selection bias and generalizability

The included studies may have had selection bias in participant recruitment, as they only selected healthy older adults who were able to complete all the tests and excluded individuals with reduced mobility or cognitive impairments. This selection bias may have led to overly optimistic results, underestimating the applicability of IMUs to the wider population. Therefore, the results of these studies should be interpreted with caution regarding their potential to generalize to the wider population.

We found that studies with larger sample sizes generally reported more consistent results, especially in terms of walking speed and stride frequency. In contrast, studies with smaller sample sizes had more fluctuating results in terms of gait variability, possibly due to statistical instability due to insufficient sample sizes. In addition, there are differences in the results of gait spatiotemporal parameters reported by institutes using different IMU brands, which may be related to the accuracy and calibration methods of the devices. Therefore, the effect of sample size and device selection on the results should be taken into account when interpreting these results.

Most studies primarily focused on community-dwelling older adults, making their results highly applicable to this population. However, the applicability of these

findings is limited when it comes to hospitalized older adults or those with severe chronic conditions, restricting the generalizability of the results. In addition, the studies included in this review were conducted in controlled laboratory settings, which may not fully reflect how IMUs perform in everyday life. Therefore, when applying these results to clinical practice, the limitations of the study setting should be taken into account and the effects of IMUs in the real world should be further validated.

Qualitative analysis

Validity analysis

Given the heterogeneity of test conditions, we conducted subgroup analyses on walking speed, IMU placement and visual conditions during static balance tests to assess their impact on reliability and validity. For most parameters, movement speed did not affect validity. However, at slower walking speeds, IMUs showed better agreement with gold standards for walking speed, step time and step length, likely due to reduced movement artifacts and signal noise, consistent with previous research (91). Subgroup analysis of test-retest reliability did not reveal significant differences across movement speeds. Regarding IMU placement, the locations did not generally impact validity. IMUs placed on the back showed higher agreement than those on the foot/ankle, except for step time validity. Surprisingly, the reliability of IMUs placed on the back was lower than those on the shank. These inconsistencies may arise from insufficient quantitative synthesis across studies to clarify the effects of IMU placement on gait and mobility outcomes. Previous studies (92, 93) suggest that foot, tibia and lumbar spine placements are suitable for reliable stride data, indicating that placement may not be a critical factor. Further research is needed to better understand the impact of IMU placement on data accuracy.

Twenty-one studies (34, 38, 39, 42, 45, 49, 54, 59, 60, 61, 62, 66, 72, 74, 75, 78, 79, 80, 87, 89) assessed the validity of IMUs for measuring spatiotemporal and biomechanical gait parameters. Of these, ten spatiotemporal outcomes met the statistical criteria for data pooling. The high heterogeneity ($I^2 = 0-98\%$) observed in gait spatiotemporal outcomes could be attributed to variability in test conditions, including differences in IMU brand, sensor placement and walking conditions (speed, distance and surface type). For the spatiotemporal outcomes – walking speed, cadence, step time, step time variability, stride time, step length and stride length – IMUs demonstrated excellent (summary ICCs ≥ 0.954) to good-to-excellent (summary $r \geq 0.869$) agreement with gold standard measurements. These results, supported by generally limited to moderate evidence, suggest that IMUs may serve as a viable alternative to motion capture systems for accurately measuring gait parameters in older adults (57).

Four studies (44, 69, 73, 88) assessed the validity of IMUs for static balance outcomes, but no data pooling was possible due to outcome inconsistencies. Moderate-to-strong validity was observed for IMUs in static balance measurement. De Groot *et al.* (44) reported moderate correlations between smartphone and force plate parameters with mean and RMS acceleration during postural stability tests. Pooranawatthanakul *et al.* (73) found moderate-to-excellent validity for RMS acceleration during double-leg stance. Ozinga *et al.* (69) showed good-to-excellent correlations between an iPad and motion capture systems for postural stability. Although no quantitative synthesis was performed, existing evidence supports the potential of IMUs for static balance testing in older adults. Further high-quality studies are needed to confirm these findings.

Nine studies (31, 36, 40, 41, 50, 64, 68, 77, 83) assessed the validity of IMUs during the STS test. Three biomechanical outcomes showed sufficient statistical qualifications for data pooling, with low-to-moderate heterogeneity ($I^2 = 0-55\%$). For the TUGT, six outcomes were extracted from five studies (41, 50, 56, 62, 84), with good-to-excellent agreement for duration (summary $r = 0.989$ for STS, $r = 0.959$ for TUGT), but moderate-to-good agreement for STS mean power and velocity, likely due to differences in calculation methods between IMUs and gold standards. In the 6 MWT, five spatiotemporal outcomes were extracted from three studies (37, 46, 57), but no data pooling was possible due to outcome inconsistencies. Despite limited evidence, existing analysis indicated good-to-excellent agreement for most outcomes, suggesting that IMUs could be useful for monitoring of older adults' functional activity during daily life.

Reliability analysis

A total of 28 spatiotemporal outcomes were extracted from eleven studies (34, 35, 38, 55, 56, 59, 60, 62, 67, 74, 75) assessing IMU reliability across different walking conditions. Fourteen outcomes met the criteria for data pooling. Walking speed, cadence, step time, stride time, stance time, swing time, step length and stride length showed excellent test-retest reliability (summary ICCs ≥ 0.941). Swing/stance time variability demonstrated good reliability (summary ICCs ≥ 0.757), while step time variability, step time asymmetry, stride time SD and stride time variability showed moderate reliability (summary ICCs ≥ 0.501). Consistent with previous research (57), time/length-related outcomes exhibited better consistency than variance-related outcomes. This may be due to mean-based gait measures being more stable and robust than variance estimates, supporting the recommendation by Lord *et al.* (94) to collect at least 12 steps.

Ten biomechanical outcomes were extracted from eight studies (32, 44, 58, 63, 70, 71, 73, 81) assessing IMU reliability during various balance tasks. Two outcomes met the criteria for data pooling. RMS acceleration during double-leg stance showed good agreement

(summary ICC = 0.774), while semi-tandem stance showed moderate agreement (summary ICC = 0.749). Traditionally, center of pressure (force plates) and center of mass (motion capture) are used for postural control measurement. IMU measurement errors are influenced by sensor algorithms, unit specifications (e.g., amplitude/frequency), placement, orientation and movement types (95). Despite these limitations, IMUs remain a rapidly advancing and valuable tool for monitoring postural control.

Twelve biomechanical outcomes from five studies (40, 51, 64, 76, 86) assessed IMU reliability during the STS test. Six outcomes met the criteria for data pooling. The individual and total STS duration, stand-up duration and maximal velocity showed good agreement (summary ICCs ≥ 0.770), while maximal jerk and peak power showed moderate agreement (summary ICCs ≥ 0.716). For the TUGT, TUGT duration from three studies (62, 76, 82, 86) demonstrated excellent agreement (summary ICC = 0.935). In the 6 MWT, walking speed and stride length showed excellent or good agreement across five outcomes from four studies (37, 46, 52, 85). Despite the generally high-reliability of IMU measures, discrepancies between IMU-derived data and gold standards (e.g. force plates or motion capture systems) in functional tests such as STS and TUGT highlight the need for standardized protocols and calibration to ensure consistent results.

In conclusion, while IMUs demonstrate strong potential for gait analysis, balance and functional mobility assessment, more robust and standardized studies are needed to optimize their application. Future research should prioritize large, well-powered studies, standardization of IMUs placement and calibration and further validation in real-world settings to ensure that these devices can be confidently utilized in clinical practice.

Limitations

This study has some limitations. First, it did not include the identification of key gait events, such as initial contact and toe-off, which are important for interpreting the results. While this review did not delve into the specifics of IMU algorithms, future studies should explore the impact of different sensor algorithms on reliability and validity, as variations in data processing approaches may substantially affect outcomes. Future research could focus on analyzing algorithmic differences between IMUs to better understand their mechanisms in sports science. Given the inconsistencies across biomechanical measures in functional mobility tests (e.g. STS and TUGT), it is important for future research to standardize outcome measures and test conditions to ensure more reliable comparisons.

Conclusion

This systematic review and meta-analysis demonstrated that IMUs exhibit excellent validity and reliability for

measuring mean spatiotemporal outcomes during gait. However, caution was warranted when using IMUs to assess spatiotemporal variability and asymmetry metrics. In the context of static balance and functional mobility tests, while biomechanical outcomes did not consistently achieve high validity and reliability, the evidence supported the use of IMUs for assessing static balance and functional mobility performance in older adults.

It was important to note that these conclusions are based on evidence from studies that are limited in number and highly heterogeneous. Therefore, factors such as IMU placement and the specific tasks being measured should be carefully considered when interpreting the results. Future research should aim to refine IMU application based on the recommendations of this review. Such efforts could improve the detection and early warning of mobility issues in real-world settings and provide valuable insights for clinicians and sports health practitioners.

Supplementary materials

This is linked to the online version of the paper at <https://doi.org/10.1530/EOR-2024-0088>.

ICMJE Statement of interest

No author has any other financial, personal or professional relationships that could be perceived as a potential conflict of interest.

Funding Statement

This work was funded by the Research and Innovation Grant for Graduate Students, Shanghai University of Sport (project no. YJSCX-2023-016).

Author contribution statement

Conceptualization was done by LY and XX. Methodology was given by RW, FL, YW and YL. LY helped with software. Validation and formal analysis was done by XX and FL. Data curation was done by LY. LY helped in writing the original manuscript. All authors have read and agreed to the published version of the manuscript.

References

- 1 United Nations. World social report 2023: leaving No one behind in an ageing world 2023. (<https://desapublications.un.org/publications/world-social-report-2023-leaving-no-one-behind-ageing-world>)
- 2 Cuevas-Trisan R. Balance problems and fall risks in the elderly. *Clin Geriatr Med* 2019 **35** 173–183. (<https://doi.org/10.1016/j.cger.2019.01.008>)
- 3 World Health Organization. *WHO Global Report On Falls Prevention In Older Age*. Geneva: World Health Organization.
- 4 Hicks C, Levinger P, Menant JC, *et al*. Reduced strength, poor balance and concern about falls mediate the relationship between knee pain and fall risk in older people. *BMC Geriatr* 2020 **20** 94. (<https://doi.org/10.1186/s12877-020-1487-2>)

- 5 Davis JC, Donaldson MG, Ashe MC, *et al.* The role of balance and agility training in fall reduction. A comprehensive review. *Eura Medicophys* 2004 **40** 211–221.
- 6 de Souza LF, Canever JB, Moreira Bd S, *et al.* Association between fear of falling and frailty in community-dwelling older adults: a systematic review. *Clin Interv Aging* 2022 **17** 129–140. (<https://doi.org/10.2147/cia.s328423>)
- 7 Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc* 1986 **34** 119–126. (<https://doi.org/10.1111/j.1532-5415.1986.tb05480.x>)
- 8 Berg KO, Wood-Dauphinee SL, Williams JI, *et al.* Measuring balance in the elderly: validation of an instrument. *Can J Public Health* 1992 **83** (Supplement 2) S7–S11.
- 9 Mathias S, Nayak US & Isaacs B. Balance in elderly patients: the “get-up and go” test. *Arch Phys Med Rehabil* 1986 **67** 387–389.
- 10 Beck Jepsen D, Robinson K, Ogliari G, *et al.* Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. *BMC Geriatr* 2022 **22** 615. (<https://doi.org/10.1186/s12877-022-03271-5>)
- 11 Tyson SF & Connell LA. How to measure balance in clinical practice. A systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions. *Clin Rehabil* 2009 **23** 824–840. (<https://doi.org/10.1177/0269215509335018>)
- 12 Miyashita TL, Cote C, Terrone D, *et al.* Detecting changes in postural sway. *J Biomech* 2020 **107** 109868. (<https://doi.org/10.1016/j.jbiomech.2020.109868>)
- 13 Montesinos L, Castaldo R & Pecchia L. Wearable inertial sensors for fall risk assessment and prediction in older adults: a systematic review and meta-analysis. *IEEE Trans Neural Syst Rehabil Eng* 2018 **26** 573–582. (<https://doi.org/10.1109/tnsre.2017.2771383>)
- 14 Subramaniam S, Faisal AI & Deen MJ. Wearable sensor systems for fall risk assessment: a review. *Front Digital Health* 2022 **4** 921506. (<https://doi.org/10.3389/fgdth.2022.921506>)
- 15 Kobsar D, Charlton JM, Tse CTF, *et al.* Validity and reliability of wearable inertial sensors in healthy adult walking: a systematic review and meta-analysis. *J NeuroEng Rehabil* 2020 **17** 62. (<https://doi.org/10.1186/s12984-020-00685-3>)
- 16 Zeng Z, Liu Y, Li P, *et al.* Validity and reliability of inertial measurement units measurements for running kinematics in different foot strike pattern runners. *Front Bioeng Biotechnol* 2022 **10** 1005496. (<https://doi.org/10.3389/fbioe.2022.1005496>)
- 17 Martino Cinnera A, Picerno P, Bisirri A, *et al.* Upper limb assessment with inertial measurement units according to the international classification of functioning in stroke: a systematic review and correlation meta-analysis. *Top Stroke Rehabil* 2023 **31** 66–85. (<https://doi.org/10.1080/10749357.2023.2197278>)
- 18 Vienne-Jumeau A, Quijoux F, Vidal PP, *et al.* Wearable inertial sensors provide reliable biomarkers of disease severity in multiple sclerosis: a systematic review and meta-analysis. *Ann Phys Rehabil Med* 2020 **63** 138–147. (<https://doi.org/10.1016/j.rehab.2019.07.004>)
- 19 Caramia C, Torricelli D, Schmid M, *et al.* IMU-based classification of Parkinson's disease from gait: a sensitivity analysis on sensor location and feature selection. *IEEE J Biomed Health Inform* 2018 **22** 1765–1774. (<https://doi.org/10.1109/jbhi.2018.2865218>)
- 20 Page MJ, Moher D, Bossuyt PM, *et al.* PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021 **372** n160. (<https://doi.org/10.1136/bmj.n160>)
- 21 Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. Finding What Works in Health Care: Standards for Systematic Reviews. Eds J Eden, L Levit, A Berg & S Morton. Washington DC, USA: National Academies Press. (<https://doi.org/17226/13059>)
- 22 Law M & MacDermid J. *Evidence-Based Rehabilitation: A Guide To Practice*. Slack Incorporated. (<https://doi.org/10.4324/9781003524106>)
- 23 Landis JR & Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977 **33** 363–374. (<https://doi.org/10.2307/2529786>)
- 24 van Tulder M, Furlan A, Bombardier C, *et al.* Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003 **28** 1290–1299. (<https://doi.org/10.1097/01.brs.0000065484.95996.af>)
- 25 Atkinson G & Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998 **26** 217–238. (<https://doi.org/10.2165/00007256-199826040-00002>)
- 26 Koo TK & Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropractic Med* 2016 **15** 155–163. (<https://doi.org/10.1016/j.jcm.2016.02.012>)
- 27 Cohen J. *Statistical Power Analysis Behavioral Sciences*, 2nd edn. Routledge. (<https://doi.org/10.4324/9780203771587>)
- 28 Cooper H, Hedges LV & Valentine JC. *The Handbook Of Research Synthesis And Meta-Analysis*. Russell Sage Foundation, 2019.
- 29 Fisher RA. Statistical methods for research workers. In *Breakthroughs in Statistics: Methodology and Distribution*, pp 66–70. Eds S Kotz & NL Johnson. New York, NY: Springer. (https://doi.org/10.1007/978-1-4612-4380-9_6)
- 30 Higgins JP, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *BMJ* 2003 **327** 557–560. (<https://doi.org/10.1136/bmj.327.7414.557>)
- 31 Adamowicz L, Karahanoglu FI, Cicalo C, *et al.* Assessment of sit-to-stand transfers during daily life using an accelerometer on the lower back. *Sensors* 2020 **20** 6618. (<https://doi.org/10.3390/s20226618>)
- 32 Alqahtani BA, Sparto PJ, Whitney SL, *et al.* Psychometric properties of instrumented postural sway measures recorded in community settings in independent living older adults. *BMC Geriatr* 2020 **20** 82. (<https://doi.org/10.1186/s12877-020-1489-0>)
- 33 Álvarez MN, Ruiz ARJ, Neira GG, *et al.* Assessing falls in the elderly population using G-STRIDE foot-mounted inertial sensor. *Sci Rep* 2023 **13** 9208. (<https://doi.org/10.1038/s41598-023-36241-x>)
- 34 Bäcklund T, Öhberg F, Johansson G, *et al.* Novel, clinically applicable method to measure step-width during the swing phase of gait. *Physiol Meas* 2020 **41** 065005. (<https://doi.org/10.1088/1361-6579/ab95ed>)
- 35 Bautmans I, Jansen B, Van Keymolen B, *et al.* Reliability and clinical correlates of 3D-accelerometry based gait analysis outcomes according to age and fall-risk. *Gait Posture* 2011 **33** 366–372. (<https://doi.org/10.1016/j.gaitpost.2010.12.003>)
- 36 Bochicchio G, Ferrari L, Bottari A, *et al.* Temporal, kinematic and kinetic variables derived from a wearable 3D inertial sensor to estimate muscle power during the 5 sit to stand test in older individuals: a validation study. *Sensors* 2023 **23** 4802. (<https://doi.org/10.3390/s23104802>)
- 37 Burton E, Hill KD, Lautenschlager NT, *et al.* Reliability and validity of two fitness tracker devices in the laboratory and home environment

- for older community-dwelling people. *BMC Geriatr* 2018 **18** 103. (<https://doi.org/10.1186/s12877-018-0793-4>)
- 38 Byun S, Han JW, Kim TH, *et al.* 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** e0158956. (<https://doi.org/10.1371/journal.pone.0158956>)
- 39 Byun S, Lee HJ, Han JW, *et al.* Walking-speed estimation using a single inertial measurement unit for the older adults. *PLoS One* 2019 **14** e0227075. (<https://doi.org/10.1371/journal.pone.0227075>)
- 40 Cerrito A, Bichsel L, Radlinger L, *et al.* Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. *Gait Posture* 2015 **41** 409–413. (<https://doi.org/10.1016/j.gaitpost.2014.11.001>)
- 41 Chan MHM, Keung DTF, Lui SYT, *et al.* A validation study of a smartphone application for functional mobility assessment of the elderly. *Hong Kong Physiother J* 2016 **35** 1–4. (<https://doi.org/10.1016/j.hknpj.2015.11.001>)
- 42 Cole MH, van den Hoorn W, Kavanagh JK, *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. (<https://doi.org/10.1371/journal.pone.0098395>)
- 43 Contreras C, Stanley EC, Deschamps-Prescott C, *et al.* Evaluation of smartphone technology on spatiotemporal gait in older and diseased adult populations. *Sensors* 2024 **24** 5839. (<https://doi.org/10.3390/s24175839>)
- 44 De Groote F, Vandevyvere S, Vanhevel F, *et al.* Validation of a smartphone embedded inertial measurement unit for measuring postural stability in older adults. *Gait Posture* 2021 **84** 17–23. (<https://doi.org/10.1016/j.gaitpost.2020.11.017>)
- 45 Digo E, Panero E, Agostini V, *et al.* Comparison of IMU set-ups for the estimation of gait spatio-temporal parameters in an elderly population. *Proc Inst Mech Eng H* 2023 **237** 61–73. (<https://doi.org/10.1177/09544119221135051>)
- 46 Donath L, Faude O, Lichtenstein E, *et al.* Mobile inertial sensor based gait analysis: validity and reliability of spatiotemporal gait characteristics in healthy seniors. *Gait Posture* 2016 **49** 371–374. (<https://doi.org/10.1016/j.gaitpost.2016.07.269>)
- 47 Ensink C, Smulders K, Warnar J, *et al.* Validation of an algorithm to assess regular and irregular gait using inertial sensors in healthy and stroke individuals. *PeerJ* 2023 **11** e16641. (<https://doi.org/10.7717/peerj.16641>)
- 48 Ferrari L, Bochicchio G, Bottari A, *et al.* Construct validity of a wearable inertial measurement unit (IMU) in measuring postural sway and the effect of visual deprivation in healthy older adults. *Biosensors* 2024 **14** 529. (<https://doi.org/10.3390/bios14110529>)
- 49 Foster JJ, Williams KL, Timmer BHB, *et al.* Concurrent validity of the garmin Vivofit®4 to accurately record step count in older adults in challenging environments. *J Aging Phys Activ* 2022 **30** 833–841. (<https://doi.org/10.1123/japa.2021-0231>)
- 50 Fudickar S, Hellmers S, Lau S, *et al.* Measurement system for unsupervised standardized assessment of timed “up & go” and five times sit to stand test in the community-A validity study. *Sensors* 2020 **20** 2824. (<https://doi.org/10.3390/s20102824>)
- 51 Greene BR, Doheny EP, McManus K, *et al.* Estimating balance, cognitive function, and falls risk using wearable sensors and the sit-to-stand test. *Wearable Technol* 2022 **3** e9. (<https://doi.org/10.1017/wtc.2022.6>)
- 52 Grimpampi E, Oesen S, Halper B, *et al.* Reliability of gait variability assessment in older individuals during a six min walk test. *J Biomech* 2015 **48** 4185–4189. (<https://doi.org/10.1016/j.jbiomech.2015.10.008>)
- 53 Hamacher D, Hamacher D, Taylor WR, *et al.* Towards clinical application: repetitive sensor position re-calibration for improved reliability of gait parameters. *Gait Posture* 2014 **39** 1146–1148. (<https://doi.org/10.1016/j.gaitpost.2014.01.020>)
- 54 Hartmann A, Luzi S, Murer K, *et al.* Concurrent validity of a trunk tri-axial accelerometer system for gait analysis in older adults. *Gait Posture* 2009 **29** 444–448. (<https://doi.org/10.1016/j.gaitpost.2008.11.003>)
- 55 Hartmann A, Murer K, de Bie RA, *et al.* Reproducibility of spatio-temporal gait parameters under different conditions in older adults using a trunk tri-axial accelerometer system. *Gait Posture* 2009 **30** 351–355. (<https://doi.org/10.1016/j.gaitpost.2009.06.008>)
- 56 Hellmers S, Izadpanah B, Dasenbrock L, *et al.* Towards an automated unsupervised mobility assessment for older people based on inertial TUG measurements. *Sensors* 2018 **18** 3310. (<https://doi.org/10.3390/s18103310>)
- 57 Kobsar D, Olson C, Paranjape R, *et al.* The validity of gait variability and fractal dynamics obtained from a single, body-fixed triaxial accelerometer. *J Appl Biomech* 2014 **30** 343–347. (<https://doi.org/10.1123/jab.2013-0107>)
- 58 Kosse NM, Caljouw S, Vervoort D, *et al.* Validity and reliability of gait and postural control analysis using the tri-axial accelerometer of the iPod touch. *Ann Biomed Eng* 2015 **43** 1935–1946. (<https://doi.org/10.1007/s10439-014-1232-0>)
- 59 Kuntapun J, Silsupadol P, Kamnardsiri T, *et al.* Smartphone monitoring of gait and balance during irregular surface walking and obstacle crossing. *Front Sports Act Living* 2020 **2** 560577. (<https://doi.org/10.3389/fspor.2020.560577>)
- 60 Maganja SA, Clarke DC, Lear SA, *et al.* Formative evaluation of consumer-grade activity monitors worn by older adults: test-retest reliability and criterion validity of step counts. *JMIR Formative Res* 2020 **4** e16537. (<https://doi.org/10.2196/16537>)
- 61 Maggio M, Ceda GP, Ticinesi A, *et al.* Instrumental and non-instrumental evaluation of 4-meter walking speed in older individuals. *PLoS One* 2016 **11** e0153583. (<https://doi.org/10.1371/journal.pone.0153583>)
- 62 Magistro D, Brustio PR, Ivaldi M, *et al.* Validation of the ADAMO Care Watch for step counting in older adults. *PLoS One* 2018 **13** e0190753. (<https://doi.org/10.1371/journal.pone.0190753>)
- 63 Mancini M, Salarian A, Carlson-Kuhta P, *et al.* ISway: a sensitive, valid and reliable measure of postural control. *J NeuroEng Rehabil* 2012 **9** 59. (<https://doi.org/10.1186/1743-0003-9-59>)
- 64 Marques DL, Neiva HP, Pires IM, *et al.* An experimental study on the validity and reliability of a smartphone application to acquire temporal variables during the single sit-to-stand test with older adults. *Sensors* 2021 **21** 2050. (<https://doi.org/10.3390/s21062050>)
- 65 Matikainen-Tervola E, Cronin N, Aartolahti E, *et al.* Validity of IMU sensors for assessing features of walking in laboratory and outdoor environments among older adults. *Gait Posture* 2024 **114** 277–283. (<https://doi.org/10.1016/j.gaitpost.2024.10.013>)
- 66 Micó-Amigo ME, Kingma I, Ainsworth E, *et al.* A novel accelerometry-based algorithm for the detection of step durations over short episodes of gait in healthy elderly. *J NeuroEng Rehabil* 2016 **13** 38. (<https://doi.org/10.1186/s12984-016-0145-6>)

- 67 Motti Ader LG, Greene BR, McManus K, *et al.* Reliability of inertial sensor based spatiotemporal gait parameters for short walking bouts in community dwelling older adults. *Gait Posture* 2021 **85** 1–6. (<https://doi.org/10.1016/j.gaitpost.2021.01.010>)
- 68 Orange ST, Metcalfe JW, Liefheith A, *et al.* Validity of various portable devices to measure sit-to-stand velocity and power in older adults. *Gait Posture* 2020 **76** 409–414. (<https://doi.org/10.1016/j.gaitpost.2019.12.003>)
- 69 Ozinga SJ & Alberts JL. Quantification of postural stability in older adults using mobile technology. *Exp Brain Res* 2014 **232** 3861–3872. (<https://doi.org/10.1007/s00221-014-4069-8>)
- 70 Pedrero-Sánchez JF, De-Rosario-Martínez H, Medina-Ripoll E, *et al.* The reliability and accuracy of a fall risk assessment procedure using mobile smartphone sensors compared with a physiological profile assessment. *Sensors* 2023 **23** 6567. (<https://doi.org/10.3390/s23146567>)
- 71 Peller A, Garib R, Garbe E, *et al.* Validity and reliability of the NIH Toolbox® standing balance test as compared to the biodex balance system SD. *Physiother Theor Pract* 2023 **39** 827–833. (<https://doi.org/10.1080/09593985.2022.2027584>)
- 72 Phillips LJ, Petroski GF & Markis NE. A comparison of accelerometer accuracy in older adults. *Res Gerontol Nurs* 2015 **8** 213–219. (<https://doi.org/10.3928/19404921-20150429-03>)
- 73 Pooranawatthanakul K & Siriphorn A. Testing the validity and reliability of a new android application-based accelerometer balance assessment tool for community-dwelling older adults. *Gait Posture* 2023 **104** 103–108. (<https://doi.org/10.1016/j.gaitpost.2023.06.016>)
- 74 Rantalainen T, Pirkola H, Karavirta L, *et al.* Reliability and concurrent validity of spatiotemporal stride characteristics measured with an ankle-worn sensor among older individuals. *Gait Posture* 2019 **74** 33–39. (<https://doi.org/10.1016/j.gaitpost.2019.08.006>)
- 75 Rantalainen T, Karavirta L, Pirkola H, *et al.* Gait variability using waist- and ankle-worn inertial measurement units in healthy older adults. *Sensors* 2020 **20** 2858. (<https://doi.org/10.3390/s20102858>)
- 76 Regterschot GRH, Zhang W, Baldus H, *et al.* Test-retest reliability of sensor-based sit-to-stand measures in young and older adults. *Gait Posture* 2014 **40** 220–224. (<https://doi.org/10.1016/j.gaitpost.2014.03.193>)
- 77 Regterschot GRH, Zhang W, Baldus H, *et al.* Accuracy and concurrent validity of a sensor-based analysis of sit-to-stand movements in older adults. *Gait Posture* 2016 **45** 198–203. (<https://doi.org/10.1016/j.gaitpost.2016.02.004>)
- 78 Rogan S, de Bie R & Douwe de Bruin E. Sensor-based foot-mounted wearable system and pressure sensitive gait analysis Agreement in frail elderly people in long-term care. *Z Gerontol Geriatr* 2017 **50** 488–497. (<https://doi.org/10.1007/s00391-016-1124-z>)
- 79 Rüdiger S, Stuckenschneider T, Abeln V, *et al.* Validation of a widely used heart rate monitor to track steps in older adults. *J Sports Med Phys Fit* 2019 **59** 1622–1627. (<https://doi.org/10.23736/s0022-4707.19.09830-x>)
- 80 Rudisch J, Jöllenbeck T, Vogt L, *et al.* Agreement and consistency of five different clinical gait analysis systems in the assessment of spatiotemporal gait parameters. *Gait Posture* 2021 **85** 55–64. (<https://doi.org/10.1016/j.gaitpost.2021.01.013>)
- 81 Saunders NW, Koutakis P, Kloos AD, *et al.* Reliability and validity of a wireless accelerometer for the assessment of postural sway. *J Appl Biomech* 2015 **31** 159–163. (<https://doi.org/10.1123/jab.2014-0232>)
- 82 Smith E, Walsh L, Doyle J, *et al.* The reliability of the quantitative timed up and go test (QTUG) measured over five consecutive days under single and dual-task conditions in community dwelling older adults. *Gait Posture* 2016 **43** 239–244. (<https://doi.org/10.1016/j.gaitpost.2015.10.004>)
- 83 Song Y, Begum M, Arthanat S, *et al.* Validation of smartphone accelerometry for the evaluation of sit-to-stand performance and lower-extremity function in older adults. *J Aging Phys Activ* 2022 **30** 3–11. (<https://doi.org/10.1123/japa.2020-0428>)
- 84 Walgaard S, Faber GS, van Lummel RC, *et al.* The validity of assessing temporal events, sub-phases and trunk kinematics of the sit-to-walk movement in older adults using a single inertial sensor. *J Biomech* 2016 **49** 1933–1937. (<https://doi.org/10.1016/j.jbiomech.2016.03.010>)
- 85 Werner C, Hezel N, Dongus F, *et al.* Validity and reliability of the Apple Health app on iPhone for measuring gait parameters in children, adults, and seniors. *Sci Rep* 2023 **13** 5350. (<https://doi.org/10.1038/s41598-023-32550-3>)
- 86 Zhang W, Regterschot GRH, Schaabova H, *et al.* Test-retest reliability of a pendant-worn sensor device in measuring chair rise performance in older persons. *Sensors* 2014 **14** 8705–8717. (<https://doi.org/10.3390/s140508705>)
- 87 Contreras C, Stanley EC, Deschamps-Prescott C, *et al.* Evaluation of smartphone technology on spatiotemporal gait in older and diseased adult populations. *Sensors* 2024 **24** 5839. (<https://doi.org/10.3390/s24175839>)
- 88 Ferrari LA-O, Bochicchio GA-O, Bottari AA-O, *et al.* Construct validity of a wearable inertial measurement unit (IMU) in measuring postural sway and the effect of visual deprivation in healthy older adults. *Biosensors* 2024 **14** 529. (<https://doi.org/10.3390/bios14110529>)
- 89 Matikainen-Tervola E, Cronin N, Aartolahti E, *et al.* Validity of IMU sensors for assessing features of walking in laboratory and outdoor environments among older adults. *Gait Posture* 2024 **114** 277–283. (<https://doi.org/10.1016/j.gaitpost.2024.10.013>)
- 90 Hughes GTG, Camomilla V, Vanwanseele B, *et al.* Novel technology in sports biomechanics: some words of caution. *Sports Biomech* 2021 **23** 393–401. (<https://doi.org/10.1080/14763141.2020.1869453>)
- 91 Napier C, Willy RW, Hannigan BC, *et al.* The effect of footwear, running speed, and location on the validity of two commercially available inertial measurement units during running. *Front Sports Active Living* 2021 **3** 643385. (<https://doi.org/10.3389/fspor.2021.643385>)
- 92 Chadeaux D, Gueguen N, Thouze A, *et al.* 3D propagation of the shock-induced vibrations through the whole lower-limb during running. *J Biomech* 2019 **96** 109343. (<https://doi.org/10.1016/j.jbiomech.2019.109343>)
- 93 Horsley BJ, Tofari PJ, Halson SL, *et al.* Does site matter? Impact of inertial measurement unit placement on the validity and reliability of stride variables during running: a systematic review and meta-analysis. *Sports Med* 2021 **51** 1449–1489. (<https://doi.org/10.1007/s40279-021-01443-8>)
- 94 Lord S, Howe T, Greenland J, *et al.* Gait variability in older adults: a structured review of testing protocol and clinimetric properties. *Gait Posture* 2011 **34** 443–450. (<https://doi.org/10.1016/j.gaitpost.2011.07.010>)
- 95 Richmond SB, Fling BW, Lee H, *et al.* The assessment of center of mass and center of pressure during quiet stance: current applications and future directions. *J Biomech* 2021 **123** 110485. (<https://doi.org/10.1016/j.jbiomech.2021.110485>)