Review

# Monocular Two-Dimensional and 3D Red-Green-Blue-Camera Marker-Less Motion Capture Systems for Human Lower-Limb Activities: Scoping Review

Behnam Gholami[1], BSc, MSc; David Hollinger[2], PhD; Xiaojun Zhai[2], PhD; Bernard Xian Wei Liew[1], PhD

[1]School of Sport, Rehabilitation and Exercise Sciences (SRES), University of Essex, Essex, England, United Kingdom
[2]School of Computer Science and Electronic Engineering (CSEE), University of Essex, Colchester, England, United Kingdom

**Corresponding Author:**

Behnam Gholami, BSc, MSc
School of Sport, Rehabilitation and Exercise Sciences (SRES), University of Essex
Wivenhoe Park, Colchester
Essex, England CO4 3SQ
United Kingdom
Phone: 44 7464273268
Email: bg24306@essex.ac.uk

## Abstract

**Background:** Single-camera markerless motion capture systems (MLMCS) offer a low-cost, portable alternative to marker-based motion capture systems (MBMCS) for lower-limb assessment, but the evidence remains fragmented across movement tasks, algorithms, and validation frameworks. Existing reviews have addressed broader health care applications, disease-specific gait technologies, or general pose-estimation use cases, but a focused synthesis of monocular two-dimensional (2D) and 3D red-green-blue (RGB)-camera MLMCS for human lower-limb activities has been lacking.

**Objective:** This scoping review aimed to map how single RGB-camera MLMCS have been used to quantify lower-limb kinematics by examining data collection protocols, participant/task characteristics, computational pipelines, human pose estimation (HPE) methods, and reference standards and validation strategies.

**Methods:** Following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidance, we searched PubMed, Scopus, IEEE Xplore, CINAHL, and SPORTDiscus from database inception to March 16, 2026. Peer-reviewed English-language studies of human lower-limb activities using a single RGB camera were eligible. Two reviewers (BG and DH) independently screened studies and appraised quality using a modified Downs and Black checklist. Findings were synthesized descriptively.

**Results:** Out of a total of 4484 identified records, 39 studies were included. The evidence base was dominated by gait-related tasks (29/39, 75%) and predominantly healthy cohorts, with knee flexion/extension and hip flexion/extension as the most frequently reported outcomes. Twenty-five distinct MLMCS methods were identified. OpenPose (Perceptual Computing Lab) was the most frequently used 2D approach. Across studies, performance was most consistent for sagittal-plane hip and knee flexion/extension, whereas ankle plantarflexion/dorsiflexion, frontal- and transverse-plane kinematics, and occluded-side estimation remained less robust and less consistently validated. Unlike broader previous reviews, this review provides the first focused evidence map of monocular RGB lower-limb applications across tasks, populations, hardware/software configurations, HPE algorithms, and benchmarking strategies. It also offers a clearer performance hierarchy, with OpenPose emerging as the strongest 2D option for sagittal knee flexion/extension and the World-grounded Humans with Accurate Motion (WHAM) method as the best-performing single-view 3D system overall in the largest benchmark.

**Conclusions:** This review advances the field by clarifying where single RGB-camera MLMCS are currently most credible and where major limitations remain. In real-world settings, these systems show promise as accessible tools for repeated movement assessment, screening, and longitudinal monitoring when the primary target is sagittal-plane hip and knee motion. However, broader clinical deployment will require more standardized protocols, richer ankle and foot keypoint definitions, pathology-specific validation, and greater robustness to occlusion, viewpoint variation, and real-world recording conditions.

***J Med Internet Res 2026;28:e91609;*** doi: 10.2196/91609

**Keywords:** Markerless; single camera; human locomotion; lower limbs; machine learning

## Introduction

Optical marker-based systems, while offering high accuracy in controlled laboratory environments [1], are often constrained by extensive setup times [2], high operational costs [3], and the potential to alter natural movement due to the placement of physical markers [4]. These limitations have spurred a growing interest in markerless motion capture systems (MLMCS) approaches that leverage monocular cameras paired with advanced computer vision (CV) algorithms to extract joint positions directly from video data [5], including both monocular Red-Green-Blue (RGB) and Red-Green-Blue-Depth (RGB-D) camera systems. Although RGB-D cameras have reported promising performance in adjacent gait-analysis applications [6,7], their susceptibility to ambient illumination and reflective surfaces [8], and depth-range constraints [9] may limit performance. As a result, monocular RGB cameras remain an attractive option where portability, robustness, and ease of deployment are critical. In particular, monocular RGB-camera-based pipelines face self-occlusion ambiguity that inflates joint-center and angle errors compared with multicamera configurations [10], underscoring the need to map heterogeneity across tasks, algorithms, and validation protocols [11,12].

Movement selection used to train and evaluate monocular systems strongly influences generalizability and error patterns. For example, slower, planar tasks typically yield smaller errors than fast, multiplanar actions [10,13]. Across mixed tasks, monocular 3D joint-center errors are commonly 40-60 mm [13], whereas multicamera reconstructions report 10-30 mm for walking and 20-40 mm for faster actions such as jumping or throwing, indicating a task- and view-count dependence of accuracy [10,13]. Yang and Park showed that 45°-anterolateral camera placement during treadmill walking reduced hip and knee flexion/extension mean absolute errors (MAE) to 2.35°–3.06° relative to a marker-based motion capture system (MBMCS) [14]. Similarly, although smartphone-based OpenCap improves accessibility, concurrent validation against MBMCS showed moderate kinematic fidelity across healthy and pathological gait, with an overall root-mean-square error (RMSE) of 5.8° [15]. Recent comparisons between the Pixmotion markerless system (TsingVA Technology) and Nokov MBMCS (Nokov Motion Capture System) suggest closer agreement for lower-limb flexion/extension during gait (root-mean-square deviation [RMSD] 5.3°-6.8°) than for internal/external rotation (RMSD 8.5°-11.1°) [16]. In high-intensity jumping, Theia3D™ (Theia Markerless Inc) showed sagittal-plane hip flexion/extension RMSD up to 5.6° versus MBMCS, but accuracy at the hip was more prone to error than at the knee or ankle [17].

Differences between algorithms and processing pipelines directly affect accuracy in single-camera settings [10]. Reviews comparing open-source human pose estimation (HPE) note that monocular 3D lifting HPE tends to accumulate depth errors of 4060 mm at joint centers compared to MBMCS unless retrained on task-specific data, which can shrink errors to 30-40 mm [10]. The Human Motion Model for Robust Pose Estimation (HuMoR) algorithm shows reduced accuracy for lumbar flexion–extension, with MAE exceeding 10°, against MBMCS, highlighting the dependency of deep learning models on the diversity of their training datasets [18]. While the OpenPose algorithm demonstrates high fidelity for proximal sagittal kinematics—achieving flexion/extension errors less than 2.0° relative to inertial sensors- its performance degrades significantly at the distal extremities [19]. They reported consistent loss of ankle key points during treadmill gait, attributing this to the system's inability to maintain tracking fidelity under frame-to-frame occlusion [19]. Reported errors in joint localization span approximately 30-60 mm, with joint angle estimation discrepancies reaching up to 30° in some cases compared to MBMCS [10,20,21]. The 3D depth prediction algorithm, acting as a 3D lifter, has achieved localization average errors as low as 40.61 mm across 21 joints of the skeleton model when validated against the Human3.6M ground-truth dataset [22]. Across reviews and comparative studies, sagittal-plane accuracy is consistently stronger than frontal/transverse planes, with hip and pelvis parameters more error-prone than knee and ankle, reflecting the combined effects of algorithmic design [10], training data [16], and the camera viewing geometry inherent to single-camera setups [17].

Validation choices and ground-truth standards also shape reported validity and hinder cross-study comparison, as studies variably use multicamera MBMCS or inertial sensors [10]. Multicamera MLMCS using OpenPose generally achieves 3D joint position errors under 30 mm relative to MBMCS, but up to approximately 10% of data points exceed 40 mm error due to two-dimensional (2D) detection artifacts, underscoring the dependence of 3D accuracy on 2D detection quality [13]. While Molteni and Andreoni [23] validated the capacity of the OpenPose library to replicate optoelectronic spatiotemporal data, their kinematic assessment prioritized discrete metrics over continuous trajectory deviations [23]. By reporting hip and knee flexion/extension errors, which remained less than 6.0°, rather than waveform errors, the study illustrates the persistent heterogeneity in MLMCS validation protocols [23]. Although OpenPose and MediaPipe (Google LLC) offer rapid gait analysis (processing 17 trials in under 3 minutes) compared to the labor-intensive Kinovea workflow (2-2.5 hours per cycle), their utility is constrained by specific anatomical failures [24]. The authors highlighted persistent ankle key point instability and hip flexion/extension biases when benchmarked against MBMCS, underscoring the trade-off between computational speed and biomechanical precision [24]. More broadly, reviews emphasize that without validation against more anatomically valid ground truths such as biplanar video radiography, the true accuracy of joint centers and rotations remains uncertain, and inconsistencies in training labels for open-source HPEs further complicate interstudy comparisons [10].

While previous reviews have investigated single-camera MLMCS in health care applications [25], disease-specific gait technologies [26], or pose estimation in broader health and performance contexts [27], a focused synthesis of single RGB-camera MLMCS for human lower-limb activities is lacking. The current review addresses this gap by specifically mapping monocular RGB-based lower-limb studies across movement tasks, participant groups, computational pipelines, HPE algorithms, and validation approaches. A preliminary search of the Cochrane Database of Systematic Reviews, Joanna Briggs Institute (JBI) Evidence Synthesis, and Open Science Framework did not identify any current or underway systematic or scoping reviews on the proposed topic in this study. To achieve the aims of the current study, the review is guided by the following research questions (RQs):

- RQ1. What methods have been used for data collection protocols, the tasks and activities performed, and the participants' conditions and ages?
- RQ2. What methods have been used for data processing, including the computational configurations, such as the software and/or hardware used, and the AI algorithms used?
- RQ3. What gold standard methods are used to compare and validate the outcomes of the AI algorithms?

Together, these questions respectively target (1) the source and characteristics of the movement data and participants, (2) the computational tools and CV algorithms used, and (3) the validation approaches and reference standards. Collectively, they provide an overview of technologies for monocular MLMCS in lower-limb activities and their implications for improving single-camera accuracy. This may help identify gaps to inform and guide future studies.

# Methods

## Study Design

This review was conducted in accordance with the latest edition of the JBI Manual for Evidence Synthesis [28,29], and the methodological framework developed by Arksey and O'Malley [30] and refined by Levac et al [31]. The framework comprises 5 stages [30]: (1) identifying the research question, (2) identifying relevant studies, (3) selecting studies, (4) extracting and charting the data, and (5) collating, summarizing, and reporting the findings. The Population (or Participants), Concept, and Context framework was used to define the review question. Reporting followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [32], and the completed checklist is provided in Checklist 1.

## Eligibility Criteria

The inclusion criteria applied in this scoping review were as follows:

1. Population: Human participants of any age, sex or gender, or health status.
2. Concept: Studies that investigate the biomechanical capture of human lower-limb movements, specifically involving the hip, knee, or ankle, using an RGB single camera, without the need for physical markers to drive the primary tracking. Eligible studies that reported either joint angles or displacements across any anatomical plane (frontal, sagittal, or transverse) were considered. Studies examining other lower-limb activities were included only if they provided angle or displacement data for at least one of the key joints-hip, knee, or ankle-as part of a broader analysis or mixed methods study. Studies using multi-RGB-camera setups or MBMCS were included only if they served as a reference for validating a markerless single-camera approach, or if the data processing pipeline treated each camera view independently as a monocular input without multiview 3D reconstruction fusion.
3. Context: The proposed scoping review included studies in any context to explore all methods that are being used for single-camera MLMCS in human lower-limb activities.
4. Publication: Full-text, peer-reviewed original research articles published in English in journals. Qualitative studies were eligible only when included as part of a mixed methods study.

Studies were excluded if they involved nonhuman populations, did not assess lower-limb activities, or were unpublished studies, conference papers, gray literature, books, or reviews. Studies were also excluded if they used only multi-RGB-camera or MBMCS without a monocular markerless analysis, used underwater, aerial, egocentric, or surveillance perspectives, or relied solely on non-RGB modalities such as depth or infrared cameras.

## Search Strategy

An initial limited search of Scopus (Elsevier) and PubMed was conducted to identify potentially relevant articles and inform the development of the final search strategy. The text words found in the titles and abstracts of relevant articles, as well as the subject headings, were used to develop a comprehensive search strategy for PubMed. The search strategy was modified for each distinct database (Multimedia Appendix 1). The search strategy was developed by the review team with subject-matter input from XZ and BXWL, and no medical librarian or information retrieval specialist was involved. Articles were extracted from PubMed, Scopus (Elsevier), IEEE Xplore, CINAHL (EBSCO), and SPORTDiscus (EBSCO). Additionally, the reference list of the final included articles was screened for additional studies. Retrieved records and references were managed using Mendeley reference management software (Elsevier). No time restrictions were applied to the start date of the evidence search. The initial search was conducted on November 25, 2024, and the final search was completed on March 16, 2026, to ensure the inclusion of newly published articles.

## Study Selection

Titles and abstracts were screened independently by 2 reviewers (BG and DH), with a third (BXWL) consulted as needed to resolve disagreements at both screening stages. Regular meetings were held between reviewers to maintain consistency in the screening process and to resolve any

disagreements through mutual discussion and agreement. When required, a third reviewer (BXWL) was involved to assist in resolving disagreements and achieving consensus during this stage of the review process. Full-text screening was conducted independently by 2 members of the review team (BG and DH). Interrater agreement during full-text screening was almost perfect (Cohen $\kappa$=0.834). Discrepancies were mainly related to differing interpretations of single-view analysis in multicamera studies and were resolved through discussion, with input from a third reviewer (BXWL) when needed. The agreed interpretation was then applied consistently across the remaining studies. The entire screening process, including title, abstract, and full-text review, was conducted using the Rayyan platform (Rayyan Platforms Inc) [33].

## Data Charting

A structured data-charting form was developed in Microsoft Word with reference to the JBI data-extraction instrument for scoping reviews [29] and adapted to the objectives of the current review. One reviewer (BG) charted the data from all included studies [34-72], and a second reviewer (DH) verified every entry against the corresponding full-text article and any accompanying supplementary materials. Disagreements or uncertainties concerning the interpretation or classification of the reported information were resolved through discussion until consensus was reached. The charting form captured study characteristics, participant data sources classified as primary, secondary, or combined primary and secondary data, participant demographics and health status, movement tasks and testing conditions, camera hardware and positioning, MLMCS methods and computational configurations, reference standards and validation procedures, and reported outcomes. Where available, joint-angle and displacement outcomes and their associated statistical measures were charted according to the joint, anatomical plane, body side, movement condition, markerless method, and reference system. When a study included multiple participant cohorts, datasets, movement tasks, camera viewpoints, or markerless methods, each relevant component was recorded separately within the corresponding study record. Information that was not reported or not applicable was recorded accordingly.

## Quality Assessment

Two authors (BG and DSH) independently evaluated the quality of the included studies [34-72] using a modified Downs and Black checklist [73]. To align with our aims, we evaluated 14 of the original 27 items, selected with reference to prior technology-focused systematic reviews [74-76]. Included items used in this study are shown in Table 1. Item 5 was scored on a three-point scale (0=no, 1=partially, 2=yes), while items 11, 15, and 16 were scored as 0 (no), 1 (yes), or Undetermined. The remaining items were scored as 0 (no) or 1 (yes). The maximum possible score to be obtained was 15. To assess internal consistency of the quality appraisal process, interrater reliability was calculated between the 2 reviewers' independent item ratings. Of the 14 items analyzed, agreement levels of interrater reliability (Cohen kappa [$\kappa$] [77]) ranged from poor to perfect. Five questions (Q1, Q2, Q6, Q13, and Q15) showed perfect agreement ($\kappa$=1.000), almost perfect agreement for Q10 ($\kappa$=0.896), substantial agreement for Q3 and Q7 ($\kappa$=0.776-0.787), moderate agreement for Q27 ($\kappa$=0.544), fair agreement for Q5 and Q16 ($\kappa$=0.330-0.381), and slight agreement for Q11, Q18, and Q20 ($\kappa$=0.000-0.187). Discrepancies stemmed mainly from differing interpretations of confounders and equipment validity. These were resolved through discussion to reach consensus, after which the agreed interpretation was applied to all remaining studies.

**Table 1.** Quality evaluation items used in this study based on a modified 14-item Downs and Black checklist [73].

| Item | Description |
|---|---|
| Q1 | Is the hypothesis/aim/objective of the study clearly described? |
| Q2 | Are the main outcomes to be measured clearly described in the introduction or methods section? |
| Q3 | Are the characteristics of the subjects included in the study clearly described? |
| Q5 | Are the distributions of principal confounders in each group clearly described? |
| Q6 | Are the main findings of the study clearly described? |
| Q7 | Does the study provide estimates of the random variability in the data for the main outcomes? |
| Q10 | Have actual probability values been reported for the main outcomes (except when <0.001)? |
| Q11 | Were the subjects asked to participate representative of the entire population from which they were recruited? |
| Q13 | Were the staff, places, and facilities where the patients were treated representative of those most patients receive? |
| Q15 | Was an attempt made to blind those measuring the main outcomes? |
| Q16 | If any results were based on data-dredging, was this made clear? |
| Q18 | Were the statistical tests used to assess the main outcomes appropriate? |
| Q20 | Were the main outcome measures used accurate (valid and reliable)? |
| Q27 | Did the study have sufficient power to detect a clinically important effect (power/sample-size calculation)? |

## Data Synthesis and Analysis

The charted data were synthesized using descriptive numerical summaries and a narrative approach in accordance with the review objectives and RQs [29,32]. Categorical data were summarized using frequencies and percentages. Where sufficient information was available, continuous descriptive variables, including participant age and sample size, were summarized using medians, IQRs, and ranges.

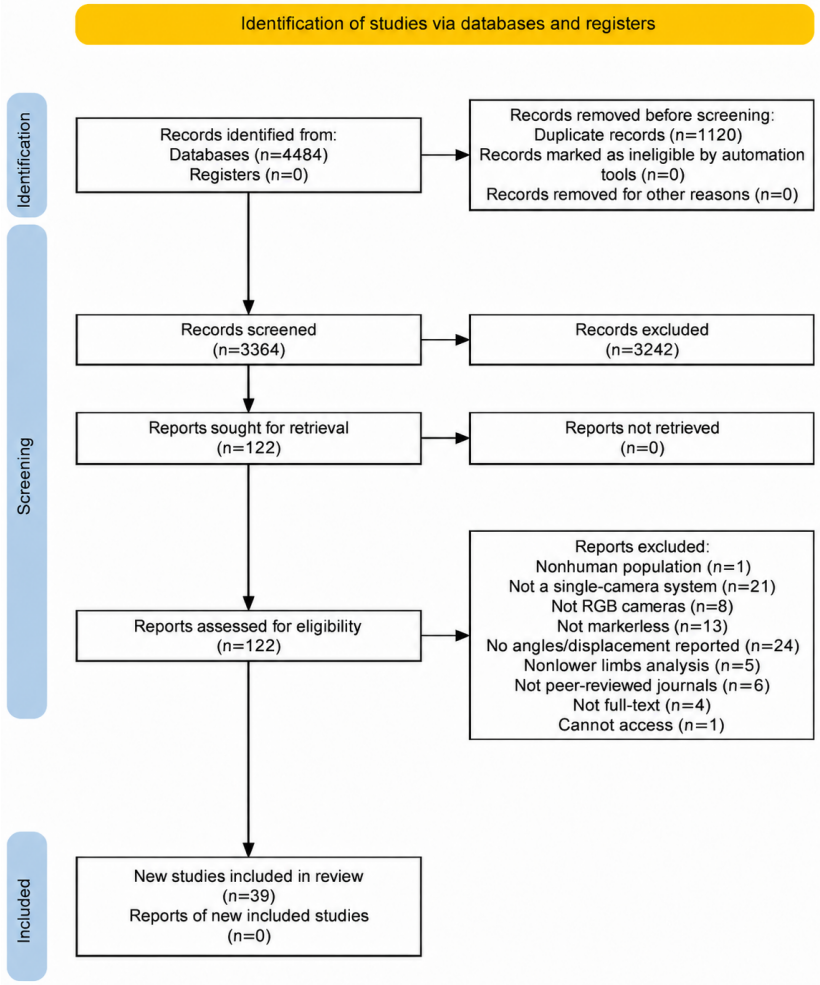
Given the heterogeneity across study populations, movement tasks, MLMCS pipelines, hardware configurations, reference standards, and reported outcome measures, the findings were synthesized descriptively rather than quantitatively. Extracted data were organized according to the review objectives and summarized across key domains, including participant and task characteristics, MLMCS methods and technical specifications, benchmarking and ground-truth approaches, and reported validity and reliability outcomes where appropriate; tables, figures, and Multimedia Appendices were used to summarize the data and facilitate cross-study comparison.

# Results

## Search Results

The search across 5 databases identified 4484 records, with no additional records identified from registers. After removal of 1120 duplicate records, 3364 records remained for title and abstract screening. Of these, 3242 records were excluded, leaving 122 reports for full-text retrieval and eligibility assessment. All 122 reports were successfully retrieved for full-text review. Following full-text assessment, 83 reports were excluded for the predefined reasons shown in Figure 1, and 39 studies [34-72] met the inclusion criteria and were retained for data extraction and synthesis in this scoping review, as presented in Multimedia Appendix 2 [78-83], which provides a full summary of the study and population characteristics of the included studies [34-72]. A summary of the search and study selection process is presented in the PRISMA flow diagram (Figure 1).

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. RGB: red-green-blue.




## Data Source Characteristics

As detailed in Multimedia Appendix 2, primary datasets (ie, data collected directly by the study authors) were used in 72% (28/39) of studies [34,35,37-44,48,49,52,54-64,66,68-70], secondary datasets (ie, preexisting or publicly available datasets) in 15% (6/39) [46,50,51,65,67,72], and a combination of primary and secondary datasets in 13% (5/39) [36,45,47,53,62]. All included studies in these datasets [34,35,39-43,45-49,53-55,58,60-66,68,69] used a cross-sectional design.

## Population Assessed

The included studies [34,35,39-43,45-49,53-55,58,60-66,68,69] examined cohorts spanning pediatric to older adult populations. Among cohorts reporting age, the median reported age was 27.4 years (IQR 23.8-36.0; range 4.9-73.3). Sample sizes varied substantially, with a median of 21 participants (IQR 14-38; range 2-1026). When stratified by data source, the median sample size was 16 participants (IQR 12-31; range 3-254) for primary data sources and 33 participants (IQR 19-79; range 2-1026) for secondary data sources. Among cohorts reporting sex, the overall reported sex distribution was 55.7% (758/1361) male and 44.3% (603/1361) female.

As summarized in Multimedia Appendix 2, 4 studies [36,59,67,72] did not report participant health status or listed it as not applicable. Healthy adults were the most frequently represented population and were included in 25 studies [34,35,39-43,45-49,53-55,58,60-66,68,69]. While most studies did not provide an explicit operational definition of "healthy," those that did generally defined it as being free from restricted physical activity due to injury or disease, neurological or cardiac disorders, or prior lower-extremity surgeries [54,69], or as having no known history of orthopedic injuries or disorders [66]. Five studies investigated athletic cohorts, including general athletic populations [37,56,57], and runners [50,70]. Seven studies examined condition-specific populations, including stroke [38], Parkinson disease [52], mixed cohorts including sarcopenia and Parkinson disease [41], autism spectrum disorder and typically developing children [44], cerebral palsy [51], high-fall-risk adults [60], and lower-limb amputees [71]. Pediatric cohorts were represented in 3 studies [44,51,56].

## Quality of the Evidence

Table 2 summarizes quality appraisal for the 39 studies [34-72]. The mean total score was 8.9/15 (SD 1.6; median 9; range 6-12). Items Q11, Q15, and Q16 were most frequently rated as "U=Undetermined" (Q11: 39/39; Q15: 38/39; Q16: 35/39). Power or sample-size justification was uncommon, with 33/39 studies [34,36,37,39-48,51-57,59,61-72] scored "No" on Q27. Items Q1, Q2, Q6, and Q13 were rated "Yes" in all 39 studies (39/39) [34-72], and Q3 was rated "Yes" in 36/39 studies [34,35,37-58,60-70,72]. For confounders (Q5), 7/39 studies [36,39,40,57,59,63,68] were fully reported, 22/39 [34,35,37,38,41,44,46,49-51,53,55,58,61,62,64,66,67,69-72] were partially reported, and 10/39 [42,43,45,47,48,52,54,56,60,65] were not reported.

**Table 2.** Risk of bias appraisal for the included studies [34-72] according to a modified version of the Downs and Black scale[a].

| First author, year [reference] | Reporting | | | | | | | External validity | | Internal validity - bias | | | | Power | Total score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q5 | Q6 | Q7 | Q10 | Q11 | Q13 | Q15 | Q16 | Q18 | Q20 | Q27 | |
| Akturk et al 2026 [34] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Aleksic et al 2024 [35] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 1 | 11 |
| Aoyagi et al 2022 [36] | 1 | 1 | 0 | 2 | 1 | 0 | 0 | U | 1 | U | 0 | 0 | 0 | 0 | 6 |
| Babouras et al 2024 [37] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Barzyk et al 2024 [38] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 1 | 11 |
| Barzyk et al 2024 [39] | 1 | 1 | 1 | 2 | 1 | 1 | 0 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Castelli et al 2015 [40] | 1 | 1 | 1 | 2 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 8 |
| Chao et al 2026 [41] | 1 | 1 | 1 | 1 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 7 |
| Dunn et al 2023 [42] | 1 | 1 | 1 | 0 | 1 | 1 | 0 | U | 1 | U | U | 1 | 1 | 0 | 8 |
| Galasso et al 2024 [43] | 1 | 1 | 1 | 0 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 6 |
| Ganai et al 2025 [44] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Goffredo et al 2010 [45] | 1 | 1 | 1 | 0 | 1 | 1 | 0 | U | 1 | U | U | 0 | 1 | 0 | 7 |
| Horsak et al 2025 [46] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Hsu et al 2025 [47] | 1 | 1 | 1 | 0 | 1 | 1 | 1 | U | 1 | U | U | 1 | 0 | 0 | 8 |
| Hulleck et al 2024 [48] | 1 | 1 | 1 | 0 | 1 | 0 | 1 | U | 1 | U | U | 1 | 1 | 0 | 8 |
| Ino et al 2023 [49] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 1 | 11 |
| Johnson et al 2022 [50] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | 1 | U | 1 | 1 | 1 | 12 |
| Kidziński et al 2020 [51] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Kim et al 2025 [52] | 1 | 1 | 1 | 0 | 1 | 1 | 0 | U | 1 | U | U | 1 | 1 | 0 | 8 |
| Li et al 2025 [53] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | 1 | 1 | 1 | 0 | 11 |
| Li et al 2024 [54] | 1 | 1 | 1 | 0 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 6 |
| Liang et al 2022 [55] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Ljubičić 2024 [56] | 1 | 1 | 1 | 0 | 1 | 1 | 1 | U | 1 | U | U | 1 | 0 | 0 | 8 |
| Mercadal-Baudart et al 2024 [57] | 1 | 1 | 1 | 2 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 8 |

| First author, year [reference] | Reporting | | | | | | | External validity | | Internal validity - bias | | | | Power | Total score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q5 | Q6 | Q7 | Q10 | Q11 | Q13 | Q15 | Q16 | Q18 | Q20 | Q27 | |
| Park et al 2026 [58] | 1 | 1 | 1 | 1 | 1 | 1 | 0 | U | 1 | U | U | 1 | 1 | 1 | 10 |
| Saboune and Charpillet 2007 [59] | 1 | 1 | 0 | 2 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 7 |
| Sakane et al 2025 [60] | 1 | 1 | 1 | 0 | 1 | 1 | 1 | U | 1 | U | 0 | 1 | 0 | 1 | 9 |
| Surer et al 2011 [61] | 1 | 1 | 1 | 1 | 1 | 1 | 0 | U | 1 | U | U | 0 | 1 | 0 | 8 |
| Usami et al 2025 [62] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Viswakumar et al 2022 [63] | 1 | 1 | 1 | 2 | 1 | 0 | 0 | U | 1 | U | U | 0 | 1 | 0 | 8 |
| Wade et al 2023 [64] | 1 | 1 | 1 | 1 | 1 | 1 | 0 | U | 1 | U | U | 1 | 1 | 0 | 9 |
| Wang et al 2025 [65] | 1 | 1 | 1 | 0 | 1 | 1 | 0 | U | 1 | U | 1 | 0 | 1 | 0 | 8 |
| Wang et al 2024 [66] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Washabaugh et al 2022 [67] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Yamamoto et al 2022 [68] | 1 | 1 | 1 | 2 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 11 |
| Yamamoto et al 2021 [69] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |
| Young et al 2023 [70] | 1 | 1 | 1 | 1 | 1 | 0 | 0 | U | 1 | U | U | 1 | 1 | 0 | 8 |
| Zhou et al 2025 [71] | 1 | 1 | 0 | 1 | 1 | 1 | 0 | U | 1 | U | U | 0 | 1 | 0 | 7 |
| Zhu et al 2023 [72] | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | 1 | U | U | 1 | 1 | 0 | 10 |

[a]Item 5 was scored on a 3-point scale (0=no, 1=partially, 2=yes); items 11, 15, and 16 were scored as 0 (no), 1 (yes), or U (undetermined). All remaining items were scored as 0 (no) or 1 (yes). The maximum possible score was 15.

## *Data Collection Protocols*

### Movement Tasks and Kinematics Outcomes

Multimedia Appendix 3 details every movement task, its description, the testing environment, and the statistical approaches used to compare MLMCS with MBMCS. Gait was the most frequently studied activity, accounting for 75% (29/39) [ 36,38,40-46,48,49,51,52,54,55,58-69,71,72]. Among these gait-focused studies, 21 clearly evaluated overground gait and 3 assessed treadmill gait [43,55,69], while 1 study included both treadmill and overground gait [52] and 4 used other or unclearly specified gait formats [36, 40,54,59]. Running [50,70], countermovement jumps (CMJs) [35,39], functional movements [53,57], and sit-to-stand [34, 59] were each examined in 2 studies. The remaining task categories were represented by single studies, including lifting [47], high jump [56], a combined drop vertical jump, single-leg hop, and single-leg squat protocol [37], and orbital gait with rotational movements [36].

Across the 39 included studies [34-72], 13 lower-limb movement variables were reported, with sagittal-plane joint angles predominating. Knee flexion/extension was the most frequently reported variable, appearing in 90% (35/39) of studies [34-44,46-49,51-55,57-60,62-72], followed by hip flexion/extension in 77% (30/39) [34-36,38-40,44-49,52-55, 57,58,60,62-69,71,72] and ankle plantarflexion/dorsiflexion in 67% (26/39) [34-36,38-44,46,48,49,52-54,57,58,61,63-66, 68,69,71]. Frontal- and transverse-plane measures were far less common. Hip abduction/adduction was reported in 10% (4/39) of studies [38,46,53,58], while frontal-plane knee angles were reported in 8% (3/39), described as knee varus/valgus in [38,57] and knee abduction/adduction in [58]. Hip internal/external rotation was reported in 3% (1/39) of studies [53]. Pelvic tilt was reported in 3% (1/39) [57], and shin angle in 5% (2/39) [50,57]. Other variables were each reported once, including rearfoot-forefoot flexion/extension [61], longitudinal knee-ankle displacement [59], and amortization-phase take-off angles at the hip, knee, and ankle [56].

### Testing Environment and Recording Conditions

Most studies were conducted indoors, with 92% (36/39) performed exclusively in indoor settings. One study was conducted outdoors [71], one included both indoor and outdoor testing [59], and 3% (1/39) did not report the task environment [56]. Clothing conditions were reported in 38% (15/39) of studies [35,36,38,40,43,46,51,55,58,59,61-64, 70], most commonly describing minimal, tight, or form-fitting attire [35,36,46,55,58,64]. Where reported, footwear and clothing conditions varied and included barefoot protocols [38,46,61], ankle socks [40,61], comfortable shoes [43], cushioned running shoes [70], normal clothes [59], identical clothing [62], and nonconventional clothing [63]. The remaining 62% (24/39) [34,37,39,41,42,44,45,47-49,52-54, 56,57,60,65-69,71,72] did not specify clothing conditions. Only 3 studies [36,40,63] reported the background conditions during data capture, describing unreliable output in dark lighting, the use of a homogeneous blue background, or testing under dark, dim, normal, and very bright lighting, including background clutter with people.

## MLMCS Configuration and Technical Specifications

### 2D and 3D MLMCS Methods and Model Types

Multimedia Appendix 4 summarizes the MLMCS, together with the corresponding technologies, hardware, and capture details reported across the included studies [34-72]. Overall, 25 distinct human MLMCS methods were identified. As MediaPipe, BlazePose, and GaitFormer algorithms were each reported in both 2D and 3D implementations, these corresponded to 28 distinct method variants, comprising 12 2D and 16 3D variants. In total, 47 MLMCS method applications were reported across the included studies [34-72], with several studies evaluating more than one MLMCS approach.

Regarding 2D MLMCS methods, OpenPose was the most frequently used approach, accounting for 40% (12/30) of all reported 2D method applications. MediaPipe [34,44,52] and DeepLabCut [50,67,71] were the next most frequent methods, each reported in 10% (3/30) of applications, while Silhouette-based Approach [40,45] and manual digitization [50,56] were each reported in almost 7% (2/30). BlazePose [70], Detectron2 [57], GaitFormer [54], Griffin [58], MMPose [35], MoveNet [67], OMGait [63], and SegM [61] were each reported once, corresponding to 3% (1/30) of 2D method applications.

For 3D lifter or reconstruction methods, BlazePose was the only approach reported in more than one study, accounting for 12% (2/17) of all reported 3D method applications [37,72]. Each of the remaining 3D methods was reported once, corresponding to 6% (1/17) each, namely 3D Pose Tracker for Gait Test [36], SmartGait [38], Sbsq-Pose [39], Customized Regular Network (RegNetY) [41], the perspective-aware human mesh recovery (HMR) model (CameraHMR) [46], BlazePose–Sequence-to-Sequence [48], single-view algorithms [53], vision transformer [54], 3D Pose Network (3DPoseNet) [55], Strided Transformer [57], interval particle filtering [59], MediaPipe [60], Sportip Motion 3D [62], Pose3DM [65], and 3D HuMoR [66].

### MLMCS RGB Cameras Hardware, Viewpoint, and Positioning

Most included studies used a single RGB camera or monocular smartphone setup for data acquisition. A small number of studies recorded participants using two-camera [42,47,58,70] or multicamera setups [41,53,57,66]; however, these configurations were used to obtain separate single-camera viewpoints rather than to perform multicamera markerless analysis. Camera types varied across studies, with smartphones being the most frequently reported, accounting for 41% (16/39) of studies [34-39,46,47,52,54,57,58,62,63,65,70]. Scientific or industrial RGB cameras were reported in 26% (10/39) [40,42,49,50,53,61,64,67-69], action cameras in 8% (3/39) [44,48,66], professional video cameras in 5% (2/39) [55,56], and virtual cameras in 5% (2/39) [36,45]. Several studies did not report an exact camera model or referred only to generic camera types [41,51,59,60,71,72].

Among the 34 studies [34,35,37-45,47-53,55,58-72] that reported a capture plane, a single sagittal view was the most common configuration, used in 53% (18/34) of studies. Sagittal views combined with a frontal plane were reported in 21% (7/34) [37,38,42,58,64,70,71], sagittal views combined with frontal and oblique planes in 9% (3/34) [41,47,53], and sagittal views combined with an oblique plane in 9% (3/34) [45,52,66]. Frontal-only capture was reported in 6% (2/34) [39,44], and one study used an oblique transverse view [55]. Among the 67% (26/39) of studies [34,35,37-39,41-44,47,48,50,53,55,57,58,60,62,63,65-70] that reported camera positioning, all used fixed camera setups, most commonly tripod-mounted or otherwise statically mounted, typically positioned approximately 1.5-4.0 (m) from the participant, with camera heights ranging from 0.60 to 2.30 (m).

### Reference Standards and Benchmarking System Configurations

Multimedia Appendix 5 [84,85] summarizes the reference measurement modalities and validation frameworks used in the included studies [34-72]. Of the 39 studies included in this scoping review, 89.7% (35/39) [34-43,45,46,48-55,57-59,61-72] reported their validation methodologies, except for 10.3% (4/39) [44,47,56,60]. Overall, validation approaches fell into 2 principal categories: marker-based techniques and nonmarker-based methods.

### Marker-Based Reference Techniques

As summarized in Multimedia Appendix 5, optical MBMCS were by far the most frequently used "ground-truth" reference, serving as the validation standard in 83% (29/35) of studies [34-36,38-42,46,48,49,51-55,57-59,61,62,64-70,72] that reported a validation framework. All of these systems used multicamera infrared arrays. Among the 22 optical MBMCS studies [34,35,38-40,48,49,51-53,55,57-59,61,62,64,66-70] that reported the marker-set model, 50% (11/22) of the included studies used anatomical-landmark-based sets [34,35,51,61,62,64,66-70], while 41% (9/22) used established biomechanical protocols [38-40,48,49,53,55,57,58]. One investigation mirrored the keypoints used in markerless models [59], and another used a manufacturer's proprietary marker configuration [52]. The marker-set model was not reported in 24% (7/29) of the optical MBMCS studies [36,41,42,46,54,65,72].

### Nonmarker-Based Reference Techniques

As detailed in Multimedia Appendix 5, of the 35 studies reporting a validation framework, 17% (6/35) used nonmarker-based references [37,43,45,50,63,71] exclusively. These comprised depth-camera systems using Microsoft Kinect [37,63], an inertial measurement unit (IMU)–based system [43], manual frame-by-frame digitization using Kinovea (developed by Joan Charmant) [50], synthetic and literature-derived biomechanical data [45], and a computer vision benchmark (CVB) [71]. One additional study supplemented marker-based validation with publicly available CVBs [54].

### Intermethod Validity and Reliability Relative to Reference Standards

Multimedia Appendices (Multimedia Appendices 6–8) summarize the intermethod relative validity, absolute validity, and relative reliability of hip, knee, and ankle joint angles for the different markerless systems relative to their respective reference standards. Intermethod relative validity reflects the degree of association between the index and reference methods [86], whereas intermethod absolute validity reflects the degree of numerical agreement between them [87]. Intermethod relative reliability reflects the extent to which the 2 methods consistently preserve the relative ranking of observations within the sample and is commonly expressed using coefficients such as the intraclass correlation coefficient (ICC) [88].

72% (28/39) of the included studies [35-43,45,46,48,49,51,52,54,55,57,58,62-64,66-70,72] reported at least one intermethod measurement property, namely relative validity, absolute validity, or relative reliability, for hip, knee, or ankle joint angles. Among studies reporting lower-limb joint-angle outcomes, knee metrics were almost universal: 93% (13/14) of studies [35,36,38-40,43,49,51,52,55,62,64,72] reporting intermethod relative validity, 100% (24/24) [35,37-43,45,46,48,49,51,52,54,57,58,62-64,66-69] reporting intermethod absolute validity, and 100% (7/7) [35,41,49,55,57,69,70] reporting intermethod relative reliability included at least 1 knee outcome, most commonly knee flexion/extension in the sagittal plane. Reporting for the hip and ankle followed a similar, though slightly less comprehensive, pattern. Hip outcomes were reported in 79% (11/14) of studies [35,36,38-40,49,52,62,64,70,72] assessing intermethod relative validity, 83% (20/24) [35,38-41,45,46,48,49,52,54,57,58,62-64,66-69] assessing intermethod absolute validity, and 71% (5/7) [35,41,49,57,69] assessing intermethod relative reliability. Corresponding values for the ankle were 64% (9/14), 79% (19/24) [ 35,38-43,46,48,49,52,54,57,58,63,64,66,68,69], and 71% (5/7) [35,41,49,57,69], respectively, generally with fewer reported outcomes per joint than for the knee.

Across studies, sagittal-plane hip and knee flexion/extension generally showed stronger intermethod associations and reliability than other joint-plane combinations, with several studies reporting very high correlations (eg, $r$=0.938-0.997) and ICC values ranging from 0.734 to 0.990. By contrast, frontal-plane hip abduction/adduction and ankle dorsiflexion/plantarflexion or inversion/eversion showed more variable and often weaker validity (eg, hip abduction/adduction: $r$=0.708-0.748 and $R^2$=0.46; ankle dorsiflexion/plantarflexion: $r$=0.27-0.984; ankle inversion/eversion: $R^2$=0.26), while ankle reliability was more heterogeneous (ICC=0.01-0.93), with some studies reporting only coefficients of variation rather than ICCs [57].

## Discussion

### Principal Findings

This scoping review identified 39 studies [34-72] of single RGB-camera MLMCS for lower-limb kinematics. Existing studies have primarily focused on gait, healthy populations, and sagittal-plane hip and knee flexion/extension. OpenPose is the most frequently used 2D method, monocular 3D approaches are less common, and optical MBMCS remains the predominant reference standard. Overall, sagittal hip and knee estimates are the most mature, whereas ankle, nonsagittal, and occluded-side measurements are less robust. Substantial heterogeneity in protocols and validation metrics limited cross-study comparison and clinical translation.

### Joint- and Plane-Specific Performance Hierarchy

The evidence demonstrates a clear hierarchy of accuracy across joints and anatomical planes under the movement conditions examined. The performance of 2D single-camera systems is strongest for sagittal-plane kinematics of the hip and knee [49,64,67]. Across studies using AI-based 2D pipelines, errors in hip and knee flexion/extension are typically lower than those observed at the ankle, particularly when the joints are assessed on the camera-facing side. This greater accuracy may be explained by the relatively large displacement of the hip and knee within the image plane during sagittal-plane movements, especially when recordings are obtained from a lateral viewpoint. By contrast, the ankle joint remains the most challenging lower-limb joint for single-camera estimation [58,68-70]. This difficulty has been attributed to its smaller anatomical size, faster movement, and greater susceptibility to occlusion by the foot [43,49,52,64]. These limitations are particularly consequential because even small errors in joint localization can result in substantial angular errors when the anatomical segments used to calculate the joint angle are relatively short.

More broadly, sagittal-plane kinematics are measured with substantially greater accuracy than frontal- or transverse-plane movements, particularly when 2D systems are used [38,64]. The lower accuracy of frontal- and transverse-plane measurements reflects a fundamental limitation of monocular imaging. Although a single camera can directly capture movement within the 2 dimensions of the image plane, estimates involving movement toward or away from the camera are more susceptible to perspective distortion and depth-related errors. These measurements may also vary according to the model-specific definitions used to determine joint centers. Consequently, frontal- and transverse-plane outcomes remain less consistent than sagittal-plane estimates. Evidence supporting transverse-plane kinematics is particularly limited. Although some monocular 3D systems have shown encouraging performance in estimating hip internal and external rotation [53,66], convincing validation of knee axial rotation remains lacking.

## Effects of Occlusion, Limb Visibility, and Camera Viewpoint

Camera positioning, limb visibility, and viewpoint are critical components of the movement conditions that influence single-camera accuracy. The occluded-side (far-side) leg consistently produces less accurate and more variable results than the camera-side (near-side) leg [49,52,64,70,72]. When the 2 limbs overlap within the image, the model may interchange corresponding keypoints, lose distal landmarks, or infer the position of a partially hidden joint. This finding has an important practical implication because bilateral estimates obtained from a single camera should not automatically be treated as equally reliable. Camera placement should therefore be considered part of the measurement method rather than a simple recording preference.

One included study [66] found that side views (90° left/right) generally yielded the smallest differences for most variables, whereas front and rear views (0°/180°) produced the largest errors. The authors attributed this pattern to the need to infer sagittal-plane motion primarily from depth rather than from direct displacement within the image plane [66]. They also found that a shorter camera-to-subject distance (2 m rather than 3 m) was generally associated with greater lower-limb kinematic accuracy [66]. Similarly, a near-frontal view was more effective for temporal gait parameters, whereas a lateral view was better suited to spatial parameters and joint-angle estimation [52]. A more recent study [53] suggested that oblique views may be advantageous in some settings, with a front-right camera positioned at approximately 45° producing the best overall performance across the evaluated tasks. Taken together, these findings indicate that single-camera systems are most reliable when the limb of interest remains clearly visible and camera placement is aligned with the target outcome. However, despite these emerging patterns, there is still no broad consensus regarding optimal camera-placement protocols. Further research is therefore needed to establish standardized, evidence-based recommendations for research and clinical settings.

## Influence of Algorithm Choice in 2D and 3D Monocular MLMCS

The current review indicates that the selection of the HPE algorithm is a critical decision that directly affects the validity and reliability of the resulting kinematic data in MLMCS. Among the open-source 2D HPE methods evaluated, OpenPose and MoveNet Thunder (MNT) produced the lowest hip flexion/extension errors, whereas OpenPose was the most accurate method for knee flexion/extension. Across hip, knee, and ankle kinematics, OpenPose and MNT achieved greater precision at the hip, whereas DeepLabCut produced substantially larger estimation disparities, with joint-tracking errors increasing markedly at both the hip and knee [49,67]. Together, these studies indicate that OpenPose is a robust and accurate option for 2D HPE-based analysis of lower-limb sagittal-plane kinematics.

The performance of the DeepLabCut algorithm varies substantially depending on whether a pretrained or custom-trained model is used. Pretrained DeepLabCut, which uses the DeepLabCut Model Zoo Full Human network without task-specific labeling, yielded the highest errors for hip and knee flexion/extension [67]. By contrast, task-specific labeling and training can substantially improve performance in focused applications such as running or prosthetic gait analysis [50,67,71]. Accordingly, algorithm comparisons should report not only the model's name but also the model version, keypoint definitions, training dataset, fine-tuning procedures, and preprocessing pipeline.

Monocular 3D MLMCS approaches offer a potential means of addressing some of the limitations of 2D methods, particularly when movements include out-of-plane components. However, their outputs remain dependent on the quality of the underlying 2D detections and on learned assumptions regarding human movement. The comparatively strong performance of World-grounded Humans with Accurate Motion (WHAM) in the largest included benchmark is encouraging, but this ranking should be interpreted within the context of the tasks, datasets, and reference methods used in that comparison [53]. It does not establish the superiority of WHAM across all clinical populations or movement types.

The mixed effects of temporal and sequence-based model enhancements further demonstrate that an improvement at one joint does not necessarily result in improvements throughout the lower limb. For example, sequence modeling may stabilize distal trajectories and improve ankle estimates while having inconsistent effects on hip or knee kinematics [37,48,55]. Similarly, transformer-based models have required task-specific fine-tuning when public training datasets did not adequately represent clinical gait or functional movements [54,57]. Algorithm development should therefore prioritize external validation across tasks and populations rather than relying solely on performance within the data used for model development.

## Heterogeneity in Validation Frameworks and Benchmarking Outcome

The heterogeneity of validation frameworks remains a major barrier to determining whether differences between systems reflect algorithmic performance or differences in study design. The included studies [34-72] used optical MBMCS, depth cameras, IMUs, manual digitization, synthetic data, and CV benchmarks as reference methods. They also reported a wide range of statistical outcomes, including correlations, ICCs, MAEs, RMSEs, and limits of agreement. These measures address different measurement properties and should not be treated as interchangeable. In particular, a strong correlation indicates that 2 methods vary together, but it does not demonstrate close numerical agreement. Optical MBMCS remain the most commonly used practical reference for lower-limb kinematic validation. However, comparisons should clearly describe marker placement, biomechanical models, joint-center definitions, synchronization, filtering, and reference-system uncertainty. IMUs can provide clinically useful comparison data, but their outputs depend

on sensor calibration, alignment, and modeling assumptions. They are therefore better interpreted as pragmatic reference methods than as definitive anatomical ground truth [89]. Greater standardization of benchmarking procedures would improve the interpretability of future research. At a minimum, studies should report camera specifications and placement, the HPE model and version, training or fine-tuning procedures, anatomical keypoint definitions, missing-data management, temporal filtering, synchronization procedures, the reference model, and whether the outcomes represent discrete values or continuous waveforms. Side-specific results should also be reported when limb visibility differs.

Several findings from the included studies [34-72], although not central to the 3 predefined objectives, provide important context for the practical deployment of single-camera MLMCS. Evidence from the included and related studies indicates that useful estimates may still be obtained under varying lighting conditions and when participants wear conventional or athletic clothing [63]. Nevertheless, most studies were conducted indoors under comparatively controlled conditions, and clothing protocols were frequently unreported. The available findings should therefore not be interpreted as evidence that recording conditions have negligible effects. Instead, they suggest that robustness to lighting, background complexity, clothing, and participant characteristics should be evaluated systematically rather than assumed. Accordingly, although many studies standardize clothing by using minimal or close-fitting attire, the available evidence suggests that moderate everyday or athletic clothing does not necessarily preclude useful kinematic measurement, although highly loose-fitting garments remain more challenging [90].

## Clinical and Future Research Implications

These findings suggest that monocular MLMCS may be useful for outpatient rehabilitation, community screening, and repeated follow-up when MBMCS are impractical because of cost, setup time, or staffing demands. Their clearest clinical role is as objective adjuncts for trend detection and treatment monitoring when sagittal-plane hip and knee flexion/extension are the main outcomes, including in type 2 diabetes [91], stroke [92], and knee osteoarthritis [93-95]. However, current pipelines should not be used as standalone tools for ankle plantarflexion/dorsiflexion during rapid or complex loading, frontal-plane hip or knee kinematics, or transverse-plane rotations, which remain less robust and less consistently validated [96-101]. Clinical use should therefore remain conservative, prioritizing the camera-facing limb, fixed camera placement, limited out-of-plane motion, repeated trials, and consistent clothing and recording conditions, because small discrete changes may fall within measurement error [102,103].

The current review can guide future systematic reviews by identifying which questions are mature for focused synthesis. Herein, we propose several RQs. First, future reviews should focus on hip, knee, and ankle sagittal plane kinematics during walking and running using monocular 2D and 3D HPE algorithms. Second, comparing different 2D algorithms (eg, OpenPose vs MNT), different 3D algorithms (eg, SmartGait vs WHAM), and 2D versus 3D algorithms should be another research focus. In contrast, questions surrounding general joint frontal- and transverse-plane kinematics, pathology-specific use, and nongait tasks (eg, CMJ) may have too few primary studies to warrant conducting a systematic review. For these questions, the priority would be conducting primary validity and reliability studies. Lastly, future reviews should investigate whether background scenes, clothing effects, and limb visibility influence the performance of different HPE algorithms on joint kinematics.

## Limitations

First, the prospectively developed review protocol was not registered in a public database. Second, by limiting our findings to only full-text articles published in English, articles reported in other languages or the gray literature may have been missed. Third, the review was purposefully limited to monocular RGB systems and lower-limb tasks; as such, evidence related to multicamera setups, depth sensors, or whole-body applications was excluded. Similarly, segment kinematics, such as pelvic kinematics, were not reported because the focus of this paper was on hip, knee, and ankle kinematics. These domains may involve distinct MLMCS algorithms and implementation strategies. Fourth, the lack of established quality assessment checklists specific to markerless data collection limited our ability to systematically evaluate how technical variables, rather than study design alone, impacted the reported accuracy. Finally, reporting across the included studies [34-72] was heterogeneous: over 40% (17/39) of studies failed to report at least one core element of participant demographics or experimental protocols, and definitions of outcome measures varied considerably, precluding the possibility of formal meta-analysis. Relatedly, potential publication bias should also be considered, as studies reporting favorable validation results may have been more likely to be published and identified. In addition, the rapid technological evolution in AI-based pose estimation means that some included algorithms, software versions, and performance estimates may be superseded quickly, which may limit the long-term generalizability of this synthesis.

## Conclusions

This scoping review provides the first focused synthesis of single RGB cameras, which specifically maps monocular lower-limb applications across movement tasks, participant groups, AI pipelines, benchmarking methods, validity, and reliability outcomes. This review offers a clearer evidence framework for understanding where single-camera MLMCS currently performs well and where important limitations remain. OpenPose emerged as the best-performing 2D HPE method for sagittal knee flexion/extension and also ranked first for sagittal hip flexion/extension, although MNT performed similarly at the hip flexion/extension. Among monocular 3D systems, WHAM was the best-performing single-view method overall, as it achieved the lowest overall RMSD in the largest benchmark of MLMCS. The clearest

performance hierarchy was therefore observed for sagittal hip and knee flexion/extension, whereas conclusions for the ankle, nonsagittal planes, and joint-specific ranking across 3D methods remain less certain. Future progress will depend on more standardized protocols, richer distal keypoint definitions including ankle and foot, pathology-specific validation, and algorithms that are more robust to occlusion, viewpoint variation, and real-world recording conditions.

## Acknowledgments

We thank everyone who offered help and all our collaborators for their active cooperation.

## Generative AI Use

The authors used ChatGPT-5 to improve the readability and linguistic clarity of the manuscript. The AI was used solely for language polishing. All intellectual content, data analysis, and conclusions were developed by the human authors, who take full responsibility for the entirety of the work.

## Funding

This work is supported in part by UK Research and Innovation (UKRI) Innovate UK (10093679) and CELTIC-NEXT Eureka (C2023/1-19). Dr B. X. W. L. is supported by the Medical Research Council (grant no. MR/Y013557/1; UKRI541). B.G. was supported by a PhD studentship jointly funded by Aparito (a wholly owned subsidiary of Eli Lilly and Company) and the University of Essex. The funding partner had no role in the conception, design, conduct, analysis, interpretation, or reporting of this review.

## Data Availability

All data generated or analyzed during this study are included in this study.

## Authors' Contributions

Conceptualization: BG, DH, BXWL, XZ
Data curation: BG, DH
Formal analysis: BG, DH
Funding acquisition: BXWL, XZ
Investigation: BG, DH, BXWL, XZ
Methodology: BG, DH, BXWL, XZ
Project administration: BXWL, XZ
Resources: BG, BXWL, XZ
Software: BG, BXWL, XZ
Supervision: BXWL, XZ
Validation: BG, DH, BXWL
Visualization: BG, DH, BXWL, XZ
Writing – original draft: BG, BXWL
Writing – review & editing: BG, DH, BXWL, XZ

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Database-specific electronic search strategies used for identifying studies included in this scoping review. Record counts are based on the search conducted up to March 16, 2026.
[DOCX File (Microsoft Word File), 22 KB-Multimedia Appendix 1]

## Multimedia Appendix 2

Summary of the study and population characteristics of the included studies.
[DOCX File (Microsoft Word File), 262 KB-Multimedia Appendix 2]

## Multimedia Appendix 3

Summary of biomechanical measurements, task characteristics, quantitative analysis methods, and testing conditions across the 39 included studies.
[DOCX File (Microsoft Word File), 237 KB-Multimedia Appendix 3]

## Multimedia Appendix 4

Summary of technical characteristics of the markerless motion capture systems (MLMCS) used across the included studies.
[DOCX File (Microsoft Word File), 265 KB-Multimedia Appendix 4]

### Multimedia Appendix 5

Overview of the ground-truth modalities and validation system characteristics used in the included studies.
[DOCX File (Microsoft Word File), 250 KB-Multimedia Appendix 5]

### Multimedia Appendix 6

Summary of the reported intermethod relative validity of markerless motion capture systems (MLMCSs) for hip, knee, and ankle joint angles in the included studies, benchmarked against corresponding ground-truth methods and presented by joint, side, and relative validity metric.
[DOCX File (Microsoft Word File), 238 KB-Multimedia Appendix 6]

### Multimedia Appendix 7

Summary of the reported intermethod absolute validity of markerless motion capture systems (MLMCSs) for hip, knee, and ankle joint angles in the included studies, benchmarked against corresponding ground-truth methods and presented by joint, side, and absolute validity metric.
[DOCX File (Microsoft Word File), 257 KB-Multimedia Appendix 7]

### Multimedia Appendix 8

Summary of the reported intermethod relative reliability of markerless motion capture systems (MLMCSs) for hip, knee, and ankle joint angles in the included studies, benchmarked against corresponding ground-truth methods and presented by joint, side, and reliability metric.
[DOCX File (Microsoft Word File), 237 KB-Multimedia Appendix 8]

### Checklist 1

PRISMA-ScR checklist.
[DOCX File (Microsoft Word File), 87 KB-Checklist 1]

## References


1. Gorton GE, Hebert DA, Gannotti ME. Assessment of the kinematic variability among 12 motion analysis laboratories. Gait Posture. Apr 2009;29(3):398-402. [doi: 10.1016/j.gaitpost.2008.10.060] [Medline: 19056271]
2. Ceseracciu E, Sawacha Z, Cobelli C. Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept. PLoS ONE. 2014;9(3):e87640. [doi: 10.1371/journal.pone.0087640] [Medline: 24595273]
3. Colyer SL, Evans M, Cosker DP, Salo AIT. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports Med Open. Jun 5, 2018;4(1):24. [doi: 10.1186/s40798-018-0139-y] [Medline: 29869300]
4. van der Kruk E, Reijne MM. Accuracy of human motion capture systems for sport applications; state-of-the-art review. Eur J Sport Sci. Jul 2018;18(6):806-819. [doi: 10.1080/17461391.2018.1463397] [Medline: 29741985]
5. Scataglini S, Abts E, Van Bocxlaer C, Van den Bussche M, Meletani S, Truijen S. 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 (Basel). Jun 6, 2024;24(11):3686. [doi: 10.3390/s24113686] [Medline: 38894476]
6. Hatamzadeh M, Busé L, Chorin F, Alliez P, Favreau JD, Zory R. A kinematic-geometric model based on ankles' depth trajectory in frontal plane for gait analysis using a single RGB-D camera. J Biomech. Dec 2022;145:111358. [doi: 10.1016/j.jbiomech.2022.111358] [Medline: 36334322]
7. Balta D, Figari G, Paolini G, et al. A model-based markerless protocol for clinical gait analysis based on a single RGB-depth camera: concurrent validation on patients with cerebral palsy. IEEE Access. 2023;11:144377-144393. [doi: 10.1109/ACCESS.2023.3340622]
8. Tychola KA, Tsimperidis I, Papakostas GA. On 3D reconstruction using RGB-D cameras. Digital. Dec 2022;2(3):401-421. [doi: 10.3390/digital2030022]
9. Khoshelham K, Elberink SO. Accuracy and resolution of Kinect depth data for indoor mapping applications. Sensors (Basel). 2012;12(2):1437-1454. [doi: 10.3390/s120201437] [Medline: 22438718]
10. Wade L, Needham L, McGuigan P, Bilzon J. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ. 2022;10:e12995. [doi: 10.7717/peerj.12995] [Medline: 35237469]
11. Afzal HMR, Louhichi B, Alrasheedi NH. Challenges in combining EMG, joint moments, and GRF from marker-less video-based motion capture systems. Bioengineering (Basel). Apr 27, 2025;12(5):461. [doi: 10.3390/bioengineering12050461] [Medline: 40428080]

12. Adlou B, Wilburn C, Weimar W. Motion capture technologies for athletic performance enhancement and injury risk assessment: a review for multi-sport organizations. Sensors (Basel). Jul 13, 2025;25(14):4384. [doi: 10.3390/s25144384] [Medline: 40732512]
13. Nakano N, Sakura T, Ueda K, et al. Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras. Front Sports Act Living. 2020;2:50. [doi: 10.3389/fspor.2020.00050] [Medline: 33345042]
14. Yang J, Park K. Improving gait analysis techniques with markerless pose estimation based on smartphone location. Bioengineering (Basel). Jan 30, 2024;11(2):141. [doi: 10.3390/bioengineering11020141] [Medline: 38391625]
15. Horsak B, Eichmann A, Lauer K, et al. Concurrent validity of smartphone-based markerless motion capturing to quantify lower-limb joint kinematics in healthy and pathological gait. J Biomech. Oct 2023;159:111801. [doi: 10.1016/j.jbiomech.2023.111801] [Medline: 37738945]
16. Huang T, Ruan M, Huang S, Fan L, Wu X. Comparison of kinematics and joint moments calculations for lower limbs during gait using markerless and marker-based motion capture. Front Bioeng Biotechnol. Mar 12, 2024;12:1280363. [doi: 10.3389/fbioe.2024.1280363] [Medline: 38532880]
17. Yang C, Wei L, Huang X, et al. Comparison of lower limb kinematic and kinetic estimation during athlete jumping between markerless and marker-based motion capture systems. Sci Rep. 2025;15(1):18552. [doi: 10.1038/s41598-025-02739-9]
18. Ueno R. Calibrationless monocular vision musculoskeletal simulation during gait. Heliyon. Jun 15, 2024;10(11):e32078. [doi: 10.1016/j.heliyon.2024.e32078] [Medline: 38868012]
19. D'Antonio E, Taborri J, Palermo E, Rossi S, Patane F. A markerless system for gait analysis based on OpenPose library. Presented at: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May 25-28, 2020; Dubrovnik, Croatia. [doi: 10.1109/I2MTC43012.2020.9128918]
20. Liao Y, Vakanski A, Xian M. A deep learning framework for assessing physical rehabilitation exercises. IEEE Trans Neural Syst Rehabil Eng. Feb 2020;28(2):468-477. [doi: 10.1109/TNSRE.2020.2966249] [Medline: 31940544]
21. Tang H, Munkasy B, Li L. Differences between lower extremity joint running kinetics captured by marker-based and markerless systems were speed dependent. J Sport Health Sci. Jul 2024;13(4):569-578. [doi: 10.1016/j.jshs.2024.01.002] [Medline: 38218372]
22. Kushwaha M, Choudhary J, Singh DP. Enhancement of human 3D pose estimation using a novel concept of depth prediction with pose alignment from a single 2D image. Comput Graph. Oct 2022;107:172-185. [doi: 10.1016/j.cag.2022.07.021]
23. Molteni LE, Andreoni G. Comparing the accuracy of markerless motion analysis and optoelectronic system for measuring gait kinematics of lower limb. Bioengineering (Basel). Apr 16, 2025;12(4):424. [doi: 10.3390/bioengineering12040424] [Medline: 40281784]
24. Menychtas D, Petrou N, Kansizoglou I, et al. Gait analysis comparison between manual marking, 2D pose estimation algorithms, and 3D marker-based system. Front Rehabil Sci. 2023;4:1238134. [doi: 10.3389/fresc.2023.1238134] [Medline: 37744429]
25. Scott B, Seyres M, Philp F, Chadwick EK, Blana D. Healthcare applications of single camera markerless motion capture: a scoping review. PeerJ. 2022;10:e13517. [doi: 10.7717/peerj.13517] [Medline: 35642200]
26. Morimoto T, Hirata H, Kobayashi T, et al. Gait analysis using digital biomarkers including smart shoes in lumbar spinal canal stenosis: a scoping review. Front Med (Lausanne). 2023;10:1302136. [doi: 10.3389/fmed.2023.1302136] [Medline: 38162877]
27. Stenum J, Cherry-Allen KM, Pyles CO, Reetzke RD, Vignos MF, Roemmich RT. Applications of pose estimation in human health and performance across the lifespan. Sensors (Basel). Nov 3, 2021;21(21):7315. [doi: 10.3390/s21217315] [Medline: 34770620]
28. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. JBI Manual for Evidence Synthesis. 2024. [doi: 10.46658/JBIMES-24-09]
29. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. JBI; 2024. [doi: 10.46658/JBIMES-24-01] ISBN: 9780648848820
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. Feb 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
31. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. Sep 20, 2010;5(1):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
32. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. Oct 2, 2018;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. Dec 5, 2016;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]

34. Akturk S, Derdiyok FB, Serbest K. Markerless joint angle estimation using MediaPipe with a rapid setup for joint moment calculation. Multimed Tools Appl. 2026;85(1):38. [doi: 10.1007/s11042-026-21256-z]
35. Aleksic J, Kanevsky D, Mesaroš D, et al. Validation of automated countermovement vertical jump analysis: markerless pose estimation vs. 3D marker-based motion capture system. Sensors (Basel). Oct 14, 2024;24(20):6624. [doi: 10.3390/s24206624] [Medline: 39460104]
36. Aoyagi Y, Yamada S, Ueda S, et al. Development of smartphone application for markerless three-dimensional motion capture based on deep learning model. Sensors (Basel). Jul 14, 2022;22(14):5282. [doi: 10.3390/s22145282] [Medline: 35890959]
37. Babouras A, Abdelnour P, Fevens T, Martineau PA. Comparing novel smartphone pose estimation frameworks with the Kinect V2 for knee tracking during athletic stress tests. Int J Comput Assist Radiol Surg. Jul 2024;19(7):1321-1328. [doi: 10.1007/s11548-024-03156-5] [Medline: 38730186]
38. Barzyk P, Boden AS, Howaldt J, et al. Steps to facilitate the use of clinical gait analysis in stroke patients: the validation of a single 2D RGB smartphone video-based system for gait analysis. Sensors (Basel). Dec 6, 2024;24(23):7819. [doi: 10.3390/s24237819] [Medline: 39686356]
39. Barzyk P, Zimmermann P, Stein M, Keim D, Gruber M. AI-smartphone markerless motion capturing of hip, knee, and ankle joint kinematics during countermovement jumps. Eur J Sport Sci. Oct 2024;24(10):1452-1462. [doi: 10.1002/ejsc.12186] [Medline: 39205332]
40. Castelli A, Paolini G, Cereatti A, Della Croce U. A 2D markerless gait analysis methodology: validation on healthy subjects. Comput Math Methods Med. 2015;2015:186780. [doi: 10.1155/2015/186780] [Medline: 26064181]
41. Chao V, Park S, Jamsrandorj A, Jung D, Kim J, Mun KR. Lower limb joint angle estimation using a single RGB camera. Measurement. Apr 2026;270:120772. [doi: 10.1016/j.measurement.2026.120772]
42. Dunn M, Kennerley A, Murrell-Smith Z, Webster K, Middleton K, Wheat J. Application of video frame interpolation to markerless, single-camera gait analysis. Sports Eng. Dec 2023;26(1):1-10. [doi: 10.1007/s12283-023-00419-3]
43. Galasso S, Carissimo C, Cerro G, et al. A novel measurement procedure for error correction in single camera gait analysis. IEEE Access. 2024;12:158052-158064. [doi: 10.1109/ACCESS.2024.3485204]
44. Ganai UJ, Ratne A, Bhushan B, Venkatesh KS. Early detection of autism spectrum disorder: gait deviations and machine learning. Sci Rep. Jan 6, 2025;15(1):873. [doi: 10.1038/s41598-025-85348-w] [Medline: 39757284]
45. Goffredo M, Bouchrika I, Carter JN, Nixon MS. Self-calibrating view-invariant gait biometrics. IEEE Trans Syst Man Cybern B Cybern. Aug 2010;40(4):997-1008. [doi: 10.1109/TSMCB.2009.2031091] [Medline: 19884085]
46. Horsak B, Simonlehner M, Quehenberger V, et al. Validity and reliability of monocular 3D markerless gait analysis in simulated pathological gait: a comparative study with OpenCap. J Biomech. Dec 2025;193(4):112986. [doi: 10.1016/j.jbiomech.2025.112986] [Medline: 41046587]
47. Hsu GSJ, Wu JS, Huang YKD, Chiu CC, Kang JH. Automatic detect incorrect lifting posture with the pose estimation model. Life (Basel). Feb 24, 2025;15(3):358. [doi: 10.3390/life15030358] [Medline: 40141703]
48. Hulleck AA, AlShehhi A, El Rich M, et al. BlazePose-Seq2Seq: leveraging regular RGB cameras for robust gait assessment. IEEE Trans Neural Syst Rehabil Eng. 2024;32:1715-1724. [doi: 10.1109/TNSRE.2024.3391908] [Medline: 38648155]
49. Ino T, Samukawa M, Ishida T, et al. Validity of AI-based gait analysis for simultaneous measurement of bilateral lower limb kinematics using a single video camera. Sensors (Basel). Dec 13, 2023;23(24):9799. [doi: 10.3390/s23249799] [Medline: 38139644]
50. Johnson CD, Outerleys J, Davis IS. Agreement between sagittal foot and tibia angles during running derived from an open-source markerless motion capture platform and manual digitization. J Appl Biomech. Apr 1, 2022;38(2):111-116. [doi: 10.1123/jab.2021-0323] [Medline: 35272264]
51. Kidziński Ł, Yang B, Hicks JL, Rajagopal A, Delp SL, Schwartz MH. Deep neural networks enable quantitative movement analysis using single-camera videos. Nat Commun. Aug 13, 2020;11(1):4054. [doi: 10.1038/s41467-020-17807-z] [Medline: 32792511]
52. Kim J, Kim R, Byun K, Kang N, Park K. Assessment of temporospatial and kinematic gait parameters using human pose estimation in patients with Parkinson's disease: a comparison between near-frontal and lateral views. PLoS ONE. Jan 1, 2025;20(1):e0317933. [doi: 10.1371/journal.pone.0317933]
53. Li Z, Shin S, Phan V, Meinders E, Halilaj E. Impact of multi-view fusion and biomechanical modeling on markerless motion tracking. IEEE Trans Biomed Eng. 2025;73(6):2054-2062. [doi: 10.1109/TBME.2025.3622032]
54. Li J, Wang Z, Wang C, Su W. GaitFormer: leveraging dual-stream spatial–temporal vision transformer via a single low-cost RGB camera for clinical gait analysis. Knowl Based Syst. Jul 2024;295:111810. [doi: 10.1016/j.knosys.2024.111810]

55. Liang S, Zhang Y, Diao Y, Li G, Zhao G. The reliability and validity of gait analysis system using 3D markerless pose estimation algorithms. Front Bioeng Biotechnol. Aug 10, 2022;10:857975. [doi: 10.3389/fbioe.2022.857975] [Medline: 36032709]
56. Ljubičić S. Kinematic analysis of high jump using the scissors technique. Studia Sportiva. Jul 11, 2024;18(1):173-182. [doi: 10.5817/StS2024-1-15]
57. Mercadal-Baudart C, Liu CJ, Farrell G, et al. Exercise quantification from single camera view markerless 3D pose estimation. Heliyon. Mar 30, 2024;10(6):e27596. [doi: 10.1016/j.heliyon.2024.e27596] [Medline: 38510055]
58. Park YS, Yu YW, Cha H, et al. Verification of markerless gait analysis: multi-camera and single-camera approaches in comparison to marker-based gait analysis. Medicina (Kaunas). Feb 22, 2026;62(2):418. [doi: 10.3390/medicina62020418] [Medline: 41752816]
59. Saboune J, Charpillet F. Markerless human motion tracking from a single camera using interval particle filtering. Int J Artif Intell Tools. Aug 2007;16(4):593-609. [doi: 10.1142/S021821300700345X]
60. Sakane N, Yamauchi K, Kutsuna I, et al. Application of machine learning for detecting high fall risk in middle-aged workers using video-based analysis of the first 3 steps. J Occup Health. Jan 7, 2025;67(1):uiae075. [doi: 10.1093/joccuh/uiae075] [Medline: 39792357]
61. Surer E, Cereatti A, Grosso E, Croce UD. A markerless estimation of the ankle-foot complex 2D kinematics during stance. Gait Posture. Apr 2011;33(4):532-537. [doi: 10.1016/j.gaitpost.2011.01.003] [Medline: 21295984]
62. Usami T, Kisohara M, Nishida K, et al. Gait analysis using an artificial intelligence-based motion capture system with a single smartphone camera. Cureus. Jul 13, 2025;17(7). [doi: 10.7759/cureus.87837.PMID:40809664]
63. Viswakumar A, Rajagopalan V, Ray T, Gottipati P, Parimi C. Development of a robust, simple, and affordable human gait analysis system using bottom-up pose estimation with a smartphone camera. Front Physiol. 2022;12:784865. [doi: 10.3389/fphys.2021.784865] [Medline: 35069246]
64. Wade L, Needham L, Evans M, et al. Examination of 2D frontal and sagittal markerless motion capture: implications for markerless applications. PLoS ONE. 2023;18(11):e0293917. [doi: 10.1371/journal.pone.0293917] [Medline: 37943887]
65. Wang C, Su W, Li J, Xu J. Bidirectional Mamba-enhanced 3D human pose estimation for accurate clinical gait analysis. Fractal Fract. 2025;9(9):603. [doi: 10.3390/fractalfract9090603]
66. Wang H, Su B, Lu L, et al. Markerless gait analysis through a single camera and computer vision. J Biomech. Mar 2024;165:112027. [doi: 10.1016/j.jbiomech.2024.112027] [Medline: 38430608]
67. Washabaugh EP, Shanmugam TA, Ranganathan R, Krishnan C. Comparing the accuracy of open-source pose estimation methods for measuring gait kinematics. Gait Posture. Sep 2022;97:188-195. [doi: 10.1016/j.gaitpost.2022.08.008] [Medline: 35988434]
68. Yamamoto M, Shimatani K, Ishige Y, Takemura H. Verification of gait analysis method fusing camera-based pose estimation and an IMU sensor in various gait conditions. Sci Rep. Oct 21, 2022;12(1):17719. [doi: 10.1038/s41598-022-22246-5] [Medline: 36271241]
69. Yamamoto M, Shimatani K, Hasegawa M, Kurita Y, Ishige Y, Takemura H. Accuracy of temporo-spatial and lower limb joint kinematics parameters using OpenPose for various gait patterns with orthosis. IEEE Trans Neural Syst Rehabil Eng. 2021;29:2666-2675. [doi: 10.1109/TNSRE.2021.3135879] [Medline: 34914592]
70. Young F, Mason R, Morris R, Stuart S, Godfrey A. Internet-of-things-enabled markerless running gait assessment from a single smartphone camera. Sensors (Basel). Jan 7, 2023;23(2):696. [doi: 10.3390/s23020696] [Medline: 36679494]
71. Zhou T, Iskandar MNS, Chiam KH. Diffusion models enable zero-shot pose estimation for lower-limb prosthetic users. PLOS Digital Health. Mar 1, 2025;4(3):e0000745. [doi: 10.1371/journal.pdig.0000745] [Medline: 40163802]
72. Zhu X, Boukhennoufa I, Liew B, et al. Monocular 3D human pose markerless systems for gait assessment. Bioengineering (Basel). May 26, 2023;10(6):653. [doi: 10.3390/bioengineering10060653] [Medline: 37370583]
73. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. Jun 1998;52(6):377-384. [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
74. Adesida Y, Papi E, McGregor AH. Exploring the role of wearable technology in sport kinematics and kinetics: a systematic review. Sensors (Basel). Apr 2, 2019;19(7):1597. [doi: 10.3390/s19071597] [Medline: 30987014]
75. Swain TA, McNarry MA, Runacres AWH, Mackintosh KA. The role of multi-sensor measurement in the assessment of movement quality: a systematic review. Sports Med. Dec 2023;53(12):2477-2504. [doi: 10.1007/s40279-023-01905-1] [Medline: 37698766]
76. Papi E, Koh WS, McGregor AH. Wearable technology for spine movement assessment: a systematic review. J Biomech. Nov 7, 2017;64:186-197. [doi: 10.1016/j.jbiomech.2017.09.037] [Medline: 29102267]
77. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22(3):276-282. [doi: 10.11613/BM.2012.031] [Medline: 23092060]

78. Lin TY, Maire M, Belongie S, et al. Microsoft COCO: common objects in context lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). Presented at: 13th European Conference on Computer Vision (ECCV 2014); Sep 6-12, 2014; Zurich, Switzerland. [doi: 10.1007/978-3-319-10602-1_48]
79. Schwartz MH, Trost JP, Wervey RA. Measurement and management of errors in quantitative gait data. Gait Posture. Oct 2004;20(2):196-203. [doi: 10.1016/j.gaitpost.2003.09.011] [Medline: 15336291]
80. von Marcard T, Henschel R, Black M, Rosenhahn B, Pons-Moll G. Recovering accurate 3D human pose in the wild using IMUs and a moving camera. Presented at: Computer Vision – ECCV 2018: 15th European Conference; Sep 8-14, 2018; Munich, Germany. [doi: 10.1007/978-3-030-01249-6_37]
81. Ionescu C, Papava D, Olaru V, Sminchisescu C. Human3.6M: large scale datasets and predictive methods for 3D human sensing in natural environments. IEEE Trans Pattern Anal Mach Intell. Jul 2014;36(7):1325-1339. [doi: 10.1109/TPAMI.2013.248] [Medline: 26353306]
82. Kwolek B, Michalczuk A, Krzeszowski T, Switonski A, Josinski H, Wojciechowski K. Calibrated and synchronized multi-view video and motion capture dataset for evaluation of gait recognition. Multimed Tools Appl. 2019;78(22):32437-32465. [doi: 10.1007/S11042-019-07945-Y/FIGURES/9]
83. Ghorbani S, Mahdaviani K, Thaler A, et al. MoVi: a large multi-purpose human motion and video dataset. PLoS ONE. 2021;16(6):e0253157. [doi: 10.1371/journal.pone.0253157] [Medline: 34138926]
84. Davis RB, Õunpuu S, Tyburski D, Gage JR. A gait analysis data collection and reduction technique. Hum Mov Sci. Oct 1991;10(5):575-587. [doi: 10.1016/0167-9457(91)90046-Z]
85. Leardini A, Sawacha Z, Paolini G, Ingrosso S, Nativo R, Benedetti MG. A new anatomically based protocol for gait analysis in children. Gait Posture. Oct 2007;26(4):560-571. [doi: 10.1016/j.gaitpost.2006.12.018] [Medline: 17291764]
86. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound in Obstet & Gyne. Jul 2003;22(1):85-93. [doi: 10.1002/uog.122]
87. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Int J Nurs Stud. Aug 2010;47(8):931-936. [doi: 10.1016/j.ijnurstu.2009.10.001]
88. Bruton A, Conway JH, Holgate ST. Reliability: What is it, and how is it measured? Physiotherapy. Feb 2000;86(2):94-99. [doi: 10.1016/S0031-9406(05)61211-4]
89. Cereatti A, Gurchiek R, Mündermann A, et al. ISB recommendations on the definition, estimation, and reporting of joint kinematics in human motion analysis applications using wearable inertial measurement technology. J Biomech. Aug 2024;173:112225. [doi: 10.1016/j.jbiomech.2024.112225] [Medline: 39032224]
90. Ray LSS, Zhou B, Suh S, Lukowicz P. A comprehensive evaluation of marker-based, markerless methods for loose garment scenarios in varying camera configurations. Front Comput Sci. Apr 5, 2024;6. [doi: 10.3389/fcomp.2024.1379925]
91. Stenholm S, Chia CW, Simonsick EM, Ferrucci L. Gait pattern alterations in older adults associated with type 2 diabetes in the absence of peripheral neuropathy—results from the Baltimore longitudinal study of aging. Gait Posture. Oct 2011;34(4):548-552. [doi: 10.1016/j.gaitpost.2011.07.014]
92. Li S, Francisco GE, Zhou P. Post-stroke hemiplegic gait: new perspective and insights. Front Physiol. Aug 2, 2018;9:389766. [doi: 10.3389/fphys.2018.01021] [Medline: 30127749]
93. Mills K, Hunt MA, Ferber R. Biomechanical deviations during level walking associated with knee osteoarthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken). Oct 2013;65(10):1643-1665. [doi: 10.1002/acr.22015] [Medline: 23554153]
94. Li H, Hu S, Zhao R, et al. Gait analysis of bilateral knee osteoarthritis and its correlation with Western Ontario and McMaster University Osteoarthritis Index assessment. Medicina (B Aires). Oct 1, 2022;58(10):1419. [doi: 10.3390/medicina58101419]
95. Toda H, Maruyama T, Fujita K, Yamauchi Y, Tada M. Self-perception of the knee is associated with joint motion during the loading response in individuals with knee osteoarthritis: a pilot cross-sectional study. Sensors (Basel). Jun 10, 2021;21(12):4009. [doi: 10.3390/s21124009] [Medline: 34200714]
96. Fox AS, Bonacci J, McLean SG, Spittle M, Saunders N. A systematic evaluation of field-based screening methods for the assessment of Anterior Cruciate Ligament (ACL) injury risk. Sports Med. May 2016;46(5):715-735. [doi: 10.1007/s40279-015-0443-3] [Medline: 26626070]
97. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. Am J Sports Med. Apr 2005;33(4):492-501. [doi: 10.1177/0363546504269591] [Medline: 15722287]
98. Wang Y, Qi Y, Ma B, et al. Three-dimensional gait analysis of orthopaedic common foot and ankle joint diseases. Front Bioeng Biotechnol. 2024;12:1303035. [doi: 10.3389/fbioe.2024.1303035]

99. Martin RL, Davenport TE, Fraser JJ, et al. Ankle stability and movement coordination impairments: lateral ankle ligament sprains revision 2021. J Orthop Sports Phys Ther. Apr 2021;51(4):CPG1-CPG80. [doi: 10.2519/jospt.2021.0302] [Medline: 33789434]
100. Rethlefsen SA, Kay RM. Transverse plane gait problems in children with cerebral palsy. J Pediatr Orthop. Jun 2013;33(4):422-430. [doi: 10.1097/BPO.0b013e3182784e16] [Medline: 23653033]
101. Guarín DL, Wong JK, McFarland NR, Ramirez-Zamora A, Vaillancourt DE. What the trained eye cannot see: quantitative kinematics and machine learning detect movement deficits in early-stage Parkinson's disease from videos. Parkinsonism Relat Disord. Oct 2024;127:107104. [doi: 10.1016/j.parkreldis.2024.107104] [Medline: 39153421]
102. Horsak B, Kainz H, Dumphart B. Repeatability and minimal detectable change including clothing effects for smartphone-based 3D markerless motion capture. J Biomech. Oct 2024;175:112281. [doi: 10.1016/j.jbiomech.2024.112281] [Medline: 39163799]
103. Horsak B, Prock K, Krondorfer P, Siragy T, Simonlehner M, Dumphart B. Inter-trial variability is higher in 3D markerless compared to marker-based motion capture: implications for data post-processing and analysis. J Biomech. Mar 2024;166:112049. [doi: 10.1016/j.jbiomech.2024.112049] [Medline: 38493576]

## Abbreviations

**2D:** two-dimensional
**3DPoseNet:** 3D Pose Network
**CMJ:** countermovement jump
**CV:** computer vision
**CVB:** computer vision benchmark
**HMR:** human mesh recovery
**HPE:** human pose estimation
**HuMoR:** Human Motion Model for Robust Pose Estimation
**ICC:** intraclass correlation coefficient
**IMU:** inertial measurement unit
**JBI:** Joanna Briggs Institute
**MAE:** mean absolute error
**MBMCS:** marker-based motion capture system
**MLMCS:** markerless motion capture system
**MNT:** MoveNet Thunder
**PRISMA-ScR:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
**RegNetY:** customized regular network
**RGB:** red-green-blue
**RGB-D:** red-green-blue-depth
**RMSD:** root-mean-square deviation
**RMSE:** root-mean-square error
**RQ:** research question
**WHAM:** World-grounded Humans with Accurate Motion

*Edited by Tiffany Leung, Rachele Hendricks-Sturrup; peer-reviewed by Adebero Samuel, Brian Horsak; submitted 16.Jan.2026; initial comments to author 11.Mar.2026; final revised version received 01.Oct.2026; accepted 02.Oct.2026; published 27.Nov.2026*

*Please cite as:*
*Gholami B, Hollinger D, Zhai X, Liew BXW*
*Monocular Two-Dimensional and 3D Red-Green-Blue-Camera Marker-Less Motion Capture Systems for Human Lower-Limb Activities: Scoping Review*
*J Med Internet Res 2026;28:e91609*
*URL: https://www.jmir.org/2026/1/e91609*
*doi: 10.2196/91609*

© Behnam Gholami, David Hollinger, Xiaojun Zhai, Bernard Xian Wei Liew. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 27.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet

Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.