Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Validated Multi-Camera Markerless Motion Capture with Real-Time Data Analytics for Gait Stability Assessment

Sajeda Al-Hammouri
The University of Arizona
Ziao Chen
The University of Arizona
Nima Toosizadeh
Rutgers University
Christopher Arellano
The University of Arizona
Kavan Hazeli
kavan.hazeli@ucr.edu

University of California

Article

Keywords:

Posted Date: May 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9347103/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Validated Multi-Camera Markerless Motion Capture with Real-Time Data Analytics for Gait Stability Assessment

Sajeda Al-Hammouri[1,6], Ziao Chen[2], Nima Toosizadeh[3], Christopher Arellano[1,4], and Kavan Hazeli[5,*]

[1]Biomedical Engineering Department, The University of Arizona, Tucson, AZ, USA
[2]Department of Psychology, The University of Arizona, Tucson, AZ, USA
[3]School of Health Professions, Rutgers University, Newark, NJ, USA
[4]Department of Orthopedic Surgery, University of Arizona, Tucson, AZ, USA
[5]Department of Mechanical Engineering, University of California, Riverside, CA, USA
[6]Biomedical Systems and Informatics Engineering Department, Yarmouk University, Irbid, Jordan
[*]Corresponding author: kavan.hazeli@ucr.edu

ABSTRACT

Marker-based motion capture systems provide highly accurate biomechanical measurements, but they are expensive, require reflective marker placement, and are typically limited to controlled laboratory environments. Their reliance on extensive post-processing also restricts real-time clinical applicability. Markerless pose estimation systems offer a more accessible alternative; however, their ability to reliably quantify clinically relevant dynamic gait stability metrics, such as anterior–posterior (AP) and mediolateral (ML) margin of stability (MoS), remains insufficiently validated. This study evaluated a cost-effective, multi-camera, three-dimensional, markerless motion-capture system based on MediaPipe to estimate gait stability. Ten participants walked on a treadmill at a constant speed of 1.1 m/s while data were collected simultaneously using a three-camera synchronized RGB setup and a gold-standard Vicon capture system. Because MoS relies heavily on accurate estimates of center of mass (CoM) motion and joint kinematics, the temporal trajectories of 14 anatomical landmarks were first compared across systems. The markerless system achieved a mean 3D Euclidean RMSE of 2.48±0.69 cm, with median and 95th-percentile errors of 1.98±0.51 cm and 4.09±1.40 cm, respectively. Using a segmental CoM model, AP MoS values were −19.6±0.46 cm (marker-based) and −18.6±0.30 cm (markerless), while ML MoS values were 4.2±0.1 cm and 3.9±0.12 cm, respectively. Errors were larger in the AP direction (RMSE = 3.8 cm) than in the ML direction (RMSE = 0.83 cm). Similar trends were observed using pelvic CoM approximation. Paired *t*-tests showed no statistically significant differences between the two systems ($p > 0.05$). These findings suggest that the proposed multi-camera markerless system can reliably estimate whole-body kinematics and gait stability metrics, support scalable clinical assessment of walking stability and fall risk beyond specialized laboratory environments.

Introduction

Falls in older adults and sports-related injuries in young athletes represent significant clinical concerns. Although these populations differ in age and activity level, both can be affected by impairments in neuromuscular control and joint function, which contributes to reduced balance and stability. Margin of stability (MoS) is a widely used measure to evaluate dynamic stability during normal and abnormal gait[1,2] and can be used to evaluate rehabilitation, performance optimization, and injury prevention by identifying people at risk of falls. It can also serve as a metric to quantify improvements in stability after training, with higher values observed compared to control groups[3–6]. MoS is defined as the minimum distance between the extrapolated center of mass (XCoM) and the boundaries of the base of support (BoS). This concept was first introduced by Arend L. Hof in 2005 as an extension of the inverted pendulum model to assess dynamic stability by combining both position and velocity of CoM[7]. While walking, the velocity of the body causes the center of mass (CoM) to move both forward and sideways. Thus, MoS is typically evaluated in two directions: the anterior-posterior (AP) and medio-lateral (ML) directions. MoS can be calculated at different phases of the gait cycle; however, most studies use heel strike[8–12] as it is considered a key event during gait in which energy is absorbed and redirection of the center of mass velocity is required for stable gait[13]. Moreover, the position where the foot lands (at heel strike) define the base of support throughout the stance phase, which is the dominant mechanism to control dynamic stability[14].

Laboratory-based motion capture systems are considered the gold standard for assessing walking performance in both stable and unstable conditions. However, their clinical use is limited due to the high cost, time-consuming operation, and the need for specialized expertise in data collection and analysis. These systems also require extensive marker placement and complex

camera setup, along with expert preparation prior to each session, making them inaccessible to most medical and sports professionals[15]. These systems typically require substantial post-processing and offline analysis, which delays result interpretation and reduces their suitability for routine clinical implementation and continuous real-world monitoring. Therefore, for wide scale deployment, a portable markerless motion capture system is needed that enables objective and accurate measurement of full-body motion and stability analysis. Markerless camera-based systems offer the most promising solution because the capture of dynamic movement and the quantification of dynamic stability can be performed with reduced setup time and lower costs, since there is no need to attach physical markers to the body[16]. These advantages make markerless systems particularly practical for clinical and research settings where efficiency, comfort, and cost-effectiveness are critical. In addition, because they capture motion directly from the surface geometry of the subject, markerless systems eliminate errors associated with variability in marker placement and skin motion artifacts that often limit marker-based approaches[17–19].

Several studies have validated markerless human pose estimation algorithms against Vicon. For example, Menychtas et al. (2023) validate three different human pose estimation algorithms (OpenPose, BlazePose and MediaPipe) against a 3D Vicon marker-based motion capture system using two cameras placed in the sagittal and frontal planes[20]. Their results showed that human pose estimation algorithms produced hip and knee angles comparable to those measured by the Vicon motion capture system; however, some variations were noted in ankle angle. In another study, Ota et al. (2021) also used a Vicon system to validate the accuracy of the OpenPose human pose algorithm to measure joint angles during treadmill walking. They showed that capturing hip, knee, and ankle angles in the sagittal plane gives higher accuracy than in the frontal plane[21], indicating that accurately capturing the 3D motion of the human body and its dynamic stability while walking still remains a significant challenge. Most existing studies have used one or two cameras to measure joint angles or basic kinematic measures. However, the ability of markerless motion capture systems to estimate clinically relevant dynamic stability metrics such as anterior posterior (AP) and mediolateral (ML) margins of stability remains to be tested and validated, particularly during continuous
walking.

In this study, we developed a cost-effective markerless motion capture system based on MediaPipe. The system was designed following an evaluation of two- and three camera configurations to assess their impact on the accuracy of 3D reconstruction. Compared to a two-camera setup, a three-camera setup demonstrated improved performance and was therefore selected and validated against a gold-standard marker-based motion capture system for subsequent analysis of dynamic gait stability. The agreement between the two systems was assessed using kinematic error metrics, and dynamic stability was quantified using the margin of stability (MoS), computed from both segmental and pelvic-based CoM models in the sagittal and frontal directions. The results support the feasibility of using multi-camera markerless motion capture for reliable and accessible assessment of gait stability.

Methods

Participants and Data Collection

Ten healthy young adults (5 males and 5 females; mean age 25 ± 5 years; BMI 18.5 to 25 kg/m²) in this study. Participants with musculoskeletal injuries, arthritis, joint implants, mobility or balance impairments, or cardiovascular/metabolic conditions were excluded. Each subject performed two normal-walking trials on a split-belt treadmill at 1.1 m/s, for a total of 20 trials. For capturing the RGB videos, we used the 4K AKASO EK7000 camera because it offers a high spatial resolution (4K) at a reasonable frame rate of 30 fps. This frame rate is sufficient to capture human movement accurately in real-time, ensuring that no important landmarks or critical transitions preceding a fall are missed. In addition, the camera is relatively low-cost compared to other high-resolution recording systems, which makes it an optimal choice for real-world applications. All experimental procedures were performed in accordance with relevant guidelines and regulations and in accordance with the Declaration of Helsinki. The experimental protocol was approved by the Institutional Review Board (IRB) of the University of Arizona (STUDY00006269). Written informed consent was obtained from all participants prior to participation in the experiment. Participants also provided written informed consent for the recording and use of video data for research purposes.

Experimental Setup

As mentioned previously, two systems with two different types of cameras were used to record the subject's motion. The first system consisted of three low-cost 4K AKASO cameras placed at three different angles. The second system consisted of an eight-camera Vicon motion capture system for tracking the subject's full-body kinematics. This full-body analysis was performed by capturing the 3D positions of reflective markers placed landmarkat anatomical positions using the plug-in gait model. Each participant underwent a standardized setup, which included the placement of 57 reflective markers for the Vicon motion capture system and the collection of anthropometric measurements, including the lengths of the left and right legs and the widths of the elbows, knees, ankles, and wrists. The participants also wore a safety harness system that was carefully fitted and securely attached to ensure their safety during walking.

Marker Based Motion Capture System

The eight-camera Vicon system recorded the 3D marker trajectories at a sampling rate of 100 Hz. Before data collection, all cameras were carefully masked and calibrated to reduce interference from unwanted reflections. Then the origin of the global coordinate system was defined using four reflective markers placed on the corners of the treadmill, as shown in Figure 1. Setting the origin plays a crucial role in ensuring that the motion capture origin is spatially aligned with the force plate's self-coordinate system, enabling accurate measurements without additional coordinate transformations. To ensure reproducibility and consistent placement of markers in all trials, a custom 3D-printed box with holes was designed and positioned at a fixed location. This ensured that the four calibration markers were placed consistently, allowing for a precise reconstruction of the movements of the participants in a three-dimensional space throughout the experiment.

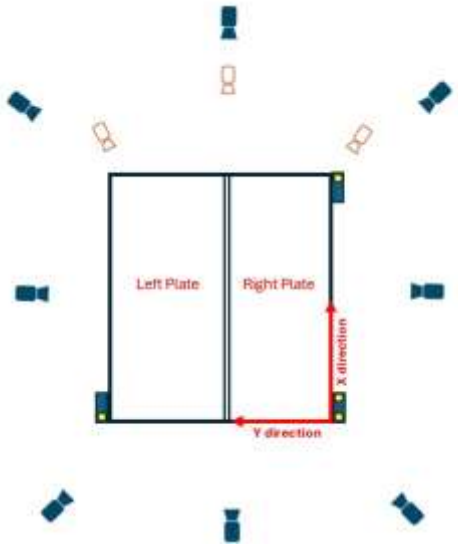


Figure 1. Top view of the experimental setup showing the arrangement of three RGB cameras and eight Vicon cameras. Yellow circles indicate reflective markers used to define the origin of the Vicon coordinate system. Rectangular boxes represent the custom 3D-printed fixture designed to ensure consistent marker positioning across trials.

Markerless Motion Capture System and Camera Configurations

Our markerless motion capture system uses the MediaPipe Pose library, a Google-developed human pose estimation framework that detects 33 anatomical landmarks in each video frame. MediaPipe Pose operates using standard RGB cameras and does not require that reflective markers be placed on the participant's body[22]. To investigate the influence of camera configuration on 3D reconstruction accuracy, two different configurations were evaluated. The first configuration used two RGB cameras positioned to capture the subject primarily from sagittal viewpoints. The second configuration extended the system by adding a third camera. One camera was placed in front of the subject, while the other two cameras were laterally placed on the left and right sides. All cameras were directed to the midpoint of the treadmill to ensure that the fields of view overlap throughout the capture volume. This non-collinear arrangement improves depth estimation and maximizes

reconstruction accuracy as it provides additional spatial coverage and reduces the likelihood of joint occlusions. MediaPipe Pose was used in this study due to its compact architecture, computational efficiency, and real-time operation. Compared with many existing open-source pose estimation frameworks, it provides reliable landmark detection while maintaining low computational cost, making it suitable for real-time biomechanical analysis and potential fall-risk monitoring applications[23–26]. In our preceding work[27–29], which sets the foundation for the present study, we introduced a framework that extracts a comprehensive set of biomechanical features in real time, including segment orientations, multi-level body centroids, trunk rotation, and joint elevation, to continuously quantify gait dynamics and detect body instability during walking. By analyzing these features simultaneously, the system captures the evolving spatiotemporal coordination of the body and enables immediate data-driven decisions regarding balance impairment and risk of falls.

Calibration is one of the most crucial steps in computer vision, particularly for 3D reconstruction applications. In this study, all cameras are individually calibrated to extract their internal parameters, followed by stereo calibration to determine the external parameters that define the spatial relationship between the two cameras. The calibration process with our three cameras starts by selecting the center camera as the reference camera, and then the calibration is performed in pairs. The resulting extrinsic matrices allow all 3D points to be expressed in the coordinate system of the reference camera. A checkerboard was used for calibration because its uniform square grid is easy to detect and reliably identify. Multiple images of the checkerboard were captured from all three cameras at different orientations to compute the projection matrices.

Data Preprocessing and Synchronization

In this study, a MediaPipe-based markerless motion capture system was validated against a gold-standard Vicon system. The validation was initially conducted by comparing the spatiotemporal trajectories of the corresponding anatomical landmarks, including bilateral shoulder, elbow, wrist, hip, knee, and ankle markers. The dynamic gait stability was then assessed by computing the margin of stability (MoS) of both systems. The 20 trials of Vicon recordings were used as the reference standard, and the corresponding MediaPipe signals were extracted for comparison. For comprehensive validation, 57 markers distributed throughout the upper and lower body were used, enabling us to assess system performance even in areas with limited camera visibility. Since the Vicon system and the RGB cameras record at two different sampling frequencies, temporal alignment was necessary. MediaPipe data were sampled up to 100 Hz using linear interpolation to match the Vicon temporal resolution. Additionally, because the two systems may not have started recording simultaneously, temporal discrepancies could distort joint trajectories. To address this, normalized cross-correlation was used to estimate the time lag between the systems. The lag corresponding to the peak correlation was then applied to shift the MediaPipe data, aligning it with Vicon data. All subsequent analyzes were performed on these synchronized datasets to ensure that observed differences reflected true performance differences rather than timing mismatches. For each trial, the first 10 consecutive gait cycles were identified based on heel-strike events. All outcome measures were computed per cycle and subsequently averaged to obtain representative values for each subject.

Global Coordinate Alignment

Different motion capture systems use distinct coordinate system origins and orientations. Therefore, even if MediaPipe accurately tracks motion, the 3D coordinates of joints such as the elbow or wrist will not numerically match those from the Vicon system. These discrepancies arise because the systems are not aligned with the same spatial reference frame. Such systematic differences are among the main sources of error that need to be addressed and corrected during data alignment and analysis. Therefore, to accurately compare the two systems, it was necessary to align them via a similarity transformation. This transformation estimates a scale factor (s), a 3×3 rotation matrix (R), and a 3×1 translation vector (t) that map markerless-system points into the coordinate frame of the Vicon system. This alignment is done using the Umeyama algorithm[30]. The similarity transformation is defined in equation 1 as

$$\mathrm{X}_{\mathrm{aligned}} = \mathrm{sRX} + \mathrm{t} \tag{1}$$

where $X_{aligned} \in R^{N\times 3}$ denotes the transformed set of 3D markerless keypoints expressed in the reference coordinate system, $s \in R$ is a uniform scale factor, $R \in R^{3\times 3}$ is the rotation matrix aligning X with the reference coordinate system, and $t \in R^{3\times 1}$ is the translation vector shifting X to match the reference origin. The Vicon coordinate system was used as the reference frame, and the rigid transformation (rotation and translation) was estimated across all frames using a robust singular value decomposition (SVD) method. The resulting transformation was then applied to the markerless data.

Margin of Stability Computation

We quantified the margin of stability by estimating the center of mass of the whole body (CoM) using a segmental method or a pelvic proxy. The segmental CoM was calculated using an anthropometric model consisting of feet, shanks, thighs,upper arms, forearms, and HAT segments. For each segment, the CoM of the segment was estimated from the proximal and distal joint positions using published fractional CoM locations[31]. The total body CoM was then calculated as the mass-weighted average of all available segment CoMs. After that, the extrapolated center of mass (XCoM) was calculated from the CoM position and the CoM velocity according to equation 2. Since the base of the support moves in the anterior direction during treadmill walking, we included the belt velocity in the equation following[32].

$$XCoM = CoM + \frac{\dot{CoM} + V_{\text{belt}}}{\omega_0} \quad (2)$$

Where $\dot{CoM}$ denotes the velocity of the center of mass, V_{belt} is the velocity of the treadmill belt (1.1 m/s), and ω_0 is the eigenfrequency of the inverted pendulum defined as

$$\omega_0 = \sqrt{\text{g/l}} \quad (3)$$

where g is the gravitational acceleration (9.81 m/s^2) and l is the effective pendulum length, typically approximated as the distance between the CoM of the whole body and the heel markers. When the pelvis was used as a proxy for CoM, the pendulum length was approximated as the distance between the pelvis and the ankle markers. The CoM velocity was computed as the numerical time derivative of the CoM trajectory. The values of (AP MOS) and (ML MoS) highly depend on how you define your (BoS); therefore, different methods to define (BoS) affect the (MoS) calculations during walking[33]. Previous studies have shown that defining the base of support using the most anterior foot marker provides the most accurate estimate of AP MoS, while the heel marker offers the best approximation of ML MoS. These markers produced the lowest root mean square error (RMSE) compared to alternative BoS definitions[34]. Another factor that can influence MoS estimates is the method used to calculate the center of mass (CoM), as different modeling approaches may lead to systematic differences in the resulting values. In the following section, MoS is calculated using these definitions to ensure consistency with previously validated approaches.

Anteroposterior Margin of Stability (AP MoS): The direction of the (AP) was defined along the x-axis, with positive values indicating forward motion. The BoS boundary was defined as the most anterior foot marker in contact with the ground (BoS_{AP}). The AP margin of stability was therefore computed as the distance between the AP BoS boundary and the AP position of the extrapolated center of mass ($XCoM_{AP}$) as shown in equation 4

$$MoS_{AP} = BoS_{AP} - XCoM_{AP} \quad (4)$$

Mediolateral Margin of Stability (ML MoS): The mediolateral (ML) direction was defined as the *y*-axis. BoS was defined using the position coordinates of the heel markers in the ML direction (BoS_{ML})[32,34]. From these data, the MoS was calculated as the distance between the $XCoM_{ML}$ and the BoS as shown in equation 5.

$$MoS_{ML} = BoS_{ML} - XCoM_{ML} \quad (5)$$

Accuracy Evaluation Metrics

To assess the accuracy of the reconstructed 3D joint trajectories obtained from the markerless framework relative to the marker-based motion capture system, the root mean square error (RMSE), median error, and 95th percentile error were used to quantify spatial discrepancies between the two systems. RMSE was selected as the primary precision measure because it summarizes the overall magnitude of the positional deviations between the corresponding joint locations. However, since RMSE can be influenced by occasional large errors, the median error was also calculated to represent the typical deviation observed in the frames. In addition, the 95th percentile error was included to describe the upper range of the error distribution and provide insight into the larger discrepancies that may occur during motion tracking. Together, these measures provide a description of the level of agreement between the markerless estimates and the reference measurements obtained from the marker system. Together, we used these metrics to provide a robust validation of both

spatial accuracy and motion consistency between the two systems. These three metrics were calculated for 14 landmarks to evaluate both the spatial precision and consistency of the markerless system compared to the gold-standard marker-based measurements. All metrics were computed following the temporal synchronization and spatial alignment of the two systems. After evaluating the accuracy of the joint reconstruction, the margin of stability (MoS) was calculated. The agreement between the marker and markerless systems was evaluated using the root mean square error (RMSE), mean absolute error, (MAE) and Bland–Altman analysis. These comparisons were made for both AP MoS and ML MoS. In addition, two CoM estimation approaches were investigated: a segmental CoM model and a pelvic-based CoM model. The pelvis-based method was included to assess its accuracy relative to the segmental CoM approach, as it is computationally less demanding and therefore may be more suitable for real-time applications.

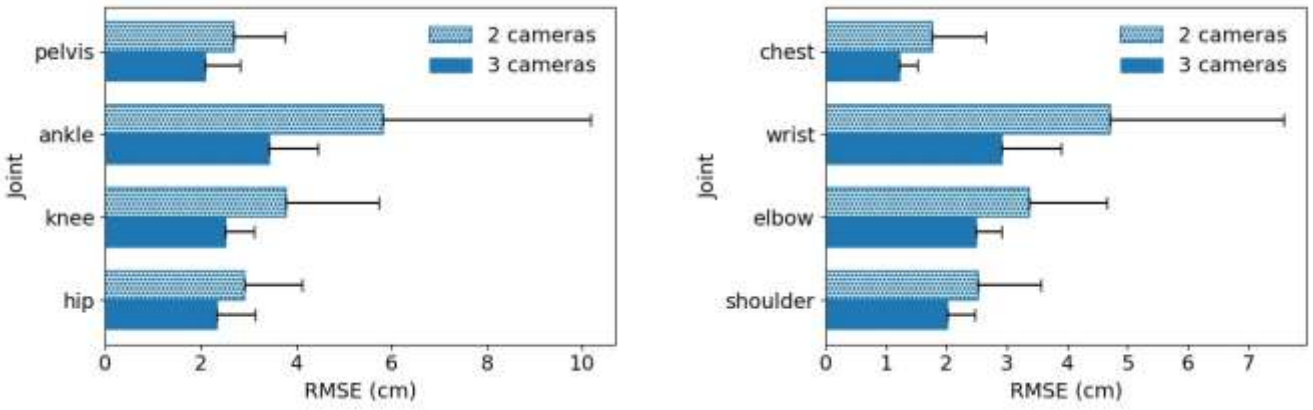
Statistical Analysis

Paired t-tests were used to assess differences between markerless and marker-based systems for the AP and ML MoS. Because the data satisfied the assumption of normality, the paired t-tests were considered appropriate. The significance level was set to α= 0.05. Data are reported as mean ± standard deviation. The exact P values are reported for all statistical tests. All statistical analyzes were performed in Python using the SciPy statistical library.

Results

Two cameras vs Three cameras setup

Figure 2 shows the comparison based on the configurations of two cameras and three cameras. Overall, the three-camera setup produced lower RMSE across all joints compared with the two-camera setup. For the lower body, the greatest improvement was observed for the ankle joint where RMSE improved from 6 cm to 3.5 cm, which means that the addition of the third camera helped reduce the ankle position error by 41 %. For the upper body, the wrist joint exhibited the largest error among all joints in the two-camera configuration, while the addition of a third camera considerably reduced RMSE by 38%. Improvements were also observed for the elbow, shoulder, and chest joints. Across both upper and lower body segments, the three-camera configuration generally resulted in smaller RMSE values and reduced variability compared to the two-camera setup. Based on these results, the three-camera configuration was used to compute the MoS for both systems, and MoS values derived from the three-camera setup were used for validation between the two systems.



(a) Lower body RMSE comparison.

(b) Upper body RMSE comparison.

Figure 2. Comparison of root-mean-square error (RMSE) between the two-camera and three-camera configurations across body joints: (a) lower body joints and (b) upper body joints. Error bars represent standard

Table 1 shows the accuracy of the markerless system model compared to the marker-based system. The root mean square error (RMSE), median and 95th percentile (P95) were calculated for each landmark as shown in Figure 3. In general, the mean RMSE across all joints was 2.48 ± 0.69 cm, with a corresponding median error of 1.98 ± 0.51 cm and a 95th percentile error of 4.09 ± 1.40 cm. Among individual joints, the chest showed the lowest RMSE (1.22 ± 0.30 cm), indicating excellent tracking accuracy for the proximal landmarks. In contrast, distal segments such as the ankles and wrists showed relatively higher RMSE

values (e.g., left ankle 3.48 ± 1.03 cm, right wrist 2.73 ± 0.87 cm), likely due to increased motion complexity, occlusion, and the limited visibility of markerless keypoint estimation in those regions. The median and P95 errors followed a similar trend, with typical errors around 2 cm and highest deviations (95th Percentile) ranging between 4 to 6 cm depending on the joint.

Landmark	RMSE Total	Median	95th Percentile
Chest	1.22±0.30	1.03±0.28	2.01±0.45
Left ankle	3.48±1.03	2.30±0.42	6.31±2.85
Left elbow	2.61±0.38	2.26±0.37	4.08±0.75
Left hip	2.50±0.96	2.28±1.00	3.58±1.13
Left knee	2.55±0.60	1.90±0.37	4.45±1.48
Left shoulder	1.91±0.42	1.73±0.43	2.87±0.56
Left wrist	3.09±1.11	2.18±0.56	5.47±2.60
Pelvis	2.10±0.73	1.89±0.78	3.14±0.81
Right ankle	3.37±1.05	2.34±0.47	6.05±2.66
Right elbow	2.38±0.45	1.99±0.46	3.92±0.87
Right hip	2.19±0.62	1.98±0.64	3.25±0.72
Right knee	2.50±0.59	1.86±0.37	4.28±1.92
Right shoulder	2.13±0.49	1.96±0.46	3.09±0.65
Right wrist	2.73±0.87	2.07±0.47	4.78±2.20
Average	2.48±0.69	1.98±0.51	4.09±1.40

Table 1. Root-mean-square error (RMSE), median error, and 95th-percentile error across anatomical landmarks (cm). Values are reported as mean ± standard deviation.

To evaluate the agreement between systems, Bland-Altman analysis was performed between marker and makerless systems for the estimation of MoS. For the AP MoS, the mean bias between the two systems was small (−1.98 cm) and the 95% limit of agreement (LoA) ranged from −6.61 cm to 2.65 cm, as shown in Figure 4 (A), while for the ML MoS the mean bias was -0.50 cm between the two systems, with 95% LoA between −1.49 cm and 0.50 cm (Figure 4 (B)).

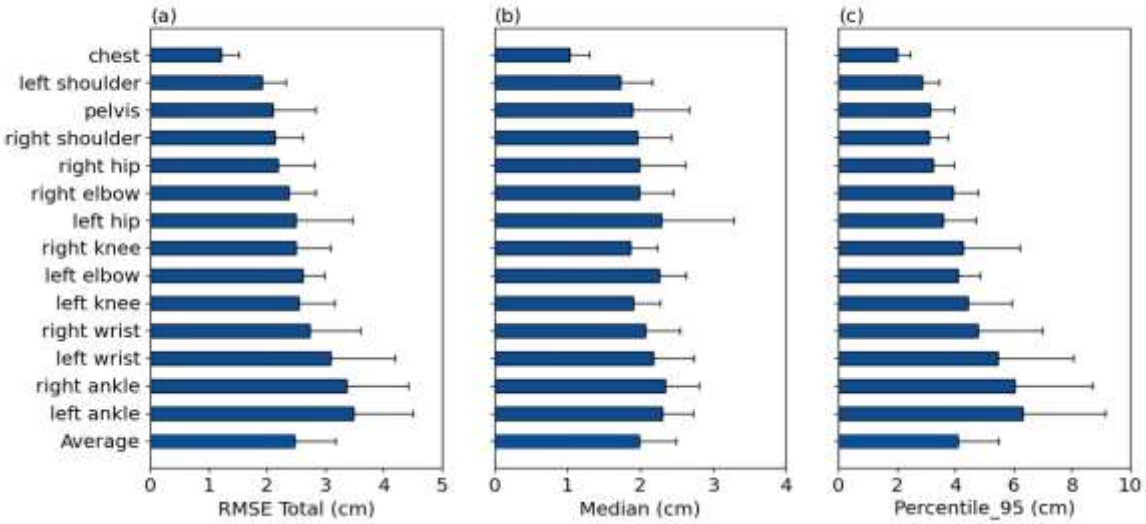


Figure 3. Pose-estimation error summaries across anatomical landmarks showing (a) root-mean-square error (RMSE), (b) median error, and (c) 95th-percentile error.

Most of the data points were within the LoA, indicating good agreement between the two capture systems. Since the pelvis can serve as a computationally simpler proxy for CoM, we evaluated the agreement between the two systems using the pelvis-based CoM estimation. As shown in Figures 4 C and D, the Bland–Altman analysis demonstrated a pattern similar to that observed with the segmental CoM method. For AP MoS, the mean bias was −2.61 cm, with 95% LoA ranging from −8.28 to 3.06 cm. For ML MoS, the mean bias was -0.80 cm, with agreement limits ranging from 0.26 to -1.87 cm. Table 2 shows that among the 10 participants, the mean (AP MoS) measured using the segmental CoM model for the marker-based system was −19.60±0.46 cm, while the markerless system estimated a slightly higher value of −18.6±0.30 cm, with an RMSE of 3.8 cm and MAE3. The paired *t*-test did not detect a statistically significant difference between the two systems (p = 0.45). The mediolateral stability showed comparable values between systems, with 4.2±0.1 cm for the marker-based system and 3.9±0.12 cm for the markerless system, and a small RMSE of 0.83 cm and MAE 0.65 (p = 0.20). In general, these results indicate good agreement between the marker-based and markerless systems to estimate MoS between participants. In addition, Table 2 presents the comparison when the COM was estimated using a pelvic marker instead of a segmental model. For AP MoS, the marker-based system produced −19.8 ± 0.5 cm, while the markerless system produced −18.3 ± 0.36 cm. RMSE and MAE were 5.2 cm and 4.5 cm, respectively, without detected statistically significant differences (p= 0.37). For ML MoS. the mean values were 3.23 ± 0.12 cm for the marker-based system and 3.3 ± 0.86 cm for the markerless system, with RMSE and MAE values of 1.03 cm and 0.77 cm, respectively, without statistically significant difference detected (p = 0.90). Compared to the pelvic-based CoM estimate, the segmental CoM model showed lower values of RMSE and MAE, for both AP and ML MoS.

Figure 5 shows the comparison of the average MoS profiles throughout the gait cycle obtained from markers and markers-based systems using segmental and pelvic COM motion approaches. Throughout the gait cycle, both systems show a similar waveform pattern for AP and ML MoS. The highest AP and ML MoS values occur during double support, specifically during the early stance after left heel strike (LHS) and are reduced during single support with the most negative values occurring near the end of the stance phase. At the next heel strike, the MoS increases again and subsequently decreases throughout the right stance phase.

CoM Model	Metric	Marker-based	Markerless	RMSE	MAE
Segmental CoM	AP	−19.6±0.46	−18.6±0.30	3.8	3.0
	ML	4.2±0.10	3.9±0.12	0.83	0.65
Pelvis CoM	AP	−19.8±0.50	−18.3±0.36	5.2	4.5
	ML	3.2±0.12	3.3±0.86	1.03	0.77

Table 2. Comparison between marker-based and markerless systems across subjects using segmental and pelvis center-of-mass (CoM) models (cm).

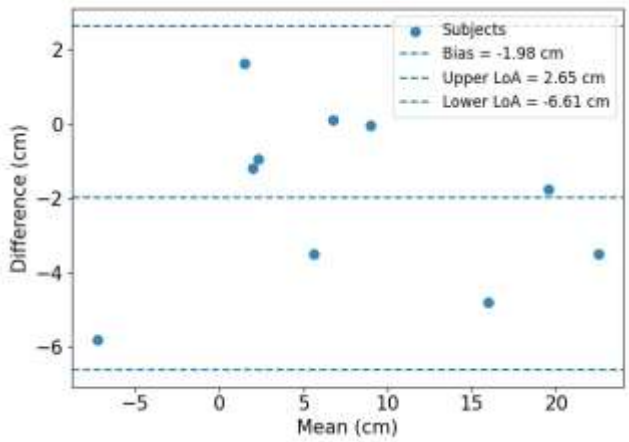
Discussion

Our findings first revealed that when the system extended to a three-camera configuration, spatial coverage was improved, 3D reconstruction was more accurate, and noise in estimating joint position was reduced. Improving these key features of our markerless motion capture system allowed for more robust analyses in the frontal and sagittal planes. Previous studies have shown agreement between marker and markerless systems[20,21]. In contrast to these studies, our approach enables a more accurate and comprehensive 3D analysis by using three synchronized and calibrated RGB cameras instead of two, representing a key advantage of our method. Another key finding was found when assessing the agreement between a marker-based motion capture system and a markerless system in estimating the margins of stability during normal walking that the AP MoS has larger RMSE values than the ML MoS. This difference can be explained by the greater sensitivity of AP

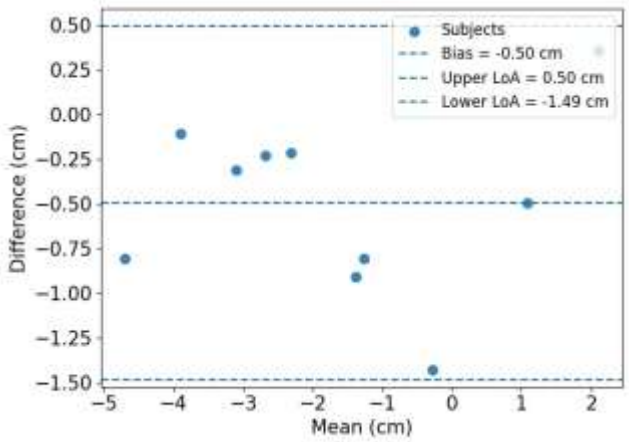
MoS to small errors in estimating the CoM velocity, as XCoM combines the position and velocity of the CoM[7]. The RMSE and MAE between the two systems were small relative to the magnitude of the AP and ML MoS values, as previously shown in Table 2, indicating that the markerless system was able to capture dynamic behavior underlying changes in stability in both directions throughout the gait cycle. In our study, CoM was calculated using the gold-standard segmental approach, which estimates total body CoM as the mass-weighted average of all individual CoM body segments[35] and compared to the pelvic approach. The segmental CoM model resulted in error metrics lower than the pelvis-based estimate, which is expected given that the segmental COM approach relies on all body segments and the motion of their masses during walking. In contrast, the pelvis marker provides a simplified approximation of CoM position, which may introduce additional variability, especially when there is significant trunk or upper body movement.

Figure 5 demonstrates that the markerless motion capture system produced MoS curves with waveform patterns similar to those of the marker-based system, with the main peaks and troughs occurring at the same gait phases and key events of the gait cycle, including left heel strike (LHS), right toe-off (RTO), right heel strike (RHS), and left toe-off (LTO). Immediately after left heel strike (LHS), both AP and ML MoS values are relatively high, indicating a more stable configuration as the BoS increases during double support. For AP MoS, the left heel becomes the new anterior boundary of the BoS at left heel strike. Immediately after contact, the CoM continues moving forward, causing the XCoM to project anteriorly relative to the BoS. During early loading response, forward CoM velocity decreases slightly due to the braking phase, shifting the XCoM closer to the stance foot. As a result, AP MoS increases slightly, producing a small upward rise in the curve immediately after left heel strike (LHS). After weight acceptance and braking, the body progresses forward over the left stance foot, causing the CoM velocity to increase again and the XCoM to move further anteriorly. As a result, the AP MoS decreases, consistent with the forward progression of the inverted-pendulum dynamics during stance. Consequently, in late stance, the XCoM moves farther beyond the anterior boundary of the BoS and is projected forward relative to the stance foot, resulting in a negative AP MoS. This does not indicate a loss of balance; rather, it reflects the normal mechanics of human gait. During forward walking, the body briefly enters a dynamically unstable phase in which the XCoM progresses ahead of the BoS, as the new stance foot forms the new BoS and moves forward relative to the CoM, and the margin values rapidly increase again as a new, more stable BoS is established. Thus, stability is restored when the swing leg is placed forward to establish the next BoS, effectively catching the CoM and generating a brake force that decelerates the forward motion of the body. ML MoS showed the expected transition from higher values at initial heel contact to more lower values during single support where the BoS decreases and the XCoM moves towards the lateral boundary of the stance foot and then increases again at RHS as the BoS increases again[8,33,36]. We found that the markerless system can both overestimate and underestimate MoS for several reasons, such as errors or inaccuracies in foot placement caused by occlusion and tracking noise. Several improvements can be applied to reduce errors by fixing joints definition mismatch between the systems, improving camera placement and calibration, which will enhance landmark tracking accuracy, especially during transition phases.

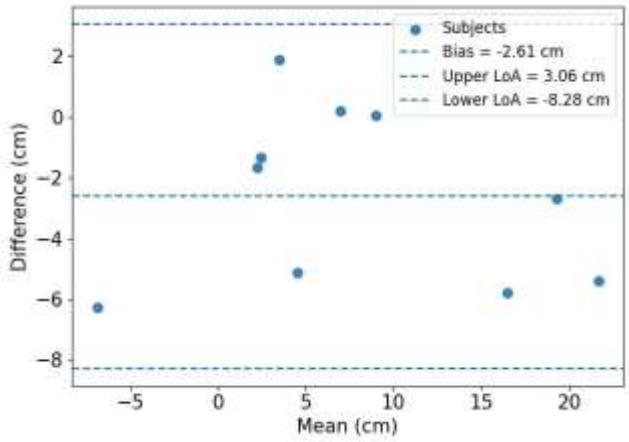
In general, the markerless system was able to capture the dynamics of whole-body motion during treadmill walking. However, minor differences between the two systems were observed, as Bland–Altman analysis indicated a small systematic bias primarily in the AP direction between marker based and markerless motion capture systems. The mean bias for AP MoS was -1.98 cm and -2.61 cm when using the segmental and pelvic CoM estimates, respectively, indicating that the markerless system slightly overestimated AP stability compared to the marker-based system, as the markerless approaches can be more sensitive to depth estimation and out of plane motion, which may introduce small spatial inaccuracies. These small spatial inaccuracies will propagate to small inaccuracies in tracking AP MoS because this metric strongly depends on step length and AP foot placement. This observation is consistent with our comparative analysis, where we observed differences in joint position between the gold-standard marker-based system and our markerless system that ranged from approximately 2-3.5 cm, even when using a three-camera setup.



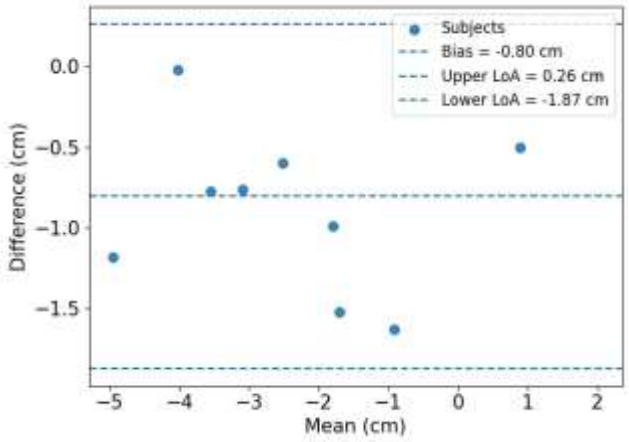
(a) AP MoS using the segmental CoM model.



(b) ML MoS using the segmental CoM model.



(c) AP MoS using the pelvis CoM model.

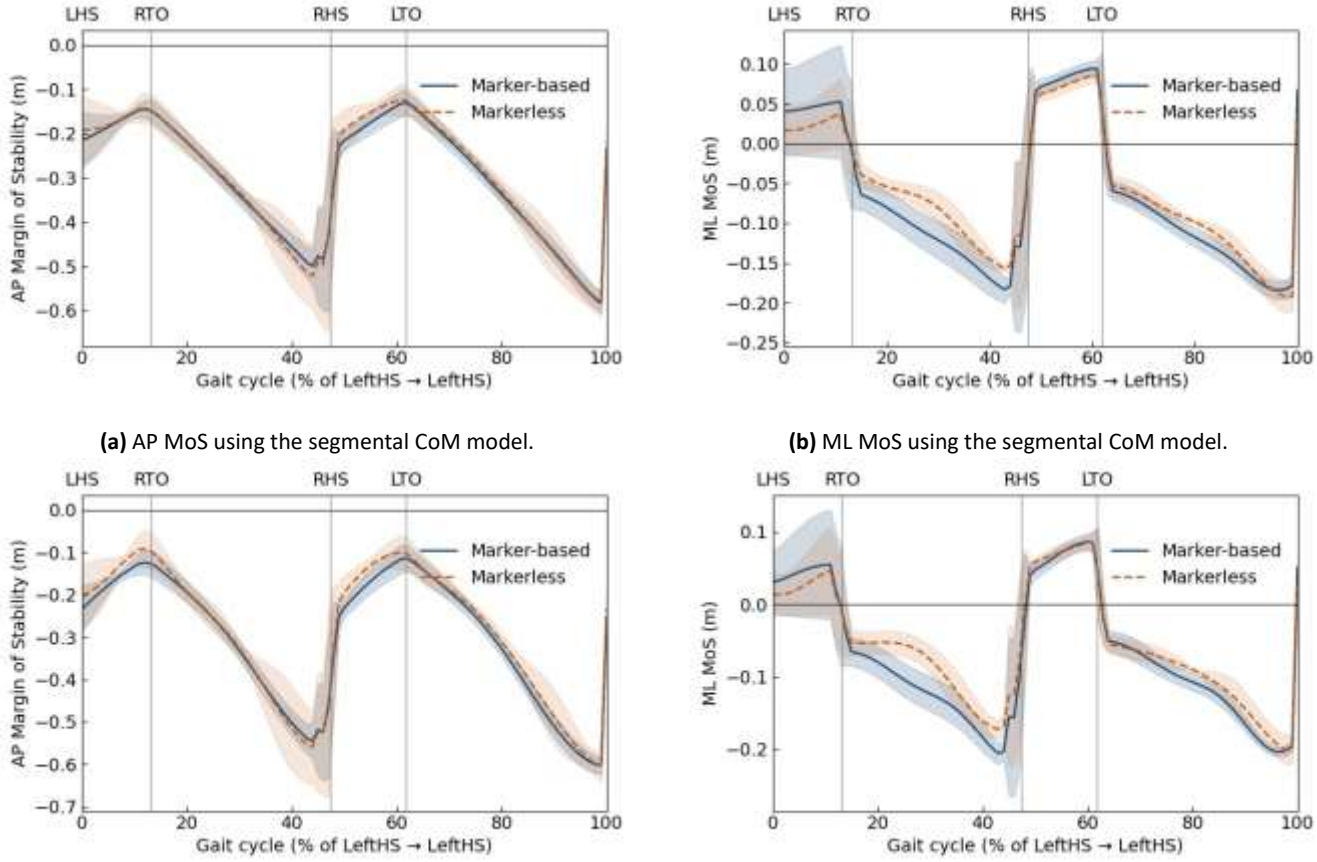


(d) ML MoS using the pelvis CoM model.

Figure 4. Bland–Altman analysis comparing margin of stability (MoS) estimates obtained from the marker-based and markerless systems. The solid central line represents the mean bias, and the dashed lines indicate the 95% limits of agreement: (a) AP MoS using the segmental CoM model, (b) ML MoS using the segmental CoM model, (c) AP MoS using the pelvis CoM model, and (d) ML MoS using the pelvis CoM model.

Previous studies have reported similar discrepancies between marker-based and markerless motion capture systems[37,38], but these small inaccuracies will likely be resolved as advances in pose estimation algorithms improve[39]. Compared to AP MoS, ML MoS showed smaller biases and more similar limits of agreement using the segmental or pelvic CoM, suggesting that the markerless system can reliably estimate lateral stability, which plays an important role in preventing falls during walking. Overall, the agreement observed in both AP and ML directions suggests that the markerless motion capture system provides comparable estimates of MoS to those derived from laboratory-based motion capture systems, supporting its potential use for gait stability assessment outside of the laboratory environment.

This study evaluated the feasibility of using a portable markerless motion capture system to estimate (AP and ML MoS) as clinically meaningful stability metrics. Because MoS is derived from the position, velocity, and BoS geometry of CoM[40], it



(a) AP MoS using the segmental CoM model.

(b) ML MoS using the segmental CoM model.

(c) AP MoS using the pelvis CoM model.

(d) ML MoS using the pelvis CoM model.

Figure 5. Comparison of average margin of stability (MoS) profiles across the gait cycle obtained from the marker-based and markerless motion capture systems. Solid lines represent the marker-based system, dashed lines represent the markerless system, and shaded bands denote standard deviation across 10 gait cycles: (a) AP MoS using the segmental CoM model, (b) ML MoS using the segmental CoM model, (c) AP MoS using the pelvis CoM model, and (d) ML MoS using the pelvis CoM model.

provides a mechanically interpretable measure of dynamic balance rather than a purely descriptive kinematic variable. For this reason, validating the MoS obtained from markerless systems is an essential next step before these tools can be applied in clinical decision making. To date, numeric thresholds for the degree of stability have not been established in the field. Interpreting MoS metrics depends on several methodological and biomechanical factors, including walking speed, direction of analysis, CoM model used, and how the BoS is defined. In general, when the XCoM moves outside the BoS, a corrective step is required to maintain balance to ensure instantaneous mechanical stability. However, because there are no strict cutoff values, researchers typically interpret the magnitude of MoS qualitatively rather than using fixed thresholds. For example, MoS has been used in different clinical applications such as the assessment of the risk of falling[41], neurological gait research[42], and rehabilitation[43]. During healthy walking, mean ML MoS at heel contact for healthy young adults generally ranges between approximately 3 and 5 cm[32]. Values that deviate from normal control groups are generally reported in several neurological populations, including individuals with stroke, multiple sclerosis, and lower limb amputation[1]. In contrast, unusually large MoS ML values often reflect compensatory widening of step width and are observed more frequently in older adults compared to younger adults[36]. In contrast, AP MoS usually ranges roughly from approximately −3 cm at a slow walking speed to −32 cm at a fast-walking speed of 1.8 m/s, which strongly depend on walking speed, with more negative values occurring at faster speeds. The negative sign is expected because the BoS typically remains posterior to the extrapolated center of mass (XCoM) during forward walking[32]. However, values that are more negative than expected for a given walking speed may

indicate reduced forward stability control. Our current understanding of stability and its thresholds is typically interpreted relative to a control group. For example, lower values of AP MoS have been reported in individuals with stroke and Parkinson's disease, reflecting impaired forward stability control[42,44], while MoS AP is often unchanged in people with vestibular disorders[45]. Individuals with lower extremity amputation typically demonstrate greater MoS ML due to increased step width, together with reduced MoS AP consistent with decreased forward walking stability[46]. In contrast, differences in ML MoS are not consistently observed in populations of stroke or Parkinson's disease. People with multiple sclerosis often exhibit more positive MoS AP values. Additionally, fallers tend to show altered MoS variability compared to non-fallers. For example, patients with multiple sclerosis who had a history of falls have a higher variability of ML MoS, lower AP MoS, and greater step-by-step variability of AP MoS compared to the control group[47].

Our results suggest that markerless motion capture provides promising estimates of AP and ML MoS that are broadly consistent with measurements obtained using laboratory-based marker systems, which could offer a pathway toward clearly defining numeric thresholds that establish the degree of stability in populations tested in clinical settings. Although portable motion capture technologies are becoming more popular and enable researchers and clinicians to capture human movement beyond controlled laboratory environments, it is important that we, as scientists, establish confidence that measurements reflect meaningful biomechanical quantities rather than system-specific artifacts. At the same time, our findings highlight the need for methodological consistency when computing MoS. Differences in marker definitions, center of mass estimation methods, and BoS boundaries can substantially influence reported values. Establishing a standardized approach to quantifying MoS is of critical importance so that MoS can be compared across studies and used to distinguish between clinical groups such as fallers and non-fallers. Previous literature has already identified this lack of standardization as a major barrier to broader clinical adoption[1], and the present results further reinforce the urgency of addressing this issue. It is important to note that the discrepancies arise from mismatches in joint definition between different systems, as measurement methods can define the same joint centers differently, resulting in systematic differences in the estimated positions for the same movement. Additionally, these differences could result from limitations in the accuracy of MediaPipe's tracking because such algorithms rely on supervised learning, which is trained with labeled data, adding small inaccuracies and noise to training datasets that result in projection errors and reduced precision of the joint position in a 3D space. Furthermore, errors may also arise from manual marker placement in the marker-based system. Our findings show that the error rate increases in distal joints, such as the ankle and wrist. This issue arises because these landmarks are small, have less visibility, and occupy fewer pixels in the camera view, meaning that slight tracking errors can translate to relatively large positional errors. The observed segment dependent error pattern indicates limitations of the human pose estimation algorithm used in tracking distal small joints. We observed that the proximal segments remain visible throughout the trials, which allows for stable and accurate triangulation. In contrast, distal joints such as the knee, ankle, and elbow are more susceptible to self-occlusion during walking, which can increase the uncertainty of pose-estimation. Overall, we find that our markerless motion capture system provides a reliable estimate of whole-body coordination and trunk motion but is less reliable for small distal joints.

Limitations of the study

This study has several limitations. First, the sample size was modest, and the study was conducted only on young, healthy participants, which may limit the generalizability of the findings to older adults, high-risk individuals, and clinical populations with impaired gait. Second, validation was performed during normal walking at a comfortable speed on a treadmill, which may restrict the applicability of the results to overground walking. Although markerless joint tracking showed good agreement with the gold-standard system, measured metrics such as AP and ML MoS remain sensitive to modeling assumptions, including CoM estimation and BoS definition.

Conclusion

This work demonstrates that markerless motion capture systems can provide reliable estimates of AP and ML MoS and therefore represent a promising tool for incorporating stability analyzes as part of everyday clinical care. Because MoS is grounded in physics-based descriptions of CoM dynamics and BoS geometry, real-time estimation using portable systems could enable objective evaluation of gait stability during routine clinical assessments. The ability to perform these

measurements outside specialized motion capture laboratories can open new opportunities to monitor balance impairments, guide rehabilitation strategies, and improve fall risk screening. In this way, markerless motion capture systems have strong potential to support the broader clinical translation of biomechanically defined stability metrics.

Author contributions statement

S.A conceived the study, developed the methodology, performed the formal analysis, curated the data, interpreted the results, and wrote the original manuscript draft. Z.C. contributed to the development of the methodology. N.T. contributed to the review and editing of the manuscript. C.A. and K.H. contributed to conceptualization, supervision, manuscript review and editing, and funding acquisition. All authors reviewed and approved the final manuscript.

Declarations

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to institutional ethical restrictions but are available from the corresponding author on reasonable request and with permission of the Institutional Review Board.

Competing interests

The authors declare no competing interests.

Funding

No Funding to report

References

1. Watson, F. *et al.* Use of the margin of stability to quantify stability in pathologic gait–a qualitative systematic review. *BMC musculoskeletal disorders* 22, 597 (2021).
2. Ohtsu, H. *et al.* Investigation of balance strategy over gait cycle based on margin of stability. *J. Biomech.* 95, 109319 (2019).
3. Pitts, J. & Bhatt, T. Domain-specific balance training reduces slip-related fall risk in young adults: A potential alternative to perturbation training. *Hum. Mov. Sci.* 98, 103294 (2024).
4. Allin, L. J. *et al.* Perturbation-based balance training targeting both slip-and trip-induced falls among older adults: a randomized controlled trial. *BMC geriatrics* 20, 205 (2020).
5. Martelli, D., Kang, J. & Agrawal, S. K. A single session of perturbation-based gait training with the a-tpad improves dynamic stability in healthy young subjects. In *2017 International Conference on Rehabilitation Robotics (ICORR)*, 479–484 (IEEE, 2017).
6. Epro, G. *et al.* Retention of gait stability improvements over 1.5 years in older adults: effects of perturbation exposure and triceps surae neuromuscular exercise. *J. neurophysiology* 119, 2229–2240 (2018).
7. Hof, A. L., Gazendam, M. & Sinke, W. The condition for dynamic stability. *J. biomechanics* 38, 1–8 (2005).
8. Young, P. M. M., Wilken, J. M. & Dingwell, J. B. Dynamic margins of stability during human walking in destabilizing environments. *J. biomechanics* 45, 1053–1059 (2012).
9. Peebles, A. T., Reinholdt, A., Bruetsch, A. P., Lynch, S. G. & Huisinga, J. M. Dynamic margin of stability during gait is altered in persons with multiple sclerosis. *J. biomechanics* 49, 3949–3955 (2016).
10. Sivakumaran, S., Schinkel-Ivy, A., Masani, K. & Mansfield, A. Relationship between margin of stability and deviations in spatiotemporal gait features in healthy young adults. *Hum. movement science* 57, 366–373 (2018).

11. Madehkhaksar, F. *et al.* The effects of unexpected mechanical perturbations during treadmill walking on spatiotemporal gait parameters, and the dynamic stability measures by which to quantify postural response. *PloS one* 13, e0195902 (2018).

12. Arampatzis, A., Karamanidis, K. & Mademli, L. Deficits in the way to achieve balance related to mechanisms of dynamic stability control in the elderly. *J. biomechanics* 41, 1754–1761 (2008).

13. Papachatzis, N., Slivka, D. R., Pipinos, I. I., Schmid, K. K. & Takahashi, K. Z. Does the heel's dissipative energetic behavior affect its thermodynamic responses during walking? *Front. bioengineering biotechnology* 10, 908725 (2022).

14. Winter, D. A. Human balance and posture control during standing and walking. *Gait & posture* 3, 193–214 (1995).

15. Moro, M., Marchesi, G., Hesse, F., Odone, F. & Casadio, M. Markerless vs. marker-based gait analysis: A proof of concept study. *Sensors* 22, 2011 (2022).

16. Scataglini, S. *et al.* 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, 3686 (2024).

17. Colyer, S. L., Evans, M., Cosker, D. P. & Salo, A. I. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open* 4, 24 (2018).

18. De Rosario, H. *et al.* Applications of using 4d scanning technologies in biomechanics. In *International Conference on Digital Human Modeling*, 98–105 (Springer, 2023).

19. Desmarais, Y., Mottet, D., Slangen, P. & Montesinos, P. A review of 3d human pose estimation algorithms for markerless motion capture. *Comput. Vis. Image Underst.* 212, 103275 (2021).

20. Menychtas, D. *et al.* Gait analysis comparison between manual marking, 2d pose estimation algorithms, and 3d markerbased system. *Front. Rehabil. Sci.* 4, 1238134 (2023).

21. Ota, M., Tateuchi, H., Hashiguchi, T. & Ichihashi, N. Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm. *Gait & Posture* 85, 290–297 (2021).

22. Lugaresi, C. *et al.* Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* (2019).

23. Kim, J.-W., Choi, J.-Y., Ha, E.-J. & Choi, J.-H. Human pose estimation using mediapipe pose and optimization method based on a humanoid model. *Appl. sciences* 13, 2700 (2023).

24. Hii, C. S. T. *et al.* Automated gait analysis based on a marker-free pose estimation model. *Sensors* 23, 6489 (2023).

25. Lugaresi, C. *et al.* Mediapipe holistic – simultaneous face, hand and pose prediction on device. https://research.google/blog/mediapipe-holistic-simultaneous-face-hand-and-pose-prediction-on-device/ (2019). Accessed: 2025-09-15.

26. Bazarevsky, V. *et al.* On-device, real-time body pose tracking with mediapipe blazepose. https://research.google/blog/on-device-real-time-body-pose-tracking-with-mediapipe-blazepose/ (2020). Accessed: 2025-09-15.

27. Hazeli, K., Jalali, B. & Roveda, J. Converting spatial information to haptic and auditory feedback (2025). US Patent 12,327,411.

28. Hazeli, K., Jalali, B., Wang, X. & Roveda, J. Fall detection and prevention system for alzheimer's, dementia, and diabetes. *US Pat. Appl.* 18-429-089 (2024). Filed August 1, 2024.

29. Al-Hammouri, S., Wung, S.-F., Chen, Z., Hazeli, K. & et al. Fall prediction algorithm with built-in instability metrics. *SSRN* DOI: 10.2139/ssrn.5218182 (2025).

30. Lawrence, J., Bernal, J. & Witzgall, C. A purely algebraic justification of the kabsch-umeyama algorithm. *J. research Natl. Inst. Standards Technol.* 124, 1 (2019).

31. Winter, D. A. *Biomechanics and motor control of human movement* (John wiley & sons, 2009).

32. Fallahtafti, F., Mohammadzadeh Gonabadi, A., Samson, K. & Yentes, J. M. Margin of stability may be larger and less variable during treadmill walking versus overground. *Biomechanics* 1, 118–130 (2021).

33. Curtze, C., Buurke, T. J. & McCrum, C. Notes on the margin of stability. *J. biomechanics* 166, 112045 (2024).

34. Dussault-Picard, C., Tisserand, R., Robidou, C. & Cherni, Y. Comparison of marker-based and center-of-pressure-based approaches for calculating the margin of stability. *Front. Sports Active Living* 7, 1571994 (2025).

35. Havens, K. L., Mukherjee, T. & Finley, J. M. Analysis of biases in dynamic margins of stability introduced by the use of simplified center of mass estimates during walking and turning. *Gait & posture* 59, 162–167 (2018).

36. Siragy, T., Russo, Y. & Horsak, B. Mediolateral margin of stability highlights motor strategies for maintaining dynamic balance in older adults. *PLoS One* 19, e0313034 (2024).

37. Stenum, J., Rossi, C. & Roemmich, R. T. Two-dimensional video-based analysis of human gait using pose estimation. *PLoS computational biology* 17, e1008935 (2021).

38. Nakano, N. *et al.* Evaluation of 3d markerless motion capture accuracy using openpose with multiple video cameras. frontiers in sports and active living. 2020; 2: 50 (2020).

39. Horsak, B. *et al.* Concurrent validity of smartphone-based markerless motion capturing to quantify lower-limb joint kinematics in healthy and pathological gait. *J. Biomech.* 159, 111801 (2023).

40. Hof, A. L. The 'extrapolated center of mass' concept suggests a simple control of balance in walking. *Hum. movement science* 27, 112–125 (2008).

41. Punt, M. *et al.* Responses to gait perturbations in stroke survivors who prospectively experienced falls or no falls. *J. biomechanics* 55, 56–63 (2017).

42. Kao, P.-C., Dingwell, J. B., Higginson, J. S. & Binder-Macleod, S. Dynamic instability during post-stroke hemiparetic walking. *Gait & posture* 40, 457–463 (2014).

43. Van Dijsseldonk, R. B. *et al.* Gait stability training in a virtual environment improves gait and dynamic balance capacity in incomplete spinal cord injury patients. *Front. neurology* 9, 963 (2018).

44. Martelli, D. *et al.* Adaptation of stability during perturbed walking in parkinson's disease. *Sci. reports* 7, 17875 (2017).

45. McCrum, C. *et al.* Deficient recovery response and adaptive feedback potential in dynamic gait stability in unilateral peripheral vestibular disorder patients. physiol rep 2: e12222 (2014).

46. Gates, D. H., Scott, S. J., Wilken, J. M. & Dingwell, J. B. Frontal plane dynamic margins of stability in individuals with and without transtibial amputation walking on a loose rock surface. *Gait & posture* 38, 570–575 (2013).

47. Peebles, A. T., Bruetsch, A. P., Lynch, S. G. & Huisinga, J. M. Dynamic balance in persons with multiple sclerosis who have a falls history is altered compared to non-fallers and to healthy controls. *J. biomechanics* 63, 158–163 (2017).