Gait & Posture 123 (2026) 110020

Contents lists available at ScienceDirect

Gait & Posture

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

How many strides are needed for reliable markerless gait analysis?

Andreia Carvalho [a,b,c,*], Jos Vanrenterghem [b], Todd C. Pataky [d], Mark A. Robinson [e], António P. Veloso [a], Vera Moniz-Pereira [a]

[a] LBMF, CIPER, Faculdade de Motricidade Humana, Universidade Lisboa, Lisboa P-1499-002, Portugal
[b] Musculoskeletal Rehabilitation Research Group, Faculty of Movement and Rehabilitation Sciences, Leuven KU, Belgium
[c] Escola Superior de Saúde de Lisboa, Instituto Politécnico de Lisboa, Lisboa 1990-096, Portugal
[d] Department of Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
[e] Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, United Kingdom

ARTICLE INFO

Keywords:
Older adults
Test-retest
Intraclass Correlation Coefficient
Standard Error of Measurement
Biomechanics

ABSTRACT

Background/Aim: Markerless motion capture is an emerging tool for gait analysis. In some populations, e.g., older adults, traditional gait analysis poses practical challenges, such as reduced assessment tolerance, and the number of strides collected can be limited. This study aimed to investigate the influence of the number of strides on test-retest reliability and measurement error of markerless gait biomechanics in older adults.

Methods: Twenty strides were extracted from 29 healthy older adults for each of two sessions. Lower-limb kinematics and kinetics were computed. Subsequently, non-consecutive random subsets of k = 2–19 strides were averaged within-subjects and within-sessions, including scenarios with unequal k between sessions. Integrated Intraclass Correlation Coefficients ($iICC_{A,k}$) and Standard Errors of Measurement (SEM) were calculated for trajectory data. $ICC_{A,k}$ [Confidence Intervals] were computed for the range of motion and peaks. Two arbitrary thresholds for the minimally acceptable number of strides were combined: (1) the smallest k that yielded an ICC within 10 % of the maximum ICC across all k, and (2) an absolute ICC threshold of 0.75. SEM$\leq$ 2º was deemed suitable for kinematics, and SEM%$\leq$ 5 % for kinetics.

Results: For joint angles and moments, iICC dropped less than 10 % from the highest iICC when using $\geq$ 7 strides with an equal number of strides across sessions, attaining iICCs$\geq$ 0.75. Reducing the number of strides in one session had less impact than reducing both equally. Lower Confidence intervals were generally $\geq$ 0.75 for discrete parameters. Kinematic SEM rarely exceeds 2º. Globally, 4 strides are needed to have a kinetics SEM %$\leq$ 5 %.

Conclusion: A minimum of 7 strides contributing to the average observation is generally sufficient to achieve reliable markerless kinematics and kinetics in older adults. These results have particular relevance to populations who may experience limited tolerance for lengthy assessments. Allowing flexibility in stride number collected across sessions, while maintaining reliability, contributes to optimizing data collection strategies.

1. Introduction

Recently, markerless motion capture has received growing interest for its potential to streamline three-dimensional gait analysis, reduce assessment time, and avoid markers being placed on the skin [1,2]. This can be particularly valuable in populations with challenging anthropometric characteristics [3] or reduced willingness to engage in diagnostic procedures [4], like older adults. Gait alterations are key contributors to activity limitations in this population [5]. However, the utility of gait assessments depends on measurement properties such as reliability.

Given the inherent variability of movement, a single stride may not be representative of the individual's gait pattern. Thus, averaging a number of strides is the typical assessment approach [6]. Averaging repetitions improves reliability (i.e., the extent to which different sources of variation influence the measurement [7]) and reduces measurement error. Using a marker-based motion capture method,

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

E-mail addresses: andreiafjcarvalho@gmail.com (A. Carvalho), jos.vanrenterghem@kuleuven.be (J. Vanrenterghem), pataky.todd.2m@kyoto-u.ac.jp (T.C. Pataky), M.A.Robinson@ljmu.ac.uk (M.A. Robinson), apveloso@fmh.ulisboa.pt (A.P. Veloso), veramps@fmh.ulisboa.pt (V. Moniz-Pereira).

https://doi.org/10.1016/j.gaitpost.2025.110020
Received 4 June 2025; Received in revised form 19 September 2025; Accepted 21 October 2025
Available online 22 October 2025

Monaghan et al. [8] advocate using ten trials for kinematic and kinetic gait variables to achieve good reliability. However, practical constraints, such as fatigue, clinical conditions, or time constraints, often limit the number of strides that can be collected, particularly in gait analysis of older adults [9].

Thus, this study aimed to test the influence of the number of strides on test-retest reliability and measurement error of averaged kinematic and kinetic markerless gait profiles in older adults.

2. Methods

Gait analysis was performed in 29 healthy older adults twice (13 females/16 males; 75.2 ± 7.7 years old), 9.3 ± 2.9 days apart, using 8 video-cameras (85 Hz) (Miqus, Qualysis, SE) and 3 force plates (850 Hz) (9283U014 - Kistler Instruments Ltd, Winterthur, Switzerland; FP4060–07&FP4060–05-PT-BERTEC - Columbus, OH, USA). Participants walked around 20 passages across the laboratory walkway (12 m long), at a self-selected speed, in their everyday clothes/shoes [10]. Video data were processed with Theia3D (Markerless Inc., CA, v2023.1.0.310) using Inverse Kinematics pose estimation. Lower-limb joint angles (XYZ Cardan) and internal moments were computed relative to the proximal segment in Visual3D (HAS-Motion Inc., CA). The Theia3D model [11], and further details are described elsewhere [12].

Twenty time-normalized strides (10right+10left) were extracted from each session of every participant and averaged. Plus, from these 20 strides, non-consecutive 2–19 were randomly selected (once per combination) and averaged (from each session/per participant), an approach acknowledged in previous work [13].

The integrated Intraclass Correlation Coefficient (iICC) absolute agreement (A) for average measurements (k) [14] and the integrated SEM were calculated for the lower-limb angles and moments, using the method developed by Pini et al. [15], and the formulae described in Carvalho et al. [12]. The $ICC_{A,k}$ [95 % confidence intervals (CI)] were calculated for kinematic and kinetic discrete parameters (range of motion [ROM] and peak values), since the expression of single ICC values may not accurately reflect population-level variability, making CI examination particularly relevant.

Two arbitrary thresholds were defined: (1) the minimally acceptable number of strides was regarded as the point from which the iICC deflected less than 10 % from the highest iICC for the corresponding variable [16]; (2) additionally, a global threshold of 0.75 was deemed an acceptable ICC, with narrower CIs reflecting greater precision. SEM ≤ 2° was considered suitable for kinematics [17], and ≤ 5 % of SEM% for kinetics [12].

To test the effect on the ICC of using an unequal number of strides contributing to the average profiles across time-points, the iICC was computed by reducing the number of averaged strides in the second session from the minimally acceptable number of strides to 2.

Variable	Number of strides 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Hip JA S	0.65	0.64	0.67	0.77	0.73	0.73	0.73	0.75	0.76	0.73	0.76	0.75	0.74	0.75	0.74	0.75	0.75	0.75	0.75
Hip JA F	0.68	0.69	0.76	0.73	0.78	0.76	0.77	0.77	0.79	0.80	0.82	0.79	0.82	0.80	0.79	0.81	0.80	0.81	0.81
Hip JA T	0.60	0.69	0.63	0.75	0.70	0.74	0.77	0.81	0.81	0.77	0.80	0.77	0.77	0.79	0.81	0.81	0.81	0.81	0.82
Knee JA S	0.75	0.81	0.76	0.81	0.79	0.79	0.79	0.83	0.82	0.82	0.83	0.82	0.83	0.83	0.82	0.82	0.83	0.83	0.83
Knee JA F	0.80	0.83	0.83	0.82	0.86	0.87	0.87	0.88	0.87	0.88	0.88	0.88	0.88	0.88	0.89	0.88	0.89	0.89	0.89
Knee JA T	0.74	0.71	0.79	0.79	0.83	0.80	0.83	0.86	0.86	0.82	0.80	0.82	0.85	0.85	0.86	0.85	0.85	0.86	0.86
Ankle JA S	0.80	0.85	0.83	0.86	0.83	0.86	0.87	0.88	0.88	0.89	0.88	0.89	0.89	0.89	0.89	0.89	0.90	0.90	0.90
Ankle JA F	0.65	0.60	0.63	0.73	0.71	0.76	0.76	0.75	0.76	0.75	0.77	0.76	0.78	0.76	0.78	0.77	0.77	0.77	0.78
Ankle JA T	0.69	0.78	0.81	0.76	0.85	0.84	0.83	0.85	0.86	0.87	0.85	0.89	0.88	0.87	0.89	0.89	0.89	0.90	0.90
Hip JM S	0.69	0.76	0.79	0.80	0.81	0.81	0.82	0.83	0.84	0.83	0.85	0.84	0.84	0.85	0.86	0.85	0.86	0.85	0.86
Hip JM F	0.69	0.75	0.79	0.77	0.80	0.83	0.84	0.82	0.82	0.84	0.84	0.87	0.86	0.86	0.85	0.85	0.85	0.86	0.86
Hip JM T	0.60	0.68	0.69	0.68	0.71	0.72	0.73	0.74	0.75	0.74	0.75	0.75	0.75	0.77	0.75	0.76	0.77	0.77	0.77
Knee JM S	0.75	0.78	0.82	0.84	0.83	0.85	0.84	0.86	0.86	0.88	0.88	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.88
Knee JM F	0.75	0.74	0.77	0.83	0.83	0.86	0.85	0.87	0.85	0.86	0.88	0.88	0.88	0.89	0.88	0.89	0.89	0.89	0.89
Knee JM T	0.73	0.81	0.84	0.87	0.84	0.89	0.89	0.87	0.89	0.89	0.90	0.90	0.90	0.91	0.90	0.91	0.91	0.91	0.91
Ankle JM S	0.75	0.80	0.85	0.85	0.87	0.88	0.87	0.87	0.88	0.88	0.89	0.90	0.89	0.89	0.90	0.89	0.90	0.90	0.90
Ankle JM F	0.61	0.73	0.76	0.70	0.73	0.79	0.83	0.84	0.80	0.85	0.84	0.85	0.87	0.84	0.85	0.88	0.87	0.87	0.87
Ankle JM T	0.67	0.79	0.80	0.85	0.86	0.87	0.87	0.89	0.90	0.89	0.90	0.90	0.92	0.92	0.91	0.92	0.92	0.92	0.92

%ICC_rMaxICC					
>26%	[25-21]%	[20-16]%	[15-11]%	[10-6]%	≤5%

Fig. 1. A tabulated color map corresponding to the percentage relative to the highest for Integrated Intraclass Correlation Coefficient (iICC) per variable (%ICC_rMaxICC), obtained from an equal number of gait strides averaged (2–20), along with the corresponding absolute iICC value (2 decimal places reported), for joint angles (JA) and joint moments (JM), for sagittal (S), frontal (F), and transverse (T) planes. The highest iICC for each variable, which serves as the reference for calculating the relative percentage, is marked with a square.

Analyses were performed in R (RStudio, v2023.03.0 Build, Posit Software, PBC), and plots were built using Python v3.11.7.

3. Results

Across all variables, %ICC dropped less than 10 % compared to the highest ICC for averages based on 7 or more strides with an equal number of strides contributing to the averages of both sessions (Fig. 1). ICCs varied between 0.77 and 0.92. Joint angles iICCs exceeded 0.75 when using 7 or more strides, except for hip sagittal (0.73–0.76) and transverse (0.74–0.82) planes (Supplementary material - Figure SP1). Joint moments generally showed iICCs above 0.75, except for hip transverse plane (0.72–0.77) (Supplementary material - Figure SP2). Overall, transverse plane kinematics required more strides to achieve ICC stability.

For kinematics, SEM exceeded 2° only in the hip sagittal plane, and in the knee and ankle transverse planes, when using less than 4, 3, and 2 strides, respectively (Figure SP1). For joint moments, generally 4 strides were needed to obtain a SEM% ≤ 5 %, with the exception of hip transverse plane, in which this threshold was exceeded even with 20 strides, and for ankle frontal plane, which needed more than 12 strides (Supplementary material - Figure SP2).

For discrete parameters, the results were generally consistent with those for iICC (Fig. 2), with two minor exceptions: hip flexion/extension ROM required only 3 strides, and peak ankle plantarflexion moment required 2 strides. CI analysis showed reduced CI widths with increasing stride numbers. Lower CI limits were generally above 0.75, except for sagittal knee ROM and peak hip extension angle.

When an unequal number of strides contributed to the averages across sessions, the results showed that using fewer strides in the retest session affects iICC for most variables to a lesser extent in comparison to when fewer strides are averaged in both sessions. The most explicit example of this can be seen for frontal plane ankle joint angles, where iICCs for the 5/7 and 6/7 stride number combinations still surpassed the 0.75 threshold, unlike the 5/5 or 6/6 scenarios (Fig. 3).

4. Discussion

This study assessed the effect of stride number on markerless gait test-retest reliability and measurement error in older adults, including when averages are obtained from an unequal number of strides across time points. Our findings suggest that a minimum of 7 strides is generally sufficient to achieve stable ICC. Importantly, our results reveal that a lower number of strides in one of the sessions may still be adequate for re-evaluation, compared to forcing an equal but smaller number of strides between sessions.

Transverse plane kinematics, particularly in the hip and ankle joints, required the highest number of strides to achieve stable reliability. This agrees with Monaghan et al. [8], who also found that kinematics were more variable than kinetics. However, in that study, the measurement of 10 strides was recommended. Notably, single-value reliability metrics, such as the ICC, may not fully capture population-level variability,

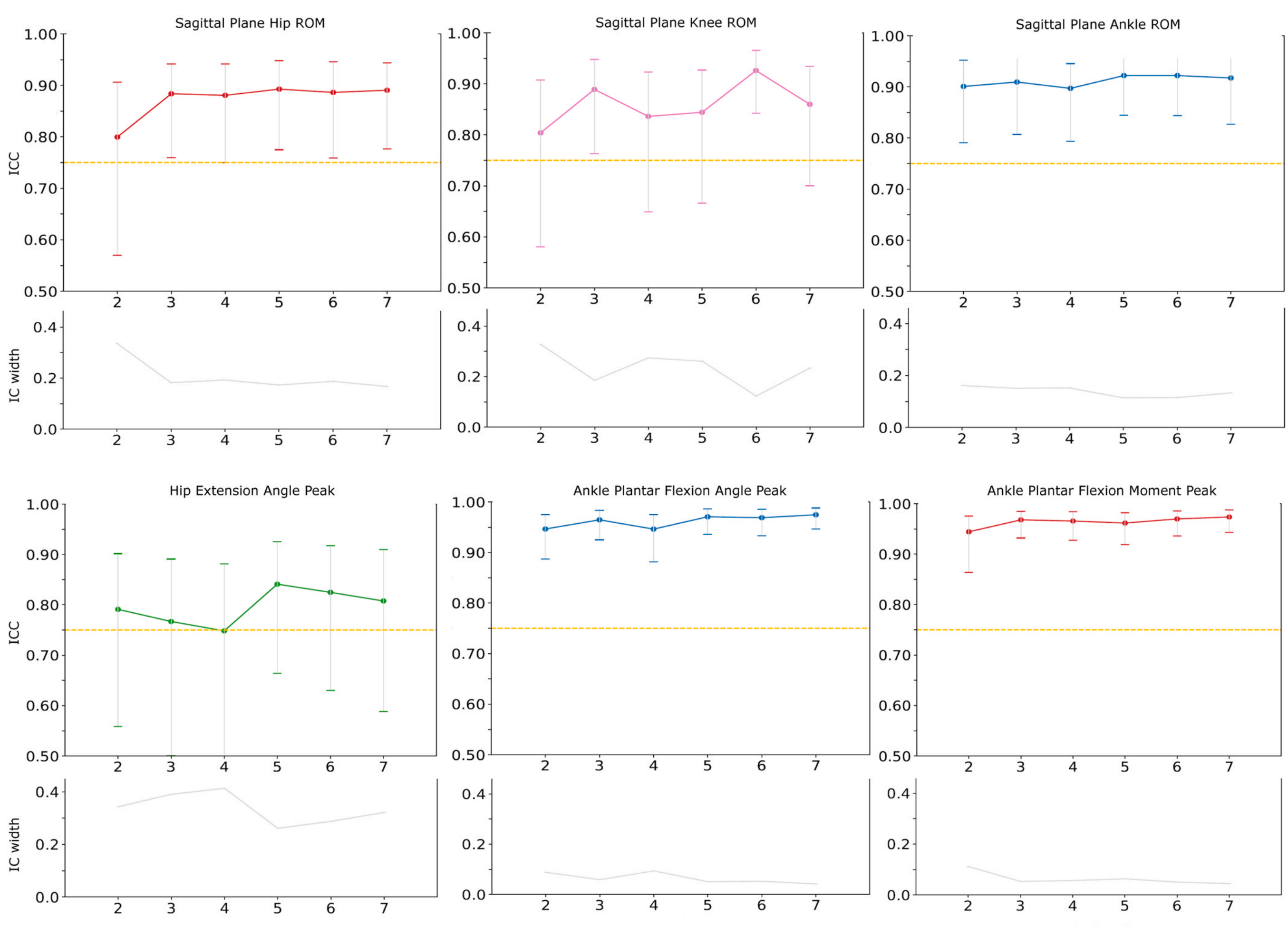

Fig. 2. Intraclass Correlation Coefficient (ICC) [95 % Confidence Interval (CI)] for key gait parameters in older adults: hip, knee, and ankle Range of Movement (ROM) in the sagittal plane, peak hip extension angle, peak ankle plantar flexion angle, and peak ankle plantarflexion moment, obtained from an equal number of gait strides averaged (2−7) across sessions. The grey line corresponds to the CI width. The yellow dotted line represents the ICC threshold.

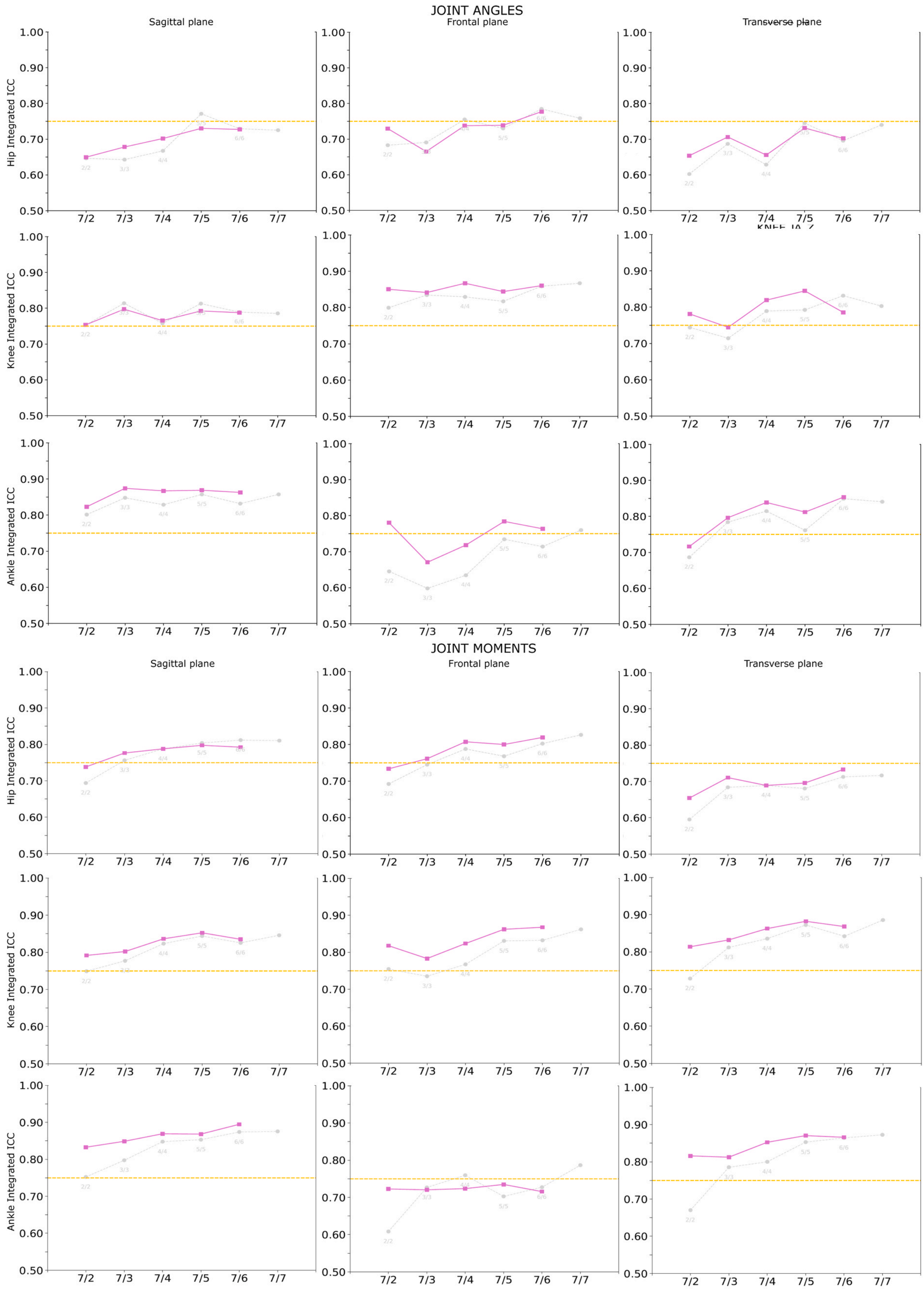

Fig. 3. Integrated Intraclass Correlation Coefficient (ICC) for hip, knee, and ankle joint angle and moments in the sagittal, frontal, and transverse plane, obtained from an unequal number of gait strides averaged (pink markers), depicted against equal scenarios (grey markers). The yellow dotted line represents the ICC threshold.

highlighting the importance of examining confidence intervals [18]. Accordingly, we extended the analysis to include ICCs [95 % CI], which globally supported the reliability observed with 7 strides, particularly for key gait parameters in older adults, demonstrating narrow CI and lower bounds above 0.75. Our results also demonstrate that the reliability of unequal combinations (e.g., 5/7 or 6/7) outperforms their "equal counterparts" (e.g., 5/5 or 6/6), with occasional exceptions. Similarly, reliable data were previously reported for unbalanced situations during one-legged hop distances [19].

The measurement error of kinematics only occasionally exceeded the established threshold (2º). Although this occurred in the hip sagittal plane, even when using 20 strides, it is important to interpret the error relative to the joint's total ROM [8]. Given the 40º-50º hip sagittal plane ROM during walking, 2º represents roughly 5 % of SEM%, which remains acceptable.

The choice of random stride selection may be perceived as a limitation. However, in our pilot analysis of sequential stride subsets (4–10 strides) [12], stability in ICC and SEM was comparable to the present results with randomly selected strides, suggesting that reliability estimates are robust and largely independent of the stride selection method.

5. Conclusion

A minimum of 7 strides contributing to the average observation is generally sufficient to achieve reliable markerless kinematics and kinetics in healthy older adults. Interestingly, even when fewer strides are averaged in one session, the reliability remains comparable to the 7 strides averaged in both sessions. This flexibility is particularly beneficial to optimizing data collection strategies in the oftentimes challenging settings, where individuals, such as older adults, may be less tolerant of assessment procedures or prone to fatigue quickly, without compromising measurement reliability.

CRediT authorship contribution statement

Andreia Carvalho: Conceptualization, Methodology; Formal analysis; Investigation; Writing - Original Draft; Visualization; Funding acquisition. **Jos Vanrenterghem:** Conceptualization, Methodology; Writing-Reviewing and Editing; Supervision. **Todd C. Pataky:** Conceptualization; Writing-Reviewing and Editing. **Mark A. Robison:** Conceptualization; Writing-Reviewing and Editing. **António P. Veloso:** Resources, Writing-Reviewing and Editing, Funding acquisition. **Vera Moniz-Pereira:** Conceptualization, Methodology, Investigation; Writing-Reviewing and Editing; Supervision, Funding acquisition.

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Acknowledgments

We gratefully acknowledge the contribution of all volunteers who participated in the study. We thank Lina Schelin and Alessia Pini for their availability to discuss the ICC R-code. Thanks to Paula Bruno for her insight into the statistical analysis. This work was supported by Fundação para a Ciência e a Tecnologia: Grant numbers DOI 10.54499/2020.07958.BD (PhD Grant) and DOI: 10.54499/UIDB/00447/2020, attributed to CIPER – Centro Interdisciplinar de Estudo da Performance Humana (unit 447).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.gaitpost.2025.110020.

References

[1] S. Scataglini, E. Abts, C. Van Bocxlaer, M. Van den Bussche, S. Meletani, S. Truijen, Accuracy, validity, and reliability of markerless camera-based 3D motion capture systems versus Marker-Based 3D motion capture systems in gait analysis: a systematic review and Meta-Analysis, Sensors 24 (2024), https://doi.org/10.3390/s24113686.
[2] L. Wade, L. Needham, P. McGuigan, J. Bilzon, Applications and limitations of current markerless motion capture methods for clinical gait biomechanics, PeerJ 10 (2022) 1–27, https://doi.org/10.7717/peerj.12995.
[3] P. JafariNasabian, J.E. Inglis, W. Reilly, O.J. Kelly, J.Z. Ilich, Aging human body: changes in bone, muscle and body fat with consequent changes in nutrient intake, J. Endocrinol. 234 (2017) R37–R51, https://doi.org/10.1530/JOE-16-0603.
[4] R. Roller-Wirnsberger, B. Thurner, C. Pucher, S. Lindner, G.H. Wirnsberger, The clinical and therapeutic challenge of treating older patients in clinical practice, Br. J. Clin. Pharm. 86 (2020) 1904–1911, https://doi.org/10.1111/bcp.14074.
[5] K.A. Boyer, R.T. Johnson, J.J. Banks, C. Jewell, J.F. Hafer, Systematic review and meta-analysis of gait mechanics in young and older adults, Exp. Gerontol. 95 (2017) 63–70, https://doi.org/10.1016/j.exger.2017.05.005.
[6] R.W. Baker, Measuring walking: a handbook of clinical gait analysis, Mac Keith Press, London, 2013.
[7] L.B. Mokkink, M. Boers, C. van der Vleuten, D.L. Patrick, J. Alonso, L.M. Bouter, H.C.W. de Vet, C.B. Terwee, 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. 20 (293) (2020).
[8] K. Monaghan, E. Delahunt, B. Caulfield, Increasing the number of gait trial recordings maximises intra-rater reliability of the CODA motion analysis system, Gait Posture 25 (2007) 303–315, https://doi.org/10.1016/j.gaitpost.2006.04.011.
[9] M. Schwenk, C. Howe, A. Saleh, J. Mohler, G. Grewal, D. Armstrong, B. Najafi, Frailty and technology: a systematic review of gait analysis in those with frailty, Gerontology 60 (2013) 79–89, https://doi.org/10.1159/000354211.
[10] S. Augustine, R. Foster, G. Barton, M.J. Lake, R. Sharir, M.A. Robinson, The inter-trial and inter-session reliability of Theia3D-derived markerless gait analysis in tight versus loose clothing, PeerJ 13 (2025) 1–15, https://doi.org/10.7717/peerj.18613.
[11] A. Carvalho, J. Vanrenterghem, S. Cabral, A.M. d'Assunção, F. Carnide, A. P. Veloso, V. Moniz-Pereira, Construct validity of markerless three-dimensional gait biomechanics in healthy older adults, Gait Posture 120 (2025) 217–225, https://doi.org/10.1016/j.gaitpost.2025.04.022.
[12] A. Carvalho, J. Vanrenterghem, S. Cabral, A. Assunção, A.P. Veloso, R. Fernandes, V. Moniz-Pereira, Markerless three-dimensional gait analysis in healthy older adults: test-retest reliability and measurement error, J. Biomech. 174 (2024) 112280, https://doi.org/10.1016/j.jbiomech.2024.112280.
[13] H.S. Yang, L.T. Atkins, C.R. James, Examination of stride-to-stride independence of selected lower extremity kinematic and temporal variables during treadmill walking, Gait Posture 50 (2016) 212–216, https://doi.org/10.1016/j.gaitpost.2016.09.010.
[14] K.O. McGraw, S.P. Wong, Forming inferences about some intraclass correlation coefficients, Psychol. Methods 1 (1996) 30–46, https://doi.org/10.1037/1082-989X.1.1.30.
[15] A. Pini, J.L. Markström, L. Schelin, Test–retest reliability measures for curve data: an overview with recommendations and supplementary code, Sport. Biomech. 21 (2022) 179–200, https://doi.org/10.1080/14763141.2019.1655089.
[16] M.L.J. Arts, S.A. Bus, Twelve steps per foot are recommended for valid and reliable in-shoe plantar pressure data in neuropathic diabetic patients wearing custom made footwear, Clin. Biomech. 26 (2011) 880–884, https://doi.org/10.1016/j.clinbiomech.2011.05.001.
[17] J.L. McGinley, R. Baker, R. Wolfe, M.E. Morris, The reliability of three-dimensional kinematic gait measurements: a systematic review, Gait Posture 29 (2009) 360–369, https://doi.org/10.1016/j.gaitpost.2008.09.003.
[18] L.G. Portney, D. Gross, Measurement revisited: reliability and validity statistics. Foundations of Clinical Research: Applications to Evidence-Based Practice, F.A. Davis, Philadelphia, 2020.
[19] L. Schelin, A. Pini, J.L. Markström, C.K. Häger, 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 (2021) 110546, https://doi.org/10.1016/j.jbiomech.2021.110546.