Applied Ergonomics 138 (2027) 104866

Contents lists available at ScienceDirect

Applied Ergonomics

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

Concurrent validity of markerless motion capture for occupational lifting: Kinematics and ergonomic risk assessment

Mengjie Zhang [a], Arne Nieuwenhuys [a], Mark Boocock [b,*], Yanxin Zhang [a,**]

[a] School of Exercise, Sport and Rehabilitation Sciences, Faculty of Science, The University of Auckland, Auckland, 1023, New Zealand
[b] Ergonomics and Human Factors Group, School of Allied Health, Auckland University of Technology, Auckland, 1142, New Zealand

ARTICLE INFO

Keywords:
Markerless motion capture
Manual handling
Work-related musculoskeletal disorders

ABSTRACT

Markerless motion capture has the potential to enhance ergonomic risk assessment, but evidence remains limited for occupational manual handling tasks of increasing complexity. This study compared a multi-view system (OpenCap) and a monocular system (MediaPipe) against a marker-based reference system (Vicon) for lifting kinematics. Outcomes were assessed using the Rapid Entire Body Assessment (REBA) risk assessment tool. Twenty healthy participants performed three two-handed lifting tasks of increasing complexity under controlled laboratory conditions. Errors increased by 15–30% from symmetrical to asymmetric multi-planar tasks, with both systems demonstrating lower error in the sagittal plane (RMSE = 9.9–31.2°) than in the frontal and transverse planes. Both systems were reliable for symmetrical lifting, including high REBA agreement, but validity declined for tasks involving trunk rotation and walking. MediaPipe showed lower error for trunk and proximal joints, whereas OpenCap performed better for distal kinematics, informing practical system selection.

1. Introduction

Work-Related Musculoskeletal Disorders (WMSDs) result from multiple occupational risk factors, including physical (e.g., repetitive motions, forceful exertion, awkward postures), psychosocial, and work organisation risk factors (Colombini and Occhipinti, 2006; Geto et al., 2025; Govaerts et al., 2021). WMSDs commonly manifest as pain or discomfort affecting muscles, bones, joints, ligaments, tendons, and nerves (Colombini and Occhipinti, 2006; Govaerts et al., 2021) and account for approximately 30–40% of occupational injuries worldwide (Govaerts et al., 2021). To identify risk factors, observational risk assessment tools such as the Ovako Working Posture Analysing System (OWAS) (Karhu et al., 1977), Rapid Upper Limb Assessment (RULA) (McAtamney and Nigel Corlett, 1993), Rapid Entire Body Assessment (REBA) (Hignett and McAtamney, 2000), and the NIOSH Lifting Equation (Lowe et al., 2019) are widely used in ergonomic assessment (Colombini and Occhipinti, 2006; Govaerts et al., 2021; Rosado et al., 2023; Takala et al., 2010). However, these methods are time-consuming, labour-intensive, and susceptible to inter-rater variability owing to their subjectivity (Colombini and Occhipinti, 2006; Govaerts et al., 2021; Rosado et al., 2023; Zayyinul Hayati Zen and Hani Sukadarin, 2024; Salisu et al., 2023). Motion capture (MoCap) systems offer a more objective alternative for assessing postural movements, quantifying human movement and posture directly (Salisu et al., 2023; Menolotto et al., 2020). Marker-based optoelectronic MoCap systems, such as Vicon (Vicon Motion Systems Ltd., Oxford, UK) (Merriaux et al., 2017), Qualisys (Qualisys AB, Gothenburg, Sweden) (Malý and Lopot, 2013), and inertial measurement unit (IMU)-based wearable systems (e.g., Xsens MVN) (Guo and Xiong, 2017) are widely regarded as gold standards for three-dimensional (3D) kinematic analysis (Teufl et al., 2019). However, these systems often require a controlled laboratory environment, expert setup, and dedicated space, and the attachment of physical markers or sensors to the body, limiting their scalability beyond such settings (Salisu et al., 2023).

Markerless motion capture has emerged as a promising alternative that uses computer vision-based pose estimation to enable automatic joint detection and tracking from video recordings without physical markers (Varcin and Boocock, 2026; Roggio et al., 2024; Cheng et al., 2025). These technologies provide low-cost solutions that are increasingly being adopted for ergonomic risk assessment applications (Yang et al., 2024). Systematic reviews have identified multi-view (e.g., OpenCap, Theia3D) and monocular (e.g., MediaPipe BlazePose)

* Corresponding author. AA Building, Level 1, 90 Akoranga Drive, Northcote, Auckland, 0627, New Zealand.
** Corresponding author. Building 907, Level 2, 368 Khyber Pass Rd, Newmarket, Auckland, 1023, New Zealand.
E-mail addresses: mark.boocock@aut.ac.nz (M. Boocock), yanxin.zhang@auckland.ac.nz (Y. Zhang).

https://doi.org/10.1016/j.apergo.2026.104866
Received 10 February 2026; Received in revised form 29 July 2026; Accepted 4 August 2026
Available online 10 August 2026
0003-6870/

approaches as suitable for human movement assessment, demonstrating reasonable accuracy for sagittal-plane kinematics (RMSE $\leq 7^\circ$), although with poor agreement for transverse-plane rotations and complex movements (Varcin and Boocock, 2026; Cheng et al., 2025). However, most existing validation studies have focused on gait and sports-related activities (e.g., running, walking and jumping), where repetitive and predictable movement patterns facilitate more reliable analysis (Varcin and Boocock, 2026; Cheng et al., 2025; Scataglini et al., 2024; D'Souza et al., 2024; Jamali et al., 2025; Outerleys et al., 2024). Evidence remains more limited for complex manual handling tasks, where asymmetric posture, trunk rotation, and self-occlusion present additional challenges.

OpenCap (Uhlrich et al., 2023) is an open-source, web-based, multi-view markerless system that reconstructs 3D joint kinematics from synchronised videos of iOS (iPhone Operating System) devices. It employs deep-learning algorithms for keypoint detection and musculoskeletal modelling based on the Rajagopal et al., 2016 full-body model (Rajagopal et al., 2016) (37 degrees of freedom) in OpenSim. Previous studies (Turner et al., 2024; Svetek et al., 2025; Yeom et al., 2025) evaluating OpenCap against marker-based systems have reported high sagittal-plane agreement but lower agreement in the frontal and transverse planes (Turner et al., 2024; Svetek et al., 2025), indicating that multi-view reconstruction and biomechanical modelling can enhance anatomical accuracy, although performance may decline for asymmetric or partially occluded multi-planar movements.

MediaPipe BlazePose (Bazarevsky et al., 2020) is a monocular, convolutional neural network (CNN)-based pipeline developed by Google, which estimates 3D human pose from a single RGB camera (Lugaresi et al., 2019), enabling highly portable setups suitable for field deployment (Scataglini et al., 2024; Lugaresi et al., 2019). However, monocular pose estimation remains fundamentally constrained by depth ambiguity, and previous studies (Cheng et al., 2025; Ceriola et al., 2024; Kim et al., 2025) have generally reported lower 3D kinematic accuracy than that of marker-based systems. Optimisation-based extensions have reduced error in representative poses, but evidence on the validity of this approach has been based primarily on quasi-static or controlled movements (Kim et al., 2023). Direct comparisons between monocular and multi-view markerless systems under the same occupational task conditions remain limited.

The suitability of incorporating markerless motion capture into occupational ergonomic assessment has received comparatively little attention. Most validation work has focused on gait and sports activities (Scataglini et al., 2024; D'Souza et al., 2024; Outerleys et al., 2024; Needham et al., 2021), where repetitive and predictable movements facilitate more reliable evaluation. In contrast, occupational manual handling, which is characterised by asymmetric posture, multi-planar trunk rotation, and self-occlusion, remains under-examined. The few studies conducted in occupational contexts have predominantly used simple or otherwise constrained lifting conditions (Kim et al., 2021; Mehrizi et al., 2018a; Wang et al., 2021; Cronin et al., 2024), providing limited evidence on validity during the complex, multi-planar movements typical of real-world manual handling (McAtamney and Nigel Corlett, 1993; Hignett and McAtamney, 2000; Turner et al., 2024; Kim et al., 2021; Mehrizi et al., 2018a; Wang et al., 2021; Cronin et al., 2024; Agostinelli et al., 2024; Salehi et al., 2026a).

Equally important is what happens to that kinematic error in practice. Most studies have reported kinematic accuracy alone, rather than whether that error propagates to the ergonomic risk classifications used in practice (Turner et al., 2024; Agostinelli et al., 2024). Observational tools such as REBA employ discrete angle thresholds (typically 10–20° increments) that may absorb small kinematic errors but can shift risk category once errors are larger (Hignett and McAtamney, 2000). Consequently, kinematic agreement and risk-classification agreement need not coincide. Whilst some studies have extended markerless motion capture to RULA and REBA (Kim et al., 2021), the relative performance of multi-view and monocular systems across lifting tasks of increasing complexity, and its effect on risk classification, has not been directly examined. A combined evaluation of kinematic validity and ergonomic risk classification would therefore provide more practice-relevant evidence for system selection in occupational ergonomic assessment.

The aim of this study was to evaluate the concurrent validity of a multi-view (OpenCap) and a monocular (MediaPipe) markerless system against a marker-based reference system (Vicon) during two-handed occupational lifting tasks. It investigated the effects of increasing task complexity on the validity of kinematic measures and their influence on the risk classification produced by the REBA. It was hypothesised that: (1) the concurrent validity of both markerless systems would decrease as task complexity increased from symmetrical sagittal-plane lifting to complex multi-planar movements involving trunk rotation and walking; (2) errors would be comparable between systems during symmetrical lifting but greater for MediaPipe than for OpenCap during complex multi-planar movements, owing to monocular depth ambiguity; and (3) REBA risk-classification agreement would decrease corresponding with task complexity.

2. Methods

A cross-sectional, within-subject experimental study was conducted to evaluate the concurrent validity of OpenCap and MediaPipe markerless motion tracking against the Vicon (Vicon, Oxford Metrics Ltd., Oxford, UK) marker-based system. Participants performed three, two-handed lifting tasks (Fig. 1) of increasing postural and multi-planar complexity:

1) Vertical Lifting (VL): symmetrical, sagittal plane lifting of a box from the floor onto a table in front of the participant, with feet remaining stationary.
2) Transverse Lifting (TL): lifting a box from the floor onto a table positioned to the right side of the participant. Participants lifted the box vertically to an upright posture before rotating the trunk to place the box, with feet remaining approximately stationary.
3) Rotational Walking-and-Lifting (RWL): lifting a box from the floor to an upright posture before rotating and walking approximately 1.5 m to place the box on a table positioned to the right of the participant.

All three motion capture systems recorded each trial concurrently, as detailed in Section 2.3.

2.1. Participants

Twenty healthy participants (12 females, 8 males) were recruited, consistent with sample sizes employed in previous markerless motion capture validation studies (Scataglini et al., 2024; Turner et al., 2024), for whom anthropometric data are presented in Table 1. As this was a concurrent-validity (agreement) study rather than a test of group differences, no formal a priori sample-size calculation was performed. Inclusion criteria required participants to be aged between 18 and 35 years (representing the adult working population whilst minimising the confounding effects of growth, musculoskeletal maturation, or age-related decline in motor control (Marras et al., 1993; Marras et al., 2006)). Participants also had to be free from current musculoskeletal pain or injury, with no neurological, vestibular, or cardiopulmonary disorders affecting their ability to perform lifting tasks. Exclusion criteria included any surgery within the previous 6 months. All participants received written information about the study and provided written informed consent prior to their participation. The experimental protocol was approved by the Auckland Health Research Ethics Committee (reference number AH29163) and adhered to the Declaration of Helsinki (World Medical Association Declaration of, 2013).

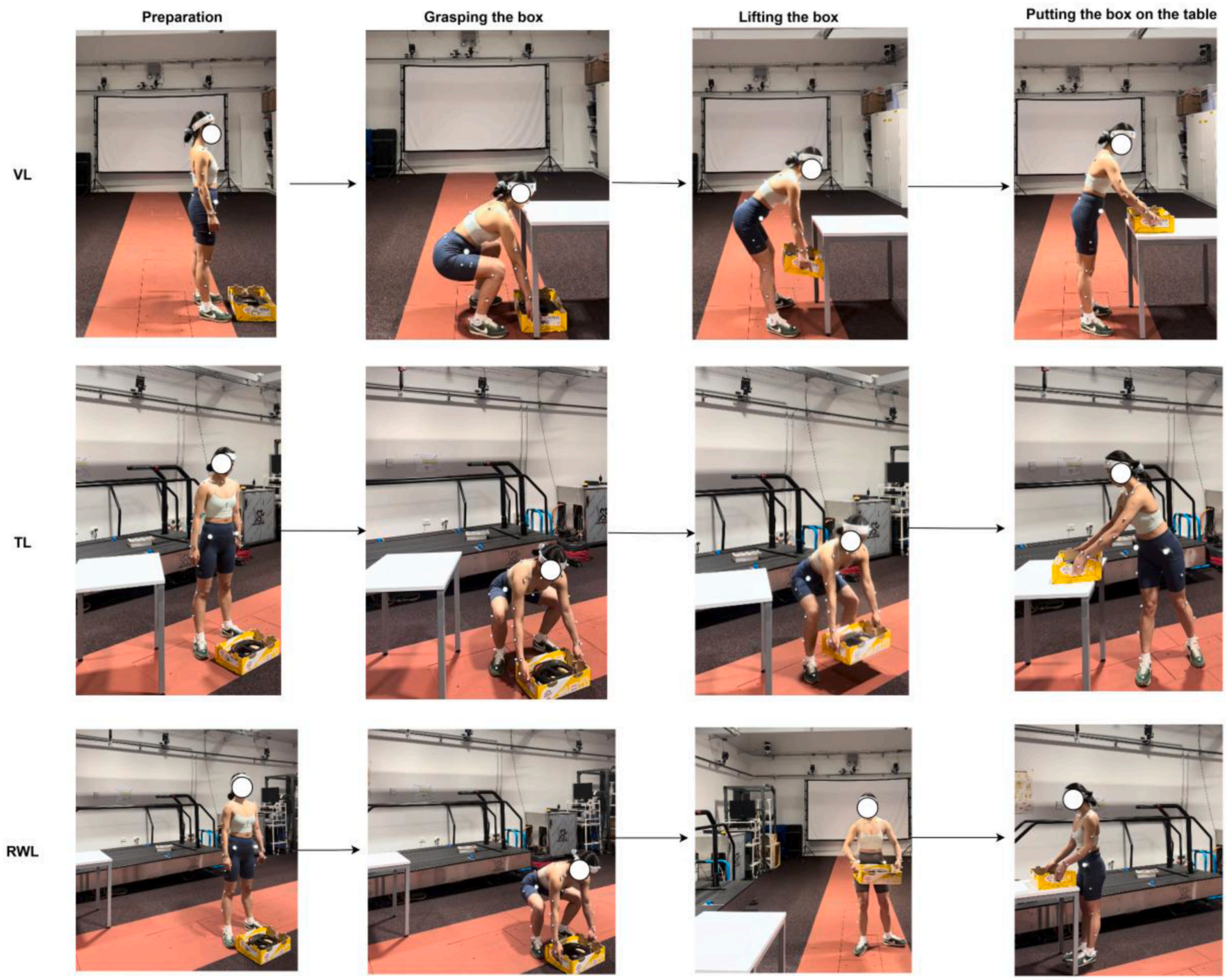

Fig. 1. Examples of the three lifting tasks (VL: vertical lifting in the sagittal plane; TL: transverse lifting with trunk rotation; RWL: rotational walking-and-lifting) depicting the different phases of each lift (preparation, grasping the box, lifting the box and box placement).

Table 1
Participant demographics (n = 20, mean ± SD).

Gender	Height (m)	Weight (kg)	BMI (kg/m^2)	Age (years)
Male (n = 8)	1.74 ± 0.06	71.13 ± 6.05	23.62 ± 3.18	27.75 ± 2.25
Female (n = 12)	1.65 ± 0.04	53.92 ± 5.38	19.87 ± 1.71	28.5 ± 2.20

2.2. Experimental protocol

The study was conducted in a large, unobstructed biomechanics laboratory (see Figs. 1 and 2). Floor tape was used to mark out the area where the lifting tasks were performed to standardise initial foot placement and ensure consistent positioning of the feet and box at the start of each task. Participants wore close-fitting clothing to allow markers to be suitably placed over anatomical landmarks and minimise occlusion due to loose fabric. They also wore flat, enclosed footwear, as this is typically the footwear worn in industry. Participants were instructed on the specific task requirements, i.e., lift origin, target placement, and self-paced timing. No information was provided about lifting technique, allowing them to adopt self-selected handling strategies. Participants lifted the box from the floor onto a table at 72 cm height (approximately waist height). A 10 kg box (dimensions: 40 cm × 30 cm × 25 cm) was used across all tasks. The load magnitude was selected in accordance with ISO 11228-1 (2021) (International Organization for, 2021) and the Revised NIOSH Lifting Equation (Waters et al., 1993) to minimise the risk of work-related musculoskeletal injury. Based on task geometry and asymmetry requirements, the estimated Lifting Index (LI) ranged between 0.2 and 0.9, below the NIOSH limit of 1.0, indicating acceptable risk for most healthy workers (Waters et al., 1993).

2.3. Data acquisition

Each participant performed the three lifting tasks: VL, TL, and RWL (Fig. 1) in a fixed order (VL→TL→RWL) to reflect the progressive increase in task complexity, from symmetrical sagittal-plane motion to asymmetric multi-planar movements involving rotation and walking. For each task, three valid trials were performed by each participant. Participants were instructed to complete each lift at a self-selected, purposeful pace, representative of workplace lifting, without rushing or compromising control. A trial was considered valid if the box was transferred continuously without adjustments to grip, pauses, drops, or deviations from the designated start/target zones; invalid trials were

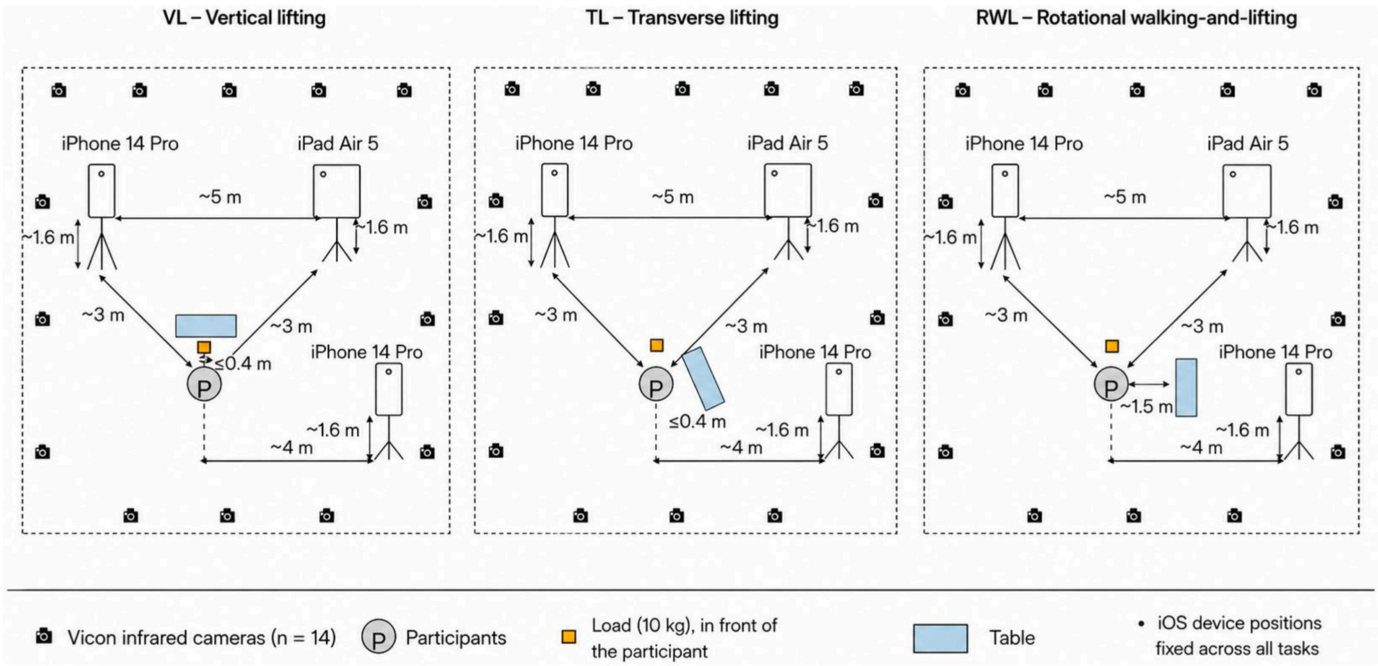

Fig. 2. Schematic diagram of the experimental setup and task-specific table positions for the three lifting conditions. For OpenCap, three iOS devices (two iPhone 14 Pro and one iPad Air 5) were placed in a fixed position for all tasks to provide multi-view coverage. For MediaPipe, camera position was task-dependent: the iPhone 14 Pro positioned to the right of the participant provided a sagittal-plane view for VL, whereas the iPad Air 5 provided an oblique view for TL and RWL.

repeated. Rest periods of 60–90 s were provided between trials, with additional rest as needed to mitigate fatigue.

Marker-based and markerless motion capture were performed simultaneously during each trial. The marker-based reference system comprised 14 infrared cameras (Vicon, Oxford Metrics Ltd., Oxford, UK), including 12 Vantage 16 cameras and 2 Vero v2.2 cameras, sampling at 100 Hz (Fig. 2). A standard volume/wand calibration was performed prior to testing to define the laboratory global coordinate system and camera geometry. These cameras were used to track 39 passive reflective markers (10 mm diameter, except for two 14 mm for the bilateral anterior superior iliac spine [R/LASI] to reduce occlusion during flexion) attached to anatomical landmarks by a trained operator. The Plug-in Gait Full Body model (Vicon Motion) defined segments (head, trunk, pelvis, bilateral upper limbs [upper arm/forearm/hand],

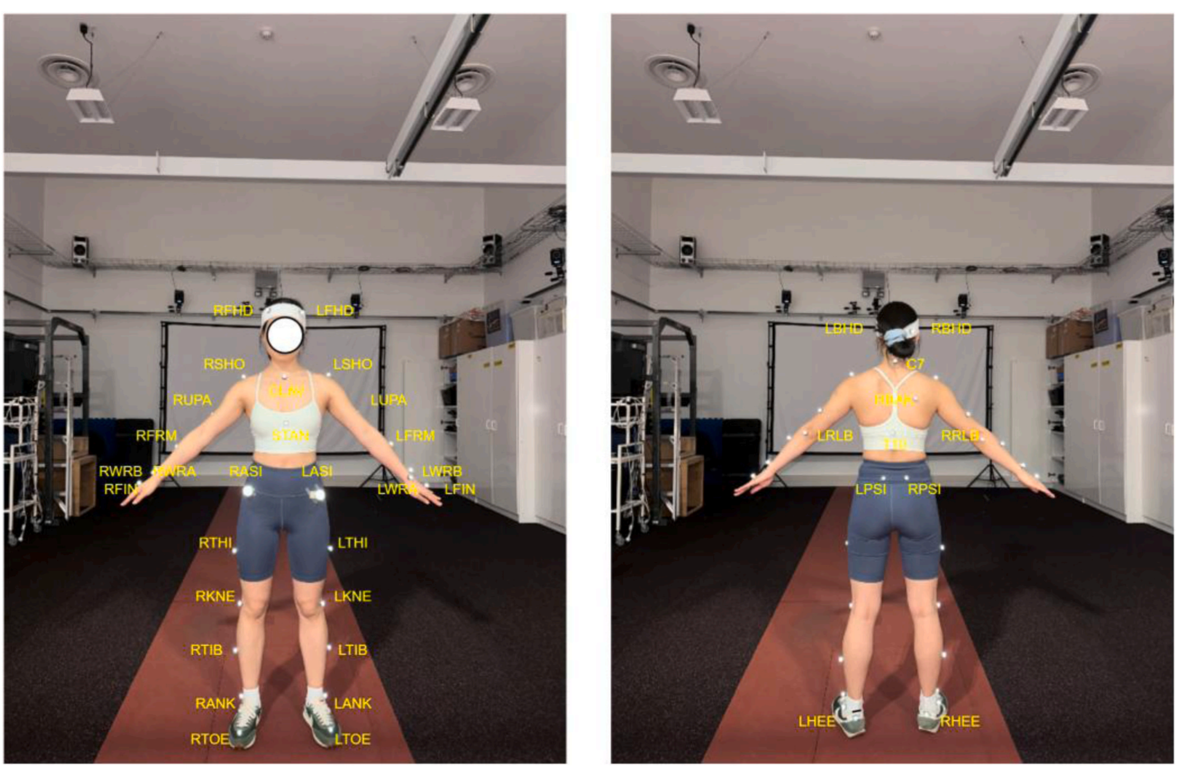

Fig. 3. Participant anatomical static calibration.

bilateral lower limbs [thigh/shank/foot]). A standing anatomical static calibration trial (T-pose; Fig. 3) was captured to map landmarks and initialise joint definitions.

For the OpenCap multi-view markerless system, three iOS devices (Apple Inc., Cupertino, CA, USA; two iPhone 14 Pro and one iPad Air 5) were positioned approximately 30–45° apart around the participant's frontal and sagittal planes on tripods at approximately 1.6 m height (Fig. 2). This configuration ensured optimal full-body visibility whilst minimising occlusion (Team). A 4×7 printed checkerboard (35 mm squares) was used for camera calibration to derive the cameras' positions and orientations for 3D reconstruction. Recordings were captured at 60 Hz and processed through the OpenCap cloud platform (Team).

For the MediaPipe monocular markerless system, RGB video recordings (60 Hz) from a single iOS device were processed using a custom pipeline based on MediaPipe BlazePose (v0.8, Google LLC) (Lugaresi et al., 2019) to extract 33 body landmarks in 3D global/world coordinates and reconstruct joint kinematics (detailed in Section 2.4). Device selection was task-dependent. For VL, the iPhone 14 Pro was positioned to the right of the participant to provide a sagittal-plane view (∼90°). For TL and RWL, the iPad Air 5 was positioned to provide an oblique view (∼45°) (Fig. 2), improving landmark visibility and reducing marker occlusion during trunk rotation and the subsequent walking transition (Kim et al., 2025; Baldinger et al., 2025; Hsu et al., 2025).

2.4. Data processing

The lifting cycle was defined from the initial standing position to the final box placement, capturing the complete manual handling sequence essential for observational risk assessments using REBA (Hignett and McAtamney, 2000). Temporal alignment across the three systems was performed using common event markers identified consistently in all recordings. Cycle onset was defined as the initiation of trunk flexion from the upright standing position, and cycle completion as the point at which the participant released the box, identified using the wrist markers in Vicon and the corresponding wrist landmarks in OpenCap and MediaPipe. All events were identified by a single researcher using the same criteria throughout. Marker-based data were labelled and processed in Vicon Nexus (Vicon Nexus v2.15 x64, Oxford Metrics Ltd, Oxford, UK) using the Plug-in Gait FullBody (Dynamics) pipeline (Fig. 4A). Marker trajectories were filtered using a zero-lag, 4th-order Butterworth low-pass filter with a cutoff frequency of 6 Hz, with processed kinematic outputs (joint angles, joint centres) exported as comma-separated values (.csv) (Vicon Motion). OpenCap recordings were processed automatically on the OpenCap cloud platform (Team) using the embedded OpenSim musculoskeletal pipeline (Rajagopal et al., 2016; Delp, 2025) (Fig. 4B), generating joint kinematics files in OpenSim motion (.mot) format with time and angle data for each trial.

For MediaPipe, pose estimation used MediaPipe BlazePose (v0.8, Google LLC) (Bazarevsky et al., 2020; Lugaresi et al., 2019) to extract the 3D positions of 33 body landmarks per frame (Fig. 4C). It is a CNN trained on large, manually annotated image datasets without an underlying biomechanical model. Joint centres were taken directly from the corresponding landmarks (shoulder, elbow, wrist, hip, knee, and ankle), with virtual mid-shoulder and mid-hip (spine base) points defined as the midpoints of the respective left and right landmarks. Local segment coordinate systems were subsequently constructed from these joint centres (**Supplementary Material A**, Table A2 and A3). Joint angles for the shoulder, elbow, hip, and trunk were computed as Z–X–Y Euler rotations (distal relative to proximal) based on the International Society of Biomechanics recommendation (Wu et al., 2005). Knee flexion/extension and ankle dorsiflexion/plantarflexion were computed as relative sagittal-plane vector angles between adjacent segments, providing stability for these predominantly uniaxial joints. Kinematics for each trial were exported as.csv files. Raw landmark positions and derived angle time series were low-pass filtered using a zero-lag, 4th-order Butterworth filter at 4 Hz. This cutoff, lower than the 6 Hz applied to the marker-based and OpenCap data, was selected to attenuate the greater high-frequency frame-to-frame jitter inherent in video-based monocular landmark detection, whilst remaining well above the fundamental movement frequencies (<1 Hz) of the lifting tasks, consistent with accepted practice for filtering manual-handling kinematics (Kim et al., 2025; Crenna et al., 2021; Williams et al., 2022).

All OpenCap and MediaPipe outputs (trunk, shoulder, elbow, hip, knee, and ankle joint angles) were remapped to the Vicon joint coordinate system, ensuring identical definitions and sign conventions (positive flexion, adduction, internal rotation; **Supplementary Material A,** Table A4). A total of 180 lifting trials were conducted (3 tasks × 3 repetitions × 20 participants). Measurement noise alone was not used as a basis for exclusion. Trials were excluded only where a system produced an outright processing failure: for OpenCap, when the cloud pipeline failed to reconstruct any valid 3D skeleton; for MediaPipe, when landmark tracking failed within individual frames, producing simultaneous, physiologically implausible joint-angle discontinuities (jumps of tens of

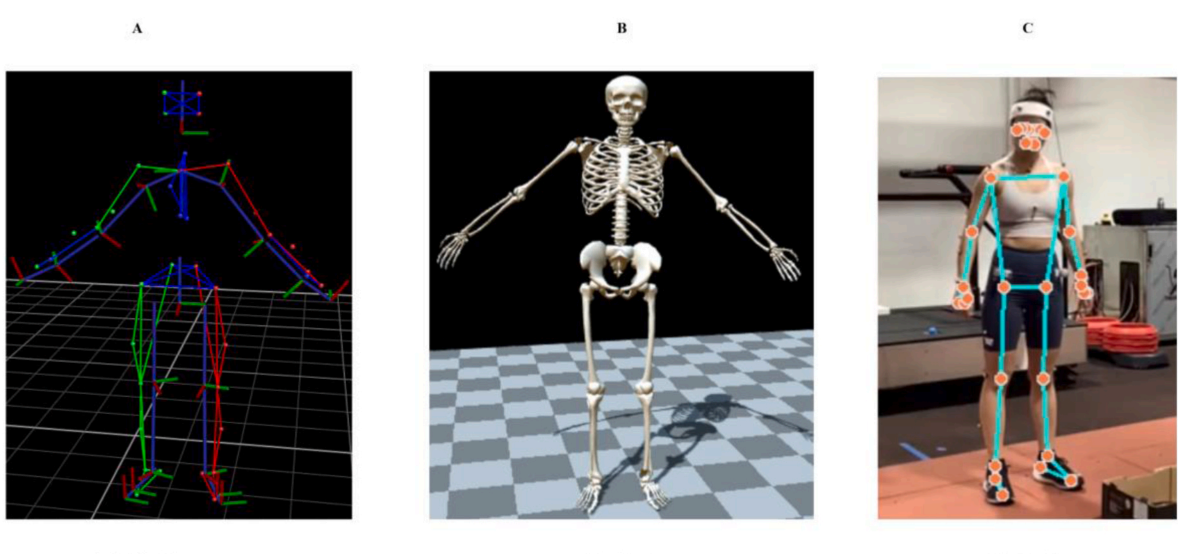

Fig. 4. Representative skeletal models from the three motion-capture systems: (A) Vicon Nexus (marker-based reference), (B) OpenCap (OpenSim musculoskeletal model), and (C) MediaPipe (33-landmark pose estimation from RGB video).

degrees within a single frame) across two or more joints. Such failures were visually unambiguous and distinct from valid but noisy measurements. In total, 137 trials (76.1%) were retained for analysis: VL (n = 50, 83.3%), TL (n = 48, 80.0%), and RWL (n = 39, 65.0%).

2.5. Data analysis

Following cycle boundary identification (Section 2.4), kinematic data from all three MoCap systems were processed in Python 3.13 for comparative analysis. For each participant, kinematic data from the three trials were time-normalised to 0–100% of the lifting cycle (101 samples) and ensemble-averaged to obtain a representative waveform for harmonisation and pointwise comparison. Concurrent validity was assessed using complementary error and agreement metrics calculated over the normalised cycle duration: 1) root-mean-square error (RMSE, °) to quantify pointwise deviation from Vicon; 2) normalised RMSE (i.e., RMSE%, expressed as a percentage of the Vicon peak-to-peak range for the same joint-plane-task combination) (Jamali et al., 2025); and 3) the coefficient of determination (R^2) from pointwise linear regression of normalised joint-angle waveforms to assess waveform similarity between systems. Because TL and RWL involved asymmetric trunk rotation (unlike the symmetrical VL), kinematic data from both left and right sides were analysed. For brevity, right-side results are presented in the main text, with left-side results presented in **Supplementary Material B** to assess bilateral consistency. This bilateral approach aligns with previous motion capture validation studies for asymmetric tasks (de Andrade and Carpes, 2024; Schurr et al., 2017).

REBA was selected because it is a posture-based assessment method that can be derived directly from the joint-angle outputs of all three motion-capture systems and is designed for whole-body assessment of dynamic manual-handling postures. The NIOSH Lifting Equation was used only to confirm task acceptability (Section 2.2). REBA total scores and risk categories were derived from peak-frame joint angles, defined as the frame of maximum sagittal-plane trunk flexion during the active lifting phase (the most demanding lifting posture) for each valid trial, consistent with other studies using automated ergonomic assessment (Agostinelli et al., 2024; Lin et al., 2022). Full REBA scores were computed from the measured joint angles using the standard scoring tables; the non-kinematic inputs (muscle-use, coupling, and activity scores) were held constant at their default lowest-risk values across all three systems, and the force/load score was set according to the 10 kg load. Consequently, between-system differences in the resulting scores reflect only differences in the measured joint angles. Agreement between markerless and Vicon-derived classifications was assessed using mean absolute error (MAE) ± standard deviation (SD) and percentage exact agreement (PEA), defined as the percentage of trials in which the markerless system assigned the same risk category as the Vicon-derived reference. Both indices were calculated separately for each system (OpenCap and MediaPipe) and task condition (see **Supplementary Material C**).

Descriptive summaries (mean ± SD) were calculated for RMSE and R^2 across all participants, and mean RMSE% values provided an indication of system performance across tasks, joints, and movement planes. All outcome metrics were tested for normality using the Shapiro-Wilk test. To examine the effect of task on system validity, a two-way (System × Task) repeated-measures ANOVA was conducted on RMSE for each joint–plane combination, using the subset of participants with complete data across all three systems and task conditions (n = 14). System (2: OpenCap, MediaPipe) and Task (3: VL, TL, RWL) were entered as within-subject factors, and partial eta-squared (ηp^2) was reported as the effect size. Main effects and the System × Task interaction were evaluated from the omnibus ANOVA; where significant, between-system differences within each task were examined using paired t-tests with Bonferroni adjustment for multiple comparisons (**Supplementary Material B**). The significance level was set at α = 0.05, applied to the adjusted (adj.) *p*-values. RMSE values were interpreted using the accuracy criteria established by Auer et al. (2024) and adopted in markerless motion-capture validation studies (Varcin and Boocock, 2026): excellent (<3°), reasonable (3–7°), and poor (>7°). R^2 values were interpreted as low (<0.50), moderate (0.50–0.74), good (0.75–0.89), and very good (≥0.90) (McGinley et al., 2009). All statistical tests were performed using IBM SPSS Statistics (Version 30.0; IBM Corp., Armonk, NY, USA).

3. Results

The concurrent validity of OpenCap and MediaPipe for right-side joints is summarised in Table 2 and Figs. 5–8; ergonomic risk classification results are presented in Table 3, whereas inferential statistics for system, task, and post-hoc comparisons are reported in Supplementary Material B.

Across all joint-plane-task combinations, RMSE values were reasonable to poor (>4°) for both systems, with waveform similarity (R^2) ranging from low to good. OpenCap exhibited RMSE of 4.4–67.1° (RMSE % = 12–121%; R^2 = 0.10–0.89), and MediaPipe exhibited RMSE of 5.9–67.1° (RMSE% = 13–150%; R^2 = 0.07–0.88). RMSE increased by approximately 15–30% from symmetrical (VL) to complex asymmetric multi-planar conditions (RWL) (Table 2; Fig. 8), with the largest task-related increases observed in transverse-plane motion (Supplementary Material B, Tables B1 and B3). Both systems showed lower RMSE in the sagittal plane (RMSE: 9.9–31.2°) than in the frontal and transverse planes (RMSE: 4.3–67.1°). Left and right sides showed similar measurement performance for both systems, with left-right differences in RMSE and RMSE% of 2–4° and 2–5%, respectively; left-side results and full ANOVA outputs are reported in Supplementary Material B.

3.1. Kinematic validity

There was a significant main effect of task on RMSE for the trunk in all three planes, for the right elbow and ankle, and for shoulder and hip rotation (Table B1), indicating that task significantly affected kinematic validity. This was largest in the transverse plane, where trunk axial-rotation RMSE increased with complexity, reaching 67.1° (OpenCap) and 58.7° (MediaPipe) during RWL (all task contrasts adj. $p < 0.001$; Tables B1, B3; Figure B5).

There was a significant main effect of system in several joint-plane combinations (Table B1), although its direction depended on joint region and movement plane. In the sagittal plane, the pattern was consistent across tasks: MediaPipe showed lower RMSE for the trunk and shoulder, whereas OpenCap showed lower RMSE for the distal elbow and ankle (adj. $p < 0.01$; Table B2), and the knee did not differ between systems (adj. $p > 0.05$). In the frontal and transverse planes, the between-system difference was more task-dependent, consistent with significant System × Task interactions for the trunk, hip, and shoulder (Table B1); the full task-by-task contrasts are reported in Table B2.

3.2. REBA risk assessment

For REBA, both systems achieved 100% category agreement (PEA) for VL, declining with task complexity (MediaPipe: TL 78.3%, RWL 70.3%; OpenCap: TL 71.7%, RWL 70.3%); REBA total scores remained within approximately two points of the marker-based reference (MAE: MediaPipe 0.47–2.03; OpenCap 0.43–1.92 points). Key agreement metrics are summarised in Table 3; full score distributions and risk-category cross-tabulations are provided in **Supplementary Material C** (Tables C1 and C2).

4. Discussion

This study evaluated the concurrent validity of a multi-view (OpenCap) and a monocular (MediaPipe) markerless system for occupational lifting tasks of increasing complexity, to provide task-specific

Table 2
Concurrent validity metrics (RMSE in degrees, RMSE% in brackets, R^2) for OpenCap and MediaPipe compared to Vicon across three lifting tasks and movement planes for right-side joints (n = 20).

Variables	Tasks	RMSE (OpenCap)	RMSE (MediaPipe)	R^2 (OpenCap)	R^2 (MediaPipe)
Flex (+)/Ext (−) (°)					
Trunk	VL	18.62 ± 9.19(29%)	9.86 ± 3.49(16%)	0.56 ± 0.18	0.85 ± 0.12
	TL	19.76 ± 10.24(29%)	12.00 ± 3.91(17%)	0.69 ± 0.13	0.86 ± 0.10
	RWL	24.12 ± 9.67(30%)	14.21 ± 5.52(18%)	0.58 ± 0.25	0.80 ± 0.12
Rhip	VL	21.10 ± 7.28(22%)	28.55 ± 6.91(30%)	0.82 ± 0.15	0.87 ± 0.10
	TL	23.61 ± 7.84(25%)	30.60 ± 9.10(32%)	0.75 ± 0.18	0.81 ± 0.16
	RWL	25.58 ± 11.14(26%)	30.34 ± 8.61(31%)	0.67 ± 0.26	0.77 ± 0.12
Rshoulder	VL	26.73 ± 5.87(36%)	16.78 ± 4.57(23%)	0.82 ± 0.12	0.83 ± 0.11
	TL	24.36 ± 10.27(33%)	15.78 ± 3.69(22%)	0.78 ± 0.16	0.76 ± 0.15
	RWL	25.40 ± 7.14(34%)	17.69 ± 4.43(24%)	0.64 ± 0.25	0.70 ± 0.15
Relbow	VL	18.45 ± 3.36(26%)	27.43 ± 3.66(39%)	0.84 ± 0.11	0.07 ± 0.06
	TL	17.73 ± 5.15(27%)	31.22 ± 7.09(47%)	0.83 ± 0.16	0.11 ± 0.10
	RWL	22.26 ± 5.46(33%)	30.86 ± 5.31(45%)	0.74 ± 0.23	0.12 ± 0.12
Rknee	VL	16.48 ± 6.61(12%)	18.45 ± 5.23(13%)	0.89 ± 0.09	0.88 ± 0.10
	TL	17.45 ± 7.81(14%)	20.12 ± 7.32(16%)	0.83 ± 0.16	0.80 ± 0.17
	RWL	25.03 ± 14.64(21%)	22.52 ± 9.39(19%)	0.64 ± 0.26	0.68 ± 0.16
Rankle	VL	9.98 ± 5.67(23%)	18.30 ± 4.98(42%)	0.81 ± 0.17	0.21 ± 0.19
	TL	15.33 ± 8.33(37%)	29.46 ± 5.56(71%)	0.55 ± 0.28	0.24 ± 0.15
	RWL	14.28 ± 6.99(32%)	24.90 ± 5.54(55%)	0.45 ± 0.21	0.20 ± 0.12
Add (+)/Abd (−) (°)					
Trunk	VL	5.90 ± 3.18(121%)	5.92 ± 1.76(122%)	0.10 ± 0.06	0.14 ± 0.16
	TL	6.69 ± 3.05(24%)	11.89 ± 5.26(42%)	0.70 ± 0.17	0.30 ± 0.23
	RWL	13.81 ± 9.51(48%)	13.13 ± 8.25(46%)	0.14 ± 0.11	0.17 ± 0.10
Rhip	VL	6.27 ± 5.15(45%)	11.28 ± 6.07 (81%)	0.52 ± 0.31	0.23 ± 0.16
	TL	7.70 ± 4.13(31%)	11.32 ± 3.04(46%)	0.49 ± 0.24	0.21 ± 0.22
	RWL	7.52 ± 3.86(31%)	10.15 ± 2.29(41%)	0.38 ± 0.24	0.10 ± 0.13
Rshoulder	VL	29.56 ± 11.42(68%)	31.41 ± 11.50(72%)	0.19 ± 0.14	0.14 ± 0.13
	TL	31.72 ± 11.81(70%)	27.98 ± 9.27(61%)	0.17 ± 0.17	0.12 ± 0.14
	RWL	29.05 ± 9.81(61%)	25.39 ± 9.98(53%)	0.07 ± 0.07	0.12 ± 0.12
Int (+)/Ext Rot (−) (°)					
Trunk	VL	4.35 ± 2.56(72%)	8.06 ± 2.10(133%)	0.15 ± 0.13	0.11 ± 0.09
	TL	26.09 ± 6.94(28%)	36.53 ± 8.85(40%)	0.86 ± 0.11	0.55 ± 0.23
	RWL	67.14 ± 9.56(52%)	58.66 ± 20.93(46%)	0.13 ± 0.13	0.83 ± 0.17
Rhip	VL	15.94 ± 8.97(62%)	13.98 ± 5.64(54%)	0.53 ± 0.29	0.41 ± 0.23
	TL	20.68 ± 10.45(76%)	19.45 ± 6.77(72%)	0.25 ± 0.18	0.15 ± 0.13
	RWL	19.45 ± 8.48(74%)	20.60 ± 7.44(78%)	0.14 ± 0.15	0.12 ± 0.12
Rshoulder	VL	32.01 ± 13.72(81%)	59.79 ± 14.53(150%)	0.26 ± 0.23	0.38 ± 0.23
	TL	25.61 ± 15.81(54%)	32.96 ± 10.41(69%)	0.24 ± 0.22	0.17 ± 0.15
	RWL	25.47 ± 12.88(53%)	31.56 ± 9.61(66%)	0.13 ± 0.13	0.19 ± 0.14

Notes: Values are presented as mean ± SD across participants. Rhip = Right Hip; Rshoulder = Right Shoulder; Relbow = Right Elbow; Rknee = Right Knee; Rankle = Right Ankle. RMSE% normalises RMSE to Vicon range-of-motion for the same joint-plane-task; percentages in parentheses indicate RMSE%. R^2 quantifies waveform similarity (time-normalised to 0–100% cycle). VL = vertical lifting; TL = transverse lifting; RWL = rotational walking-and-lifting. Flex (+)/Ext (−) = Flexion/Extension; Add (+)/Abd (−) = Adduction/Abduction; Int (+)/Ext Rot (−) = Internal/External Rotation. The (+)/(−) symbols denote the joint-angle sign convention (Table A4); RMSE values are reported as absolute (unsigned) errors.

guidance for system selection in ergonomic assessment. Overall, both systems demonstrated reasonable validity for symmetrical lifting, but performance declined substantially with increasing multi-planar demands. The following sections discuss kinematic concurrent validity, ergonomic risk assessment performance, and study limitations in relation to the study hypotheses.

4.1. Concurrent validity of markerless systems

Consistent with the first study hypothesis, task complexity moderated system performance for both systems, with RMSE increasing by approximately 15–30% from VL to RWL and the most pronounced increases in the transverse plane. Performance during symmetrical lifting (VL) was comparable to that reported in gait studies (RMSE ranged from 5° to 10° with $R^2 > 0.80$) (Cheng et al., 2025; Scataglini et al., 2024; D'Souza et al., 2024; Jamali et al., 2025; Outerleys et al., 2024; McGinley et al., 2009; Yoma et al., 2025), likely due to the constrained, symmetrical nature of the task, minimal self-occlusion, and predictable movement patterns. As task complexity increased to involve trunk rotation and walking, kinematic agreement declined substantially, particularly in the transverse plane: trunk flexion–extension retained moderate waveform agreement even during RWL (OpenCap $R^2 = 0.58$), whereas trunk axial rotation exhibited very poor agreement (OpenCap $R^2 = 0.13$), most likely stemming from depth-axis and occlusion-related challenges in markerless pose estimation (Yeom et al., 2025; Kim et al., 2025; Cronin et al., 2024). The reduction in the number of suitable trials, from 83.3% at VL to 65.0% at RWL, further indicated that complex asymmetric postures and dynamic movement patterns increase the likelihood of pose-estimation failure (McGuirk et al., 2022).

These error magnitudes should be interpreted against task demands. The 3–7° accuracy thresholds applied here derive from gait and cyclical functional tasks (Varcin and Boocock, 2026; Auer et al., 2024) and may be overly stringent for occupational lifting, which involves asymmetric postures, multi-planar rotation, and external load handling. Much of the error classified as poor on this basis may therefore overstate the systems' practical limitations.

System differences followed a consistent proximal-distal pattern in the sagittal plane, with neither system showing an overall advantage, whereas frontal- and transverse-plane differences were more sensitive to task complexity. MediaPipe showed lower RMSE for trunk and shoulder, whilst OpenCap demonstrated lower RMSE for distal joints (elbow and ankle). These contrasting patterns likely reflect differences in 3D pose-

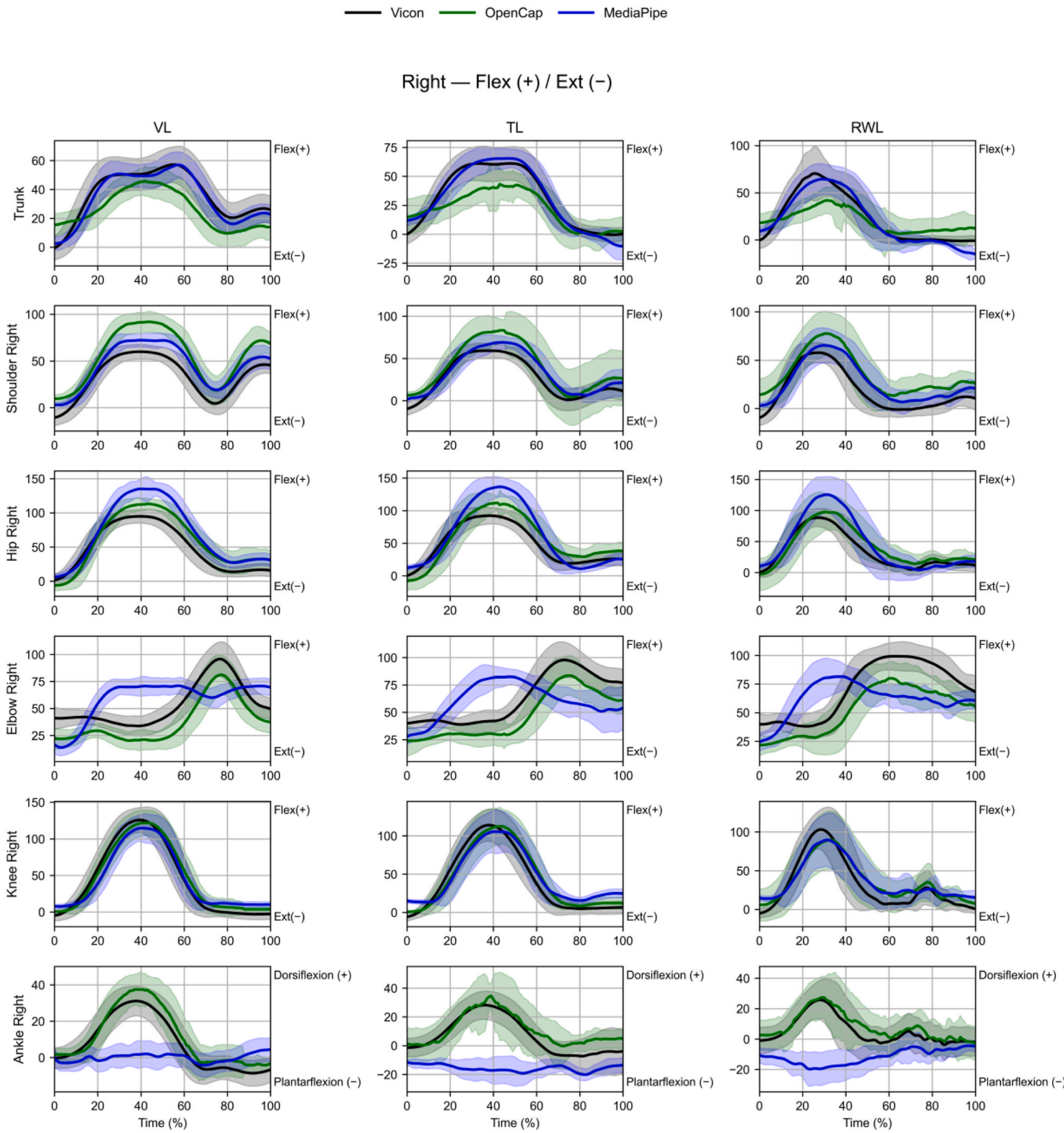

Fig. 5. Sagittal-plane joint-angle waveforms (mean ± SD; 0–100% cycle). Flex (+)/Ext (−) = flexion/extension; Dorsi (+)/Plantar (−) = dorsiflexion/plantarflexion.

estimation architecture. MediaPipe estimates 3D landmarks from a single RGB video using statistical patterns learned from large image datasets (e.g., COCO and MPII) (Scataglini et al., 2024; Lugaresi et al., 2019), which may stabilise well-represented proximal joints but increase error at distal joints, where occupational lifting postures are underrepresented (Scataglini et al., 2024; Kim et al., 2023, 2025; Yoma et al., 2025). By contrast, OpenCap constrains 3D positions through multi-view triangulation and musculoskeletal modelling, which may improve distal-joint estimation when multiple views remain visible (Uhlrich et al., 2023; Rajagopal et al., 2016). Yeom et al. (2025) similarly reported proximal errors exceeding 36° but distal errors below 5° in multi-view systems. Knee flexion–extension showed no system-level difference, likely reflecting its uniaxial kinematics and consistent landmark visibility (Scataglini et al., 2024; Outerleys et al., 2024). These findings partially support the second study hypothesis: between-system error was comparable during VL, but the joint-specific advantages persisted across tasks rather than increasing with task complexity, as hypothesised. System selection should therefore be guided primarily by the joint regions and movement planes of primary interest.

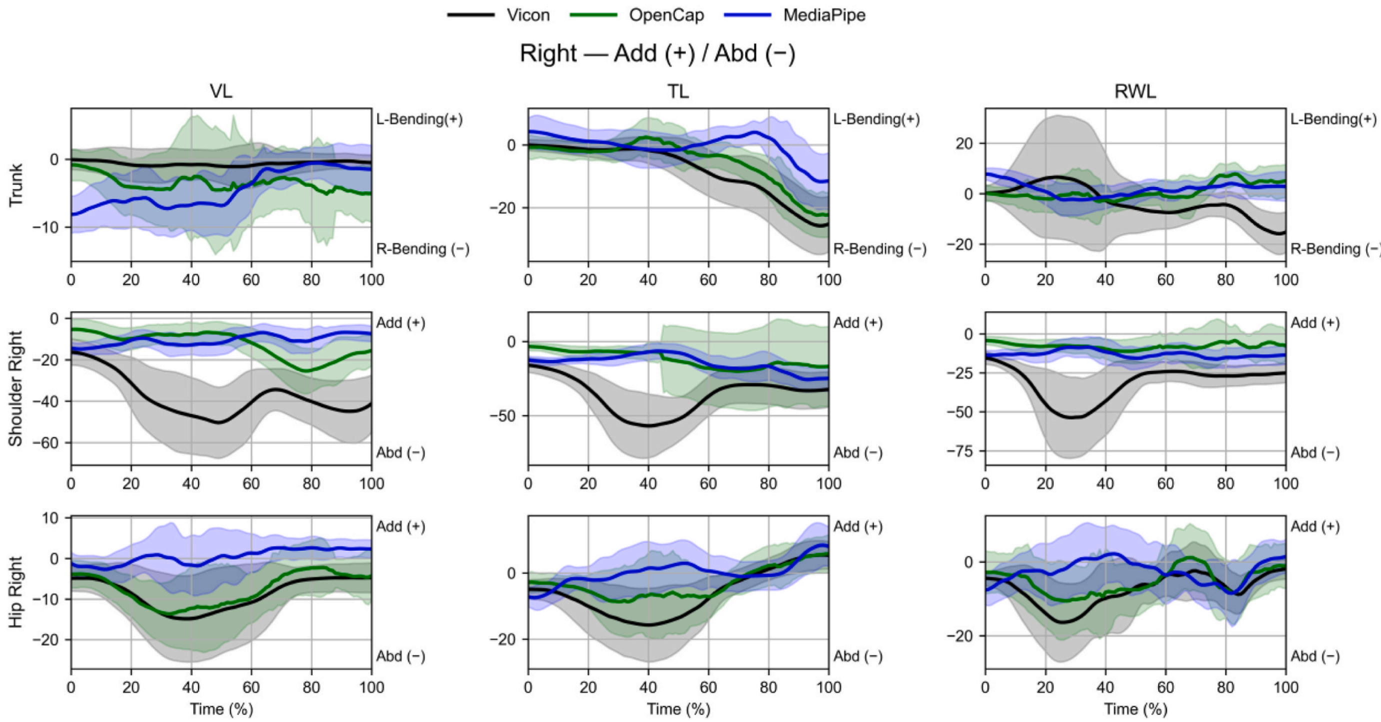

Fig. 6. Frontal-plane joint-angle waveforms (mean ± SD, 0–100% cycle). Add (+)/Abd (−) = Adduction/Abduction; L-Bending/R-Bending = left/right lateral bending.

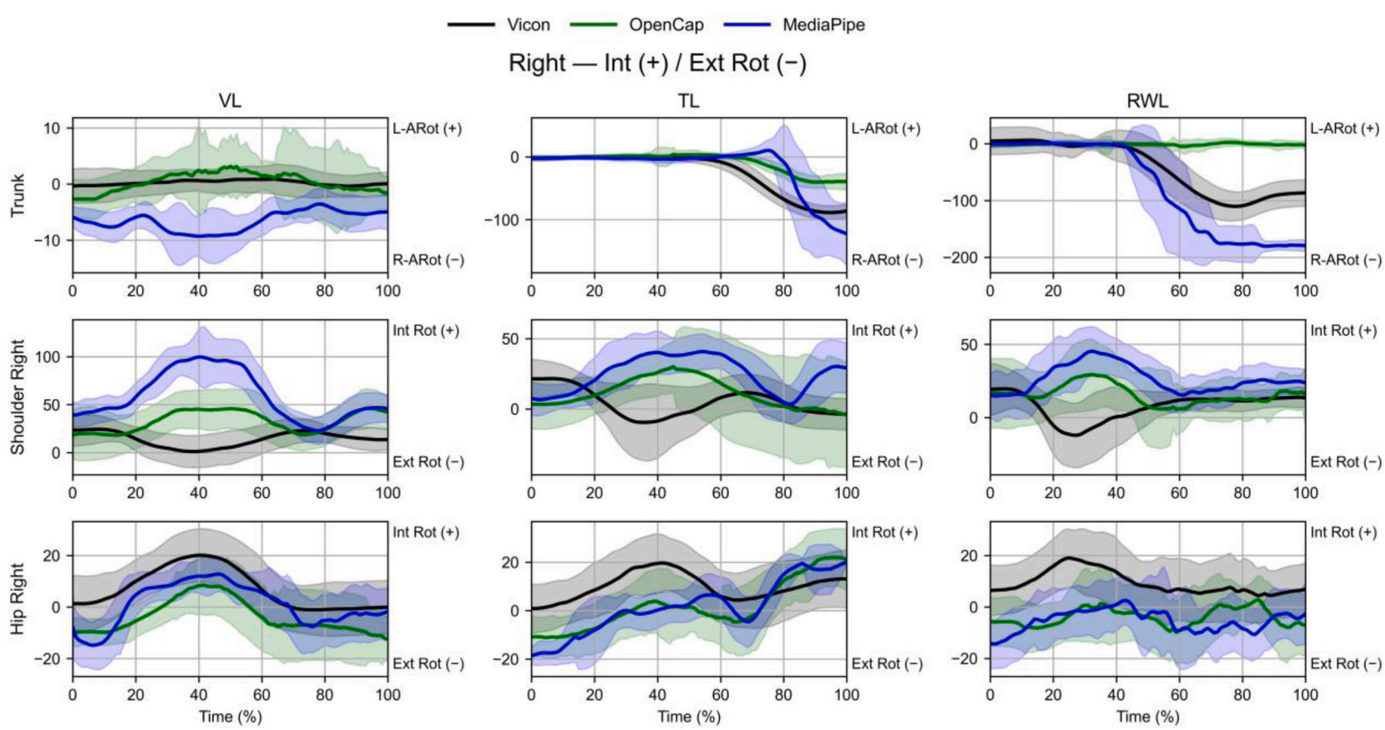

Fig. 7. Axial-rotation joint-angle (mean ± SD, 0–100% cycle). Int (+)/Ext Rot (−) = Internal/External Rotation; L-ARot/R-ARot = left/right axial rotation.

4.2. Ergonomic risk assessment

Consistent with the third study hypothesis, REBA risk classification agreement declined progressively with increasing task complexity for both systems. For symmetrical lifting (VL), both systems achieved 100% REBA risk-category agreement with the marker-based reference. As task complexity increased to involve trunk rotation and walking, agreement declined substantially: REBA PEA fell to approximately 70% during RWL for both systems. Trial retention also declined from VL (83.3%) to RWL (65.0%), indicating that complex multi-planar tasks increased the likelihood of tracking failure (McGuirk et al., 2022), which further constrains the reliability of risk classifications under these conditions.

These findings extend previous work on automated ergonomic risk scoring from markerless systems (Agostinelli et al., 2024; Abobakr et al.,

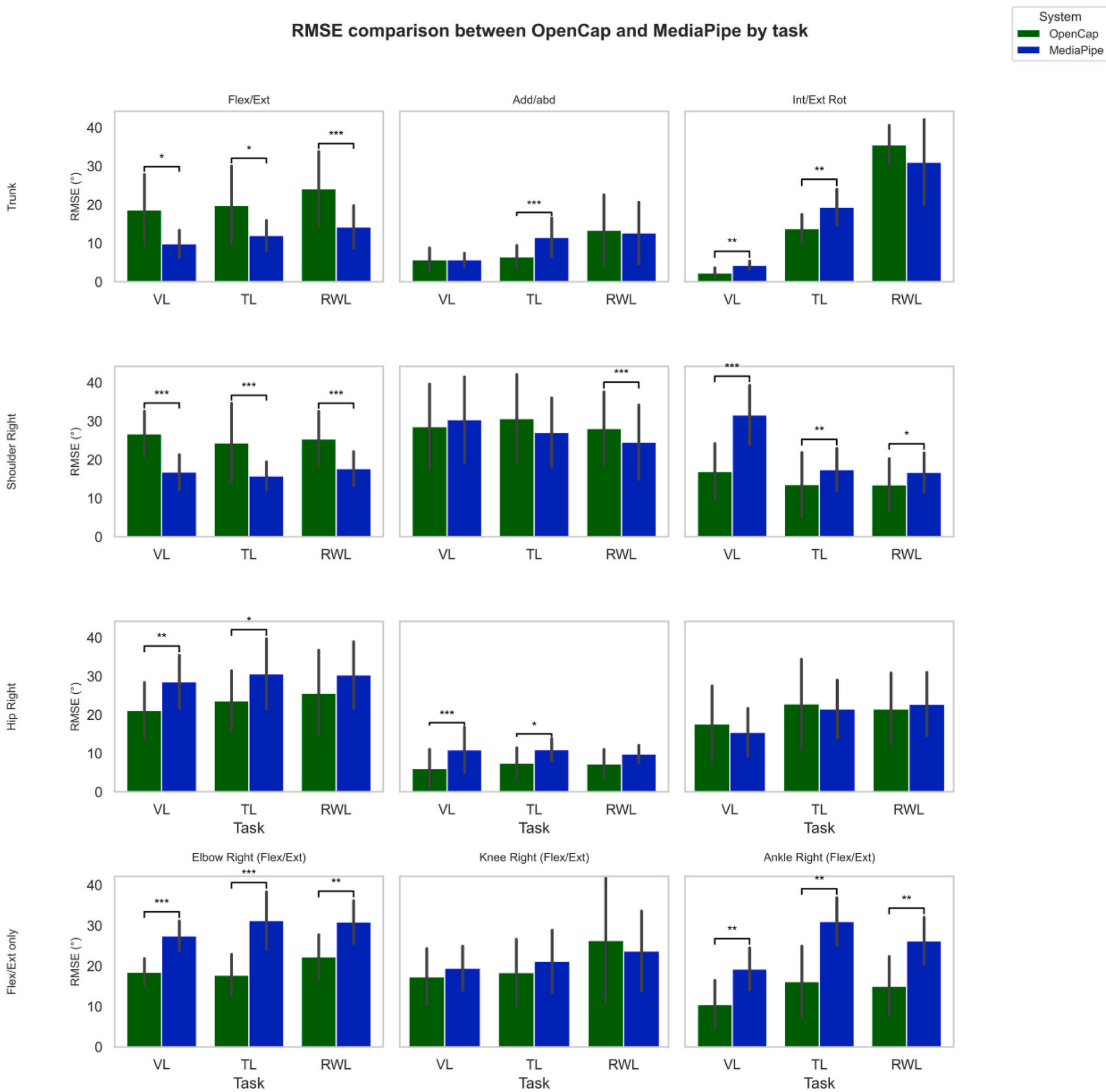

Fig. 8. RMSE comparison between OpenCap and MediaPipe by task and movement plane for right-side joints. Bars show mean ± SD across participants. Brackets indicate significant between-system differences within each task. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Flex/Ext = Flexion/Extension; Add/Abd = Adduction/Abduction; Int/Ext Rot = Internal/External Rotation.

2019; Bonakdar et al., 2025), showing that performance is task-dependent and that both systems offer acceptable classification accuracy for symmetrical tasks. REBA agreement was complete for symmetrical lifting because the postures that drive the REBA score lie largely in the sagittal plane, which both systems reproduce accurately (Hignett and McAtamney, 2000). In the transverse plane during TL and RWL, kinematic errors exceeded both the 10–20° angular thresholds used by REBA (Hignett and McAtamney, 2000) and the ISO 11226 criterion for trunk axial rotation (>10° as a relevant risk factor) (Delleman et al., 2000). As more trials crossed these thresholds with increasing complexity, errors increasingly altered the risk category, accounting for the decline in classification agreement. This decline was concentrated where REBA is most demanding: the trunk dominates the REBA score and is most affected by transverse-plane (axial-rotation) error, so the plane in which both systems were least accurate was precisely the one that governs the risk category in complex lifts. Agreement therefore depended less on overall kinematic accuracy than on accuracy in the specific postures that REBA weights most heavily. For practitioners, markerless REBA screening is most dependable for sagittal-dominant lifting; its reliability for tasks involving substantial trunk rotation cannot be inferred from sagittal-plane validity alone.

For practical deployment, both systems represent viable alternatives to subjective observational methods for preliminary screening of symmetrical lifting tasks, and may provide more consistent classifications

Table 3
REBA risk classification agreement (MAE ± SD in points and PEA) for MediaPipe and OpenCap compared to the Vicon reference across three lifting tasks.

Task	MAE (points)		PEA	
	OpenCap	MediaPipe	OpenCap	MediaPipe
VL	0.43 ± 0.76	0.47 ± 0.77	100.0%	100.0%
TL	1.74 ± 2.06	1.52 ± 2.15	71.7%	78.3%
RWL	1.92 ± 2.10	2.03 ± 2.22	70.3%	70.3%

Notes: MAE = mean absolute error (lower is better). PEA = percentage exact agreement (higher is better). VL = vertical lifting; TL = transverse lifting; RWL = rotational walking-and-lifting. Trial counts per task condition: VL n = 49, TL n = 46, RWL n = 37. Full score distributions and risk-category cross-tabulations are provided in **Supplementary Material C** (Tables C1 and C2).

than inexperienced observers (Balogh et al., 2025). MediaPipe offers a portable single-camera solution suitable for field deployment, whilst OpenCap provides broader joint coverage under multi-planar conditions. For tasks involving substantial trunk rotation and walking, both systems should be applied with caution; more advanced markerless systems (e.g., Theia3D) or marker-based systems remain preferable when definitive risk assessment is required.

4.3. Limitations and future work

This study has several limitations. The controlled laboratory protocol (a fixed 10 kg load, a healthy adult sample, and standardised foot placement and pace) limits generalisation to the variable conditions and externally paced work rates of real workplaces. Methodologically, the marker-based reference is itself subject to soft-tissue artefact during trunk rotation (Leardini et al., 2005), so the reported values reflect between-system agreement rather than absolute accuracy, and the lower 4 Hz filter applied to MediaPipe (versus 6 Hz) (Kim et al., 2025; Crenna et al., 2021; Williams et al., 2022) may have modestly reduced its RMSE.

Notwithstanding these limitations, this study provides the first direct comparison of multi-view and monocular markerless systems across occupational lifting tasks of increasing complexity, combining kinematic validity with ergonomic risk classification to offer practice-relevant guidance for system selection. Future work should address the key technical challenges identified here, particularly robustness to self-occlusion, depth ambiguity, and tracking failure during complex multi-planar movements — through transformer-based temporal pose estimation (Zheng et al., 2021), multi-view fusion architectures (Mehrizi et al., 2018b), and hybrid biomechanical constraint frameworks (Kim et al., 2023; Pilla-Barroso et al., 2025). Domain-specific model adaptation for occupational lifting, as demonstrated for OpenCap, offers a promising direction (Salehi et al., 2026a, 2026b). Future studies should examine system performance across diverse worker populations, real workplace conditions, and a broader range of manual handling tasks to support wider deployment in occupational ergonomic practice.

5. Conclusion

This study demonstrated that markerless motion capture systems are viable alternatives for ergonomic risk assessment in symmetrical occupational lifting, though validity declined substantially for more complex multi-planar tasks involving trunk rotation and walking. For both OpenCap (multi-view) and MediaPipe (monocular), RMSE increased by approximately 15–30% from VL to RWL and was consistently lower in the sagittal plane than in the frontal and transverse planes, suggesting that task complexity is a primary determinant of markerless system validity. MediaPipe showed lower error for proximal joints (trunk and shoulder), whilst OpenCap performed better for distal joints (elbow and ankle), providing task- and joint-specific guidance for system selection. Both systems achieved 100% REBA risk-category agreement for symmetrical lifting, indicating their practical utility for preliminary ergonomic screening; agreement declined with increasing task complexity. Marker-based systems remain preferable for definitive risk assessment of complex multi-planar manual handling tasks. Future work should evaluate system performance under real-world workplace conditions and explore advanced pose estimation approaches to extend the utility of markerless systems in occupational settings.

Disclosure statement

The authors report there are no competing interests to declare.

Data availability

The data that support the findings of this study are available on request from the corresponding author, YZ. The data is not publicly available due to its containing information that could compromise the privacy of research participants.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author(s) used Claude (Anthropic) and Codex (OpenAI) to assist with language editing, condensing the text to meet the word limit, and improving the readability and structure of the manuscript. The author(s) reviewed and edited the output as needed and take full responsibility for the content of the published article.

Funding

None.

CRediT authorship contribution statement

Mengjie Zhang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Arne Nieuwenhuys:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Mark Boocock:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Yanxin Zhang:** Conceptualization, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

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.

Acknowledgements

The authors would like to acknowledge the financial support provided by the China Scholarship Council (CSC) for living expenses during the research. Gratitude is also extended to the participants for their time and willingness to contribute to this study, as well as to the laboratory assistants who supported data collection efforts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apergo.2026.104866.

References

Abobakr, A., et al., 2019. RGB-D ergonomic assessment system of adopted working postures. Appl. Ergon. 80, 75–88. https://doi.org/10.1016/j.apergo.2019.05.004.

Agostinelli, T., Generosi, A., Ceccacci, S., Mengoni, M., 2024. Validation of computer vision-based ergonomic risk assessment tools for real manufacturing environments. Sci. Rep. 14 (1). https://doi.org/10.1038/s41598-024-79373-4, pp. 27785–19.

Auer, S., Süß, F., Dendorfer, S., 2024. Using markerless motion capture and musculoskeletal models: an evaluation of joint kinematics. Technol. Health Care 32 (5), 3433–3442. https://doi.org/10.3233/thc-240202.

Baldinger, M., Reimer, L.M., Senner, V., 2025. Influence of the camera viewing angle on OpenPose validity in motion analysis. Sensors 25 (3), 799. https://doi.org/10.3390/s25030799.

Balogh, D., Cui, X., Mayer, M., Koehncke, N., Dueck, R., Lang, A.E., 2025. Reliability and agreement during the rapid entire body assessment: comparing rater expertise and artificial intelligence. PLoS One 20 (5), e0323262. https://doi.org/10.1371/journal.pone.0323262.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., Grundmann, M., 2020. BlazePose: on-device real-time body pose tracking. arXiv preprint, 2020/06/17.

Bonakdar, A., et al., 2025. Validation of markerless vision-based motion capture for ergonomics risk assessment. Int. J. Ind. Ergon. 107, 103734. https://doi.org/10.1016/j.ergon.2025.103734.

Ceriola, L., Taborri, J., Donati, M., Rossi, S., Patanè, F., Mileti, I., 2024. Comparative analysis of markerless motion capture systems for measuring human kinematics. IEEE Sens. J. 24 (17), 28135–28144. https://doi.org/10.1109/jsen.2024.3431873.

Cheng, X., Jiao, Y., Meiring, R.M., Sheng, B., Zhang, Y., 2025. Reliability and validity of current computer vision based motion capture systems in gait analysis: a systematic review. Gait Posture 120, 150–160. https://doi.org/10.1016/j.gaitpost.2025.04.016.

Colombini, D., Occhipinti, E., 2006. Preventing upper limb work-related musculoskeletal disorders (UL-WMSDS): new approaches in job (re)design and current trends in standardization. Appl. Ergon. 37 (4), 441–450. https://doi.org/10.1016/j.apergo.2006.04.008.

Crenna, F., Rossi, G.B., Berardengo, M., 2021. Filtering biomechanical signals in movement analysis. Sensors 21 (13), 4580 [Online]. Available: https://www.mdpi.com/1424-8220/21/13/4580.

Cronin, N.J., et al., 2024. Feasibility of OpenPose markerless motion analysis in a real athletics competition. Front. Sports Act. Living 5. https://doi.org/10.3389/fspor.2023.1298003.

D'Souza, S., Siebert, T., Fohanno, V., 2024. A comparison of lower body gait kinematics and kinetics between Theia3D markerless and marker-based models in healthy subjects and clinical patients. Sci. Rep. 14 (1). https://doi.org/10.1038/s41598-024-80499-8.

de Andrade, A.G.P., Carpes, F.P., 2024. A new approach to quantify asymmetry in human movement. J. Biomech. 170, 112158. https://doi.org/10.1016/j.jbiomech.2024.112158.

Delleman, N., Boocock, M., Kapitaniak, B., Schaefer, P., Schaub, K., 2000. ISO/FDIS 11226: evaluation of static working postures. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 44 (35). https://doi.org/10.1177/154193120004403512, pp. 6–442-6-443.

Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G., 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans. Biomed. Eng. 54 (11), 1940–1950. https://doi.org/10.1109/TBME.2007.901024. https://simtk.org/.

Geto, A.K., Daba, C., Desye, B., Berihun, G., Aweke, L., 2025. Prevalence of work-related musculoskeletal disorder and its associated factors among weavers in low-and middle-income countries: a systematic review and meta-analysis. BMJ Open 15, 93124. https://doi.org/10.1136/bmjopen-2024-093124.

Govaerts, R., et al., 2021. Prevalence and incidence of work-related musculoskeletal disorders in secondary industries of 21st century Europe: a systematic review and meta-analysis. BMC Muscoskelet. Disord. 22 (1), 751. https://doi.org/10.1186/s12891-021-04615-9.

Guo, L., Xiong, S., 2017. Accuracy of base of support using an inertial sensor based motion capture system. Sensors (Basel, Switzerland) 17 (9), 2091. https://doi.org/10.3390/s17092091.

Hignett, S., McAtamney, L., 2000. Rapid entire body assessment (REBA). Appl. Ergon. 31 (2), 201–205. https://doi.org/10.1016/S0003-6870(99)00039-3.

Hsu, G.-S.J., Wu, J.S., Huang, Y.-K.D., Chiu, C.-C., Kang, J.-H., 2025. Automatic detect incorrect lifting posture with the pose estimation model. Life 15 (3), 358. https://doi.org/10.3390/life15030358.

International Organization for, S., 2021. ISO 11228-1:2021 Ergonomics — Manual Handling — Part 1: Lifting, Lowering and Carrying. International Organization for Standardization, Geneva, Switzerland. ISO 11228-1:2021, [Online]. Available: https://www.iso.org/standard/76820.html.

Jamali, P., Chou, L.-S., Catena, R.D., 2025. Whole-cycle and time-specific validation of a GUI-based ground reaction force estimation tool for clinical gait analysis without a force plate. Med. Eng. Phys. 141, 104366. https://doi.org/10.1016/j.medengphy.2025.104366.

Karhu, O., Kansi, P., Kuorinka, I., 1977. Correcting working postures in industry: a practical method for analysis. Appl. Ergon. 8 (4), 199–201. https://doi.org/10.1016/0003-6870(77)90164-8.

Kim, J.-W., Choi, J.-Y., Ha, E.-J., Choi, J.-H., 2023. Human pose estimation using MediaPipe pose and optimization method based on a humanoid model. Applied sciences 13 (4), 2700. https://doi.org/10.3390/app13042700.

Kim, T., Jung, M.-C., Mo, S.-M., 2025. Suggestion for camera location in monocular markerless 3D motion capture system: focused on accuracy comparison with marker-based system for upper limb joints. IEEE Access 13, 47605–47616. https://doi.org/10.1109/access.2025.3547715.

Kim, W., Sung, J., Saakes, D., Huang, C., Xiong, S., 2021. Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose). Int. J. Ind. Ergon. 84. https://doi.org/10.1016/j.ergon.2021.103164.

Leardini, A., Chiari, L., Della Croce, U., Cappozzo, A., 2005. Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation. Gait Posture 21 (2), 212. https://doi.org/10.1016/j.gaitpost.2004.05.002.

Lin, P.-C., Chen, Y.-J., Chen, W.-S., Lee, Y.-J., 2022. Automatic real-time occupational posture evaluation and select corresponding ergonomic assessments. Sci. Rep. 12 (1). https://doi.org/10.1038/s41598-022-05812-9.

Lowe, B.D., Dempsey, P.G., Jones, E.M., 2019. Ergonomics assessment methods used by ergonomics professionals. Appl. Ergon. 81. https://doi.org/10.1016/j.apergo.2019.102882, 102882-102882.

Lugaresi, C., et al., 2019. Mediapipe: a Framework for Building Perception Pipelines. https://doi.org/10.48550/arxiv.1906.08172.

Malý, P., Lopot, F., 2013. QUALISYS system applied to industrial testing. Appl. Mech. Mater. 486 (Experimental Stress Analysis 51), 135–140. https://doi.org/10.4028/www.scientific.net/AMM.486.135.

Marras, W.S., et al., 1993. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders. Spine 18 (5), 617–628. Philadelphia, Pa. 1976.

Marras, W.S., Parakkat, J., Chany, A.M., Yang, G., Burr, D., Lavender, S.A., 2006. Spine loading as a function of lift frequency, exposure duration, and work experience. Clin. Biomech. 21 (4), 345–352. https://doi.org/10.1016/j.clinbiomech.2005.10.004.

McAtamney, L., Nigel Corlett, E., 1993. RULA: a survey method for the investigation of work-related upper limb disorders. Appl. Ergon. 24 (2), 91–99. https://doi.org/10.1016/0003-6870(93)90080-S.

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 (3), 360–369. https://doi.org/10.1016/j.gaitpost.2008.09.003.

McGuirk, T.E., Perry, E.S., Sihanath, W.B., Riazati, S., Patten, C., 2022. Feasibility of markerless motion capture for three-dimensional gait assessment in community settings. Front. Hum. Neurosci. 16. https://doi.org/10.3389/fnhum.2022.867485.

Mehrizi, R., Peng, X., Tang, Z., Xu, X., Metaxas, D., Li, K., 2018b. Toward marker-free 3D pose estimation in lifting: a deep multi-view solution. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pp. 485–491. https://doi.org/10.1109/FG.2018.00078.

Mehrizi, R., Peng, X., Xu, X., Zhang, S., Metaxas, D., Li, K., 2018a. A computer vision based method for 3D posture estimation of symmetrical lifting. J. Biomech. 69, 40–46. https://doi.org/10.1016/j.jbiomech.2018.01.012.

Menolotto, M., Komaris, D.-S., Tedesco, S., O'Flynn, B., Walsh, M., 2020. Motion capture technology in industrial applications: a systematic review. Sensors 20 (19), 5687. https://doi.org/10.3390/s20195687.

Merriaux, P., Dupuis, Y., Boutteau, R., Vasseur, P., Savatier, X., 2017. A study of vicon system positioning performance. Sensors (Basel, Switzerland) 17 (7), 1591. https://doi.org/10.3390/s17071591.

Needham, L., et al., 2021. The accuracy of several pose estimation methods for 3D joint centre localisation. Sci. Rep. 11 (1). https://doi.org/10.1038/s41598-021-00212-x.

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.

Pilla-Barroso, M., Ruiz-Ruiz, L., Jiménez-Ruiz, A.R., Seco, F., Jiménez-Martín, A., 2025. In: Assessing the Impact of Depth Data on Human Kinematics: a Hybrid RGB-D and Mediapipe Method. IEEE, pp. 1–6. https://doi.org/10.1109/memea65319.2025.11068070, 10.1109/memea65319.2025.11068070.

Rajagopal, A., Dembia, C.L., DeMers, M.S., Delp, D.D., Hicks, J.L., Delp, S.L., 2016. Full-body musculoskeletal model for muscle-driven simulation of human gait. IEEE Trans. Biomed. Eng. 63 (10), 2068–2079. https://doi.org/10.1109/TBME.2016.2586891.

Roggio, F., Trovato, B., Sortino, M., Musumeci, G., 2024. A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: a narrative review. Heliyon 10 (21), e39977. https://doi.org/10.1016/j.heliyon.2024.e39977.

Rosado, A.S., Baptista, J.S., Guilherme, M.N.H., Guedes, J.C., 2023. Economic impact of work-related musculoskeletal disorders—A systematic review. Studies in Systems, Decision and Control). Springer International Publishing, Cham, pp. 599–613.

Salehi, M., Taheri, A., Choi, S., Kim, J.H., 2026a. Evaluation of a markerless motion capture to measure 3D joint kinematics during occupational lifting tasks using mobile devices. Appl. Ergon. 134, 104743. https://doi.org/10.1016/j.apergo.2026.104743.

Salehi, M., Taheri, A., Kim, J.H., 2026b. Estimation of dynamic spinal loads during manual lifting using smartphone-based markerless motion capture. Appl. Ergon. 137, 104823. https://doi.org/10.1016/j.apergo.2026.104823.

Salisu, S., Ruhaiyem, N.I.R., Eisa, T.A.E., Nasser, M., Saeed, F., Younis, H.A., 2023. Motion capture technologies for ergonomics: a systematic literature review. Diagnostics 13 (15), 2593. https://doi.org/10.3390/diagnostics13152593.

Scataglini, S., Abts, E., Van Bocxlaer, C., Van Den Bussche, M., Meletani, S., Truijen, S., 2024. 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 (11), 3686. https://doi.org/10.3390/s24113686.

Schurr, S.A., Marshall, A.N., Resch, J.E., Saliba, S.A., 2017. TWO-DIMENSIONAL video analysIS IS comparable to 3D motion capture in lower extremity movement assessment. Int. J. Sports Phys. Ther. 12 (2), 163–172.

Svetek, A., Morgan, K., Burland, J., Glaviano, N.R., 2025. Validation of OpenCap on lower extremity kinematics during functional tasks. J. Biomech. 183, 112602. https://doi.org/10.1016/j.jbiomech.2025.112602.

Takala, E.-P., et al., 2010. Systematic evaluation of observational methods assessing biomechanical exposures at work. Scand. J. Work. Environ. Health 36 (1), 3–24. https://doi.org/10.5271/sjweh.2876.

O. Team. "Opencap: Open-Source Platform for Human Movement Analysis." OpenCap Inc. https://www.opencap.ai/(accessed January/22, 2025).

Teufl, W., Miezal, M., Taetz, B., Fröhlich, M., Bleser, G., 2019. Validity of inertial sensor based 3D joint kinematics of static and dynamic sport and physiotherapy specific movements. PLoS One 14 (2), e0213064. https://doi.org/10.1371/journal.pone.0213064.

Turner, J.A., Chaaban, C.R., Padua, D.A., 2024. Validation of OpenCap: a low-cost markerless motion capture system for lower-extremity kinematics during return-to-sport tasks. J. Biomech. 171, 112200. https://doi.org/10.1016/j.jbiomech.2024.112200.

Uhlrich, S.D., et al., 2023. OpenCap: human movement dynamics from smartphone videos. PLoS Comput. Biol. 19 (10). https://doi.org/10.1371/journal.pcbi.1011462 e1011462-e1011462.

Varcin, F., Boocock, M., 2026. The accuracy, validity and reliability of Theia3D markerless motion capture for studying the biomechanics of human movement: a systematic review. Artif. Intell. Med. 173. https://doi.org/10.1016/j.artmed.2025.103332, 103332-103332.

S. Vicon Motion. "Plug-in Gait Kinematic Variables." Vicon Motion Systems Ltd. https://help.vicon.com/space/Nexus216/11611411/Plug-in+Gait+kinematic+variables (accessed January/22, 2025).

Wang, H., Xie, Z., Lu, L., Li, L., Xu, X., 2021. A computer-vision method to estimate joint angles and L5/S1 moments during lifting tasks through a single camera. J. Biomech. 129, 110860. https://doi.org/10.1016/j.jbiomech.2021.110860.

Waters, T.R., Putz-Anderson, V., Garg, A., Fine, L.J., 1993. Revised NIOSH equation for the design and evaluation of manual lifting tasks. Ergonomics 36 (7), 749–776. https://doi.org/10.1080/00140139308967940.

Williams, J.M., Frey, M., Breen, A., De Carvalho, D., 2022. Systematic analysis of different low-pass filter cut-off frequencies on lumbar spine kinematics data and the impact on the agreement between accelerometers and an optoelectronic system. J. Biomech. 145, 111395. https://doi.org/10.1016/j.jbiomech.2022.111395.

World medical association declaration of helsinki. JAMA 310 (20), 2013, 2191. https://doi.org/10.1001/jama.2013.281053.

Wu, G., et al., 2005. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion–Part II: shoulder, elbow, wrist and hand. J. Biomech. 38 (5), 981–992. https://doi.org/10.1016/j.jbiomech.2004.05.042.

Yang, Z., Song, D., Ning, J., Wu, Z., 2024. A systematic review: advancing ergonomic posture risk assessment through the integration of computer vision and machine learning techniques. IEEE Access 12, 180481–180519. https://doi.org/10.1109/access.2024.3509447.

Yeom, S., et al., 2025. Validation of markerless motion capture system (OpenCap) for overhead squat kinematics. Meas. Phys. Educ. Exerc. Sci. 1–14. https://doi.org/10.1080/1091367x.2025.2512361.

Yoma, M., Llurda-Almuzara, L., Herrington, L., Jones, R., 2025. Reliability and validity of lower extremity and trunk kinematics measured with markerless motion capture during sports-related and functional tasks: a systematic review. J. Sports Sci. 43 (17), 1703–1730. https://doi.org/10.1080/02640414.2025.2518359.

Zayyinul Hayati Zen, M.W., Hani Sukadarin, Ezrin, 2024. Ergonomics risk assessment tools: a systematic review. Malaysian Journal of Medicine and Health Sciences 20 (5), 289–300. https://doi.org/10.47836/mjmhs20.5.35.

Zheng, C., Zhu, S., Mendieta, M., Yang, T., Chen, C., Ding, Z., 2021. 3D human pose estimation with spatial and temporal transformers. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 11636–11645. https://doi.org/10.1109/ICCV48922.2021.01145, 10-17.