Journal of Biomechanics 174 (2024) 112280

Contents lists available at ScienceDirect

Journal of Biomechanics

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

Markerless three-dimensional gait analysis in healthy older adults: test–retest reliability and measurement error

Andreia Carvalho[a,b,c,*], Jos Vanrenterghem[b], Sílvia Cabral[a], Ana Assunção[a], Rita Fernandes[a,d,e], António P. Veloso[a], Vera Moniz-Pereira[a]

[a] Universidade Lisboa, Faculdade de Motricidade Humana, CIPER, LBMF, P-1499-002 Lisboa, Portugal
[b] Musculoskeletal Rehabilitation Research Group, Faculty of Movement and Rehabilitation Sciences, Leuven KU, Belgium
[c] Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa, 1990-096 Lisboa, Portugal
[d] Instituto Politécnico de Setúbal, Escola Superior de Saúde, Campus do Instituto Politécnico de Setúbal, ESCE, Estefanilha, Edifício 2914-503 Setúbal, Portugal
[e] Comprehensive Health Research Centre, Nova Medical School, 1150-190 Lisboa, Portugal

ARTICLE INFO

Keywords:
Walking
Biomechanics
Older population
Measurement properties
Pointwise indices

ABSTRACT

In older adults, gait analysis may detect changes that signal early disease states, yet challenges in biomechanical screening limit widespread use in clinical or community settings. Recently, a markerless method from multi-camera video data has become accessible, making screenings less challenging. This study evaluated the test–retest reliability and measurement error of markerless gait kinematics and kinetics in healthy older adults. Twenty-nine healthy older adults performed gait analysis on two occasions, at preferred walking speed, using their everyday clothes. Lower limb angles and moments were averaged from 8 gait cycles. Integrated pointwise indices [Intraclass Correlation Coefficient ($ICC_{A,K}$) and Standard Error of Measurement (SEM)] were calculated for curve data, as well as $ICC_{A,K,}$ and SEM [95 % confidence intervals] for selected peaks. Generally, kinematic ICCs were good (>0.75) and reasonably stable throughout the gait cycle, except for the hip kinematics during the swing phase in the sagittal plane and pelvis tilt and rotation. The integrated and peaks SEM were $<2.4°$. The reliability of kinetics was similar (ICC>0.75), except for the transverse hip moment and abduction peak, fluctuating more during the swing than through the stance phase. SEM were < 0.07Nm/Kg. In conclusion, these results showed good overall test–retest reliability for markerless gait kinematics and kinetics for the hip, knee, and ankle joints, moderate for the pelvis angles, and error levels of $\leq 5°$, and SEM%$\leq$5% for the sagittal plane. This supports this method's use in assessing gait in healthy older adults, including kinetics, for which reliability data from markerless systems is difficult to find reported.

1. Introduction

Three-dimensional (3D) gait analysis (3DGA) represents a fundamental approach to understanding joint mechanics during walking (Kubota et al., 2012). In older adults, alterations in biomechanics during activities of daily living can be the early signs of various diseases and disabilities and thus 3DGA can inform future therapeutic decision-making related to these dysfunctions (de Campos et al., 2022; Wade et al., 2022).

In several settings, the widespread use of 3DGA (using biplanar videoradiography or markerbased optoelectronic systems) has been restricted due to limitations, such as exposure to radiation and/or the need for expert assessors (Wade et al., 2022). Whilst markerbased systems have been used in clinical populations, these systems require a laboratory setting and the placement of markers on the skin with the participant wearing minimal clothing, which is known to interfere with the participant's natural movement pattern (Sandau et al., 2014). By allowing to collect data in a real-world environment, wearable devices, such as inertial measurement units (IMUs), have often been chosen in gait analysis studies with older adults (Gausden et al., 2018; Jeon et al.,

---

* Corresponding author at: Laboratório de Biomecânica e Morfologia Funcional, Faculdade de Motricidade Humana, Estrada da Costa, 1499-002 Cruz Quebrada, Dafundo, Portugal.

E-mail addresses: andreiafjcarvalho@gmail.com (A. Carvalho), jos.vanrenterghem@kuleuven.be (J. Vanrenterghem), scabral@fmh.ulisboa.pt (S. Cabral), aassuncao@edu.ulisboa.pt (A. Assunção), rita.fernandes@ess.ips.pt (R. Fernandes), apveloso@fmh.ulisboa.pt (A.P. Veloso), veramps@fmh.ulisboa.pt (V. Moniz-Pereira).

https://doi.org/10.1016/j.jbiomech.2024.112280
Accepted 12 August 2024

Fig. 1. Self-selected clothing and sports shoes worn by 11 representative participants, randomly selected, in session 1 (top row) and session 2 (bottom row).

Table 1

Integrated ICC (Interclass Correlation Coefficient), Integrated SEM (Standard Error of Measurement), SEM percentage, MDC (Minimal Detectable Change) and MDC percentage, as well as ICC [(95% Confidence Interval (CI)] and SEM [(95% Confidence Interval (CI)] for peak values of Flex(Df)/Ad/Int Rot [flexion (dorsal flexion), adduction, internal rotation] and Ext(Pf)/Abd/Ext Rot [extension (plantar flexion), abduction, external rotation] of lower-limb joint angles and joint moments, for the hip, knee, and ankle joints in the sagittal, frontal, and transverse planes. For the pelvis, refers to the segment (seg) angle.

			Integrated			**Tilt/Obliquity Down/Int Rot Peak**		**Tilt/Obliquity Up/Ex. Rot Peak**	
			ICC	**SEM|%**	**MDC|%**	**ICC [CI]**	**SEM [CI]**	**ICC [CI]**	**SEM [CI]**
Seg. Angle (°)	**Pelvis**	**Sagittal**	0.53	1.8|141%	4.9|391%	0.52[-0.03,0.77]	1.7[1.4,2.3]	0.52[-0.02,0.77]	1.6[1.3,2.2]
		Frontal	0.82	1.0|21%	2.9|59%	0.76[0.47,0.89]	1.0[0.8,1.3]	0.87[0.73,0.94]	0.9[0.7,1.2]
		Transverse	0.72	1.8|20%	5.0|56%	0.69[0.35,0.85]	2.1[1.7,2.9]	0.72[0.41,0.87]	2.0[1.6,2.8]
						Flex(Ankle Df)/Ad/Int Rot Peak		**Ext(Ankle Pf)/Abd/Ex. Rot Peak**	
			ICC	**SEM|%**	**MDC|%**	**ICC [CI]**	**SEM [CI]**	**ICC [CI]**	**SEM [CI]**
Joint Angle (°)	**Hip**	**Sagittal**	0.74	2.1|5%	5.8|12%	0.71[0.41,0.87]	1.9[1.5,2.7]	0.81[0.60,0.91]	2.1[1.6,2.8]
		Frontal	0.81	1.6|20%	4.4|54%	0.76[0.49,0.88]	1.6[1.3,2.2]	0.83[0.63,0.92]	1.6[1.2,2.1]
		Transverse	0.80	1.5|26%	4.1|72%	0.77[0.51,0.89]	1.3[1.0,1.8]	0.91[0.80,0.96]	1.3[2.0,1.7]
	Knee	**Sagittal**	0.84	1.7|3%	4.7|8%	0.91[0.81,0.96]	1.2[0.1,1.7]	0.64[0.24,0.83]	1.5[1.2,2.0]
		Frontal	0.87	1.2|47%	3.2|132%	0.84[0.66,0.93]	1.2[0.9,1.6]	0.88[0.74,0.94]	1.0[0.8,1.4]
		Transverse	0.82	2.2|39%	6.1|102%	0.83[0.63,0.92]	1.8[1.4,2.5]	0.80[0.57,0.90]	2.4[1.9,3.3]
	Ankle	**Sagittal**	0.86	1.1|4%	3.2|11%	0.85[0.68,0.93]	1.1[0.9,1.5]	0.97[0.94,0.99]	1.1[0.8,1.4]
		Frontal	0.82	1.0|15%	2.9|42%	0.76[0.50,0.90]	1.2[0.9,1.7]	0.89[0.77,0.95]	1.0[0.8,1.4]
		Transverse	0.88	1.6|12%	4.3|33%	0.89[0.76,0.95]	1.3[1.0,1.8]	0.81[0.58,0.91]	1.9[1.5,2.8]
Joint Moment (Nm/kg)	**Hip**	**Sagittal**	0.82	0.06|4%	0.16|12%	0.87[0.72,0.94]	0.07[0.06,0.10]	0.90[0.76,0.95]	0.07[0.05,0.09]
		Frontal	0.84	0.05|4%	0.13|12%	0.85[0.69,0.93]	0.03[0.03,0.04]	0.72[0.38,0.87]	0.07[0.05,0.10]
		Transverse	0.74	0.02|8%	0.05|21%	0.63[0.19,0.83]	0.02[0.02,0.04]	0.71[0.38,0.87]	0.03[0.02,0.04]
	Knee	**Sagittal**	0.84	0.04|4%	0.11|21%	0.85[0.68,0.93]	0.04[0.03,0.06]	0.87[0.72,0.94]	0.07[0.06,0.10]
		Frontal	0.86	0.03|5%	0.08|14%	0.89[0.76,0.95]	0.01[0.01,0.02]	0.87[0.72,0.94]	0.04[0.03,0.06]
		Transverse	0.89	0.01|4%	0.03|10%	0.92[0.82,0.96]	0.02[0.01,0.03]	0.92[0.82,0.96]	0.01[0.01,0.02]
	Ankle	**Sagittal**	0.88	0.03|2%	0.09|6%	0.90[0.78,0.95]	0.03[0.02,0.04]	0.95[0.89,0.98]	0.04[0.03,0.06]
		Frontal	0.86	0.01|6%	0.04|16%	0.86[0.71,0.94]	0.02[0.01,0.02]	0.88[0.74,0.94]	0.04[0.03,0.05]
		Transverse	0.88	0.01|5%	0.02|13%	0.92[0.84,0.96]	0.005[0.004,0.007]	0.89[0.76,0.95]	0.02[0.01,0.02]

2019). However, the validity of the IMUs is questionable when assessing complex and outside the sagittal plane movements (Poitras et al., 2019), in addition to other constraints, like sensor drift (Gu et al., 2023). Thus, to try to overcome these constraints with existing methods, a novel markerless motion capture system using a deep-learning anatomical landmark prediction from video data – *Theia3D*, has recently become commercially available (Kanko et al., 2021b).

Regardless of the system, the usefulness of measurements depends on the extent to which professionals can rely on data. First, to derive meaningful clinical insights from the movement analyses, it is essential to explore the consistency of the measurements taken at different time points. This property, known as test–retest reliability, is defined as the proportion of the total variance in the measurement which is due to true differences between individuals. Second, the errors inherent to the measurement process should be considered acceptable. Here one refers to measurement error which is defined as the systematic and random error of a subject's score that is not attributed to true changes (Prinsen et al., 2018). Therefore, both parameters are important to assess (de Vet et al., 2006a, de Vet et al., 2006b).

Recent research has investigated the reliability of adult gait acquired with *Theia3D* markerless motion capture (Kanko et al., 2021a; Riazati et al., 2022). However, to our knowledge, there are no studies specifically performed on the older population. Older adults have been shown to exhibit greater gait variability than younger adults (Kowalski et al., 2022), which may have an effect on test–retest reliability. Since the reliability is expected to be population-specific (de Vet et al., 2006a, de

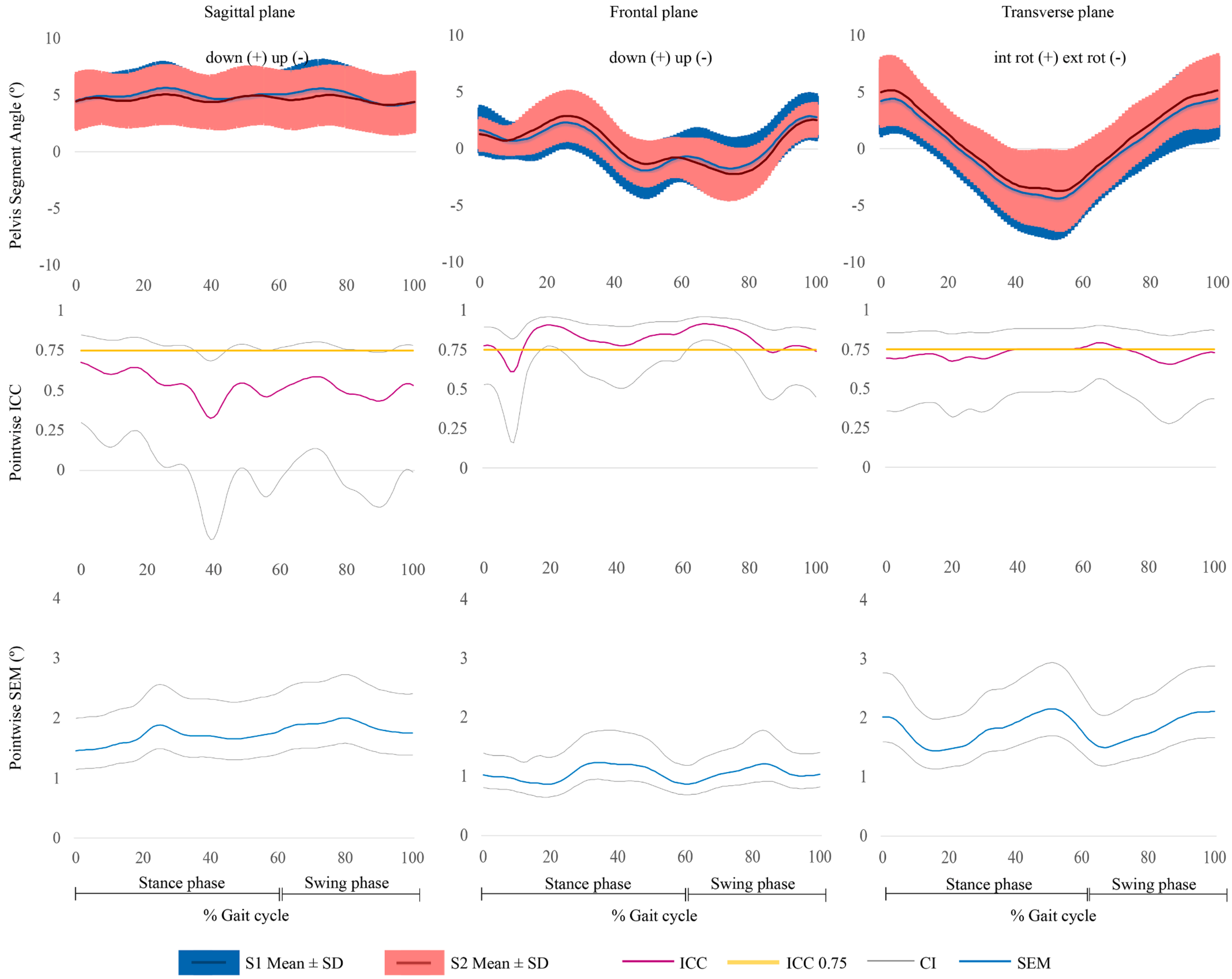

Fig. 2. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **pelvis segmental angle** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; int rot, internal rotation; ext rot, external rotation.

Vet et al., 2006b; Weir, 2005), this study aimed to evaluate test–retest reliability and measurement error of 3D gait kinetics and kinematics, measured with a markerless motion capture system, in healthy older adults.

2. Methods

2.1. Study design

A prospective within-examiner test–retest study was conducted.

2.2. Participants

For a predefined 5 % level of significance with 80 % power and a desired reliability coefficient of 0.90 with a minimum reliability of 0.75, a sample size of 29 was determined as the minimum required (Kraemer and Thiemann, 1987). To account for potential non-attenders, a convenience sample of 30 healthy older adults were recruited. One participant was excluded due to unavailability to perform the second assessment. Participants were eligible if they were aged ≥65; were independent community-dwelling; and could walk 100 m independently, without walking aids. Exclusion criteria included: to have been diagnosed with a clinical condition and/or currently taking medication that may affect the ability to walk; to indicate having any pain/symptom in the lower limb; to have cognitive impairment [Montreal Cognitive Assessment (MoCA)<23 (Carson et al., 2018)] and/or obesity (BMI>30Kg/m^2). Written informed consent was signed. The study was approved by the Faculty Ethics Committee (CEIFMHn°1/2022).

2.3. Data collection procedure and experimental setup

Gait analysis was conducted twice for each participant in sessions 1 and 2 separated by 9.3±2.9 days on average and carried out at the same time of the day (less than 3 h difference) in 27 of the 29 participants, in a laboratory setting. During the first session, participants' socio-demographic-health information was verified and MoCA (Freitas et al., 2011) was administered. The Composite Physical Function Scale (CPF) (Moniz-Pereira et al., 2023) was auto-filled and body mass and height were measured.

Gait data was collected, through Qualisys Track Manager (v2021.03.1 Qualisys AB, Gothenburg, Sweden), with 8 Miqus video cameras (Qualisys AB, Sweden) synchronized in time and space with 3 force plates (9283U014, Kistler Instruments Ltd, Winterthur, Switzerland; FP4060-07&FP4060-05-PT, BERTEC, Columbus OH, USA),

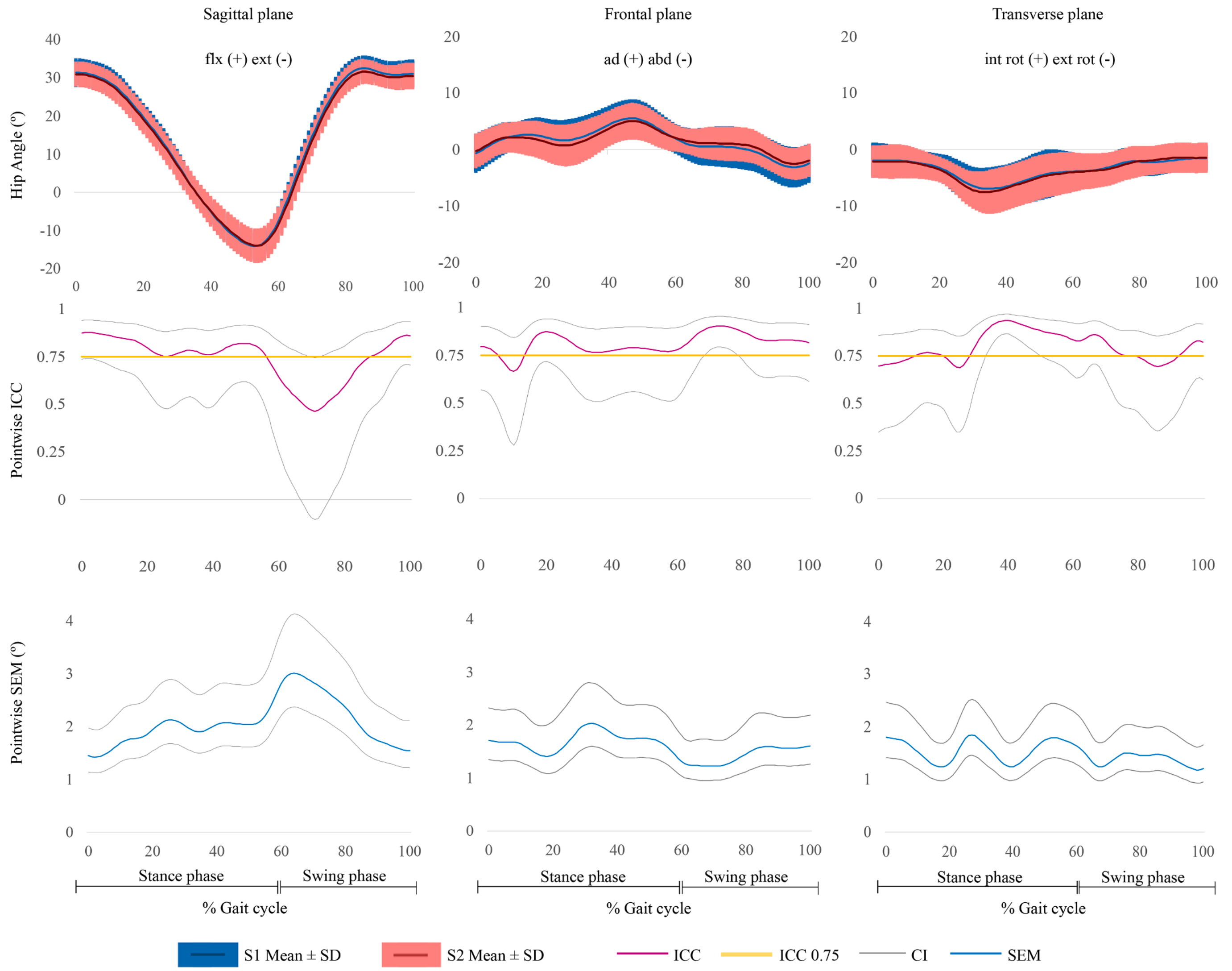

Fig. 3. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **hip joint angle** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation.

sampling at 85 Hz (1080p) and 850 Hz, respectively. Participants were instructed to continuously walk back and forth across the diagonal of the laboratory (12 m long), at their normal comfortable walking speed, during periods not exceeding 1–2 min., to prevent fatigue, following our laboratory testing method and until around ten valid kinetic gait cycles were obtained with each side. Overall, around 20 passages in the walkway were performed. The same assessor conducted all procedures. Participants wore their everyday clothing and sports shoes (Fig. 1).

2.4. Data processing

Video data were processed with Theia3D [v2023.1.0.310(patch: 1)–TMBatch, Theia Markerless Inc., Kingston, ON, Canada], using the default Inverse Kinematics 3D pose-estimation, with 6 degrees of freedom (DOF) at the pelvis and 3 DOF at the hip, knee, and ankle joints. The Woltring generalized cross-validity quintic-spline (GVCSPL) filter set at 8 Hz cut-off frequency was selected based on a residual analysis. The resulting 4x4 pose matrices, for each frame, were exported to c3d format.

Time-normalized right gait cycles were randomly selected from each session of every participant and averaged. Since no systematic difference was expected between the lower limbs in healthy subjects (Meldrum et al., 2014)], only the right gait cycles were used, for simplicity. Eight cycles were chosen based on a sensitivity analysis (supplementary material 1). Each gait cycle was determined using heel-strike events on the force plates (threshold of 20 N).

Gait speed was computed, as stride length divided by stride time.

Analog signals were low-pass filtered, using a 4th-order Butterworth filter (8 Hz) (Kristianslund et al., 2012). Lower limb joint angles were calculated using an XYZ Cardan sequence, and ZYX for pelvis segment angle (Baker, 2001). Internal joint moments were determined through Newton-Euler inverse dynamics, normalized to subjects' body mass, and computed relative to the proximal segment. These computations were performed in Visual3D (HAS-Motion Inc., Kingston, ON).

2.5. Data analysis

The pointwise Intraclass Correlation Coefficient ($ICC_{A,K}$) (McGraw and Wong, 1996) and its 95 % confidence interval (CI), based on a 2-way mixed-effects model, and the corresponding integrated pointwise index ICC, were calculated for the lower limb angles and moments, for entire time-series data, using an adaptation of the formula to compute $ICC_{(A,1)}$ (McGraw and Wong, 1996) developed by Pini et al. (2022) (supplementary material 2).

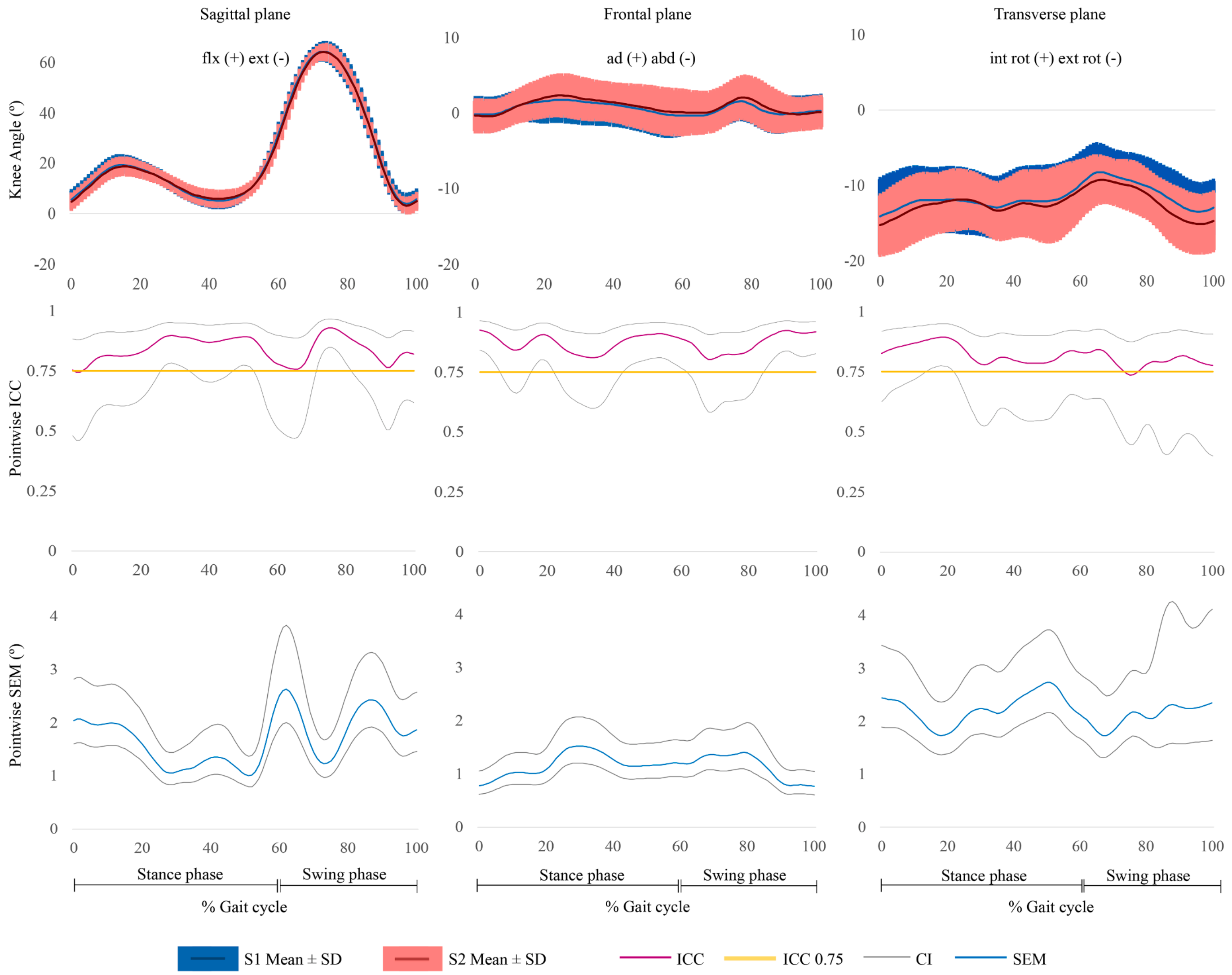

Fig. 4. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **knee joint angle** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation.

The pointwise Standard Error of Measurement (SEM) (95 % CI) and the integrated pointwise index SEM were computed for the same parameters using code supplied by Pini et al. (2022) (supplementary material 2). The ICC (95 % CI) and SEM (95 % CI) were calculated for kinematic and kinetic peak values (maximum angle peaks were extracted from the swing phase, except for the pelvis/hip frontal plane, and ankle sagittal/transverse plane, which were extracted from the stance phase; minimum angle peaks were extracted from the stance phase, with the same exceptions as just mentioned, that were extracted from the swing phase; all peak joint moments were extracted from the stance phase), using the former methods. These analyses were performed using R (RStudio,v2023.03.0 Build, Posit Software, PBC) (R Core Team, 2019). Minimal Detectable Change (MDC) was calculated from the integrated pointwise index SEM, using: $\text{MDC}=1.96\text{x}\sqrt{2}\text{xSEM}$ (de Vet et al., 2006a, de Vet et al., 2006b). SEM% was computed as $\text{SEM\%}=(\text{SEM}/\overline{\text{X}})\text{x}100$ [where *SEM* is the integrated pointwise index SEM and $\overline{X}$ is the amplitude of the mean for all observations from both sessions (Nair et al., 2012)] and MDC% similarly by replacing the SEM with MDC.

ICC above 0.75 indicated good reliability (excellent above 0.90), moderate between 0.5–0.75, and poor below 0.5 (Portney and Gross, 2020). ICCs >0.75, and absolute errors <5° for kinematics, were considered acceptable (McGinley et al., 2009). For kinetics errors, an arbitrary SEM% of ≤5 % was defined, based on SEMs/Root Mean Square Difference (RMSD) of previous studies (Fernandes et al., 2016; Kanko et al., 2024). RMSD was computed and provided in the (supplementary material 3) for comparison with other studies.

The gait speed in each session was compared using a paired *t*-test, in IBM SPSS Statistics 29, for an $\alpha = 0.05$. Plots were built in Microsoft Excel (v2307 Build16.0.16626.20170).

3. Results

Twenty-nine participants (13 females/16 males; 75.2±7.7 years old; Body Mass Index 25.0±2.6 kg/m^2), with MoCA score of 25.7±1.9, and advanced physical function (23.7±0.9 points) according to the CPF Scale, were included in this study. The average walking speed was not significantly different between sessions (1.28±0.15 vs 1.25±0.14 m/s, t (29) = 1.701, $p = 0.1$).

Kinematic integrated pointwise index ICCs ranged between 0.53–0.88, showing overall good reliability, except for the sagittal hip joint and sagittal and transverse pelvic segmental angle, (integrated pointwise ICC=0.74, 0.53 and 0.72, respectively) (Table 1).

Regarding ICCs of angle curves, hip joint sagittal plane motion showed good reliability during the stance and moderate during the

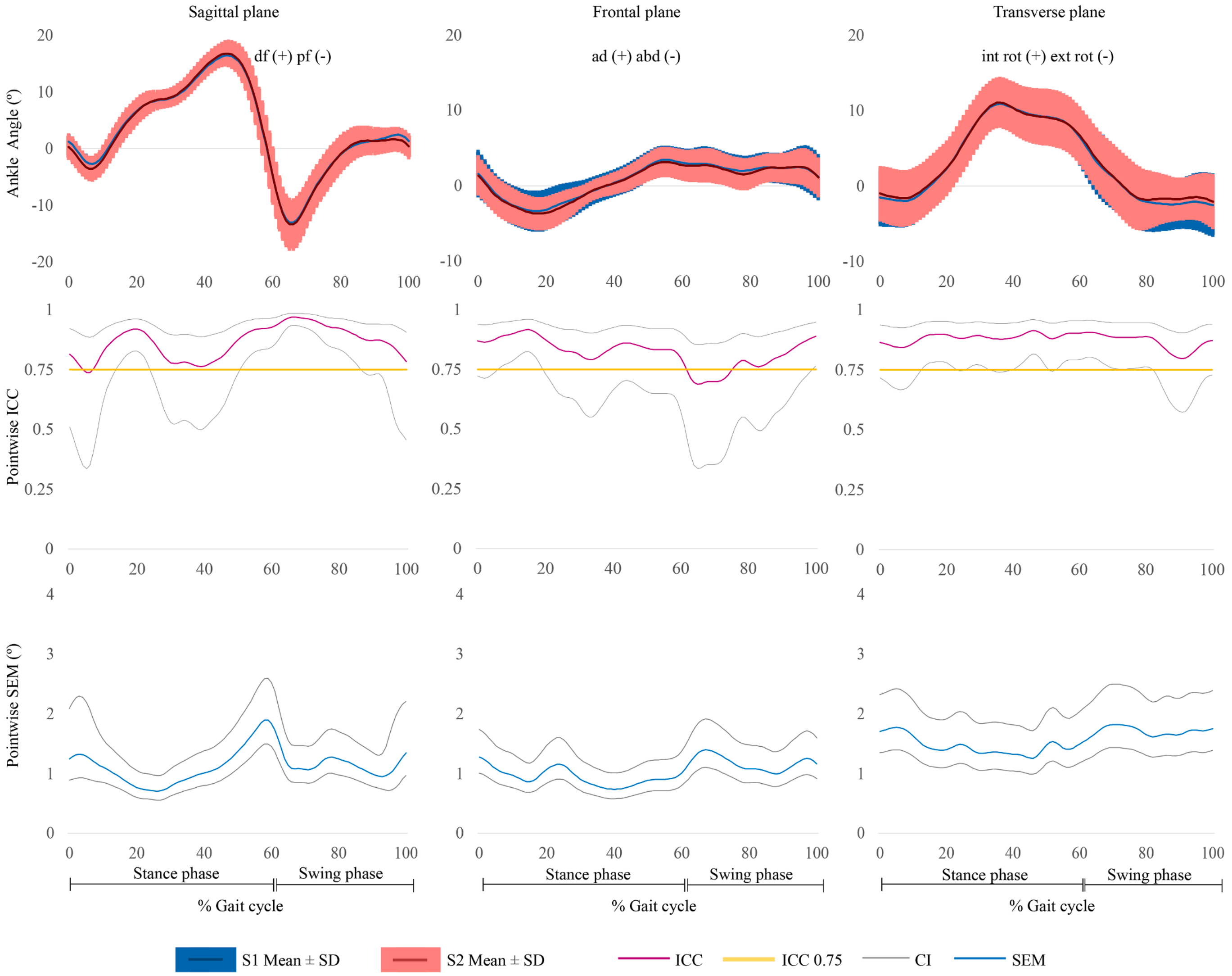

Fig. 5. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **ankle joint angle** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation; df, dorsal flexion; pf, plantar flexion.

swing phase, with some pointwise ICCs in the early-swing phase <0.5. Pointwise ICCs for the pelvis angle on the sagittal and transverse plane were globally <0.75 In the remaining planes, as well as for the knee and the ankle joints, the ICCs indicated, in general, good reliability. The ankle on the sagittal plane exhibited excellent reliability (>0.90) in the ~60–80 % of the gait cycle. (Figs. 2-5, middle row). For the peak signals, ICCs were >0.75, with exception of the hip flexion, knee extension, and peak pelvic angles on sagittal and transverse planes. Hip external rotation, knee flexion, and ankle plantar flexion peaks showed excellent reliability (>0.90) (Table 1).

Visual inspection of the mean ± standard deviation (SD) graphs showed the greatest overlap between session curves in the sagittal plane, except for the pelvic angle, more markedly during the middle-swing phase. Greater variability was observed in the frontal plane, for the hip, and in the transverse plane, for the 3 joints (mainly for the ankle joint) and transverse pelvic angle (Figs. 2-5, top row).

The pointwise index SEMs varied between 1.0°-2.2°, with the lowest SEM% (<5%) for the hip, knee, and ankle, in the sagittal plane, and the highest for the knee joint in the frontal (47 %) and transverse plane (39 %), which was similar for MDC%. The biggest exception was the SEM% in the sagittal plane for the pelvis angle (141 %) (Table 1). The RMSD ranged between 1.4°-2.8° (supplementary material 3).

The SEM curves varied between gait cycle phases, with the biggest fluctuation during the swing of the hip (sagittal plane), and knee joint (sagittal and transverse plane) (Figs. 2-5, bottom row). The MDC ranged from 2.9° to 6.1°.

The integrated ICCs for kinetics showed good reliability (0.74–0.89) in the three joints, except for the hip internal/external rotation (ICC=0.74) (Table 1).

Globally the pointwise ICC fluctuated more during the swing phase than the stance, showing good reliability with 3 main exceptions: for hip flexion/extension moments during the swing, for hip internal/external rotation moments during the late-stance and swing, and knee flexion/extension moments during the late-stance phase (Figs. 6-8, middle row).

The ICCs for the peaks were all >0.75, except for the peak moments at the hip in the transverse plane and abduction moment. Knee rotation moments, ankle plantar flexion moments, and ankle internal rotation peaks showed excellent reliability (>0.90).

Regarding waveform overlap, the greatest variability was found in the transverse plane moments, for the three joints (Figs. 6-8, top row).

The pointwise index SEMs varied between 0.01–0.06Nm/Kg and the MDC between 0.03–0.16Nm/Kg. Kinetic SEMs% were ≤5 %, except for

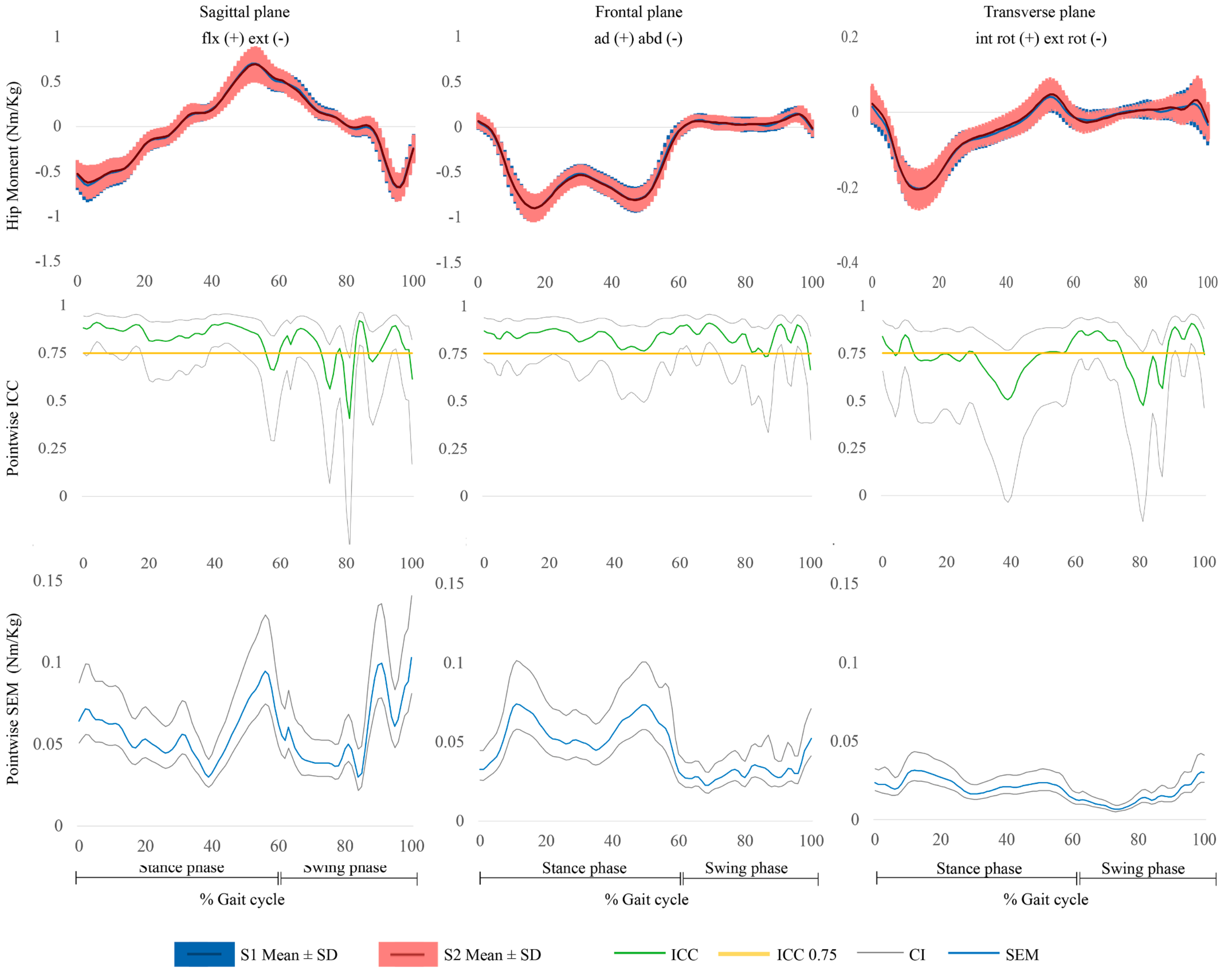

Fig. 6. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **hip joint moment** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation.

the transverse plane for the hip and the frontal plane for the ankle. The pointwise SEM during the gait cycle showed the highest values in the late stance phase in all kinetic parameters (Figs. 6-8, bottom row). The RMSD ranged between 0.01–0.08 Nm/Kg (supplementary material 3).

4. Discussion

To our knowledge, this study represents the first assessment of the test–retest reliability and measurement error of markerless 3D kinematic and kinetic gait time-series data in older adults. Generally, the results revealed good test–retest reliability (ICCs >0.75), and acceptable levels of error (<2.4°) for lower limb joint angles. Lower reliability (ICCs <0.75) was identified on the sagittal and transverse pelvis segmental angles. An overall good test–retest reliability was also found for the lower limb kinetics and measurement errors below 0.07Nm/kg (SEM% ≤5%, except for the hip in the transverse plane and the ankle in the frontal plane) for joint moments. MDC for lower limb joint angles and moments varied between 2.9°-5.8° and 0.02–0.16Nm/kg, respectively. These findings demonstrate that markerless gait kinematics and kinetics are reliable (capable of detecting changes over time in older adults) and have acceptable errors (≤5° for kinematics and SEM% ≤5% for kinetics). The reported MDC values will allow gait analysts to better interpret "real" changes, i.e., changes beyond the systematic and random error. This can help clinicians to monitor gait changes over time in this population and to detect early signs of diseases and dysfunctions.

The most reliable kinematic parameter was the sagittal plane ankle joint angle, showing good reliability, and low absolute and relative measurement error. The worst kinematic reliability was found for the sagittal plane pelvis segmental angle. While the sagittal plane hip joint angles did not reach the ICC cut-off for good reliability, this only applies to part of the swing phase. Thus, although a low ICC can mask good trial-to-trial consistency when between-subject variability is low (Weir, 2005) and generally the older adults had greater hip flexion variability through the early stance phase (Kowalski et al., 2022), the variability during the swing phase differed only slightly compared to the stance phase, which does not seem to justify the lower ICCs obtained. While the comparison of ICCs is limited due to the scarcity of investigations reporting these parameters within older populations, it is noteworthy that, in a young adult's markerless reliability study (Tamura et al., 2020), the ICCs of the sagittal plane hip joint angle were overall low, and lower during the swing phase compared to the stance. The low ICCs detected for the pelvis on the sagittal plane, along with the smaller overlap between the 2 session curves in the swing phase, observed in our study, may be contributing to the lower reliability in the hip in this

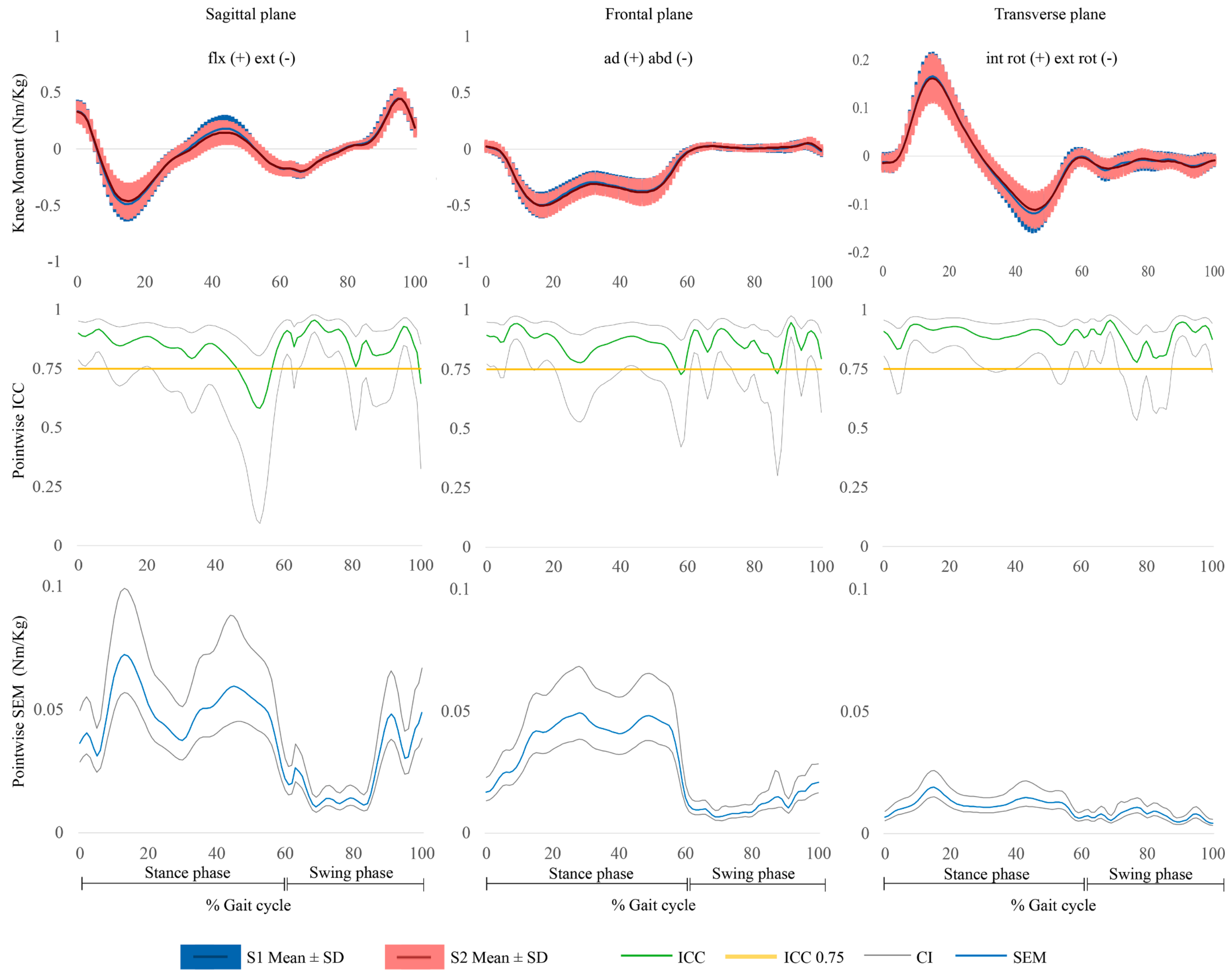

Fig. 7. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **knee joint moment** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation.

movement plane. In fact, in a study comparing different clothing conditions in running data using the Theia3D (Kanko et al., 2024), the pelvis segment angle in the mediolateral axis was one of those that showed the least consistency (but not the thigh angle around the same axis). Low reliability of pelvic tilt had already been reported for data acquired with markerbased systems (Kadaba et al., 1989; Meldrum et al., 2014) and associated with its low range of motion. In the other movement planes, a high variability seems to explain the higher ICCs obtained [e.g. hip transverse plane SD curve-average (Stergiou, 2004) of 2.8° for an amplitude of the mean of 5.7°].

Laroche et al. (2011) also reported ICCs below 0.75 for the hip sagittal plane and higher reliability for the frontal plane, as well as generally higher reliability for the osteoarthritis (OA) side, in marker-based data of OA patients. Considering this, our good reliability in healthy individuals (where the variability may be lower) can be indicative of promising results in pathological populations. Actually, in a study involving 10 patients with knee OA using Theia3D, a high repeatability of gait kinematics was described (Outerleys et al., 2024). Our results for peak joint angles were partly consistent with Fernandes et al. (2016), who analyzed the test–retest reliability of makerbased 3D gait kinematic and kinetic peak data in healthy young adults. Both studies found similarly high ICC values for the ankle joint, but our study showed generally higher ICCs for peak hip and knee joint angles. These higher ICCs in an older population thus seem good results, considering that one of the main difficulties of motion capture is the challenge of landmark identification covered by more lax muscle/skin and increased fat mass, common characteristics in older people (Moniz-Pereira et al., 2014), possibly even for landmark detection by markerless systems.

Our joint angle SEMs were all equal or below 3° during the gait cycle. However, the SEM% and MDC% (i.e., the relative magnitude of change) were considerably higher for the frontal and transverse planes in all joints (except for the sagittal pelvis angle), especially for the knee joint, meaning that higher differences are needed to imply a real change in these parameters. This higher relative magnitude of change was highly expected considering the lower range of movement in these planes in comparison to the one that occurs in the sagittal plane (excluding the pelvis). This results are in line with a previous Theia3D report (Kanko et al., 2021a) which found a between-session variability throughout the gait cycle mostly below 5° and the largest between-session variation in transverse planes. Our joint angle SEMs of peaks were also similar to those reported by Riazati et al. (2022), with an exception for the ankle joint for the sagittal and frontal plane, and for transverse plane peak angles, which were lower in our study (with the RMSD values following the same tendency). In fact, in our investigation, none of the participants

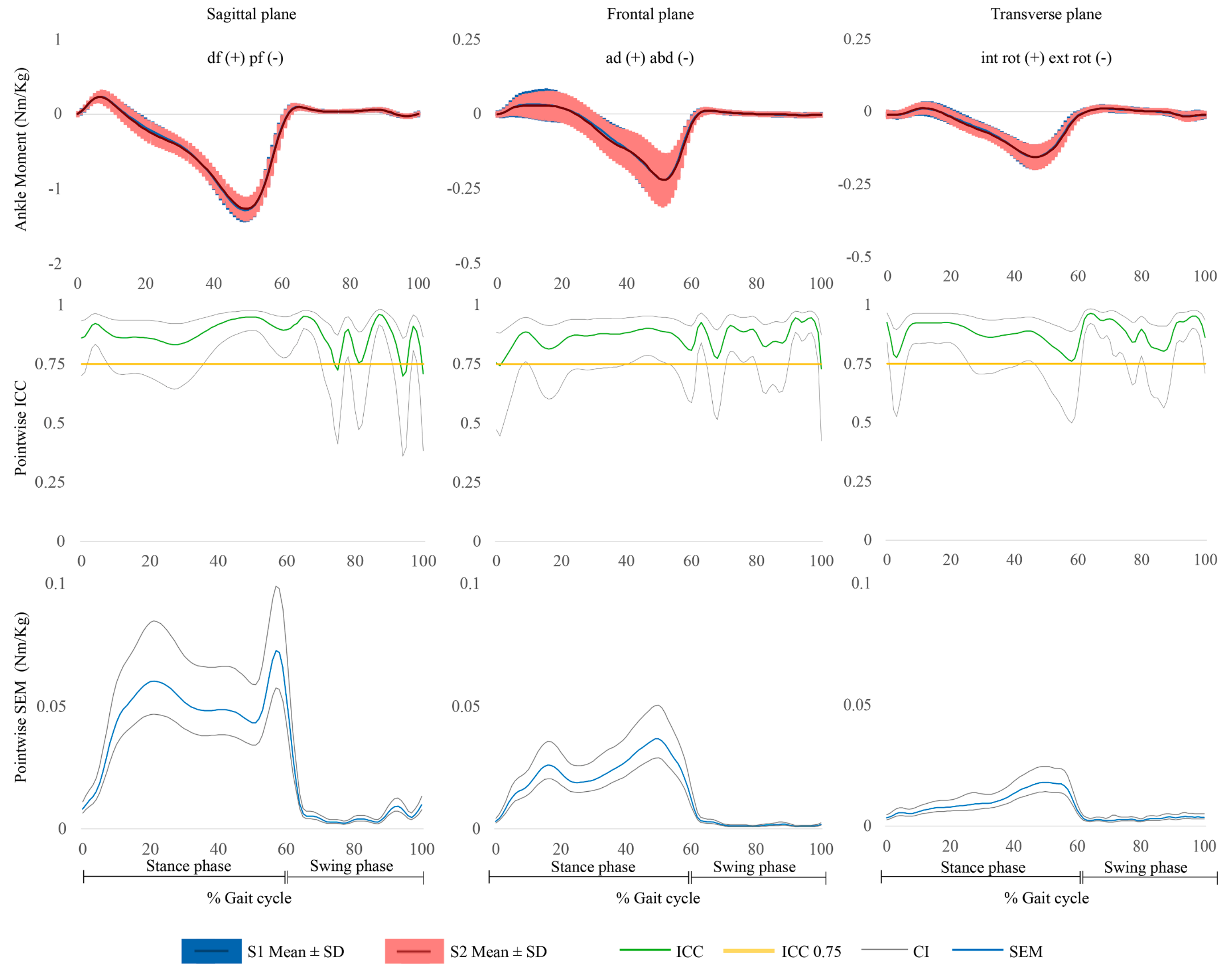

Fig. 8. Gait cycle waveforms (top row), pointwise ICC (middle row), and pointwise SEM (bottom row) and their 95% confidence intervals for **ankle joint moment** in the sagittal, frontal, and transverse plane. Abbreviations: ICC, Interclass Correlation Coefficient; SEM, Standard Error of Measurement; S1, Session 1; S2, Session 2; SD, standard deviation; flx, flexion; ext, extension; ad, adduction; abd, abduction; int rot, internal rotation; ext rot, external rotation; df, dorsal flexion; pf, plantar flexion.

wore skirts, which would be expected to better facilitate feature identification. Our SEMs and MDCs for peak joint angles were also generally lower compared to the Fernandes et al (2016) study. Since measurement error provides an absolute index of an instrument (de Vet et al., 2006a, de Vet et al., 2006b) irrespective of the group of subjects being studied, this seems to suggest that the measurement errors of peak joint kinematics obtained with a markerless system are smaller.

Similarly to kinematics, the most reliable kinetic parameter was the sagittal plane ankle joint moment, along with the transverse plane knee joint moment. Lower reliability was found for the transverse plane hip moment, together with high SEM% (>5%), and hip abduction peak moment. The greatest fluctuation in ICCs for all joint moments during the swing phase may be explained by their low magnitudes during this phase. Still, slightly better reliability was revealed for the kinetics, compared to kinematics. To our knowledge, no reliability on 3D gait kinetics from this markerless system has been reported yet. Compared to a markerbased study (Fernandes et al., 2016), globally, the hip and the knee joint moments ICCs were higher in our investigation. Our findings thus suggest that markerless reliability is higher compared with the markerbased, despite the interference of clothing. The SEMs and MDCs of joint kinetics were also generally lower than the ones found in the aforementioned study, indicating that this system is good for evaluating changes over time in lower limb kinetic parameters in an older population, particularly concerning the relevant ankle dorsi/plantar flexion moments (Boyer et al., 2017).

This study emphasizes the importance of testing the reliability of the entire gait cycle, especially when the curve data is not stable (Schelin et al., 2021), rather than just observing integrated indices (which do not have a validated method for confidence intervals) or discrete (peak) values. Our ICCs for sagittal plane ankle joint angles highlight this, showing considerable differences between integrated and peak plantar flexion ICCs (0.86 vs 0.97), and pointwise ICC reaching values close to 0.75 in the early-to-mid stance phase.

Some limitations of this study should be discussed. No specific clothing instructions were given to participants, which may have interfered with the pose estimation between sessions, despite the studies indicating that clothing likely only has a minimal impact on meaningful clinical interpretations (Keller et al., 2022), However, allowing participants to wear their preferred clothing reflects the markerless system's advantage, making our results valid for the most liberal implementation in the clinic. The data collection was performed in a laboratory setting with natural light restrictions, which can affect the quality of videos, although the light conditions were similar between sessions. Lastly, the linear registration of data in the time domain may ignore temporal variations, which could affect the pointwise ICC/SEM. While we do not believe this to be an important factor in normal gait data with clearly identifiable events, a non-linear data registration approach could be considered in further studies (Pataky et al., 2022).

In conclusion, our findings demonstrated good test–retest reliability, with an acceptable level of error, for markerless gait kinematics and kinetics for the hip, knee, and ankle joints, except for the pelvis tilt and rotation, the hip sagittal plane kinematics during the swing, and the hip kinetics in the transverse plane. Given the overall good results of reliability, this tool allows measurements with a low level of associated measurement error. Furthermore, if a pathological group has similar variability to the current group, the SEM and MDC values may also be used to monitor changes over time. Considering the applicability of this markerless system for 3DGA in the older adult population, particularly as it allows participants to wear their usual clothing in their natural environment, we are confident that it will have a role in future clinical contexts.

CRediT authorship contribution statement

Andreia Carvalho: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization, Funding acquisition. **Jos Vanrenterghem:** Conceptualization, Methodology, Writing – reviewing & editing, Supervision. **Sílvia Cabral:** Methodology, Investigation, Writing – reviewing & editing. **Ana Assunção:** Investigation, Writing – reviewing & editing. **Rita Fernandes:** Conceptualization, Writing – reviewing & editing. **António P. Veloso:** Resources, Writing – reviewing & editing. **Vera Moniz-Pereira:** Conceptualization, Methodology, Investigation, Writing – reviewing & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful to all the subjects who volunteered to participate in this study.

We thank Lina Schelin and Alessia Pini for their help and support regarding statistical analysis in this paper.

Thanks also to Camilla De Bleecker for the help with R-studio.

This work was supported by Fundação para a Ciência e a Tecnologia: Grant numbers DOI 10.54499/2020.07958.BD (PhD Grant), DOI: 10.54499/UIDB/00447/2020, and DOI: 10.54499/UIDP/00447/2020) attributed to CIPER–Centro Interdisciplinar de Estudo da Performance Humana (unit 447).

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jbiomech.2024.112280.

References

Baker, R., 2001. Pelvic angles: A mathematically rigorous definition which is consistent with a conventional clinical understanding of the terms. Gait Posture 13, 1–6. https://doi.org/10.1016/S0966-6362(00)00083-7.

Boyer, K.A., Johnson, R.T., Banks, J.J., Jewell, C., Hafer, J.F., 2017. Systematic review and meta-analysis of gait mechanics in young and older adults. Exp. Gerontol. 95, 63–70. https://doi.org/10.1016/j.exger.2017.05.005.

Carson, N., Leach, L., Murphy, K.J., 2018. A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. Int. J. Geriatr. Psychiatry 33, 379–388. https://doi.org/10.1002/gps.4756.

de Campos, D.da.S.F., Shokur, S., de Lima-Pardini, A.C., Runfeng, M., Bouri, M., Coelho, D.B., 2022. Kinematics predictors of spatiotemporal parameters during gait differ by age in healthy individuals. Gait Posture 96, 216–220. https://doi.org/10.1016/j.gaitpost.2022.05.034.

de Vet, H.C., Terwee, C.B., Ostelo, R.W., Beckerman, H., Knol, D.L., Bouter, L.M., 2006a. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. Health Qual. Life Outcomes 4, 3–7. https://doi.org/10.1186/1477-7525-4-54.

de Vet, H.C.W., Terwee, C.B., Knol, D.L., Bouter, L.M., 2006b. When to use agreement versus reliability measures. J. Clin. Epidemiol. 59, 1033–1039. https://doi.org/10.1016/j.jclinepi.2005.10.015.

Fernandes, R., Armada-da-Silva, P., Pool-Goudzwaard, A., Moniz-Pereira, V., Veloso, A. P., 2016. Three dimensional multi-segmental trunk kinematics and kinetics during gait: Test-retest reliability and minimal detectable change". Gait Posture 48, 18–25. https://doi.org/10.1016/j.gaitpost.2016.05.010.

Freitas, S., Simões, M.R., Alves, L., Santana, I., 2011. Montreal Cognitive Assessment (MoCA): normative study for the portuguese population. J. Clin. Exp. Neuropsychol. 33, 989–996. https://doi.org/10.1080/13803395.2011.589374.

Gausden, E.B., Sin, D., Levack, A.E., Wessel, L.E., Moloney, G., Lane, J.M., Lorich, D.G., 2018. Gait analysis after intertrochanteric hip fracture: does shortening result in gait impairment? J. Orthop. Trauma 32, 554–558. https://doi.org/10.1097/bot.0000000000001283.

Gu, C., Lin, W., He, X., Zhang, M., 2023. IMU-based motion capture system for rehabilitation applications: A systematic review. Biomim. Intell. Robot. 3, 1–13. https://doi.org/10.1016/j.birob.2023.100097.

Jeon, Y.T., Kim, B.R., Han, E.Y., Nam, K.W., Lee, S.Y., Park, Y.G., Suh, M.J., Kim, J.H., 2019. Post-operative physical performance factors associated with gait speed in patients surgically treated for hip fracture: A cross-sectional study. Ann. Rehabil. Med. 43, 570–580. https://doi.org/10.5535/arm.2019.43.5.570.

Kadaba, M.P., Ramakrishnan, H.K., Wootten, M.E., Gainey, J., Gorton, G., Cochran, G.V. B., 1989. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. J. Orthop. Res. 7, 849–860. https://doi.org/10.1007/978-1-4471-5451-8_101.

Kanko, R.M., Laende, E., Selbie, W.S., Deluzio, K.J., 2021a. Inter-session repeatability of markerless motion capture gait kinematics. J. Biomech. 121, 110422 https://doi.org/10.1016/j.jbiomech.2021.110422.

Kanko, R.M., Laende, E.K., Davis, E.M., Selbie, W.S., Deluzio, K.J., 2021b. Concurrent assessment of gait kinematics using marker-based and markerless motion capture. J. Biomech. 127, 110665 https://doi.org/10.1016/j.jbiomech.2021.110665.

Kanko, R.M., Outerleys, J.B., Laende, E.K., Selbie, W.S., Deluzio, K.J., 2024. Comparison of concurrent and asynchronous running kinematics and kinetics from marker-based motion capture and markerless motion capture under two clothing conditions. J. Appl. Biomech. 40, 129–137. https://doi.org/10.1123/jab.2023-0069.

Keller, V.T., Outerleys, J.B., Kanko, R.M., Laende, E.K., Deluzio, K.J., 2022. Clothing condition does not affect meaningful clinical interpretation in markerless motion capture. J. Biomech. 141, 111182 https://doi.org/10.1016/j.jbiomech.2022.111182.

Kowalski, E., Catelli, D.S., Lamontagne, M., 2022. Gait variability between younger and older adults: An equality of variance analysis. Gait Posture 95, 176–182. https://doi.org/10.1016/j.gaitpost.2022.04.022.

Kraemer, H.C., Thiemann, S., 1987. Correlation Coefficients. In: How Many Subjects? Statistical Power Analysis in Research. Sage Publications Inc, California.

Kristianslund, E., Krosshaug, T., Van den Bogert, A.J., 2012. Effect of low pass filtering on joint moments from inverse dynamics: Implications for injury prevention. J. Biomech. 45, 666–671. https://doi.org/10.1016/j.jbiomech.2011.12.011.

Kubota, M., Uchida, K., Kokubo, Y., Shimada, S., Matsuo, H., Yayama, T., Miyazaki, T., Takeura, N., Yoshida, A., Baba, H., 2012. Changes in gait pattern and hip muscle strength after open reduction and internal fixation of acetabular fracture. Arch. Phys. Med. Rehabil. 93, 2015–2021. https://doi.org/10.1016/j.apmr.2012.01.016.

Laroche, D., Duval, A., Morisset, C., Beis, J.N., D'Athis, P., Maillefert, J.F., Ornetti, P., 2011. Test-retest reliability of 3D kinematic gait variables in hip osteoarthritis patients. Osteoarthr. Cartil. 19, 194–199. https://doi.org/10.1016/j.joca.2010.10.024.

McGinley, J.L., Baker, R., Wolfe, R., Morris, M.E., 2009. The reliability of three-dimensional kinematic gait measurements: A systematic review. Gait Posture 29, 360–369. https://doi.org/10.1016/j.gaitpost.2008.09.003.

McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30–46. https://doi.org/10.1037/1082-989X.1.1.30.

Meldrum, D., Shouldice, C., Conroy, R., Jones, K., Forward, M., 2014. Test-retest reliability of three dimensional gait analysis: including a novel approach to visualising agreement of gait cycle waveforms with bland and altman plots. Gait Posture 39, 265–271. https://doi.org/10.1016/j.gaitpost.2013.07.130.

Moniz-Pereira, V., Cabral, S., Carnide, F., Veloso, A.P., 2014. Sensitivity of joint kinematics and kinetics to different pose estimation algorithms and joint constraints in the elderly. J. Appl. Biomech. 30, 446–460. https://doi.org/10.1123/jab.2013-0105.

Moniz-Pereira, V., Silva, T.F., Cruz, E.B., Carnide, F., 2023. The european portuguese version of the composite physical function scale: cross-cultural adaptation and validation in community-dwelling older adults. J. Aging Phys. Act. 31, 860–865. https://doi.org/10.1123/japa.2022-0135.

Nair, P.M., George Hornby, T., Louis Behrman, A., 2012. Minimal detectable change for spatial and temporal measurements of gait after incomplete spinal cord injury. Top. Spinal Cord Inj. Rehabil. 18, 273–281. https://doi.org/10.1310/sci1803-273.

Outerleys, J., Mihic, A., Keller, V., Laende, E., Deluzio, K., 2024. Markerless motion capture provides repeatable gait outcomes in patients with knee osteoarthritis. J. Biomech. 168, 112115 https://doi.org/10.1016/j.jbiomech.2024.112115.

Pataky, T.C., Robinson, M.A., Vanrenterghem, J., Donnelly, C.J.W., 2022. Simultaneously assessing amplitude and temporal effects in biomechanical trajectories using nonlinear registration and statistical nonparametric mapping. J. Biomech. 136, 111049 https://doi.org/10.1016/j.jbiomech.2022.111049.

Pini, A., Markström, J.L., Schelin, L., 2022. Test–retest reliability measures for curve data: an overview with recommendations and supplementary code. Sport. Biomech. 21, 179–200. https://doi.org/10.1080/14763141.2019.1655089.

Poitras, I., Bielmann, M., Campeau-lecours, A., Mercier, C., Bouyer, L.J., Roy, J., 2019. Validity and reliability of wearable sensors for joint. Sensors 19, 1–17. https://doi.org/10.3390/s19071555.

Portney, L.G., Gross, D., 2020. Measurement Revisited: Reliability and Validity Statistics. In: Foundations of Clinical Research: Applications to Evidence-Based Practice. F.A. Davis, Philadelphia.

Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., Terwee, C.B., 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual. Life Res. 27, 1147–1157. https://doi.org/10.1007/s11136-018-1798-3.

R Core Team, 2019. R: A Language and Environment for Statistical Computing.

Riazati, S., McGuirk, T.E., Perry, E.S., Sihanath, W.B., Patten, C., 2022. Absolute reliability of gait parameters acquired with markerless motion capture in living domains. Front. Hum. Neurosci. 16, 1–11. https://doi.org/10.3389/fnhum.2022.867474.

Sandau, M., Koblauch, H., Moeslund, T.B., Aanæs, H., Alkjær, T., Simonsen, E.B., 2014. Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane. Med. Eng. Phys. 36, 1168–1175. https://doi.org/10.1016/j.medengphy.2014.07.007.

Schelin, L., Pini, A., Markström, J.L., Häger, C.K., 2021. Test-retest reliability of entire time-series data from hip, knee and ankle kinematics and kinetics during one-leg hops for distance: Analyses using integrated pointwise indices. J. Biomech. 124, 110546 https://doi.org/10.1016/j.jbiomech.2021.110546.

Stergiou, N., 2004. Innovative analyses of human movement: Analytical tools for movement research. Human Kinetics, Champaign, IL.

Tamura, H., Tanaka, R., Kawanishi, H., 2020. Reliability of a markerless motion capture system to measure the trunk, hip and knee angle during walking on a flatland and a treadmill. J. Biomech. 109, 109929 https://doi.org/10.1016/j.jbiomech.2020.109929.

Wade, L., Needham, L., McGuigan, P., Bilzon, J., 2022. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ 10, 1–27. https://doi.org/10.7717/peerj.12995.

Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Cond. Res. 19, 231–240.