

Review

# Digital Technologies and Biomarkers for Locomotor Capacity Assessment in Older Adults: Systematic Review

Shuhan Zhou<sup>1\*</sup>, PhD; Jundan Huang<sup>1\*</sup>, PhD; Jia Yu<sup>1</sup>, MSN; Xiaoyang Li<sup>1</sup>, PhD; Chi Zhang<sup>1</sup>, PhD; Yinan Zhao<sup>1</sup>, PhD; Mingyue Hu<sup>1</sup>, PhD; Hui Feng<sup>1,2,3</sup>, PhD

<sup>1</sup>Xiangya School of Nursing, Central South University, Changsha, China

<sup>2</sup>Oceanwide Health Management Institute, Central South University, Changsha, China

<sup>3</sup>Hunan Engineering Research Center for Intelligent Medical Care, Central South University, Changsha, China

\*these authors contributed equally

## Corresponding Author:

Hui Feng, PhD  
Xiangya School of Nursing  
Central South University  
172 Tongzipo Road, Yuelu District  
Changsha 410013  
China  
Phone: 86 15173121969  
Email: [fenghui0365@163.com](mailto:fenghui0365@163.com)

## Abstract

**Background:** Locomotor capacity, encompassing endurance, balance, muscle strength, muscle function, muscle power, and joint function of the body, is a key determinant of functional ability in older adults. Assessment tools based on digital technologies for objectively assessing locomotor capacity are increasingly being developed, but their reliability, validity, and clinical potential remain underexplored.

**Objective:** This systematic review aims to evaluate the current state of digital technologies, assess their validity and reliability for assessing locomotor capacity, and facilitate their effective implementation in clinical settings.

**Methods:** Systematic literature searches were performed in 6 electronic databases from inception to March 7, 2025. Citation lists from the included studies and gray literature from Google Scholar were additionally searched. Studies focusing on the reliability and validity of digital technologies for assessing locomotor capacity in general older adults were included. Standardized forms were used to extract information on study characteristics, participant demographics, digital technology details, and validity and reliability results. Methodological quality assessment and rating of measurement properties were conducted in accordance with the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) guidelines.

**Results:** A total of 14 studies were included, of which 13 assessed balance using inertial measurement units, smartphones, balance boards, and force plates, and 1 assessed muscle power using smartphones. Fifty-one digital biomarkers were identified, including 47 for balance and 4 for muscle power assessment. Test-retest reliability coefficients ranged from 0.016 to 0.97, and validity was context specific. Overall, 13 studies demonstrated sufficient test-retest reliability and validity, whereas 1 study was rated as insufficient for convergent validity. Methodological quality was rated as “doubtful” or “inadequate” in 11 studies.

**Conclusions:** This review provides a comprehensive summary of digital technologies for assessing locomotor capacity in older adults and identifies 51 digital biomarkers with generally acceptable reliability and validity. Unlike previous studies that focused on specific sensor types or disease-specific populations, this review integrates evidence across technologies within general older populations, providing insights into the clinical application potential of digital biomarkers as well as the key translational barriers limiting their real-world implementation. Specifically, existing digital technologies show considerable promise for early detection of functional decline, longitudinal monitoring, and informing personalized interventions. However, their clinical applicability remains constrained by limited assessment of certain locomotor components and by methodological shortcomings across current studies. Future research should prioritize rigorous, high-quality investigations that expand evaluation to a broader range of locomotor components in real-world settings while developing age-friendly tools with enhanced clinical interpretability.

**Trial Registration:** PROSPERO CRD420251074143; <https://www.crd.york.ac.uk/PROSPERO/view/CRD420251074143>

**Keywords:** locomotor capacity; mobility; digital technology; digital biomarkers; assessment tools

## Introduction

With an aging population, physical disability among older adults is becoming a pressing public health issue [1]. Locomotor capacity is a key factor for maintaining independence in daily activities [2]. The World Health Organization defines locomotor capacity as “a state (static or dynamic over time) of the musculoskeletal system that encompasses endurance, balance, muscle strength, muscle function, muscle power, and joint function of the body” [3]. In older adults, locomotion limitations increase the risk of adverse outcomes, such as frailty, falls, hospitalization, and death [4]. The prevalence of positive screenings for locomotion limitation varies from 2.8% to 52% depending on the setting [5].

Traditional methods for evaluating locomotor capacity, such as the Short Physical Performance Battery [6], exhibit several limitations, including the requirement for specialized assessor training, susceptibility to assessor bias, and ceiling effects [7,8]. Given these limitations, there is growing interest in developing assessment tools using digital technologies that offer more objective, quantifiable, and frequent monitoring of locomotor capacity [9-11]. Moreover, these frequent measures can capture short-term intraindividual fluctuations that may be early and sensitive indicators of negative trends and critical transitions [12]. For example, Kumar et al [13] used a tri-axial accelerometer worn by older adults over 48 hours to identify continuous walking bouts of 20, 30, 40, 50, and 60 seconds without pauses. They found that gait variability, asymmetry, and irregularity produced a sensitivity of 76.8% and specificity of 80% between robust and physical prefrail or frail individuals. By identifying these early signs from an individual’s typical mobility pattern, digital biomarkers have the potential to contribute to early risk prediction models for functional decline and disability, offering insights that conventional assessments may miss [14]. Digital technologies for measuring digital biomarkers can be categorized into 3 types [15]. First, wearable devices, such as wrist-worn accelerometers and portable sensors, provide continuous monitoring of movement patterns. Second, smartphones or applications are used for assessment. Third, nonvisual technologies enable passive and unobtrusive measurement in sensor-equipped environments.

Growing empirical evidence highlights the potential of digital technologies for assessing locomotor capacity. For instance, Burq et al [16] developed smartwatch-based unsupervised active tests of motor function in Parkinson disease and established the analytical validity of associated digital measures. Zhu et al [10] explored the use of inertial measurement units (IMUs) for assessing lower limb muscle strength during sit-to-stand transfers in older adults. Smartphone-based accelerometry data measured during sit-to-stand tests were also useful to identify subtle changes in locomotor capacity [17]. Several recent reviews [18-20] have summarized the application of digital technologies for assessing

locomotor capacity. However, these reviews primarily focused on specific sensor types or disease populations, thereby overlooking other viable technologies and providing limited insight into assessments conducted in general older adult populations. In addition, the lack of robust evidence regarding the reliability and validity of digital assessment tools significantly limits their clinical utility [21]. The gap between research use and clinical application remains insufficiently established.

Therefore, this systematic review aimed to summarize digital technologies and biomarkers and evaluate their performance in assessing locomotor capacity in general older adults. The objectives were to (1) describe the types and mounting locations of digital technologies or devices used to measure overall or specific aspects of locomotor capacity; (2) identify digital biomarkers and corresponding assessment tasks derived from these technologies; and (3) assess the reliability and validity of digital technologies for detecting locomotor capacity.

## Methods

### *Review Protocol and Amendments*

This systematic review (registered in PROSPERO [International Prospective Register of Systematic Reviews]; CRD420251074143) followed the guidelines outlined in the PRISMA-COSMIN (Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Consensus-Based Standards for the Selection of Health Measurement Instruments) for Outcome Measures Instruments in Systematic Reviews (OMIs) (Checklist 1) [22]. The only amendments to the registered review protocol were the inclusion of additional searches from Google Scholar and citation lists to ensure a more comprehensive review of the relevant literature.

### *Search Strategies*

The literature search was reported according to PRISMA-S to promote transparency of the search process [23]. The search terms were constructed using the PICO (Population, Intervention, Control, and Outcomes) framework, as displayed in Table S1 in [Multimedia Appendix 1](#). Six databases, including PubMed (National Library of Medicine), Web of Science (Clarivate), Cochrane Library (Wiley), IEEE Xplore (IEEE), Scopus (Elsevier), and Embase (Elsevier), were systematically searched on March 7, 2025, without any restrictions or search filters. Major terms related to older adults, digital technology, assessment, and locomotor capacity were combined for the initial search, according to the review by Honvo et al [24]. The search strategy was developed in consultation with an experienced librarian and experts in geriatric medicine and digital health. In addition, we searched citation lists from included studies by browsing reference lists and gray literature on Google Scholar on November 17, 2025, limiting the search to the first 100

publications, as search results beyond this number rapidly lost relevance to the topic of the review [25]. No other online resources, such as conference proceedings or print journals, were specifically browsed for this review. No study registries were searched, and no additional studies were sought by contacting authors, manufacturers, or others. The search strategy is detailed in Table S2 in [Multimedia Appendix 1](#). The search process was not updated after the initial search date.

## Eligibility Criteria

Following the PICO framework, the inclusion criteria were (1) study population (P): studies enrolling older adults as research subjects or separately reporting results on older people; (2) interventions (I): studies focusing on digital biomarkers (defined as objective, quantifiable physiological and behavioral measurements acquired through portable, wearable, implantable, or ingestible digital devices [15]), derived from digital technologies; (3) comparators (C): any or none; (4) outcomes (O): studies on the assessment of locomotor capacity (ie, endurance, balance, muscle strength, muscle function, muscle power, joint function) while examining the measurement properties of reliability and validity; and (5) studies written in English.

Reliability focused on 2 types [26,27]: (1) test-retest reliability: the extent to which scores for patients who have not changed are the same for repeated measurements over time; and (2) measurement error: the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.

Validity prioritized 2 categories [27]: (1) criterion validity: the degree to which the scores of a device are an adequate reflection of a "gold standard"; and (2) hypothesis testing for construct validity: the degree to which the scores of a device are consistent with hypotheses, including convergent validity (measured by relationships to scores of other instruments) and discriminative validity (assessed through differences between relevant groups).

The exclusion criteria included the following: (1) systematic reviews and literature reviews; (2) books and other non-peer-reviewed literature; (3) studies conducted on populations with specific medical conditions (eg, Parkinson disease, stroke); (4) studies focusing on digital measures during rehabilitation exercises (eg, supervised therapeutic movements, postsurgical rehabilitation protocols, physical therapy interventions, and exercise programs designed for recovery or strength training); (5) studies lacking data on measurement properties; and (6) studies without full text.

## Study Selection

After the literature search, all identified studies were imported into EndNote 21, and duplicates were removed. Two authors independently screened the remaining titles and abstracts for potential eligibility and then retrieved the full-text articles of the selected studies. Any discrepancies regarding study selection were resolved through consultation with a third reviewer.

## Data Extraction

Two authors independently extracted data using standardized forms (Table S3 in [Multimedia Appendix 1](#)), and discrepancies were resolved through discussion among the 3 authors. Data items comprised: (1) characteristics of studies (first author, year of publication, and study design); (2) participant demographic information (sample size, age, sex distribution, and setting); (3) digital technology- or device-related information (type, number, assessment tasks, and digital biomarkers); and (4) results of validity and reliability.

## Quality Assessment

### Overview

Quality assessment was independently undertaken by 2 authors based on the COSMIN guidelines [28], which comprised 3 substeps: the methodological quality of each study for a measurement property, the rated result of each study for a measurement property, and the pooled results of all available studies for a measurement property. Given the substantial heterogeneity in locomotion protocols and variables across the included studies, pooled results were not performed. In cases of disagreement during the evaluation process, a third researcher was consulted to reach a consensus.

### Methodological Quality Assessment

For the methodological quality of the included studies, 4 measurement properties from the consensus-based COSMIN checklist [29] were evaluated: test-retest reliability, measurement error, criterion validity, and hypothesis testing for construct validity. As this checklist was originally developed to assess measurement properties of a patient-reported outcome measure, the extended version of COSMIN was used to replace the original boxes on test-retest reliability or measurement error [30]. The methodological quality of each study for a measurement property was categorized as very good, adequate, doubtful, or inadequate. The "worst score counts" principle was applied to determine the overall methodological quality of each measurement property. When a study reported a range of values or multiple values for a measurement property, the best value was selected for the measurement property rating [24].

### Measurement Properties Quality Assessment

We first assessed each study for its measurement property quality according to the updated COSMIN criteria for good measurement properties [28]. Each property's quality was rated as sufficient (+), insufficient (−), or indeterminate (?).

When at least 2 studies used the same protocols and variables, we synthesized the results. If the ratings were consistent across studies, the results from different studies for one measurement property were qualitatively summarized or pooled using meta-analysis, ultimately assigning a rating of "+" or "−."

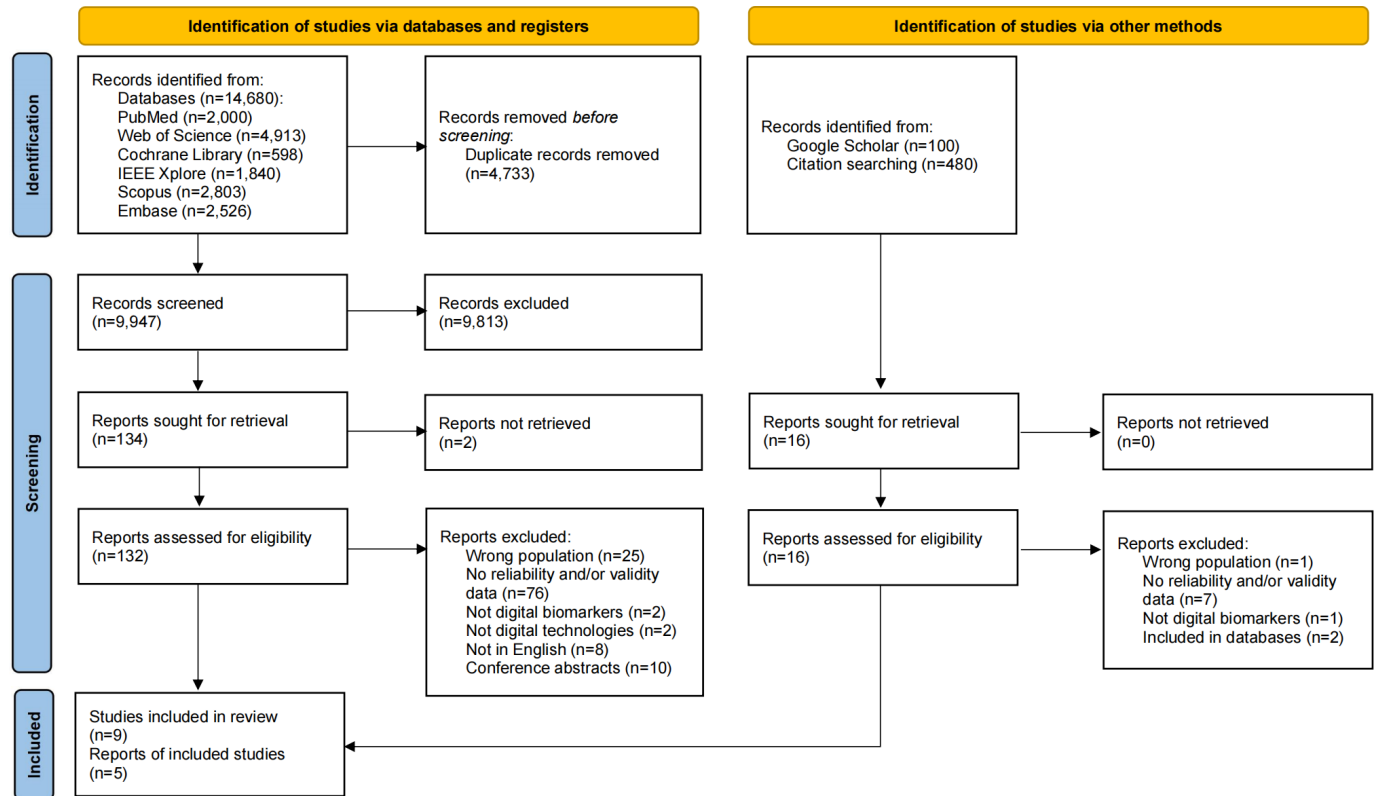
## Results

### Literature Search Result

A total of 14,680 articles were identified from 6 databases, and 4,733 duplicates were removed. Following the initial

screening of titles and abstracts, 123 articles were fully evaluated, with 9 of them eligible for inclusion in the review [31-39]. An additional 5 publications were obtained from reference lists [40,41] and gray literature [42-44]. The study selection process is presented in Figure 1.

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the study selection process. This diagram illustrates the study selection process for the systematic review of digital measures of locomotor capacity in older adults, detailing each stage (from record identification through screening, eligibility assessment, and final inclusion) and providing explicit reasons for exclusion at each stage.



### Study Characteristics

The main characteristics of the included studies and populations are summarized in Table 1. These studies were published between 2013 and 2024 and were conducted across 10 countries, primarily in the United States (n=3). Thirteen studies employed a cross-sectional design, while 1 adopted a pretest-posttest design. Across all studies, balance was the predominant assessment target (n=13), with

only 1 study evaluating muscle power. No eligible studies assessed endurance, muscle strength, or joint function, and the exclusion reasons included incorrect population [45], lack of reliability and/or validity data [46], studies not involving digital biomarkers or technologies [47,48], and conference abstracts [49]. The sample sizes ranged from 12 to 248, with the majority from community-dwelling populations.

**Table 1.** Characteristics of the included studies.

Study	Country	Study design	Assessment target	Sample (count; age <sup>a</sup> (y); male participants)	Population sources
Cerrito et al [31] (2015)	Switzerland	Cross-sectional study	Muscle power	16; 73.5 (10.4); 16	Retirement homes and community-dwelling
Chang et al [42] (2013)	China	Pretest-posttest design	Balance	20; 67.32 (3.43); NR <sup>b</sup>	NR
da Silva et al [32] (2013)	Brazil	Cross-sectional study	Balance	28; 69 (5); 8	Community-dwelling
De Groot et al [33] (2021)	Belgium	Cross-sectional study	Balance	97; 50-90; 42	Community-dwelling
Greene et al [34] (2022)	Ireland	Cross-sectional study	Balance	168; 75.0 (7.2); 51	Community-dwelling
Harro and Garascia [40] (2019)	United States	Cross-sectional study	Balance	46; 67.7 (5.1); 22	Community-dwelling
Kuntapun et al [35] (2020)	Thailand	Cross-sectional study	Balance	12; 75.6 (5.6); 3	NR

Study	Country	Study design	Assessment target	Sample (count; age <sup>a</sup> (y); male participants)	Population sources
Levy et al [36] (2018)	United States	Cross-sectional study	Balance	49; 71.3 (7.3); 23	Community-dwelling
McManus et al [37] (2022)	Ireland	Cross-sectional study	Balance	248; 74.9 (6.5); 91	Community-dwelling
Okada et al [38] (2024)	Japan	Cross-sectional study	Balance	27; 74.7 (7.1); 7	Community-dwelling
Olsen et al [43] (2023)	New Zealand	Cross-sectional study	Balance	34; 42-94; 14	Community-dwelling
Pooranawatthanakul and Siriphorn [44] (2023)	Thailand	Cross-sectional study	Balance	20; 70.85 (4.09); 3	Community-dwelling
Scaglioni-Solano and Aragón-Vargas [41] (2013)	Costa Rica	Cross-sectional study	Balance	37; 69 (8); NR	Exercise groups
Zhou et al [39] (2021)	United States	Cross-sectional study	Balance	15; 71.4 (5.9); 7	Community-dwelling

<sup>a</sup>Age values are given as mean (SD) or range.

<sup>b</sup>NR: not reported.

## Balance Assessment

### Device Type and Number

Table 2 provides an overview of the devices, tasks, variables, missing data, and data denoising methods. In terms of the assessment of balance, the most common device type was a

smartphone (n=5), followed by a force plate (n=3), an IMU (n=2), a balance board (n=2), and an infrared depth sensor (n=1). All 13 studies used a single device for data collection. Two studies further assessed the reliability and validity across different sensor placement locations.

**Table 2.** Description of devices, tasks, variables, and data processing in the included studies.

Study	Devices (type; number)	Task settings	Device variables	Missing data and handling methods	Data denoising methods
Balance					
Chang et al [42] (2013)	Balance board; n=1	Three 10-s trials of stance with EO <sup>a</sup> , stance with EC <sup>b</sup> , and one-leg stance with EO	Center of gravity displacement	NR <sup>c</sup>	NR
da Silva et al [32] (2013)	Force plate; n=1	Three 30-s trials of the one-leg stance	COP <sup>d</sup> area, COP sway RMS <sup>e</sup> amplitude AP <sup>f</sup> , COP sway RMS amplitude ML <sup>g</sup> , mean velocity AP, mean velocity ML, mean frequency AP, and mean frequency ML	Complete data; NA <sup>h</sup>	Low-pass second-order Butterworth filter
De Groot et al [33] (2021)	Smartphone; n=1	Four 35-s trials of bipodal stance with EO, bipodal stance with EC, a dual task consisting of bipodal stance with EO and a cognitive task, and semi-tandem stance with EO	RMS acceleration ML, RMS acceleration AP, mean acceleration ML, and mean acceleration AP	Complete data; NA	Savitzky-Golay filter
Greene et al [34] (2022)	IMU <sup>i</sup> ; n=1 (on thigh or sternum)	Five times sit-to-stand test	Balance model: including total time, coefficient of variation of sit-stand-sit time, mean time to stand up, spectral entropy angular velocity, mean Z-axis acceleration at stand start, and mean Z-axis acceleration at sit end	Complete data; NA	Butterworth IIR digital filter; signal validity checks
Harro and Garascia [40] (2019)	Force plate; n=1	Three standardized tests: limits of stability, motor control test, and sensory organization test	Limits of stability average end point excursion, motor control test average latency, sensory organization test composite	NR	NR

Study	Devices (type; number)	Task settings	Device variables	Missing data and handling methods	Data denoising methods
Kuntapun et al [35] (2020)	Smartphone; n=1 (attached horizontally at the level of the third lumbar vertebrae or rested on the right hip)	Walking 10 m at self-selected comfortable speed under three conditions: level walking, irregular surface walking, and obstacle crossing	equilibrium, and sensory organization test vestibular ratio scores COM <sup>j</sup> AP displacement, COM ML displacement, and COM SI <sup>k</sup> displacement	NR	Butterworth fourth-order low-pass filter
Levy et al [36] (2018)	Balance tracking system (consisting of a force plate and software); n=1	Six 20-s static standing trials of 3 with EO and 3 with EC	COP sway ML and COP sway AP	Complete data; NA	Dual-pass Butterworth filter
McManus et al [37] (2022)	IMU; n=1	Two 30-s trials of semi-tandem stance with EO and a narrow stance with EC	Balance score: including RMS acceleration magnitude, RMS acceleration ML, and RMS acceleration AP	Complete data; NA	Fourth order low-pass Butterworth filter
Okada et al [38] (2024)	Infrared depth sensor; n=1	20-s stepping-in-place test	Total movement distance, knee movement distance, maximum movement displacement, sway index, and step number	NR	NR
Olsen et al [43] (2023)	Smartphone; n=1	Four 30-s standing tasks: firm surface with EO, firm surface with EC, compliant surface with EO, and compliant surface with EC; and two 6-s walking tasks with EO, each performed four times: looking straight ahead and turning head	Postural stability, postural stability ML, postural stability AP, walking speed, mean step length, mean step time, step length variability, step time variability, step length asymmetry, and step time asymmetry	NR	NR
Pooranawatthanakul and Siriphorn [44] (2023)	Smartphone; n=1	Three standardized tests: the Modified Clinical Test of Sensory Interaction in Balance (ie, standing on a flat surface and a foam surface with EO or EC), a single-leg stance test, and a limit of stability test	RMS acceleration	Complete data; NA	Fourth-order Butterworth filter
Scaglioni-Solano and Aragón-Vargas [41] (2013)	Balance board; n=1	Four 30-s standing tasks (firm surface with EO, firm surface with EC, compliant surface with EO, and compliant surface with EC) and 10-s tandem stance test	COP displacement	Complete data; NA	Eighth-order Butterworth filter
Zhou et al [39] (2021)	Smartphone; n=1	Three 30-s trials of EO, EC, and dual task standing	Two-dimensional path length and RMS acceleration	Complete data; NA	Low-pass filter
Muscle power Cerrito et al [31] (2015)	Smartphone; n=1	Sit-to-stand movement	Total movement duration, peak force, rate of force development, and peak power	Complete data; NA	The signal was digitally low-pass filtered

<sup>a</sup>EO: eyes open.

<sup>b</sup>EC: eyes closed.

<sup>c</sup>NR: not reported.

<sup>d</sup>COP: center of pressure.

<sup>e</sup>RMS: root mean square.

<sup>f</sup>AP: anteroposterior.

<sup>g</sup>ML: mediolateral.

<sup>h</sup>NA: not applicable.

<sup>i</sup>IMU: inertial measurement unit.

<sup>j</sup>COM: center of mass.

<sup>k</sup>SI: superior-inferior.

## Assessment Tasks

The included studies employed a diverse array of tasks designed to assess different aspects of balance. Nine studies [32,33,36,37,39,41-44] applied static standing tasks with eyes open and eyes closed, including bipodal, semi-tandem, narrow, dual task, and one-leg stance. Dynamic balance tasks were evaluated in 6 studies [34,35,38,40,43,44] using tasks such as walking, the 5-time sit-to-stand test, stepping-in-place test, limits of stability, motor control test, and sensory organization test. Two studies [43,44] conducted both static and dynamic standing tasks.

## Device Variables

As presented in Table 3, a total of 47 distinct variables related to balance were identified from 13 studies, categorized as temporal (8, 17%), spatial (12, 25.5%), linear acceleration (11, 23.4%), angular velocity (1, 2.1%), position (11, 23.4%), and energy (4, 8.5%). Five studies [32,34,39,40,43] simultaneously assessed multiple categories of variables, primarily temporal, linear acceleration, and position variables.

**Table 3.** Classification and definitions of digital biomarkers of balance assessment.

Category and variable	Definition (unit)	Study
<b>Temporal</b>		
Walking speed	The mean of the ratios of step length to step time (m/s)	Olsen et al [43] (2023)
Mean step time	The mean of the time between two consecutive initial contacts of alternative feet (s)	Olsen et al [43] (2023)
Step time variability	The RMS <sup>a</sup> of the SD of left step times and the SD of right step times (%)	Olsen et al [43] (2023)
Step time asymmetry	The percentage difference between left and right mean step times compared to the overall mean step time (%)	Olsen et al [43] (2023)
Total time to complete five times sit-to-stand test	The total time to complete five times sit-to-stand test (s)	Greene et al [34] (2022)
Coefficient of variation of sit-stand-sit time	The coefficient of variation of sit-stand-sit time across repetitions (%)	Greene et al [34] (2022)
Mean time to stand up	The mean time of the stand-up phase across repetitions (s)	Greene et al [34] (2022)
Motor control test average latency	The average time between the force plate translation and the individual's active force responses in each leg (ms)	Harro and Garascia [40] (2019)
<b>Spatial</b>		
Mean step length	The mean of the AP <sup>b</sup> distance between two consecutive initial contacts of alternative feet (m)	Olsen et al [43] (2023)
Step length variability	The RMS of the SD of left step lengths and the SD of right step lengths (%)	Olsen et al [43] (2023)
Step length asymmetry	The percentage difference between left and right mean step lengths compared to the overall mean step length (%)	Olsen et al [43] (2023)
Periodicity index	The step symmetry between the right and left step within a stride and the gait regularity across strides (%)	Olsen et al [43] (2023)
COM <sup>c</sup> AP displacement	The average of minimum to maximum displacement in the AP direction across all step cycles (cm)	Kuntapun et al [35] (2020)
COM ML <sup>d</sup> displacement	The average of minimum to maximum displacement in the ML direction across all step cycles (cm)	Kuntapun et al [35] (2020)
COM SI <sup>e</sup> displacement	The average of minimum to maximum displacement in the SI direction across all step cycles (cm)	Kuntapun et al [35] (2020)
Total movement distance	The total linear distance determined from the sum total of step-by-step changes in head position from its initial to its final position (m)	Okada et al [38] (2024)
Knee movement distance	The mean of the sum of the three-dimensional movements of each knee joint point during stepping (m)	Okada et al [38] (2024)
Maximum movement displacement	The maximum displacement of the head along the movement path from its initial position (m)	Okada et al [38] (2024)

Category and variable	Definition (unit)	Study
Sway index	The ratio of total movement distance to maximum movement displacement	Okada et al [38] (2024)
Step number in 20-s stepping-in-place test	The total number of foot contacts beginning with the return of the right foot to the ground after raising the right leg to initiate the test (steps)	Okada et al [38] (2024)
<b>Linear acceleration</b>		
RMS acceleration	RMS of acceleration in AP and ML directions ( $\text{mm/s}^2$ )	Pooranawatthanakul and Siriphorn [44] (2023); Zhou et al [39] (2021)
RMS acceleration magnitude	RMS of the magnitude of the 3-axis acceleration	McManus et al [37] (2022)
RMS acceleration AP	RMS of COM acceleration in AP direction ( $\text{m/s}^2$ )	De Groote et al [33] (2021); McManus et al [37] (2022)
RMS acceleration ML	RMS of COM acceleration in ML direction ( $\text{m/s}^2$ )	De Groote et al [33] (2021); McManus et al [37] (2022)
Mean acceleration AP	Average absolute value of COM acceleration in AP directions ( $\text{m/s}^2$ )	De Groote et al [33] (2021)
Mean acceleration ML	Average absolute value of COM acceleration in ML direction ( $\text{m/s}^2$ )	De Groote et al [33] (2021)
Mean Z-axis acceleration at stand-start	The average Z-axis (vertical) accelerometer value recorded at the detected stand-start time across repetitions ( $\text{m/s}^2$ )	Greene et al [34] (2022)
Mean Z-axis acceleration at sit-end	The average Z-axis (vertical) accelerometer value recorded at the detected sit-end time across repetitions ( $\text{m/s}^2$ )	Greene et al [34] (2022)
Postural stability	The negative natural logarithm of the mean of the absolute acceleration along mediolateral, anterior–posterior, and vertical axes resultant vector ( $-\ln[\text{m/s}^2]$ )	Olsen et al [43] (2023)
Postural stability ML	The negative natural logarithm of the absolute acceleration along ML axis resultant vector ( $-\ln[\text{m/s}^2]$ )	Olsen et al [43] (2023)
Postural stability AP	The negative natural logarithm of the absolute acceleration along AP axis resultant vector ( $-\ln[\text{m/s}^2]$ )	Olsen et al [43] (2023)
<b>Angular velocity</b>		
Spectral entropy angular velocity	The spectral entropy of the angular-velocity signal	Greene et al [34] (2022)
<b>Position</b>		
Center of gravity displacement	The average displacement of the center of gravity (cm)	Chang et al [42] (2013)
2D path length	The total length of the acceleration trajectories (mm)	Zhou et al [39] (2021)
COP <sup>f</sup> displacement	The distance traveled by the COP (cm)	Scaglioni-Solano and Aragón-Vargas [41] (2013)
COP sway RMS amplitude AP	RMS amplitude of COP sway in AP direction (cm)	da Silva et al [32] (2013)
COP sway RMS amplitude ML	RMS amplitude of COP sway in ML direction (cm)	da Silva et al [32] (2013)
Mean velocity AP	Mean velocity of COP in AP direction (cm/s)	da Silva et al [32] (2013)
Mean velocity ML	Mean velocity of COP in ML direction (cm/s)	da Silva et al [32] (2013)
COP area	95% confidence ellipse area of COP ( $\text{cm}^2$ )	da Silva et al [32] (2013)
Limits of stability average end point excursion	The average ratio of the initial distance traveled by the center of gravity on the primary attempt to reach each of the 8 multidirectional visual targets (%)	Harro and Garascia [40] (2019)
Sensory organization test composite equilibrium	The average equilibrium score (a measure of anterior/posterior sway during each trial compared to a theoretical stability limit of $12.5^\circ$ ) from conditions with fixed surface and normal vision, and fixed surface with eyes closed, added to the equilibrium scores from conditions with fixed surface and sway referenced vision, sway-referenced surface and normal vision, sway-referenced surface and EC, and sway-referenced surface and sway-referenced vision, then divided by the total number of trials	Harro and Garascia [40] (2019)
Sensory organization test vestibular ratio scores	The ratio of the equilibrium score in the sway-referenced surface, eyes closed condition to the fixed surface, normal vision condition	Harro and Garascia [40] (2019)
<b>Energy</b>		
Mean frequency AP	Mean frequency of COP in AP direction (Hz)	da Silva et al [32] (2013)
Mean frequency ML	Mean frequency of COP in ML direction (Hz)	da Silva et al [32] (2013)
COP sway AP	15.5 multiplied by the result of subtracting the sum of the force sensor data from the bottom left and bottom right corners from the sum of the	Levy et al [36] (2018)

Category and variable	Definition (unit)	Study
COP sway ML	force sensor data from the top left and top right corners and then dividing that result by the total sum of force sensor data from all four corners 24.25 multiplied by the result of subtracting the sum of the force sensor data from the top left and bottom left corners from the sum of the force sensor data from the top right and bottom right corners and then dividing that result by the total sum of force sensor data from all four corners	Levy et al [36] (2018)

<sup>a</sup>RMS: root mean square.

<sup>b</sup>AP: anteroposterior.

<sup>c</sup>COM: center of mass.

<sup>d</sup>ML: mediolateral.

<sup>e</sup>SI: superior-inferior.

<sup>f</sup>COP: center of pressure.

The device-derived variables extracted in the reviewed studies varied and were largely determined by the employed technology and the specific balance task. Five of the 7 studies utilizing smartphones or IMUs reported linear acceleration-based variables, with 4 studies [33,37,39,44] quantifying root mean square (RMS) acceleration. Among the 5 studies using force plates or balance boards, 4 reported position-based variables, primarily center of pressure (COP) metrics [32,40-42]. Additionally, 5 of the 6 studies involving dynamic tasks extracted temporal and spatial variables. Notably, 2 studies integrated multiple device-derived variables into composite balance scores [34,37].

## Data Processing

Eight studies reported complete datasets and therefore did not require missing data handling, whereas 5 studies did not report the occurrence or management of missing data. Data denoising methods were described in 9 studies and primarily involved signal-filtering techniques, including low-pass Butterworth, Savitzky-Golay, and dual-pass filters.

## Reliability

Table S4 in [Multimedia Appendix 1](#) displays the findings on reliability and validity statistics. All studies investigated the test-retest reliability of balance assessment using the intraclass correlation coefficient (ICC). The reported ICC values varied considerably, ranging from 0.016 to 0.99, depending on the device types, tasks, or variable of choice. Chang et al [42] reported the highest test-retest reliability for the average displacement of gravity derived from a balance board during stance with eyes closed. Measurement error was reported in 6 studies, with standard error of measurement (SEM) values ranging from 0.02 to 16.1 and minimal detectable change (MDC) values ranging from 0.3 to 44.6.

When examined by device type and task setting, smartphones reported in 5 studies demonstrated ICCs from 0.016 to 0.97, with higher values in static tasks (0.50-0.97) [33,39,43,44] and wider variability during dynamic tasks (0.016-0.96) [35,43,44]. IMUs showed ICC values from 0.30 to 0.95, with more consistent reliability in static tasks (0.83-0.95) [37] and lower, broader values in dynamic tasks (0.30-0.91) [34]. Across both smartphones and IMUs, static tasks produced better and more consistent reliability than dynamic tasks. Force plates reported ICCs from 0.40 to 0.85 for static

tasks [32,36] and 0.710 to 0.898 for dynamic tasks [40]. Balance boards were used exclusively for static assessments and yielded ICCs from 0.64 to 0.99 [41,42], whereas infrared depth sensors were applied only for dynamic assessments, with ICCs between 0.64 and 0.96 [38]. Regarding variable choice, temporal parameters showed ICCs from 0.30 to 0.91 [34,40,43], spatial parameters from 0.016 to 0.96 [35,38,43], linear acceleration metrics from 0.50 to 0.95 [33,34,37,39,43,44], angular velocity metrics from 0.67 to 0.81 [34], position-related metrics from 0.40 to 0.99 [32,39-42], and energy parameters from 0.72 to 0.83 [32,36]. As the reported SEM values corresponded to heterogeneous outcome variables with different units and scales, we did not perform comparisons across studies.

## Validity

Of the 7 studies that investigated criterion validity, 4 used a force plate as the gold standard [33,36,39,41] and 3 used 3D motion capture [35,43,44]. In smartphone-based assessments compared with motion capture, spatial variables measured during dynamic tasks showed substantial variability, with correlations ranging from -0.51 to 0.98 [43]. In contrast, linear acceleration variables demonstrated consistently excellent correlations in static tasks, with coefficients of 0.91 to 0.98 [43,44], whereas dynamic tasks yielded a low, nonsignificant correlation of 0.37 [44]. When compared with force plates, smartphone-derived linear acceleration measures showed correlations of 0.14 to 0.82 in static tasks [33] and 0.38 to 0.60 in dynamic tasks [39]. Energy variables derived from force plates yielded Pearson correlations of 0.82 to 0.89 [36], and position variables from balance boards produced regression coefficients between 0.92 and 1.05 across static tasks [41].

Four studies [32,37,40,42] assessed convergent validity. Position variables derived from balance boards in static tasks showed significant positive correlations with smart balance master systems, ranging from 0.58 to 0.86 [42]. IMU-derived linear acceleration variables from static tasks had significant, fair correlations of 0.30 to 0.34 with Time Up and Go time [37]. In contrast, da Silva et al [32], using static tasks, and Harro et al [40], using dynamic tasks, reported generally low and nonsignificant correlations between position, energy, and temporal variables from force plates and clinical testing.

Discriminative validity was reported in 5 studies [32,34,37,38]. Two studies [32,42] compared young and older adults, and 2 [37,38] compared individuals with normal versus impaired balance, all demonstrating statistically significant between-group differences across multiple parameters, including position parameters from balance boards and force plates, and linear acceleration parameters from IMUs in static tasks, and spatial parameters from infrared depth sensors in dynamic tasks. One study [34] developed classification models using IMU data during dynamic tasks on the sternum and thigh, achieving 76.76% and 81.69% accuracy, respectively.

### Muscle Power Assessment

One study assessed muscle power using a smartphone during a sit-to-stand task [31]. The derived variables, including total movement duration, peak force, rate of force development, and peak power, were measured. Complete datasets were obtained, and the raw signals were filtered before analysis. Test-retest reliability of these variables ranged from 0.43 to

0.92. Measurement error, expressed as SEM, varied between 3.1 and 26.1. Convergent validity, evaluated by comparing smartphone measures with force plate data, showed a correlation coefficient ranging from 0.69 to 0.98, with peak power demonstrating a statistically significant correlation.

## Quality Assessment

### Methodological Quality Assessment

The methodological quality and evaluation results of measurement properties are summarized in Table 4. For test-retest reliability, 9 of 14 studies reported “doubtful,” primarily because repeated assessments were conducted without familiar tests or adequate rest time, which may introduce variations and affect participant stability on the construct to be measured. For measurement error, 3 of 6 studies received a “doubtful” rating due to the absence of familiarization procedures, while the remaining studies were rated as very good.

**Table 4.** Methodological quality of studies and evaluation results of measurement properties based on COSMIN.

Study	Test-retest reliability		Measurement error		Criterion validity		Convergent validity		Discriminative validity	
	MQ <sup>a</sup>	ER <sup>b</sup>	MQ	ER	MQ	ER	MQ	ER	MQ	ER
<b>Balance</b>										
Chang et al [42] (2013)	D <sup>c</sup>	+ <sup>d</sup>	NR <sup>e</sup>	NR	NR	NR	V <sup>f</sup>	+	D	+
da Silva et al [32] (2013)	V	+	V	? <sup>g</sup>	NR	NR	V	+	V	+
De Groote et al [33] (2021)	V	+	NR	NR	V	+	NR	NR	NR	NR
Greene et al [34] (2022)	D	+	NR	NR	NR	NR	NR	NR	D	+
Harro and Garascia [40] (2019)	V	+	V	?	NR	NR	I <sup>h</sup>	+	NR	NR
Kuntapun et al [35] (2020)	D	+	NR	NR	V	+	NR	NR	NR	NR
Levy et al [36] (2018)	D	+	D	?	V	+	NR	NR	NR	NR
McManus et al [37] (2022)	D	+	NR	NR	NR	NR	V	+	D	+
Okada et al [38] (2024)	D	+	NR	NR	NR	NR	NR	NR	V	+
Olsen et al [43] (2023)	D	+	D	?	V	+	NR	NR	NR	NR
Pooranawatthanakul and Siriphorn [44] (2023)	D	+	D	?	V	+	NR	NR	NR	NR
Scaglioni-Solano and Aragón-Vargas [41] (2013)	V	+	V	?	V	+	NR	NR	NR	NR
Zhou et al [39] (2021)	D	+	NR	NR	V	- <sup>i</sup>	NR	NR	NR	NR
<b>Muscle power</b>										
Cerrito et al [31] (2015)	V	+	V	?	NR	NR	I	+	NR	NR

<sup>a</sup>MQ: methodological quality.

<sup>b</sup>ER: evaluation results.

<sup>c</sup>D: doubtful.

<sup>d</sup>+: sufficient.

<sup>e</sup>NR: not reported.

<sup>f</sup>V: very good.

<sup>g</sup>?: indeterminate.

<sup>h</sup>I: inadequate.

<sup>i</sup>-: insufficient.

Four studies evaluated criterion validity, and all were assessed as “very good.” Convergent validity was assessed in 5 studies, with 2 studies poorly describing the construct and measurement properties of the comparator instrument,

resulting in an “inadequate” methodological quality rating. Of the 5 studies involving discriminative validity, 3 were rated “doubtful” owing to insufficient description of important subgroup characteristics.

## Quality of Measurement Properties

As shown in Table 4, test-retest reliability, convergent validity, and discriminative validity were rated as sufficient (+) in 14, 5, and 5 studies, respectively. Measurement error was rated as indeterminate (?) in all 6 studies, as none defined a minimal important change (MIC). Criterion validity was rated as insufficient (–) in 1 study due to a correlation coefficient below 0.70 with the gold standard.

## Discussion

### Overview

This systematic review summarized digital technologies and biomarkers for assessing locomotor capacity in general older adults, including their validity and reliability. Fourteen studies were included, identifying 5 types of digital technologies and 51 digital biomarkers (47 for balance and 4 for muscle power). Twelve studies reported generally acceptable test-retest reliability and validity for balance, indicating promising potential. One study evaluated muscle power and demonstrated good reliability and validity, though further research is needed to validate these findings. However, caution is warranted when interpreting these results, as reliability and validity vary by context, and the overall strength of evidence remains limited. Future research should address other locomotor components and focus on developing assessments that simultaneously ensure suitability, reliability, and validity [50,51].

### Reliability

The test-retest reliability reported across included studies was acceptable, suggesting that digital technologies can produce stable measurements of locomotor capacity among older adults. Nevertheless, there was considerable variability in reported reliability coefficients, ranging from 0.016 to 0.97. This broad range highlights the nascent stage of the field and could be attributed to differences in device modalities, task settings, and the specific digital biomarkers extracted. Specifically, force plates generally exhibited more consistent test-retest reliability than smartphone- and IMU-based systems, which is commonly attributed to their direct measurement of ground reaction forces and COP trajectories and the absence of variability introduced by wearable sensor placement and soft-tissue artifacts [52,53]. Consistent with previous literature [54], portable sensors yielded higher and more consistent reliability for static tasks than for dynamic tasks [43,44], likely because dynamic conditions introduce increased movement variability, sensor micro-slippage, and more complex signal processing requirements. Variable choice also markedly influenced reliability. For example, in a wearable sensor-based sit-to-stand assessment, linear acceleration and angular velocity variables demonstrated higher test-retest reliability than temporal variables [34], likely because they are directly derived from continuous sensor signals, whereas temporal measures rely on event detection. Overall, these findings underscore the need for standardized sensor fixation and cautious use of integrated

spatial metrics when employing mobile devices and depth sensors in locomotor assessment of older adults.

Only about half of the studies provided measurement error, such as SEM and MDC, which limited the assessment of whether observed differences reflect true change or fall within the bounds of random measurement error. Among those that reported measurement error, none defined a MIC, which prevents evaluation of whether the reported MDC values correspond to clinically meaningful change [55]. Therefore, the rating results of measurement error for relevant studies remained indeterminate. Future studies should systematically report SEM, MDC, and MIC to enable robust evaluation of measurement performance and to support meaningful clinical interpretation.

### Validity

Most studies showed acceptable validity, supporting the use of digital technologies to evaluate locomotor capacity. Regarding criterion validity, smartphone-derived linear acceleration measures showed strong correlations with 3D motion capture systems in static tasks but substantially weaker associations during dynamic conditions [44], highlighting the increased difficulty of accurately capturing complex movement patterns with smartphone sensors. In the studies comparing digital biomarkers with force plates, the correlation coefficients of linear acceleration variables derived from smartphones were generally lower than those of the energy variables from force plates and position variables from balance boards [33,36,41]. This could be attributed to inherent differences in the constructs measured; specifically, smartphones leveraged acceleration signals, whereas force plates and balance boards recorded ground reaction forces and COP.

As for convergent validity, the correlation between IMUs and Timed Up and Go tests was significant but low. This weak correlation may be because Timed Up and Go tests mainly assess rougher motor ability, whereas digital tools can measure balance dynamics and subtle changes more precisely [56]. Furthermore, the correlation between the force plate and clinical tests was generally low and nonsignificant in most cases. Clinical tests can measure functional balance abilities, whereas force plates can directly analyze balance related to reactive postural strategies and sensory integration [40,42].

The results of discriminative validity indicated that device-based measures were valid in differentiating balance performance between younger and older participants, which was consistent with a previous systematic review [57]. This not only reinforces the well-established physiological understanding of age-related declines in balance control, but more importantly, it positions digital biomarkers as sensitive tools for objectively quantifying these subtle yet significant changes associated with the aging process. The devices also distinguished between groups with impaired and normal balance. However, the criteria used to define balance groups varied across studies. This heterogeneity makes it difficult to compare the discriminative ability of different device-based measures.

## Digital Biomarkers

In this systematic review, we identified several digital biomarkers that demonstrated excellent reliability and validity for assessing locomotor capacity in older adults. Aligning with previous studies [58,59], COP parameters, including COP displacement, mean frequency anteroposterior (AP), COP sway AP, and COP sway mediolateral (ML), measured from balance boards and force plates during static standing tasks, exhibited strong reliability and validity for assessing balance. Furthermore, RMS acceleration, derived from smartphones, was also identified as a reliable and valid indicator of static balance. It has been frequently utilized in portable sensor-based balance assessments [33,37,39]. Beyond static balance measures, dynamic balance indicators such as walking speed, mean step time, and mean step length showed great reliability and validity for assessing dynamic balance. These gait-related biomarkers were also used to enhance the discrimination of gait disturbances and disease severity in individuals with Parkinson disease [56]. Additionally, total movement distance and step number in the 20-second stepping test measured by infrared depth sensors exhibited strong test-retest reliability and differentially discriminated between high- and low-balance subgroups. These 2 variables appeared to contribute to the differentiation of disordered balance and to suggest directions for future research.

Among studies focused on muscle power, Cerrito et al [31] demonstrated that smartphone-based peak power during sit-to-stand was a robust digital biomarker for evaluating lower-body muscle power in older adults. The use of peak power as a valid indicator of muscle power was further supported by studies employing linear transducers [60]. Furthermore, sit-to-stand muscle power was found to be a stronger predictor of falls, fractures, and physical independence than grip strength [61,62].

Integrating multiple variables can address multidimensional deficits that single measures miss [63] and has shown utility in frailty classification [64]. There is a growing trend toward constructing composite indices for balance impairment. Although 2 studies [34,37] integrating IMU-derived variables reported encouraging discriminative performance, these indices remain exploratory given the absence of rigorous validation and external replication. To establish their clinical applicability in balance assessment, future studies should focus on systematic performance evaluation and external validation.

## Quality of Evidence

Eleven studies were rated as having doubtful or inadequate methodological quality, leading to inflated or underestimated reliability and validity. Specifically, many studies lacked details about the familiarization process or the rest periods between repeated measurements. These omissions may introduce variation due to participant fatigue, learning effects, or other extraneous factors.

In addition, only 3 included studies reported a sample size greater than 50, and none used a sample size

calculation. Small sample sizes may inflate the impact of individual variability, reduce statistical power, and result in unstable estimates of reliability and validity. In heterogeneous older adult populations, insufficient sample sizes further limit the generalizability of findings and increase uncertainty in measurement performance. Future research should prioritize adequately powered studies with predefined sample size considerations tailored to the intended measurement properties.

Only De Groote et al [33] met both good methodological quality and an adequate sample size, but the selected indicators, including RMS acceleration ML, RMS acceleration AP, mean acceleration ML, and mean acceleration AP, did not demonstrate satisfactory reliability and validity. Although many studies have reported good reliability and validity for other parameters, the methodological flaws in these studies necessitate caution when interpreting these findings.

## The Gap Between Research and Clinical Application

The digital transformation of healthcare in recent years has significantly advanced digital technologies, facilitating the transition of digital biomarkers from research contexts to clinical settings. Nevertheless, incorporating digital technology into healthcare to assess locomotor capacity in older adults comes with its own set of challenges.

One major challenge lies in balancing cost-effectiveness and measurement accuracy when selecting appropriate technologies to assess locomotor capacity. Force plates that measure postural sway by calculating COP are considered the gold standard for assessing balance [65]. However, their high cost and limited portability restrict their widespread clinical application [36]. More affordable and available alternatives, such as smartphones, IMUs, balance boards, and infrared depth sensors, have been suggested to demonstrate acceptable reliability and validity for assessing balance and locomotor performance [66]. In addition, emerging technologies are providing new perspectives for monitoring locomotor function. For example, Teel et al [67] explored the VR-based balance module for detecting persistent balance deficits in clinical concussion care, reporting a sensitivity of 85.7% and a specificity of 87.8%. As the field advances, hybrid technology integration may enable the capture of different aspects of locomotor function, thereby offering a more nuanced and holistic understanding of an individual's physical health [61]. Manupibul et al [68] demonstrated that the integration of force-sensitive resistance sensors and IMUs provided complementary data, high accuracy, and enhanced performance to estimate the gait parameters.

The successful clinical translation of digital technologies for locomotor capacity requires careful consideration of user experience and usability, particularly ease of use among older adults with digital exclusion [69,70]. Tools or systems with complex interfaces or that require frequent manual adjustments may pose substantial barriers to adoption, compliance, and long-term use in older populations. Consistent findings

from qualitative studies highlight ease of use as a key facilitator of technology adoption in this population [71,72]. Current research primarily involved participants performing specific, structured tasks under controlled conditions, resulting in potential missing data [33,39]. To achieve clinical effectiveness, future developments in locomotor capacity assessment tools should emphasize unobtrusive sensing, voice-based guidance, and minimal user interaction [73].

Context-specific performance also represents a key challenge for clinical application. Most of the studies included in this review collected data during the performance of predefined tasks under controlled conditions [35, 40,44]. While certain digital biomarkers may demonstrate strong reliability and validity in specific tasks, their performance may vary across different tasks or settings. For example, RMS acceleration derived from smartphones in static tasks exhibited excellent criterion validity, but its performance deteriorated when applied to dynamic tasks [44]. Its performance can also be influenced by factors such as sensor placement, sampling frequency, and configured range [74]. Furthermore, the lack of standardized definitions and measurement protocols for these biomarkers complicates their reproducibility and generalization across studies.

To facilitate clinical application, interpretability represents an important consideration for future research. Interpretability can be considered from two perspectives: the clinical meaningfulness of digital biomarkers and the transparency of algorithms. The interpretability of digital biomarkers depends on how they reflect locomotor capacity, including the establishment of meaningful cut-off points that correspond to declines or improvements in mobility. Minimal clinically important difference is defined as the smallest change in a clinical outcome measure that is considered beneficial and meaningful to patients. By establishing minimal clinically important difference thresholds, clinicians can determine whether changes in digital biomarkers are clinically significant, thereby supporting decision-making [75]. With the integration of digital biomarkers, artificial intelligence-based algorithms may be employed, raising concerns about “black-box” systems, and clinical implementation should prioritize explainable artificial intelligence approaches to improve healthcare professionals’ and patients’ understanding of the measurement mechanism [76].

## Limitations

This systematic review offers a comprehensive overview of digital technologies for assessing locomotor capacity in older adults, providing integrated insights into digital health and physical function assessment. This systematic review has several limitations to note. First, we included only articles published in English, which may introduce publication bias. Second, although our search strategy was intentionally broad, the included studies predominantly focused on balance assessment, with other key components of locomotor capacity remaining underrepresented, reflecting a gap in the existing literature on other locomotor capacity components. Third, the included studies exhibited considerable heterogeneity, with variable study quality and assessment protocols.

This variability is due to the emerging and expanding nature of this research field, which has yet to establish optimal assessment protocols and highlights a valuable strength of this review. Fourth, this systematic review predominantly focused on general older adults and excluded populations with specific diseases, such as Parkinson disease. The scope of our findings may be limited, and the generalizability to populations with specific diseases requires further investigation. Finally, this review may not have captured all emerging technologies, such as virtual reality and artificial intelligence, which are rapidly evolving and hold considerable potential for locomotor capacity assessment.

## Implications

Despite these limitations, the sufficient test-retest reliability and validity of the device-based assessment support its potential use in examining older adults’ balance and muscle power. Several recommendations need to be formulated. First, future studies should enhance methodological quality and adhere to COSMIN terminologies and recommendations [27, 77]. Second, researchers should thoroughly evaluate both reliability and validity and compare the relevant measurement properties for various digital tools to facilitate the development of the optimal system. Third, alongside rigorous validation of measurement properties, future research should prioritize developing user-friendly digital tools. Current studies often employ task-specific protocols, limiting their widespread adoption in clinical practice and community settings. Fourth, variability in reliability and validity results indicated that no single device or method can be universally recommended for locomotor capacity assessment at this stage. Further studies need to standardize protocols and assess the robustness of these technologies in diverse real-world settings. Lastly, given the current diversity of technologies, the studies included in this review do not cover all existing tools, such as virtual reality and artificial intelligence applications. Future research is expected to explore the integration of these emerging technologies for a more precise and fine-grained assessment of locomotor capacity.

## Conclusion

This review offers an integrated understanding of the potential and limitations of digital technologies for assessing locomotor capacity in older adults, summarizing 51 digital biomarkers with generally acceptable reliability and validity. Unlike previous studies that predominantly targeted specific sensor types or disease-specific populations, our review incorporates diverse technologies and biomarkers across general older populations, offering a novel perspective on the potential for scalable, objective, and remote monitoring of changes in locomotor capacity related to aging. Our findings pave the way for developing digital biomarkers that enable the early detection of subtle declines in locomotor capacity, which is critical for timely intervention. However, the current evidence reveals several limitations, particularly task dependency and limited investigation of key locomotor components. The methodological quality of the included studies was generally doubtful, limiting the immediate clinical applicability of digital assessments. Future

studies should expand the evaluation to a broader range of locomotor components and improve methodological quality with predefined sample size calculations. Furthermore, age-friendly assessments with enhanced clinical interpretability should be explored to facilitate clinical

implementation. By addressing these challenges, digital biomarkers could become essential tools for enhancing the monitoring and management of older adults' locomotor capacity, contributing to healthy aging strategies.

---

### Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) to improve grammar and spelling, and to rephrase sentences for enhanced clarity. According to GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was ChatGPT-5.2. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

---

### Funding

This systematic review was funded by the National Key Research & Development Program of China (grant 2023YFC3604602, 2023YFC3605204), Fundamental Research Funds for the Central Universities of Central South University (grant 2025ZZTS0171), and the Hunan Provincial Innovation Foundation for Postgraduate (grant CX20250407). The funder had no involvement in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

---

### Data Availability

Data sharing was not applicable to this systematic review as no datasets were generated or analyzed.

---

### Authors' Contributions

Conceptualization: SZ, JH

Data curation: SZ, JH, CZ, YZ

Methodology: SZ, JH

Supervision: MH, HF

Writing – original draft: SZ, JH, JY, XL

Writing – review & editing: MH, HF

All authors read and approved the final manuscript.

HF and MH contributed equally to this work and share the co-corresponding authorship.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Search terms, search strategies, data extraction form, and description of reliability and validity in the included studies.

[\[DOCX File \(Microsoft Word File\), 46 KB-Multimedia Appendix 1\]](#)

---

### Checklist 1

PRISMA-COSMIN checklist.

[\[DOCX File \(Microsoft Word File\), 29 KB-Checklist 1\]](#)

---

### References

1. Wang S, Yang Z, Tan X, Lai F, Luo L, Ding Y. Association between standing height and physical disability among U.S. adults aged 60 years and older: findings from NHANES 2015-2018. *BMC Geriatr*. Jun 18, 2024;24(1):529. [doi: [10.1186/s12877-024-05100-3](#)] [Medline: [38890578](#)]
2. Braun T, Thiel C, Peter RS, et al. Association of clinical outcome assessments of mobility capacity and incident disability in community-dwelling older adults - a systematic review and meta-analysis. *Ageing Res Rev*. Nov 2022;81:101704. [doi: [10.1016/j.arr.2022.101704](#)] [Medline: [35931411](#)]
3. Veronese N, Honvo G, Amuthavalli Thiyagarajan J, et al. Attributes and definitions of locomotor capacity in older people: a World Health Organisation (WHO) locomotor capacity working group meeting report. *Aging Clin Exp Res*. Mar 2022;34(3):481-483. [doi: [10.1007/s40520-022-02080-5](#)] [Medline: [35133612](#)]
4. Wang G, Zhou Y, Zhang L, et al. Prevalence and incidence of mobility limitation in Chinese older adults: evidence from the China Health and Retirement Longitudinal Study. *J Nutr Health Aging*. Mar 2024;28(3):100038. [doi: [10.1016/j.jnha.2024.100038](#)] [Medline: [38280833](#)]
5. de Oliveira VP, Ferriolli E, Lourenço RA, González-Bautista E, de Souto Barreto P, de Mello RGB. The sensitivity and specificity of the WHO's ICOPE screening tool, and the prevalence of loss of intrinsic capacity in older adults: a scoping review. *Maturitas*. Nov 2023;177:107818. [doi: [10.1016/j.maturitas.2023.107818](#)] [Medline: [37542782](#)]

6. López-Ortiz S, Lista S, Peñín-Grandes S, et al. Defining and assessing intrinsic capacity in older people: a systematic review and a proposed scoring system. *Ageing Res Rev.* Aug 2022;79(101640):101640. [doi: [10.1016/j.arr.2022.101640](https://doi.org/10.1016/j.arr.2022.101640)] [Medline: [35569785](https://pubmed.ncbi.nlm.nih.gov/35569785/)]
7. Jung HW, Roh H, Cho Y, et al. Validation of a multi-sensor-based kiosk for short physical performance battery. *J Am Geriatr Soc.* Dec 2019;67(12):2605-2609. [doi: [10.1111/jgs.16135](https://doi.org/10.1111/jgs.16135)] [Medline: [31441514](https://pubmed.ncbi.nlm.nih.gov/31441514/)]
8. Bergquist R, Weber M, Schwenk M, et al. Performance-based clinical tests of balance and muscle strength used in young seniors: a systematic literature review. *BMC Geriatr.* Jan 9, 2019;19(1):9. [doi: [10.1186/s12877-018-1011-0](https://doi.org/10.1186/s12877-018-1011-0)] [Medline: [30626340](https://pubmed.ncbi.nlm.nih.gov/30626340/)]
9. Piau A, Steinmeyer Z, Cesari M, et al. Intrinsic capacity monitoring by digital biomarkers in integrated care for older people (ICOPE). *J Frailty Aging.* 2021;10(2):132-138. [doi: [10.14283/jfa.2020.51](https://doi.org/10.14283/jfa.2020.51)] [Medline: [33575701](https://pubmed.ncbi.nlm.nih.gov/33575701/)]
10. Zhu Y, Li H, Wu X, Chen N. Accuracy validation of a sensor-based inertial measurement unit and motion capture system for assessment of lower limb muscle strength in older adults—a novel and convenient measurement approach. *Sensors (Basel).* Sep 18, 2024;24(18):6040. [doi: [10.3390/s24186040](https://doi.org/10.3390/s24186040)] [Medline: [39338786](https://pubmed.ncbi.nlm.nih.gov/39338786/)]
11. Estévez-Pedraza ÁG, Parra-Rodríguez L, Martínez-Méndez R, Portillo-Rodríguez O, Ronzón-Hernández Z. A novel model to quantify balance alterations in older adults based on the center of pressure (CoP) measurements with a cross-sectional study. *PLoS ONE.* 2021;16(8):e0256129. [doi: [10.1371/journal.pone.0256129](https://doi.org/10.1371/journal.pone.0256129)] [Medline: [34398918](https://pubmed.ncbi.nlm.nih.gov/34398918/)]
12. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. *J Med Internet Res.* Aug 30, 2019;21(8):e12785. [doi: [10.2196/12785](https://doi.org/10.2196/12785)] [Medline: [31471958](https://pubmed.ncbi.nlm.nih.gov/31471958/)]
13. Pradeep Kumar D, Toosizadeh N, Mohler J, Ehsani H, Mannier C, Laksari K. Sensor-based characterization of daily walking: a new paradigm in pre-frailty/frailty assessment. *BMC Geriatr.* May 6, 2020;20(1):164. [doi: [10.1186/s12877-020-01572-1](https://doi.org/10.1186/s12877-020-01572-1)] [Medline: [32375700](https://pubmed.ncbi.nlm.nih.gov/32375700/)]
14. Huang J, Zhou S, Xie Q, Yu J, Zhao Y, Feng H. Digital biomarkers for real-life, home-based monitoring of frailty: a systematic review and meta-analysis. *Age Ageing.* Mar 28, 2025;54(4):afaf108. [doi: [10.1093/ageing/afaf108](https://doi.org/10.1093/ageing/afaf108)] [Medline: [40251836](https://pubmed.ncbi.nlm.nih.gov/40251836/)]
15. Sun YM, Wang ZY, Liang YY, Hao CW, Shi CH. Digital biomarkers for precision diagnosis and monitoring in Parkinson's disease. *NPJ Digit Med.* Aug 21, 2024;7(1):218. [doi: [10.1038/s41746-024-01217-2](https://doi.org/10.1038/s41746-024-01217-2)] [Medline: [39169258](https://pubmed.ncbi.nlm.nih.gov/39169258/)]
16. Burq M, Rainaldi E, Ho KC, et al. Virtual exam for Parkinson's disease enables frequent and reliable remote measurements of motor function. *NPJ Digit Med.* May 23, 2022;5(1):65. [doi: [10.1038/s41746-022-00607-8](https://doi.org/10.1038/s41746-022-00607-8)] [Medline: [35606508](https://pubmed.ncbi.nlm.nih.gov/35606508/)]
17. Hayek R, Gutman I, Baranes G, Nudelman Y, Springer S. Smartphone-based sit-to-stand analysis for mobility assessment in middle age. *Innov Aging.* 2024;8(10):igae079. [doi: [10.1093/geroni/igae079](https://doi.org/10.1093/geroni/igae079)] [Medline: [39391811](https://pubmed.ncbi.nlm.nih.gov/39391811/)]
18. Polvorinos-Fernández C, Sigcha L, Borzi L, et al. Evaluating motor symptoms in Parkinson's disease through wearable sensors: a systematic review of digital biomarkers. *Appl Sci.* 2024;14(22):10189. [doi: [10.3390/app142210189](https://doi.org/10.3390/app142210189)]
19. Woelfle T, Bourguignon L, Lorscheider J, Kappos L, Naegelin Y, Jutzeler CR. Wearable sensor technologies to assess motor functions in people with multiple sclerosis: systematic scoping review and perspective. *J Med Internet Res.* Jul 27, 2023;25:e44428. [doi: [10.2196/44428](https://doi.org/10.2196/44428)] [Medline: [37498655](https://pubmed.ncbi.nlm.nih.gov/37498655/)]
20. Ghislieri M, Gastaldi L, Pastorelli S, Tadano S, Agostini V. Wearable inertial sensors to assess standing balance: a systematic review. *Sensors (Basel).* Sep 20, 2019;19(19):4075. [doi: [10.3390/s19194075](https://doi.org/10.3390/s19194075)] [Medline: [31547181](https://pubmed.ncbi.nlm.nih.gov/31547181/)]
21. Lipsmeier F, Taylor KI, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord.* Aug 2018;33(8):1287-1297. [doi: [10.1002/mds.27376](https://doi.org/10.1002/mds.27376)] [Medline: [29701258](https://pubmed.ncbi.nlm.nih.gov/29701258/)]
22. Elsmann EBM, Mokkink LB, Terwee CB, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *J Clin Epidemiol.* Sep 2024;173:111422. [doi: [10.1016/j.jclinepi.2024.111422](https://doi.org/10.1016/j.jclinepi.2024.111422)] [Medline: [38849061](https://pubmed.ncbi.nlm.nih.gov/38849061/)]
23. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev.* Jan 26, 2021;10(1):39. [doi: [10.1186/s13643-020-01542-z](https://doi.org/10.1186/s13643-020-01542-z)] [Medline: [33499930](https://pubmed.ncbi.nlm.nih.gov/33499930/)]
24. Honvo G, Sabico S, Veronese N, et al. Measures of attributes of locomotor capacity in older people: a systematic literature review following the COSMIN methodology. *Age Ageing.* Oct 28, 2023;52(Suppl 4):iv44-iv66. [doi: [10.1093/ageing/afad139](https://doi.org/10.1093/ageing/afad139)] [Medline: [37902521](https://pubmed.ncbi.nlm.nih.gov/37902521/)]
25. Mohsen F, Al-Absi HRH, Yousri NA, El Hajj N, Shah Z. A scoping review of artificial intelligence-based methods for diabetes risk prediction. *NPJ Digit Med.* Oct 25, 2023;6(1):197. [doi: [10.1038/s41746-023-00933-5](https://doi.org/10.1038/s41746-023-00933-5)] [Medline: [37880301](https://pubmed.ncbi.nlm.nih.gov/37880301/)]

26. Johnston W, O'Reilly M, Argent R, Caulfield B. Reliability, validity and utility of inertial sensor systems for postural control assessment in sport science and medicine applications: a systematic review. *Sports Med*. May 2019;49(5):783-818. [doi: [10.1007/s40279-019-01095-9](https://doi.org/10.1007/s40279-019-01095-9)] [Medline: [30903440](https://pubmed.ncbi.nlm.nih.gov/30903440/)]
27. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. Jul 2010;63(7):737-745. [doi: [10.1016/j.jclinepi.2010.02.006](https://doi.org/10.1016/j.jclinepi.2010.02.006)] [Medline: [20494804](https://pubmed.ncbi.nlm.nih.gov/20494804/)]
28. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. May 2018;27(5):1147-1157. [doi: [10.1007/s11136-018-1798-3](https://doi.org/10.1007/s11136-018-1798-3)] [Medline: [29435801](https://pubmed.ncbi.nlm.nih.gov/29435801/)]
29. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. May 2018;27(5):1171-1179. [doi: [10.1007/s11136-017-1765-4](https://doi.org/10.1007/s11136-017-1765-4)] [Medline: [29260445](https://pubmed.ncbi.nlm.nih.gov/29260445/)]
30. Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol*. Dec 3, 2020;20(1):293. [doi: [10.1186/s12874-020-01179-5](https://doi.org/10.1186/s12874-020-01179-5)] [Medline: [33267819](https://pubmed.ncbi.nlm.nih.gov/33267819/)]
31. Cerrito A, Bichsel L, Radlinger L, Schmid S. Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. *Gait Posture*. Feb 2015;41(2):409-413. [doi: [10.1016/j.gaitpost.2014.11.001](https://doi.org/10.1016/j.gaitpost.2014.11.001)] [Medline: [25467428](https://pubmed.ncbi.nlm.nih.gov/25467428/)]
32. da Silva RA, Bilodeau M, Parreira RB, Teixeira DC, Amorim CF. Age-related differences in time-limit performance and force platform-based balance measures during one-leg stance. *J Electromyogr Kinesiol*. Jun 2013;23(3):634-639. [doi: [10.1016/j.jelekin.2013.01.008](https://doi.org/10.1016/j.jelekin.2013.01.008)] [Medline: [23403137](https://pubmed.ncbi.nlm.nih.gov/23403137/)]
33. De Groote F, Vandevyvere S, Vanhevel F, Orban de Xivry JJ. Validation of a smartphone embedded inertial measurement unit for measuring postural stability in older adults. *Gait Posture*. Feb 2021;84:17-23. [doi: [10.1016/j.gaitpost.2020.11.017](https://doi.org/10.1016/j.gaitpost.2020.11.017)] [Medline: [33260077](https://pubmed.ncbi.nlm.nih.gov/33260077/)]
34. Greene BR, Doheny EP, McManus K, Caulfield B. Estimating balance, cognitive function, and falls risk using wearable sensors and the sit-to-stand test. *Wearable Technol*. 2022;3:e9. [doi: [10.1017/wtc.2022.6](https://doi.org/10.1017/wtc.2022.6)] [Medline: [38486905](https://pubmed.ncbi.nlm.nih.gov/38486905/)]
35. Kuntapun J, Silsupadol P, Kamnardsiri T, Lugade V. Smartphone monitoring of gait and balance during irregular surface walking and obstacle crossing. *Front Sports Act Living*. 2020;2:560577. [doi: [10.3389/fspor.2020.560577](https://doi.org/10.3389/fspor.2020.560577)] [Medline: [33345119](https://pubmed.ncbi.nlm.nih.gov/33345119/)]
36. Levy SS, Thralls KJ, Kviatkovsky SA. Validity and reliability of a portable balance tracking system, BTrackS, in older adults. *J Geriatr Phys Ther*. 2018;41(2):102-107. [doi: [10.1519/JPT.000000000000111](https://doi.org/10.1519/JPT.000000000000111)] [Medline: [27893566](https://pubmed.ncbi.nlm.nih.gov/27893566/)]
37. McManus K, Greene BR, Ader LGM, Caulfield B. Development of data-driven metrics for balance impairment and fall risk assessment in older adults. *IEEE Trans Biomed Eng*. Jul 2022;69(7):2324-2332. [doi: [10.1109/TBME.2022.3142617](https://doi.org/10.1109/TBME.2022.3142617)] [Medline: [35025734](https://pubmed.ncbi.nlm.nih.gov/35025734/)]
38. Okada S, Takeshima N, Fujita E, Kohama T, Kusunoki M, Brechue WF. The stepping test, and infrared depth sensor, provide reliable measures of balance in community-dwelling older adults. *J Phys Ther Sci*. Jan 2024;36(1):9-20. [doi: [10.1589/jpts.36.9](https://doi.org/10.1589/jpts.36.9)] [Medline: [38186969](https://pubmed.ncbi.nlm.nih.gov/38186969/)]
39. Zhou J, Jiang X, Yu W, et al. A smartphone app-based application enabling remote assessments of standing balance during the COVID-19 pandemic and beyond. *IEEE Internet Things J*. 2021;8(21):15818-15828. [doi: [10.1109/JIOT.2021.3064442](https://doi.org/10.1109/JIOT.2021.3064442)]
40. Harro CC, Garascia C. Reliability and validity of computerized force platform measures of balance function in healthy older adults. *J Geriatr Phys Ther*. 2019;42(3):E57-E66. [doi: [10.1519/JPT.000000000000175](https://doi.org/10.1519/JPT.000000000000175)] [Medline: [29324510](https://pubmed.ncbi.nlm.nih.gov/29324510/)]
41. Scaglioni-Solano P, Aragón-Vargas LF. Validity and reliability of the Nintendo Wii Balance Board to assess standing balance and sensory integration in highly functional older adults. *Int J Rehabil Res*. Jun 2014;37(2):138-143. [doi: [10.1097/MRR.000000000000046](https://doi.org/10.1097/MRR.000000000000046)] [Medline: [24445863](https://pubmed.ncbi.nlm.nih.gov/24445863/)]
42. Chang WD, Chang WY, Lee CL, Feng CY. Validity and reliability of Wii Fit Balance Board for the assessment of balance of healthy young adults and the elderly. *J Phys Ther Sci*. Oct 2013;25(10):1251-1253. [doi: [10.1589/jpts.25.1251](https://doi.org/10.1589/jpts.25.1251)] [Medline: [24259769](https://pubmed.ncbi.nlm.nih.gov/24259769/)]
43. Olsen S, Rashid U, Allerby C, et al. Smartphone-based gait and balance accelerometry is sensitive to age and correlates with clinical and kinematic data. *Gait Posture*. Feb 2023;100:57-64. [doi: [10.1016/j.gaitpost.2022.11.014](https://doi.org/10.1016/j.gaitpost.2022.11.014)] [Medline: [36481647](https://pubmed.ncbi.nlm.nih.gov/36481647/)]
44. 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*. Jul 2023;104:103-108. [doi: [10.1016/j.gaitpost.2023.06.016](https://doi.org/10.1016/j.gaitpost.2023.06.016)] [Medline: [37379735](https://pubmed.ncbi.nlm.nih.gov/37379735/)]
45. Arai T, Obuchi S, Shiba Y, Omuro K, Nakano C, Higashi T. The feasibility of measuring joint angular velocity with a gyro-sensor. *Arch Phys Med Rehabil*. Jan 2008;89(1):95-99. [doi: [10.1016/j.apmr.2007.07.051](https://doi.org/10.1016/j.apmr.2007.07.051)] [Medline: [18164337](https://pubmed.ncbi.nlm.nih.gov/18164337/)]

46. Huang CH, Nihey F, Fukushi K, Kajitani H, Nozaki Y, Nakahara K. Constructing and testing a lightweight model of converting single stride of in-shoe-motion-sensor-measured foot motion to TUG-represented mobility. *IEEE Sens Lett*. Aug 2023;7(8):1-4. [doi: [10.1109/LSENS.2023.3297797](https://doi.org/10.1109/LSENS.2023.3297797)]
47. Huet J, Nordez A, Sarcher A, Mathieu M, Cornu C, Boureau AS. Concordance of freehand 3D ultrasound muscle measurements with sarcopenia parameters in a geriatric rehabilitation ward. *J Cachexia Sarcopenia Muscle*. Feb 2025;16(1):e13648. [doi: [10.1002/jcsm.13648](https://doi.org/10.1002/jcsm.13648)] [Medline: [39575643](https://pubmed.ncbi.nlm.nih.gov/39575643/)]
48. Romero Morales C, Calvo Lobo C, Rodríguez Sanz D, Sanz Corbalán I, Ruiz Ruiz B, López López D. The concurrent validity and reliability of the leg motion system for measuring ankle dorsiflexion range of motion in older adults. *PeerJ*. 2017;5:e2820. [doi: [10.7717/peerj.2820](https://doi.org/10.7717/peerj.2820)] [Medline: [28070457](https://pubmed.ncbi.nlm.nih.gov/28070457/)]
49. Alemayoh TT, Lee JH, Okamoto S. A neural network-based lower extremity joint angle estimation from insole data. Presented at: 2023 20th International Conference on Ubiquitous Robots (UR); Jun 25-28, 2023; Honolulu, HI, USA. URL: <https://ieeexplore.ieee.org/abstract/document/10202438> [Accessed 2026-03-28] [doi: [10.1109/UR57808.2023.10202438](https://doi.org/10.1109/UR57808.2023.10202438)]
50. Pires IM, Denysyuk HV, Villasana MV, et al. Development technologies for the monitoring of six-minute walk test: a systematic review. *Sensors (Basel)*. Jan 12, 2022;22(2):581. [doi: [10.3390/s22020581](https://doi.org/10.3390/s22020581)] [Medline: [35062542](https://pubmed.ncbi.nlm.nih.gov/35062542/)]
51. Storm FA, Cesareo A, Reni G, Biffi E. Wearable inertial sensors to assess gait during the 6-minute walk test: a systematic review. *Sensors (Basel)*. May 6, 2020;20(9):2660. [doi: [10.3390/s20092660](https://doi.org/10.3390/s20092660)] [Medline: [32384806](https://pubmed.ncbi.nlm.nih.gov/32384806/)]
52. Ancillao A, Tedesco S, Barton J, O'Flynn B. Indirect measurement of ground reaction forces and moments by means of wearable inertial sensors: a systematic review. *Sensors (Basel)*. Aug 5, 2018;18(8):2564. [doi: [10.3390/s18082564](https://doi.org/10.3390/s18082564)] [Medline: [30081607](https://pubmed.ncbi.nlm.nih.gov/30081607/)]
53. Prisco G, Pirozzi MA, Santone A, et al. Validity of wearable inertial sensors for gait analysis: a systematic review. *Diagnostics (Basel)*. Dec 27, 2024;15(1):39795564. [doi: [10.3390/diagnostics15010036](https://doi.org/10.3390/diagnostics15010036)] [Medline: [39795564](https://pubmed.ncbi.nlm.nih.gov/39795564/)]
54. van Helden JFL, Martinez-Valdes E, Strutton PH, Falla D, Chiou SY. Reliability of high-density surface electromyography for assessing characteristics of the thoracic erector spinae during static and dynamic tasks. *J Electromyogr Kinesiol*. Dec 2022;67:102703. [doi: [10.1016/j.jelekin.2022.102703](https://doi.org/10.1016/j.jelekin.2022.102703)] [Medline: [36096034](https://pubmed.ncbi.nlm.nih.gov/36096034/)]
55. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*. Aug 22, 2006;4(54):54. [doi: [10.1186/1477-7525-4-54](https://doi.org/10.1186/1477-7525-4-54)] [Medline: [16925807](https://pubmed.ncbi.nlm.nih.gov/16925807/)]
56. Ortega-Bastidas P, Gómez B, Aqueveque P, Luarte-Martínez S, Cano-de-la-Cuerda R. Instrumented Timed Up and Go Test (iTUG)-more than assessing time to predict falls: a systematic review. *Sensors (Basel)*. Mar 24, 2023;23(7):3426. [doi: [10.3390/s23073426](https://doi.org/10.3390/s23073426)] [Medline: [37050485](https://pubmed.ncbi.nlm.nih.gov/37050485/)]
57. Baker N, Gough C, Gordon SJ. Inertial sensor reliability and validity for static and dynamic balance in healthy adults: a systematic review. *Sensors (Basel)*. Jul 30, 2021;21(15):5167. [doi: [10.3390/s21155167](https://doi.org/10.3390/s21155167)] [Medline: [34372404](https://pubmed.ncbi.nlm.nih.gov/34372404/)]
58. Lo PY, Su BL, You YL, Yen CW, Wang ST, Guo LY. Measuring the reliability of postural sway measurements for a static standing task: the effect of age. *Front Physiol*. 2022;13:850707. [doi: [10.3389/fphys.2022.850707](https://doi.org/10.3389/fphys.2022.850707)] [Medline: [35634138](https://pubmed.ncbi.nlm.nih.gov/35634138/)]
59. van Gulick DJJ, van der Leeden M, Lucas C, Stuijver MM. Reliability of static and dynamic balance assessments using the Footwork Pro pressure® platform in a clinical setting. *J Sci Med Sport*. Dec 2025;28(12):1034-1041. [doi: [10.1016/j.jsams.2025.08.006](https://doi.org/10.1016/j.jsams.2025.08.006)] [Medline: [40866129](https://pubmed.ncbi.nlm.nih.gov/40866129/)]
60. Balachandran AT, Orange ST, Wang Y, Lustin R, Vega A, Quiles N. Comparison of two popular transducers to measure sit-to-stand power in older adults. *PLoS ONE*. 2024;19(8):e0308808. [doi: [10.1371/journal.pone.0308808](https://doi.org/10.1371/journal.pone.0308808)] [Medline: [39133754](https://pubmed.ncbi.nlm.nih.gov/39133754/)]
61. Gao S, Chen J, Xia Y, et al. Wearable technologies for assisted mobility in the real world. *Nat Commun*. Dec 8, 2025;16(1):10988. [doi: [10.1038/s41467-025-67126-4](https://doi.org/10.1038/s41467-025-67126-4)] [Medline: [41361181](https://pubmed.ncbi.nlm.nih.gov/41361181/)]
62. Hetherington-Rauth M, Magalhães JP, Alcazar J, et al. Relative sit-to-stand muscle power predicts an older adult's physical independence at age of 90 yrs beyond that of relative handgrip strength, physical activity, and sedentary time: a cross-sectional analysis. *Am J Phys Med Rehabil*. Nov 1, 2022;101(11):995-1000. [doi: [10.1097/PHM.0000000000001945](https://doi.org/10.1097/PHM.0000000000001945)] [Medline: [35034060](https://pubmed.ncbi.nlm.nih.gov/35034060/)]
63. Milner T, Brown MRG, Jones C, Leung AWS, Brémault-Phillips S. Multidimensional digital biomarker phenotypes for mild cognitive impairment: considerations for early identification, diagnosis and monitoring. *Front Digit Health*. 2024;6:1265846. [doi: [10.3389/fdgh.2024.1265846](https://doi.org/10.3389/fdgh.2024.1265846)] [Medline: [38510280](https://pubmed.ncbi.nlm.nih.gov/38510280/)]
64. Fan S, Ye J, Xu Q, et al. Digital health technology combining wearable gait sensors and machine learning improve the accuracy in prediction of frailty. *Front Public Health*. 2023;11:1169083. [doi: [10.3389/fpubh.2023.1169083](https://doi.org/10.3389/fpubh.2023.1169083)] [Medline: [37546315](https://pubmed.ncbi.nlm.nih.gov/37546315/)]
65. Kelly D, Esquivel KM, Gillespie J, et al. Feasibility of sensor technology for balance assessment in home rehabilitation settings. *Sensors (Basel)*. Jun 28, 2021;21(13):4438. [doi: [10.3390/s21134438](https://doi.org/10.3390/s21134438)] [Medline: [34203571](https://pubmed.ncbi.nlm.nih.gov/34203571/)]

66. Zhong R, Rau PLP. Are cost-effective technologies feasible to measure gait in older adults? A systematic review of evidence-based literature. *Arch Gerontol Geriatr.* 2020;87:103970. [doi: [10.1016/j.archger.2019.103970](https://doi.org/10.1016/j.archger.2019.103970)] [Medline: [31743825](https://pubmed.ncbi.nlm.nih.gov/31743825/)]
67. Teel EF, Gay MR, Arnett PA, Slobounov SM. Differential sensitivity between a virtual reality balance module and clinically used concussion balance modalities. *Clin J Sport Med.* Mar 2016;26(2):162-166. [doi: [10.1097/JSM.0000000000000210](https://doi.org/10.1097/JSM.0000000000000210)] [Medline: [26505696](https://pubmed.ncbi.nlm.nih.gov/26505696/)]
68. Manupibul U, Tanthuwapathom R, Jarumethitanont W, Kaimuk P, Limroongreungrat W, Charoensuk W. Integration of force and IMU sensors for developing low-cost portable gait measurement system in lower extremities. *Sci Rep.* Jun 30, 2023;13(1):10653. [doi: [10.1038/s41598-023-37761-2](https://doi.org/10.1038/s41598-023-37761-2)] [Medline: [37391570](https://pubmed.ncbi.nlm.nih.gov/37391570/)]
69. Mannheim I, Wouters EJM, Köttl H, van Boekel LC, Brankaert R, van Zaalén Y. Ageism in the discourse and practice of designing digital technology for older persons: a scoping review. *Gerontologist.* Aug 24, 2023;63(7):1188-1200. [doi: [10.1093/geront/gnac144](https://doi.org/10.1093/geront/gnac144)] [Medline: [36130318](https://pubmed.ncbi.nlm.nih.gov/36130318/)]
70. De Santis KK, Mergenthal L, Christianson L, Busskamp A, Vonstein C, Zeeb H. Digital technologies for health promotion and disease prevention in older people: scoping review. *J Med Internet Res.* Mar 23, 2023;25:e43542. [doi: [10.2196/43542](https://doi.org/10.2196/43542)] [Medline: [36951896](https://pubmed.ncbi.nlm.nih.gov/36951896/)]
71. Ding J, Ren Y, Xu J, Hu Q, Chu T. Experience of older adults using smart devices and mHealth apps in nursing homes in Chinese megacities: a descriptive qualitative study. *Digit Health.* 2025;11:20552076251353334. [doi: [10.1177/20552076251353334](https://doi.org/10.1177/20552076251353334)] [Medline: [40547432](https://pubmed.ncbi.nlm.nih.gov/40547432/)]
72. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. *JMIR Aging.* Mar 21, 2023;6:e43197. [doi: [10.2196/43197](https://doi.org/10.2196/43197)] [Medline: [36943333](https://pubmed.ncbi.nlm.nih.gov/36943333/)]
73. Sharma N, Brinke JK, Van Gemert-Pijnen JEW, Braakman-Jansen LMA. Implementation of unobtrusive sensing systems for older adult care: scoping review. *JMIR Aging.* Oct 6, 2021;4(4):e27862. [doi: [10.2196/27862](https://doi.org/10.2196/27862)] [Medline: [34612822](https://pubmed.ncbi.nlm.nih.gov/34612822/)]
74. Teng S, Kim JY, Jeon S, et al. Analyzing optimal wearable motion sensor placement for accurate classification of fall directions. *Sensors (Basel).* Oct 4, 2024;24(19):6432. [doi: [10.3390/s24196432](https://doi.org/10.3390/s24196432)] [Medline: [39409472](https://pubmed.ncbi.nlm.nih.gov/39409472/)]
75. Trigg A, Ratitch B, Kruesmann F, Majumder M, Parfionovas A, Krahn U. Interpretation of change in novel digital measures: a statistical review and tutorial. *Digit Biomark.* 2025;9(1):52-66. [doi: [10.1159/000543899](https://doi.org/10.1159/000543899)] [Medline: [40103940](https://pubmed.ncbi.nlm.nih.gov/40103940/)]
76. Sadeghi Z, Alizadehsani R, Cifci MA, et al. A review of explainable artificial intelligence in healthcare. *Comput Electr Eng.* Aug 2024;118:109370. [doi: [10.1016/j.compeleceng.2024.109370](https://doi.org/10.1016/j.compeleceng.2024.109370)]
77. Gagnier JJ, Lai J, Morkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res.* Aug 2021;30(8):2197-2218. [doi: [10.1007/s11136-021-02822-4](https://doi.org/10.1007/s11136-021-02822-4)] [Medline: [33818733](https://pubmed.ncbi.nlm.nih.gov/33818733/)]

## Abbreviations

**AP:** anteroposterior

**COP:** center of pressure

**COSMIN:** Consensus-Based Standards for the Selection of Health Measurement Instruments

**ICC:** intraclass correlation coefficient

**IMU:** inertial measurement unit

**MDC:** minimal detectable change

**MIC:** minimal important change

**ML:** mediolateral

**OMI:** Outcome Measures Instruments in Systematic Reviews

**PICO:** Population, Intervention, Control, and Outcomes

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**PROSPERO:** International Prospective Register of Systematic Reviews

**RMS:** root mean square

**SEM:** standard error of measurement

*Edited by Stefano Brini; peer-reviewed by Charlotte Ahmadu, Jinyu Guo, Wenhao Qi; submitted 09.Sep.2025; final revised version received 07.Mar.2026; accepted 09.Mar.2026; published 14.Apr.2026*

*Please cite as:*

*Zhou S, Huang J, Yu J, Li X, Zhang C, Zhao Y, Hu M, Feng H*

*Digital Technologies and Biomarkers for Locomotor Capacity Assessment in Older Adults: Systematic Review*

*J Med Internet Res* 2026;28:e83814  
URL: <https://www.jmir.org/2026/1/e83814>  
doi: [10.2196/83814](https://doi.org/10.2196/83814)

© Shuhan Zhou, Jundan Huang, Jia Yu, Xiaoyang Li, Chi Zhang, Yinan Zhao, Mingyue Hu, Hui Feng. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 14.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.