Journal of Biomechanics 157 (2023) 111747

Contents lists available at ScienceDirect

Journal of Biomechanics

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

Validity and reliability of two-dimensional video-based assessment to measure joint angles during running: A systematic review and meta-analysis

Gustavo Leporace [a,b,*], Leonardo Metsavaht [a,b], Felipe F. Gonzalez [a,b,c], Fabio Arcanjo de Jesus [d], Mariana Machado [a], Eliane Celina Guadagnin [a], Mansueto Gomes-Neto [d]

[a] Instituto Brasil de Tecnologias da Saúde (IBTS), Rio de Janeiro, Brazil
[b] Departamento de Diagnóstico por Imagem, Escola Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil
[c] Division of Sports Medicine, Midwest Orthopaedics at Rush, Rush University Medical Center, Chicago, IL, USA
[d] Programa de Pós-Graduação em Medicina e Saúde, Universidade Federal da Bahia (UFBA), Bahia, Brazil

ARTICLE INFO

Keywords:
Kinematics
Biomechanics
Motion capture
Running
Imaging, three-dimensional

ABSTRACT

Two-dimensional video analysis systems (2DVAS) are commonly used by clinicians and researchers to determine angles during running. The aim of this systematic review (PROSPERO: CRD42022322798) was to synthesize the literature on the criterion validity and reliability of 2DVAS for measuring angles during running compared to three-dimensional motion analysis systems (3DMAS). We searched for articles on MEDLINE/Pubmed, EMBASE, SciELO, and LILACS up to October/2022. We included studies that evaluated the validity of 2DVAS (when compared to 3DMAS) and/or the reliability of 2DVAS measurements of lower limb and trunk angles during running. Qualitative and quantitative analyses were performed. Seven hundred and five studies were found and 17 were included. Ten studies analysed criterion validity between 2DVAS and 3DMAS and the results ranged from poor to excellent, with most of the parameters assessed presenting poor or moderate validity. Inter-rater reliability of 2DVAS was assessed in nine studies and most of the parameters investigated had good to excellent reliability. Intra-rater reliability (between-day processing) of angular running parameters – investigated in ten studies – was considered excellent for most of the parameters analysed. Inter-session reliability was assessed in three studies and was defined as good or excellent for most of the variables assessed. 2DVAS is a reliable method for measuring joint angles during running. However, the validity of 2DVAS compared to 3DMAS ranges from low to moderate for most running parameters. Therefore, based on the available evidence, caution should be taken when applying 2DVAS, particularly for frontal and transverse plane angles.

1. Introduction

Running biomechanics has become an important area of study because movement pattern can be considered a surrogate measure of joint and muscle overload (Noehren et al., 2007) and has been associated with different injuries in runners (Vannatta et al., 2020). Three-dimensional motion analysis systems (3DMAS), which employ two or more synchronized cameras to track markers in different planes, are considered the gold standard for quantitative kinematic biomechanical analyses in clinical practice. However, the high costs make them inaccessible to many clinical sports medicine settings. Conversely, two-dimensional video analysis systems (2DVAS), utilizing a single camera to track markers in one plane, are cost-effective, user-friendly, and frequently employed in clinical settings, both quantitatively and qualitatively (Mousavi et al., 2020). Nevertheless, before 2DVAS can be widely recommended, it is imperative to investigate whether 2DVAS produces measurements that are valid when compared to 3DMAS. Moreover, the 2DVAS method is often rater-dependent, given that analyses are performed manually on the recorded videos (Peebles et al., 2021; Weber and McClinton, 2020). Consequently, in addition to the validity of 2DVAS, understanding the intra- and inter-rater reliability of 2DVAS in measuring angles during running is also critical.

A systematic review with meta-analysis to test the validity and reliability of 2DVAS during squatting and landing tasks was conducted and

* Corresponding author at: Instituto Brasil de Tecnologias da Saúde, Rua Visconde de Pirajá, 407, Room 905, Ipanema, Rio de Janeiro, RJ, Brazil.
E-mail address: gustavo@biocinetica.com.br (G. Leporace).

https://doi.org/10.1016/j.jbiomech.2023.111747
Accepted 27 July 2023
Available online 29 July 2023
0021-9290/

concluded that the evidence does not support the use of 2DVAS for accurately measuring trunk and lower extremity frontal plane angles (Lopes et al., 2018). In a systematic review, limited evidence to suggest that 2DVAS is valid to measure step rate and foot strike pattern during running was found (de Oliveira et al., 2019). However, the authors found strong evidence supporting the reliability of 2DVAS for these two variables. With respect to angular running parameters, a systematic review recently published (Hensley et al., 2022) showed that the intra-rater reliability of 2D video analysis ranged from moderate to excellent, while the inter-rater reliability and validity varied from poor to excellent. However, a meta-analysis was not performed in this study. Thus, this systematic review with meta-analysis aimed to analyze and synthesize qualitatively and quantitatively the published studies that have investigated the criterion validity and reliability of 2DVAS to determine angles during running.

2. Methods

This systematic review with meta-analysis was completed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021) and was registered in the PROSPERO (CRD42022322798).

2.1. Eligibility criteria

This systematic review included studies that investigated the criterion validity and/or reliability of 2DVAS to measure trunk and lower-extremity angles during running. Studies were eligible if they met the following criteria: (a) included adult participants (mean age ⩾18 years); (b) analysed the lower-extremity and/or trunk kinematics during running; (c) analysed the intra-rater and/or inter-rater reliability (intra- or inter-session) and/or the criterion validity of the 2DVAS to measure lower-extremity and/or trunk angles during running; (d) criterion validity studies should have analysed the validity of 2DVAS by comparison to a 3DMAS gold standard. Studies that only evaluated gait or that utilized only 3DMAS were excluded.

2.2. Search methods and study selection

We searched for references on MEDLINE/Pubmed, EMBASE, Scientific Electronic Library Online (SciELO), and Lilacs up to October 2022 without language restrictions. We used a standard protocol for this search and, whenever possible, a controlled vocabulary (Mesh term for MEDLINE). In the search strategy, we used the following groups of keywords and their synonymous: biomechanics, running/runners, reliability, and validity. The strategy developed by (Higgins and Green, 2006) was used for the identification of studies in PUBMED. The search strategy for MEDLINE via PUBMED is presented in Supplementary Material 1. To identify the studies in the other databases we adopted a search strategy using similar terms, and they are available upon request. We checked the reference lists of the articles included in this systematic review to identify other potentially eligible studies.

The list of titles and abstracts from each database were independently evaluated by two authors. If at least one of the reviewers considered one study eligible, the full text was obtained for complete analysis. Then, two reviewers independently assessed the full text of the selected studies to verify if they met the eligibility criteria. Discrepancies were solved by a third reviewer.

2.3. Data extraction

Two authors independently extracted data from the studies using the same structured form, created specifically for this study. The following data were extracted: characteristics of the participants, equipment utilized, characteristics of data collection, methods to determine the reliability and the validity, biomechanical parameters determined, and results found.

2.4. Methodological quality of studies

The methodological quality of studies included in this systematic review and meta-analysis was scored by two researchers using the clinical evaluation tool (CAT) scale (Brink and Louw, 2012). The CAT scale contains 13 evaluation items. Five of the 13 items refer to validity and reliability issues, four of the 13 items refer to validity issues only, and the other four of the 13 items refer to reliability issues only. Each study was classified for each item as "yes" when information was described in sufficient detail or "no" when there was not enough information for clarification. A final percentage (%) evaluation column was added based on the items that each study achieved [% = (Items "yes" x 100)/number of items scored]. Studies were considered as having high quality if they scored equal to or above 70%.

2.5. Statistical analysis

For the meta-analysis, when the study reported results from either the left and right legs or both the dominant and non-dominant legs, we extracted data from the right or the dominant leg as these were more commonly reported in accordance with Lopes et al. (Lopes et al., 2018). For reliability studies, when authors reported results of more than one rater, data from the most experienced rater or the first rater were used (Lopes et al., 2018). Furthermore, when more than two test sessions were reported for reliability, the results comparing the first two sessions were used for analysis (Lopes et al., 2018). When two conditions were presented (e.g., different shoes, different speeds/intensities) and when the angles were determined in two ways, the mean of both was utilized in the analyses. For reliability, if two ratings were performed, the mean of both was utilized.

The Comprehensive Meta-Analysis software (v4; Biostat Inc., Englewood, NJ, USA) was used for the meta-analysis. A random-effects model was used, and the heterogeneity was assessed through the inconsistency test (I^2: low heterogeneity: $I^2 \leq 25\%$; moderate: $26\% < I^2 \leq 50\%$; and high: $I^2 > 50\%$). The level of significance considered was $p < 0.05$.

Regarding the reliability indicators used in this systematic review, we started with the intraclass correlation coefficients (ICC), which is usually the main indicator of reliability (Jamovi, 2020). Some studies reported the r or the weighted kappa. Thus, for reliability, we entered the mean ICC, r or weighted kappa value, and the sample size in the analyses. The Pearson Product Correlation Coefficients (r) or ICC values and the sample size were used for criterion validity.

According to Koo and Li (Koo and Li, 2016), an ICC was classified as follows: <0.5 poor (identified in red in Supplementary Material 3); between 0.5 and 0.75: moderate (identified in orange in Supplementary Material 3); between 0.75 and 0.9: good (identified in yellow in the Supplementary Material 3); >0.90 excellent (identified in green in the Supplementary Material 3). Correlations were classified as follows: <0.1: trivial (identified in red in Supplementary Material 3); 0.1–0.3 small (identified in red in Supplementary Material 3); 0.3–0.5 moderate (identified in orange in Supplementary Material 3); 0.5–0.7 large (identified in yellow in the Supplementary Material 3); 0.7–0.9 very large (identified in green in the Supplementary Material 3); >0.9 nearly perfect (identified in green in the Supplementary Material 3) (Hopkins et al., 2009). Weighted Kappa was classified as 0–0.40 fair to poor (identified in red in Supplementary Material 3), 0.41–0.60 moderate (identified in orange in Supplementary Material 3), 0.61–0.80 substantial (identified in yellow in Supplementary Material 3), and 0.81–1.00 almost perfect (identified in green in the Supplementary Material 3) (Landis and Koch, 1977).

3. Results

3.1. Yield

The initial search identified 705 studies, from which 70 full texts were considered potentially relevant. From these, 17 studies were included. Fig. 1 shows the PRISMA flow diagram of studies assessed and included in this review.

3.2. Methodological quality

Thirteen studies (71%) were considered as having high quality (above 70%; mean of all studies: 80%; mean of the high-quality studies: 85%). The results of the Risk of Bias Evaluation performed according to the CAT scale are presented in Table 1.

3.3. Study characteristics

The number of participants in the studies included ranged from 9 (Macpherson et al., 2016) to 32 (Grau et al., 2000) (Table 2). The mean age of participants ranged from 18.7 (Dingenen et al., 2018b) to 36.2 (Murray et al., 2018) years old (Table 2).

Ten studies reported data on intra-rater reliability of the 2DVAS (intra-session; analysis of the same video) and three reported data on intra-rater inter-session reliability. Nine analysed the inter-rater reliability of the 2DVAS, and ten reported data on the criterion validity between the 2DVAS and 3DMAS (Table 2). One of the included studies (Pipkin et al., 2016) assessed the intra- and inter-rater reliability through a qualitative (categorical) analysis.

Most studies analysed the lower limb kinematics, and four studies assessed the trunk kinematics. Sagittal plane parameters of running and frontal plane parameters of running were investigated in eleven studies, while transverse plane variables were analysed in three studies.

Treadmill running and overground running were assessed in 14 and three studies, respectively. Self-selected speed was utilized in 12 studies, while a fixed speed or a normalized speed was utilized in five studies (range: 8.5 to 15.3 km/h). The 2D assessment was performed using one or two cameras, mostly positioned in a sagittal plane or a frontal plane view. Sampling frequency ranged from 25 Hz to 300 Hz. Two studies utilized depth-sensing cameras and most of the studies determined the 2D parameters using markers. The general characteristics of the participants and the methods of the included studies are summarized in Table 2.

3.4. Criterion validity

The criterion validity of 2D measurements was investigated in ten of the included studies (Dingenen et al., 2018b; Macpherson et al., 2016; Maykut et al., 2015; McClay and Manal, 1998; Mousavi et al., 2020; Neal et al., 2020; Ota et al., 2021; Peebles et al., 2021; Pfister et al., 2014;

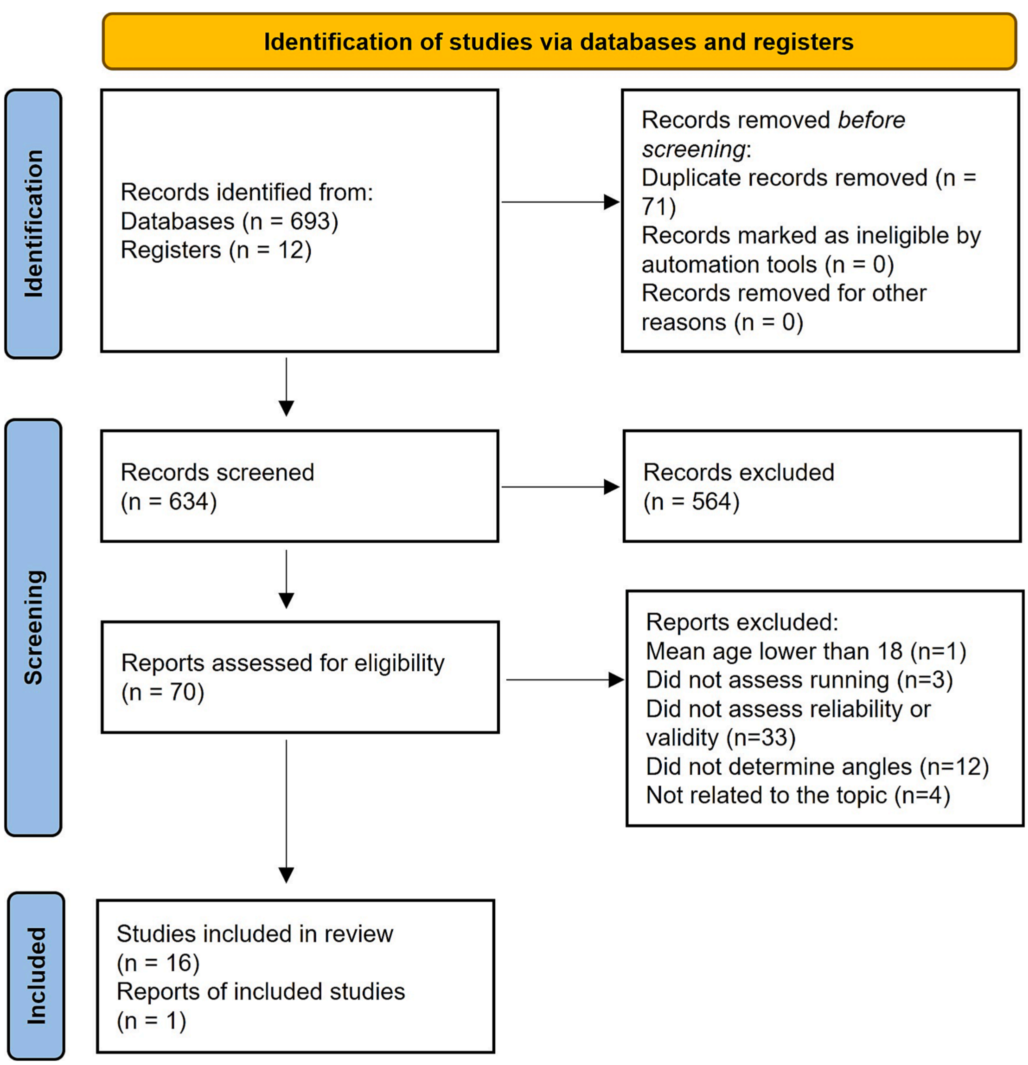

Fig. 1. PRISMA flow diagram illustrating the search and selection process of studies included in this systematic review and meta-analysis.

Table 1
Methodological quality of the included studies.

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Atkins et al. (Atkins et al., 2014)	✔	✔	n/a	✖	✖	✔	n/a	✔	n/a	✔	n/a	✔	✔	78%
Damsted et al. (Damsted et al., 2015)	✔	✔	n/a	✔	✖	✖	n/a	✔	n/a	✔	n/a	✔	✔	78%
Dingenen et al. (Dingenen et al., 2018b)	✔	✖	✔	✔	✖	✖	✔	✔	✔	✔	✔	✔	✔	77%
Dingenen et al. (Dingenen et al., 2018a)	✔	✖	n/a	n/a	✖	✖	n/a	✔	n/a	✔	n/a	✔	✔	63%
Grau et al. (Grau et al., 2000)	✔	✖	n/a	n/a	✖	✖	n/a	✔	n/a	✔	n/a	✔	✔	63%
Machperson et al. (Macpherson et al., 2016)	✔	✖	✔	n/a	n/a	n/a	✔	n/a	✔	✔	✔	✔	✔	89%
Maykut et al. (Maykut et al., 2015)	✔	✖	✔	n/a	✖	✖	✔	✔	✔	✔	✔	✔	✔	75%
McClay and Manal (McClay and Manal, 1998)	✖	✖	✔	n/a	n/a	n/a	✔	n/a	✔	✖	✔	✔	✔	67%
Mousavi et al. (Mousavi et al., 2020)	✔	✔	✔	✔	✖	✖	✔	✔	✔	✔	✔	✔	✔	85%
Murray et al. (Murray et al., 2018)	✔	✔	n/a	✔	✔	✖	n/a	✔	n/a	✔	n/a	✔	✔	89%
Neal et al. (Neal et al., 2020)	✔	✖	✔	✖	✖	✖	✔	✔	✔	✔	✔	✔	✔	69%
Ota et al. (Ota et al., 2021)	✔	✖	✔	n/a	n/a	n/a	✔	n/a	✔	✔	✔	✔	✔	89%
Peebles et al. (Peebles et al., 2021)	✔	✖	✔	n/a	n/a	n/a	✔	n/a	✔	✔	✔	✔	✔	89%
Pfister et al. (Pfister et al., 2014)	✔	✖	✔	n/a	n/a	n/a	✔	n/a	✔	✔	✔	✔	✔	89%
Pipkin et al. (Pipkin et al., 2016)	✖	✔	n/a	✔	✔	✔	n/a	✔	n/a	✔	n/a	✔	✔	89%
Reinking et al. (Reinking et al., 2018)	✔	✔	n/a	✔	✔	✔	n/a	✔	n/a	✔	n/a	✔	✔	100%
Weber and McClinton (Weber and McClinton, 2020)	✔	✔	✔	✔	✔	✖	✔	✔	✔	✖	✖	✔	✔	77%

✔ yes; ✖ no; n/a: not applicable.

Weber and McClinton, 2020). The results ranged from poor to excellent. However, most of the parameters assessed presented poor or moderate criterion validity (28 from 42 assessments). The detailed results found by each study and the classification of the results found are presented in Supplementary Materials 2 and 3.

The most assessed variables were the hip and knee sagittal plane angles and hip frontal plane angles, measured at different instants during the gait cycle. Trunk angles were assessed in two studies (Macpherson et al., 2016; Weber and McClinton, 2020). For both the frontal and the transverse plane, the correlation between 2D and 3D was very large or nearly perfect. In one of the studies (Weber and McClinton, 2020) an overhead camera setup was employed and they defined the trunk segment utilizing markers on the acromioclavicular joints. Meanwhile, in the other study (Macpherson et al., 2016), a depth-sensing camera (KinectTM for Windows, Version 1, Microsoft, USA) was utilized. Notably, neither of these methodologies validated 2D transverse trunk angles against traditional 3D Trunk Euler Angles, which are based on 3D joint coordinate systems as defined by the International Society of Biomechanics (ISB) (Wu et al., 2002).

Pelvic obliquity was assessed in four studies (Dingenen et al., 2018b; Macpherson et al., 2016; Maykut et al., 2015; Ota et al., 2021), with moderate or large correlation (0.3 to 0.7) and poor ICC (<0.5). One study assessed the pelvis transverse plane ROM and found a large correlation ($r = 0.5$ to 0.7) between 2D and 3D (Macpherson et al., 2016).

Hip angles were assessed in the sagittal and frontal planes. ICCs ranged from poor to good and correlation ranged from trivial or small to large. None of the parameters assessed presented excellent or very large criterion validity.

Sagittal plane knee angles criterion validity ranged from poor to excellent, with most of the parameters presenting poor or moderate validity. Frontal plane knee angle was assessed by one study, with a trivial or small correlation between 2D and 3D. Considering the ankle, angles were assessed only in the sagittal plane and the ICCs ranged from poor to good. Foot angles criterion validity in the sagittal and frontal planes ranged from moderate to excellent.

For the same parameters measured at the same run instant and the same parameter measured at different running instants, the meta-analysis revealed that there was a trivial to nearly perfect relationship between the 2DVAS and 3DMAS (Table 3). Meta-analysis was performed for 11 variables. For 27% of them, the pooled correlation was very large or nearly perfect, for 18% it was large, for 9% it was moderate, and for 45% it was trivial or small. For most parameters, there was a low heterogeneity between the studies.

3.5. Inter-rater reliability

Inter-rater reliability for angular parameters during running was of interest in nine of the included studies (Atkins et al., 2014; Damsted et al., 2015; Dingenen et al., 2018b; Mousavi et al., 2020; Murray et al., 2018; Neal et al., 2020; Pipkin et al., 2016; Reinking et al., 2018; Weber and McClinton, 2020). The reliability ranged from poor to excellent, with most parameters (21 of 37) showing excellent inter-rater reliability. However, there was a substantial number of parameters presenting poor to good reliability (16 of 37). The results of each study and the classification of ICC, correlation, and weighted kappa are presented in Supplementary Materials 2 and 3.

The qualitative analysis performed in the study of Pipkin et al. (Pipkin et al., 2016) showed that the weighted kappa was moderate (from 0.41 to 0.60) for most of the variables analysed.

Inter-rater reliability of frontal plane parameters was most common among the studies, followed by the sagittal and transverse angles and they presented similar reliability (Supplementary Materials 2 and 3).

For the same parameters measured at the same run instant and the same parameter measured at different running instants, the meta-

Table 2
Characteristics of the included studies.

Study	Participants		Methods					
	n	Characteristics	Running assessment	Focus	2D assessment	3D assessment	Reliability	Parameters assessed
Ota et al., 2021 (Ota et al., 2021)	24	26.1 ± 4.6y; 7F/17 M; Healthy; H: 167.9 ± 7.7 cm; M: 61.1 ± 10.8 kg	Treadmill; Fixed speed: 8.5 km/h; 4 cycles	Validity	2 digital cameras; Position: back and right; 60 Hz; Analysis: OpenPose (human pose tracking algorithms); Cycle defined visually by heel contact	8 cameras (Vicon); 60 Hz; Cycle defined visually by heel contact	NA	Peak: pelvic elevation, pelvic depression, hip flex, hip ext, hip abd, hip add, knee flex, knee ext, ankle dorsiflexion, ankle plantar flex; ROM: pelvic elevation–depression, hip flex-ext, hip abd-add, knee flex-ext, ankle dorsiflexion-plantar flex (determined during each cycle; LL: NI)
Peebles et al., 2021 (Peebles et al., 2021)	20	21.9 ± 2.4y; 10F/10 M; Healthy uninjured; Recreationally active; H: 175 ± 8 cm; M: 70.5 ± 12.6 kg	Treadmill; Preferred speed (mean: 9.79 km/h); 1 min	Validity	1 video camera (Hero 6 Black, GoPro); Position: right; 120 Hz; Analysis: Markers automatically tracked by Automated video Analysis for Dynamic Systems; IC and TO defined by 1st and 2nd peak knee ext	10 cameras (Oqus 700, Qualisys); 120 Hz; IC and TO defined by 1st and 2nd peak knee ext	NA	Knee, ankle, and foot angle at IC (SP); Peak knee flex during stance; Knee flex ROM (IC to peak) (LL: NI)
Neal et al., 2020 (Neal et al., 2020)	21	32.1 ± 12.9y; 11F/10 M; With retropatellar or peripatellar pain; H: 169.1 ± 45.2 cm; M: 69.8 ± 19.6 kg	Overground (indoors); Self-selected speed (mean: NI); Neutral shoes (Asics Nimbus); 5 trials of 10 *m*	Validity Intra-rater Inter-rater	2 smartphone cameras (iPhone 6, Apple Corporation); 240 Hz; Position: SP (right or left to record the knee with pain) and FP; Analysis: Hudl Technique	4 cameras (CX-1 Codamotion); 200 Hz; IC and TO defined from force plate	Intra: 1 rater; twice; 24 h; 1st trial Inter: 2 raters; 1st trial; 1st analysis of rater 1 vs. rater 2 Exp: NI	Peak hip add and knee flex during stance (one LL per subject: the limb with pain; if bilateral pain, the LL with the most strong pain; if the pain was similar, the preferred LL)
Mousavi et al., 2020 (Mousavi et al., 2020)	20	28 ± 4y; 20F; Recreational runners; Without injuries; H: 168 ± 5 cm; M: 61 ± 6 kg	Treadmill; Self-selected speed (mean: NI); Own shoes; 10 consecutive strides from a 25 s analysis	Validity Intra-rater Inter-rater	2 video cameras (Canon, G16); 60 Hz; Position: SP (right) and FP (back); Analysis: Coach's Eye; IC and TO visually identified by heel and toe contact	3 cameras (Bonita, Vicon); 200 Hz; IC and TO defined by machine learning using kinematics	Intra: 1 rater; twice; 5d; 1st trial Inter: 2 raters; 1st trial Exp: >6y	Hip, knee and ankle angle at IC and TO (SP); Rearfoot in/eversion angle at IC and TO (FP) [LL: R (SP) and R and L (FP)]
Weber and McClinton, 2020 (Weber and McClinton, 2020)	16	25 ± 4.5y; 8F/8M; Runners (competing or with a history of competition in a collegiate team or recreational); Without injuries; H: 174 ± 6.5 cm; M: 65.9 ± 6.7 kg	Treadmill; Self-selected speed (mean: 11.9 km/h) to run 8 km; Preferred shoes; 12–15 s of analysis of 5 consecutive strides	Validity Intra-rater Inter-rater	1 camera (Hudl Technique on an Apple Ipad Air 2, front camera); 60 Hz; Position: superior (TP); Analysis: Kinovea to identify peak angles and ImageJ to measure the angle	10 cameras (Motion Analysis Corporation); 120 Hz	Intra*: 2 raters; twice; 7d Inter*: 2 raters; twice; 7 days Exp: 1 with 1y and 1 with no exp	Peak right and left upper trunk rotation (during the stride)[@]
Dingenen et al., 2018 (Dingenen et al., 2018b)	15	18.7 ± 1.62y; 9F/6M; Elite athletics athletes; Without injury and pain; H: 178 ± 8 cm; M: 67.31 ± 11.02	Overground; Preferred speed (mean: NI); 4 trials of 10 m; Neutral shoes (Kelme Indoor Copa)	Validity Intra-rater Inter-rater	1 digital video camera (Sony); 50 Hz; Position: frontal plane; Analysis: Dartifsh 7.0	10 cameras (T-10, Vicon); 100 Hz; IC and TO defined by a force plate	Intra*: 2 raters; twice; 7d Inter*: 2 raters; twice; 7d Exp: NI	3D: Contralateral pelvic drop angle; Hip adduction; Knee abduction 2D: Contralateral pelvic drop angle; Femoral adduction angle; Knee valgus angle (determined at the deepest landing position, at midstance) (LL: R)
Macpherson et al., 2016 (Macpherson et al., 2016)	9	29.2 ± 4.2y; 9 M; Without injuries; H: 182.9 ± 7.3 cm; M: 84.5 ± 10.4 kg	Treadmill; Speed defined as 70% (10.9–14 km/h) and 90% (14–18 km/h) of the 30–15 Intermittent Fitness Test; 180 s each speed	Validity	1 depth-sensing camera (KinectTM for Windows, v1); 30 Hz upsampled to 100 Hz; Position: rear; Analysis: NI; Beginning of the gait cycle identified at every 2nd point of inflexion on the	6 cameras (MX13, Vicon); 100 Hz	NA	Pelvis and trunk ROM (FP); Pelvis and trunk ROM (TP) (during the stride)

(*continued on next page*)

Table 2 (*continued*)

Study	Participants		Methods					
	n	Characteristics	Running assessment	Focus	2D assessment	3D assessment	Reliability	Parameters assessed
					superior-inferior time-series for the pelvis			
Maykut et al., 2015 (Maykut et al., 2015)	24	19.9 ± 1.3y; 10F/14 M; Collegiate cross-country runners; Not under medical supervision; H: 167.8 ± 23.1 cm; M: 59.7 ± 6.6 kg	Treadmill; Self-selected speed (mean: NI) for a 20-min easy run; 20 consecutive steps analysed	Validity Intra-rater	Camera type and position: NI; 60 Hz; Analysis: Dartifsh	10 cameras (Eagle; Motion Analysis Corporation); 240 Hz	Intra: 1 rater; twice; 7d Exp: NI	Peak contralateral pelvic drop (FP), Peak hip adduction angle (FP), Peak knee abduction angle (FP) (LL: L and R, at midstance)
Pfister et al., 2014 (Pfister et al., 2014)	20	27.4 ± 10.0y; 11F/9M; Healthy, able for walking/jogging; Regularly participating in moderate-to-vigorous activity; H: 169.4 ± 10.9 cm; M: NI	Treadmill; Speed: 8.85 km/h; 10–40 steps per leg	Validity	1 camera (Xbox KinectTM Microsoft); 30–37 Hz; Position: left to the subject, 45° to the treadmill; Analysis: Brekel Kinect TM	10 cameras (MX, Vicon); 120 Hz	NA	Maximum hip flex, hip ext, knee flex, and knee ext (LL: R and L, for each stride)
McClay & Manal, 1998 (McClay and Manal, 1998)	18	18–40 years; Recreational runners; 9 with normal eversion and 9 with excessive rearfoot eversion during running	Treadmill; Speed: 12 km/h; 5 contact phases	Validity	NI	4 cameras (Motion Analysis Corp.); 200 Hz	NA	Eversion at IC, eversion at TO, peak eversion during stance (LL: R, except when the L had greater eversion)
Murray et al., 2018 (Murray et al., 2018)	28	36.2 ± 11.7y; 11F/17 M; Recreational runners participating in a 12-km race; Without injuries; H: 172.1 ± 8.6; M: 73.6 ± 12.5	Overground (outdoors); Perceived race pace (mean: 15.3 km/h); 1 footstrike of three 15 m trials before and after a race; Own running shoes	Intra-rater Inter-rater	1 digital camera (Cyber-shot DSC-RX10 II); 240 Hz; Position: sagittal right; Analysis: Siliconcoach Pro8 (The Tarn Group); IC visually identified	NA	Intra*: 2 raters; twice; 7d Inter*: 2 raters; twice; 7d Exp: >2y	Foot-strike angle (LL: R)
Reinking et al., 2018 (Reinking et al., 2018)	10	23.0 ± 4.6y; 5F/5M; Recreational and competitive runners (high school and college cross country runners); Without injuries; H: 170.6 ± 7.4 cm; M: 59.8 ± 8.5 kg	Treadmill; Self-selected speed (mean: NI); 4 min	Intra-rater Inter-rater	1 camera (EX FH25, Casio America Inc.); 240 Hz; Position: right SP, left SP, full posterior (FP), and posterior leg/heel (FP)	NA	Intra[#]: 6 raters; twice; at least 7d Inter[&]: 6 raters; twice; at least 7d Exp: 1 with 12y, 1 with 10y and 4 with no exp	Angle of the shoe to the treadmill at IC; Angle of the lower leg at IC; Knee flex at IC, midst, and midswing; Hip add angle at IC and midst; Rearfoot angle at IC and midst (LL: R and L)
Atkins et al., 2014 (Atkins et al., 2014)	10	26 ± 3.5y; 5F/5M; Recreational or competitive runners, except one subject; Healthy; BMI: 23.8 ± 1.5 kg/m^2	Treadmill; Preferred speed plus 10%; 2 strides from each of 2 videos of 30 s	Intra-rater Inter-rater	1 camera (Panasonic HDC-TM 900); 60 Hz; Position: posterior (FP); Analysis: using a flat-screen computer monitor and the iHandy Level Free application on an Apple iPhone; IC and peak angles visually identified	NA	Intra: 6 raters; twice; 7d Inter[§]: 6 raters Exp: no exp	Contralateral pelvic drop and hip adduction at IC; Contralateral pelvic drop and hip adduction excursion (stance phase; from IC to maximum) (LL: R and L)
Pipkin et al., 2016 (Pipkin et al., 2016)	15	Age: NI; 7F/8M; 10 with a running-related injury and 5 without injury	Treadmill; Speed: self-selected (11.4 km/h); one stride from each of 3 videos of 5 s	Intra-rater Inter-rater	1 camera (Casio EX-FH25); 120 Hz; Position: FP (whole-body and feet) and SP (whole-body); Analysis: visual and categorically rated	NA	Intra: 3 raters; twice; 7–10d Inter: 3 raters; twice; 7–10d Exp: at least 1y Stride videos and still-frame images (IC SP, midstance SP, and FP, mid-flight SP) defined by another investigator	Trunk sidebend, lateral pelvic drop, rearfoot position, and forefoot position (FP, at midst); Tibial inclination and knee flex angle (SP, IC); Knee flex angle, and ankle dorsiflexion angle (SP, midst); Forward trunk lean (SP, full gait cycle) (LL: NI)
Damsted et al., 2015 (Damsted et al., 2015)	25	35 ± 9y; 13F/12 M; Recreational runners; Without injuries; H: 175.8 ±	Treadmill; Self-selected speed (10.14 km/h); Own shoes; 30 s	Intra-rater (intra- and	1 video camera (Exilim EX-F1, Casio); 300 Hz; Position: SP; 30 s;	NA	Intra[+]: 2 raters, twice; at least 14d (between day processing and	Hip flexion angle at foot strike; Knee flexion angle at foot strike (LL: L)

(continued on next page)

Table 2 (*continued*)

Study	Participants		Methods					
	n	Characteristics	Running assessment	Focus	2D assessment	3D assessment	Reliability	Parameters assessed
		10.5 cm; M: 76.7 ± 19 kg		inter-session) Inter-rater (intra- and inter-session)	Analysis: Kinovea 0.8.15		between day test–retest) Inter+: 2 raters, twice; at least 14d (between day processing and between day test–retest) Exp: experienced Test-retest: 7d	
Grau et al., 2000 (Grau et al., 2000)	32	FP: healthy 30.4 ± 4.3y; CA 36.8 ± 6.4y. SP: healthy 30.2 ± 4.1y; CA 36.1 ± 6.9y; 32 M; Runners; FP: healthy 180.6 ± 6.8 cm and 69.1 ± 6.0 kg; CA 178.2 ± 6.9 cm and 73.9 ± 9.8 kg. SP: healthy 180.0 ± 6.4 cm and 70.6 ± 6.8 kg; CA 178.8 ± 6.7 cm and 73.0 ± 9.2 kg	Treadmill; Speed: 80% of the individual anaerobic threshold; Two shoes: own and neutral (Nike Talaria); 10-step cycles	Intra-rater (inter-session)	1 camera (Panasonic F10); 25–50 Hz;	NA	Intra-rater (between day test–retest): two measurement days (mean of 7d between them; range 5–9d) Number of raters: NI Exp: NI Test-retest: 7d (range 5–9d)	Maximum knee flex and minimum ankle angle (stance phase) Total pronation and total hindfoot pronation (LL: NI)
Dingenen et al., 2018 (Dingenen et al., 2018a)	21	28.1 ± 8.3 years; 12F/9M; Recreational runners; Without injuries; H: 172.9 ± 9.7 cm; M: 67.0 ± 12.2 kg	Treadmill; Speed: preferred (mean: 10.2 km/h); Own shoes; 30 s	Intra-rater (inter-session)	2 cameras (tablets iPad Air); 120 Hz; Position: FP and SP (left and right); Analysis: Kinovea 0.8.15; Deepest landing position, IC, and midstance defined visually	NA	Two raters; twice; at least 7 days (between day test–retest) Experience: NI One rater assessed 10 and the other assessed 11 participants Test-retest: 7d	Lateral trunk position angle, contralateral pelvic drop angle, femoral add angle, hip add angle, ankle dorsiflexion angle, and knee flex angle at midst; foot inclination angle and tibia inclination angle at IC (LL: R and L)

F: female; M: male; y: years; d: days; H: height; W: weight; ROM: range of motion; Flex: flexion; Ext: extension; Abd: abduction; Add: adduction; SP: sagittal plane; FP: frontal plane; TP: transverse plane; NA: not applicable; IC: initial contact; TO: toe off; NI: not informed; Exp: experience; *both assessments of rater 1 and rater 2 were utilized; #data from the most experienced rater; &data from the comparison between the two most experienced raters; +analyses performed within and between day; LL: lower limb; R: right; L: left; @data from one single trial; BMI: Body Mass Index; Midst: midstance; §data from the 2nd analysis; CA: chronic achillodynia.

analysis revealed that there was very large to nearly perfect inter-rater reliability (Table 4). Meta-analysis was performed for nine parameters. For 44% of them, the pooled correlation was very large and for 56% it was nearly perfect. For most parameters, there was high heterogeneity between the studies.

3.6. Intra-rater reliability (between day processing)

Intra-rater reliability (between day processing) of running angular parameters was investigated in ten studies (Atkins et al., 2014; Damsted et al., 2015; Dingenen et al., 2018b; Maykut et al., 2015; Mousavi et al., 2020; Murray et al., 2018; Neal et al., 2020; Pipkin et al., 2016; Reinking et al., 2018; Weber and McClinton, 2020). Intra-rater reliability was considered excellent for most of the parameters analysed (31 from 42). The results found by each study and the classification of the results found are presented in Supplementary Materials 2 and 3.

For the same parameters measured at the same run instant and the same parameter measured at different running instants, the meta-analysis revealed that there was very large to nearly perfect inter-rater reliability (Table 5). Meta-analysis was performed for nine parameters. For 44% of them, the pooled correlation was very large and for 56% it was nearly perfect. For most parameters, there was high heterogeneity between the studies.

3.7. Intra-rater reliability (inter-session)

Intra-rater inter-session reliability was assessed in three of the included studies (Damsted et al., 2015; Dingenen et al., 2018a; Grau et al., 2000). The reliability was good or excellent for most of the variables assessed. A few parameters had moderate reliability. The results found by each study and the classification of the results found are presented in Supplementary Materials 2 and 3.

4. Discussion

The main findings of our systematic review and meta-analysis were the evidence suggesting that the reliability of 2DVAS for determining lower-extremity angles during running varies from good to excellent. Criterion validity showed poor or moderate results for most parameters, except for knee and ankle angles in the sagittal plane during the stance phase. Intra-rater reliability demonstrated excellent agreement, followed by good to excellent inter-rater reliability and good intra-rater inter-session reliability.

Kinematic analysis of running is considered an important part of a global functional assessment and holds significant clinical relevance as it allows for the identification of factors contributing to injury occurrence and facilitates the development of targeted interventions. While three-dimensional motion analysis systems (3DVAS) are considered the gold standard in kinematic assessment, their high cost limits their widespread use in clinical practice, leading to the common utilization of two-

Table 3
Criterion validity meta-analyses for the parameters with at least two studies.

	Studies	Correlation	Classification	p	95% CI	Fisher's Z	I²
Same Parameter							
Pelvic obliquity ROM	Ota et al. (2021) Macpherson et al. (2016)	0.25		0.368	-0.287-0.662	0.251	39%
Peak hip flexion (cycle)	Ota et al. (2021) Pfister et al. (2014)	0.01		0.947	-0.364-0.387	0.013	35%
Peak hip extension (cycle)	Ota et al. (2021) Pfister et al. (2014)	0.10		0.539	-0.215-0.395	0.100	0%
Peak knee flexion (cycle)	Ota et al. (2021) Pfister et al. (2014)	0.91		<0.001*	0.807-0.956	1.505	32%
Peak knee extension (cycle)	Ota et al. (2021) Pfister et al. (2014)	0.64		0.149	-0.267-0.946	0.761	90%
Peak knee flexion (stance)	Peebles et al. (2021) Neal et al. (2020)	0.38		0.018	0.068-0.623	0.399	0%
Knee SP angle (IC)	Mousavi et al. (2020) Peebles et al. (2021)	0.63		<0.001*	0.384-0.792	0.741	0%
Ankle SP angle (IC)	Mousavi et al. (2020) Peebles et al. (2021)	0.71		<0.001*	0.449-0.855	0.880	28%
Similar Parameters							
Peak pelvic drop (cycle) Peak contralateral pelvic drop (midstance)	Ota et al. (2021) Maykut et al. (2015)	0.19		0.203	-0.105-0.461	0.197	0%
Peak hip adduction (cycle) Peak hip adduction (stance) Peak hip adduction (midstance)	Ota et al. (2021) Neal et al. (2020) Maykut et al. (2015	0.37		0.044*	0.011-0.644	0.388	55%
Knee flex-ext ROM (stride/step) Knee flex ROM (IC to peak)	Ota et al. (2021) Peebles et al. (2021)	0.86		0.007*	0.343-0.976	1.285	88%

*p<0.05; IC: initial contact; SP: sagittal plane; FP: frontal plane; Red color: correlation <0.1 or 0.1-0.3 (trivial or small, respectively); Orange color: correlation 0.3-0.5 (moderate); Yellow color: correlation 0.5-0.7 (large); Green color: correlation 0.7-0.9 or >0.9 (very large or nearly perfect, respectively.

*p < 0.05; IC: initial contact; SP: sagittal plane; FP: frontal plane; Red color: correlation < 0.1 or 0.1–0.3 (trivial or small, respectively); Orange color: correlation 0.3–0.5 (moderate); Yellow color: correlation 0.5–0.7 (large); Green color: correlation 0.7–0.9 or > 0.9 (very large or nearly perfect, respectively.

Table 4
Inter-rater reliability meta-analyses for the parameters with at least two studies.

	Studies	Correlation	Classification	p	95% CI	Fisher's Z	I²
Same Parameter							
Femoral adduction angle (midstance)	Dingenen et al. (2018b) Reinking et al. (2018)	0.96		<0.001*	0.902-0.983	1.932	0%
Hip adduction angle (IC)	Reinking et al. (2018) Atkins et al. (2014)	0.99		0.037*	0.148-1.000	2.567	95%
Knee flexion angle (midstance)	Reinking et al. (2018) Pipkin et al. (2016)	0.87		0.321	-0.857-0.999	1.312	97%
Knee SP angle (IC)	Mousavi et al. (2020) Reinking et al. (2018) Pipkin et al. (2016)	0.83		<0.001*	0.563-0.941	1.193	63%
Foot SP angle (IC)	Murray et al. (2019) Reinking et al. (2018)	0.95		<0.001*	0.709-0.991	1.781	78%
Rearfoot angle (midstance)	Reinking et al. (2018) Pipkin et al. (2016)	0.84		<0.001*	0.648-0.932	1.221	0%
Rearfoot in/eversion angle FP (IC)	Mousavi et al. (2020) Reinking et al. (2018)	0.96		<0.001*	0.833-0.988	1.884	60%
Similar Parameters							
Contralateral pelvic drop (midstance) Contralateral pelvic drop (IC) Lateral pelvic drop (midstance)	Dingenen et al. (2018b) Atkins et al. (2014) Pipkin et al. (2016)	0.96		0.037*	0.120-0.999	1.919	96%
Angle of the lower leg (IC) Tibial inclination (IC)	Reinking et al. (2018) Pipkin et al. (2016)	0.75		0.007*	0.256-0.933	0.971	57%

*p<0.05; IC: initial contact; SP: sagittal plane; FP: frontal plane; Green color: correlation 0.7-0.9 or >0.9 (very large or nearly perfect, respectively)

*p < 0.05; IC: initial contact; SP: sagittal plane; FP: frontal plane; Green color: correlation 0.7–0.9 or > 0.9 (very large or nearly perfect, respectively).

Table 5
Intra-rater (intra-session) reliability meta-analyses performed for the parameters with at least two studies.

	Studies	Correlation	Classification	p	95% CI	Fisher's Z	I²
Same Parameter							
Peak contralateral pelvic drop (midstance)	Maykut et al. (2015) Dingenen et al. (2018b)	0.96		<0.001*	0.930-0.982	1.999	0%
Peak hip adduction (midstance)	Maykut et al. (2015) Reinking et al. (2018)	0.91		<0.001*	0.624-0.979	1.504	70%
Hip adduction angle (IC)	Reinking et al. (2018) Atkins et al. (2014)	0.96		0.007*	0.491-0.997	1.934	86%
Knee flexion angle (midstance)	Reinking et al. (2018) Pipkin et al. (2016)	0.83		0.001*	0.450-0.955	1.187	57%
Knee SP angle (IC)	Mousavi et al. (2020) Reinking et al. (2018) Pipkin et al. (2016)	0.95		<0.001*	0.715-0.993	1.846	87%
Foot strike SP angle (IC)	Murray et al. (2019) Reinking et al. (2018)	0.95		<0.001*	0.709-0.991	1.781	78%
Rearfoot angle (midstance)	Reinking et al. (2018) Pipkin et al. (2016)	0.87		<0.001*	0.714-0.946	1.345	0%
Rearfoot in/eversion angle FP (IC)	Mousavi et al. (2020) Reinking et al. (2018)	0.96		<0.001*	0.897-0.983	1.924	20%
Similar Parameters							
Contralateral pelvic drop (IC) Lateral pelvic drop (midstance)	Atkins et al. (2014) Pipkin et al. (2016)	0.95		0.014*	0.361-0.998	1.867	90%
Peak hip adduction (stance) Peak hip adduction (midstance)	Neal et al. (2020) Maykut et al. (2015) Reinking et al. (2018)	0.84		0.001*	0.467-0.962	1.237	83%
Angle of the lower leg (IC) Tibial inclination (IC)	Reinking et al. (2018) Pipkin et al. (2016)	0.86		<0.001*	0.688-0.941	1.293	0%

*p<0.05; IC: initial contact; SP: sagittal plane; FP: frontal plane; Green color: correlation 0.7-0.9 or >0.9 (very large or nearly perfect, respectively)
*$p < 0.05$; IC: initial contact; SP: sagittal plane; FP: frontal plane; Green color: correlation 0.7–0.9 or > 0.9 (very large or nearly perfect, respectively).

dimensional motion analysis systems (2DMAS). Our findings reveal a low correlation between 2D and 3D measurements in the frontal and transverse planes, emphasizing that 2D angles represent distinct variables from those assessed in three-dimensional analyses. While the validity of 2D measures is limited for most parameters, their good reliability suggests their potential utility in evaluating the effects of interventions. However, it is crucial to acknowledge that 2D biomechanical assessments should not be considered simplified representations of 3D measurements. Longitudinal studies are needed to explore the true applicability of 2D measures in differentiating between injured and healthy runners and their potential as injury predictors.

Several factors can influence the results of a 2DMAS assessment. Lighting conditions and marker colours play a crucial role in facilitating marker identification during angle determination and reducing measurement errors. Reflective markers are generally recommended for improved visibility. The clinician's level of experience in data collection and processing are also important considerations. During data collection, the placement of markers (especially for test-retest analyses) and camera positioning are rater-dependent factors. In terms of data analysis, accurate determination of gait events and precise angle measurements are essential for a reliable 2DMAS assessment.

In this context, the experience of the raters may impact the reliability of the 2DMAS. In the analysis of gait and running kinematics, this is due to the rater-dependent nature of identifying the instants when angular parameters are determined, as well as drawing the corresponding angles. Some of the included studies (Atkins et al., 2014; Murray et al., 2018; Reinking et al., 2018) reported that raters underwent training to enhance data processing and standardize procedures. These studies demonstrated good to excellent reliability for all assessed parameters. Such training can improve the identification of gait events and the determination of angular parameters, significantly influencing intra- and inter-rater reliability assessments. Among the included studies, raters' experience ranged from no experience to over 12 years. However, five out of the twelve reliability studies did not specify the experience of the raters.

The sampling frequency is a critical aspect that can significantly affect the analysis, particularly if it is too low for the movement being analysed. According to the Nyquist theorem, the sampling rate should be at least twice the maximum frequency of the signal being recorded in order to accurately reconstruct the signal (Hebert-Losier et al., 2022; Winter, 1982). Moreover, for optimal results, the sampling rate should ideally be four to six times the maximum frequency of the signal (Giakas, 2004). Consequently, a minimum sampling frequency of 100 Hz is recommended for collecting kinematics data during running. For higher speeds, it is advisable to use a higher sampling frequency of at least 200 Hz, especially if the analysis involves capturing the position of foot markers, which tend to move at relatively higher velocities. In the studies included in this review, the sampling rates employed for 2DVAS ranged from 30–240 Hz, 25–300 Hz, and 50–300 Hz for criterion validity, intra-rater, and inter-rater studies, respectively. The most common sampling rate used in the studies was 60 Hz. In comparison, 3DMAS analyses for running typically employed a sampling frequency ranging from 100 to 240 Hz.

Correct camera positioning (perpendicular to the plane being analysed) is also essential to guarantee reliable and valid results. However, in a 2DMAS analysis, some movements can be done out-of-plane, leading to perspective errors and incorrect angular values. This is very difficult to control, once the human movements are not in a unique plane of movement. This fact can influence a 2DMAS assessment and explain part of the low criterion validity found here. Literature on running injuries proposes that most of the potential biomechanical risk factors are related to frontal and transverse plane movements (Willwacher et al., 2022), exactly where the greatest errors for 2DMAS occur. Inadequate diagnosis of a movement dysfunction could lead to incorrect interventions and contribute to its ineffectiveness. To fully avoid this source of error, multiple cameras, and a 3D reconstruction are necessary (Peebles et al., 2021).

A qualitative analysis of movement was performed in one of the included studies (Pipkin et al., 2016). In this study, kinematic variables were rated using a categorical scale at specific events of the gait cycle, instead of joint angles measured in quantitative studies. Compared to the quantitative studies, the intra- and inter-rater reliability was, in general, worse, mainly for the inter-rater reliability, although a reasonable sample frequency was utilized (120 Hz) and minimum one-year raters' experience.

Due to the high variability between the methods utilized in the included studies, grouping and comparing their results is still difficult. Angular parameters of running can be determined at different instants along the running cycle. Each of them has a specific clinical importance considering the risk of injury and the performance. Here, the authors determined the angles at specific instants (e.g., initial contact, toe-off), the peak angle or ROM during different phases (e.g., stance phase, midstance), and the peak angle or ROM along the entire running cycle. In addition, the way the parameters were determined, considering the placement of the markers, the angles drawn, and the identification of running events, had high variability among the studies. Due to the low number of studies that analysed the same parameter at the same instant analysed in the meta-analyses and the consequent high heterogeneity among them, there is a necessity for new studies with the same methodological procedures.

5. Conclusion

2DVAS proves to be a reliable method for measuring joint angles during running. However, the criterion validity of 2DVAS in comparison to 3DMAS (gold standard) ranges from low to moderate for most of the running parameters. Notably, knee flexion during stance and ankle angles in the sagittal plane at initial contact exhibit relatively higher validity. Consequently, the application of 2DVAS should be undertaken with caution, particularly when assessing the trunk, pelvis, and hip. Furthermore, given the variation in methodologies adopted across different studies, it is essential that standardized procedures and training be integral to future research involving 2DVAS.

CRediT authorship contribution statement

Gustavo Leporace: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Leonardo Metsavaht:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Conceptualization. **Felipe F. Gonzalez:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Fabio Arcanjo de Jesus:** . **Mariana Machado:** Writing – review & editing, Writing – original draft, Visualization, Investigation. **Eliane Celina Guadagnin:** . **Mansueto Gomes-Neto:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Conceptualization.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiomech.2023.111747.

References

Atkins, L.T., James, C.R., Sizer, P.S., Jonely, H., Brismee, J.M., 2014. Reliability and concurrent criterion validity of a novel technique for analyzing hip kinematics during running. Physiother. Theory Pract. 30, 210–217.

Brink, Y., Louw, Q.A., 2012. Clinical instruments: reliability and validity critical appraisal. J. Eval. Clin. Pract. 18, 1126–1132.

Damsted, C., Nielsen, R.O., Larsen, L.H., 2015. Reliability of video-based quantification of the knee- and hip angle at foot strike during running. Int. J. Sports Phys. Ther. 10, 147–154.

de Oliveira, F.C.L., Fredette, A., Echeverria, S.O., Batcho, C.S., Roy, J.S., 2019. Validity and Reliability of 2-Dimensional Video-Based Assessment to Analyze Foot Strike Pattern and Step Rate During Running: A Systematic Review. Sports Health 11, 409–415.

Dingenen, B., Barton, C., Janssen, T., Benoit, A., Malliaras, P., 2018a. Test-retest reliability of two-dimensional video analysis during running. Phys. Ther. Sport 33, 40–47.

Dingenen, B., Staes, F.F., Santermans, L., Steurs, L., Eerdekens, M., Geentjens, J., Peers, K.H.E., Thysen, M., Deschamps, K., 2018b. Are two-dimensional measured frontal plane angles related to three-dimensional measured kinematic profiles during running? Phys. Ther. Sport 29, 84–92.

Giakas, G., 2004. Power Spectrum Analysis and Filtering. In: Stergiou, N. (Ed.), Innovative Analyses of Human Movement. Human Kinetics, Champaign.

Grau, S., Muller, O., Baurle, W., Beck, M., Krauss, I., Maiwald, C., Baur, H., Mayer, F., 2000. Limits and possibilities of 2D video analysis in evaluating physiological and pathological foot rolling motion in runners. Sportverletz. Sportschaden 14, 107–114.

Hebert-Losier, K., Dai, B., Nunome, H., Kong, P.W., Hobara, H., Hsu, W.C., Bradshaw, E. J., Fong, D.T.P., Vanwanseele, B., 2022. Reporting guidelines for running biomechanics and footwear studies using three-dimensional motion capture. Sports Biomech. 1–12.

Hensley, C.P., Kontos, D., Feldman, C., Wafford, Q.E., Wright, A., Chang, A.H., 2022. Reliability and validity of 2-dimensional video analysis for a running task: A systematic review. Phys. Ther. Sport 58, 16–33.

Higgins, J.P.T., Green, S., 2006. Cochrane handbook for systematic reviews of interventions 4.2.6. John Wiley & Sons, Chichester.

Hopkins, W.G., Marshall, S.W., Batterham, A.M., Hanin, J., 2009. Progressive statistics for studies in sports medicine and exercise science. Med. Sci. Sports Exerc. 41, 3–13.

Jamovi, 2020. The Jamovi Project. Jamovi (Version 1.2) [Computer Software]. Retrieved from https://www.jamovi.org.

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 15, 155–163.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Lopes, T.J.A., Ferrari, D., Ioannidis, J., Simic, M., Micolis de Azevedo, F., Pappas, E., 2018. Reliability and Validity of Frontal Plane Kinematics of the Trunk and Lower Extremity Measured With 2-Dimensional Cameras During Athletic Tasks: A Systematic Review With Meta-analysis. J. Orthop. Sports Phys. Ther. 48, 812–822.

Macpherson, T.W., Taylor, J., McBain, T., Weston, M., Spears, I.R., 2016. Real-time measurement of pelvis and trunk kinematics during treadmill locomotion using a low-cost depth-sensing camera: A concurrent validity study. J. Biomech. 49, 474–478.

Maykut, J.N., Taylor-Haas, J.A., Paterno, M.V., DiCesare, C.A., Ford, K.R., 2015. Concurrent validity and reliability of 2d kinematic analysis of frontal plane motion during running. Int. J. Sports Phys. Ther. 10, 136–146.

McClay, I., Manal, K., 1998. The influence of foot abduction on differences between two-dimensional and three-dimensional rearfoot motion. Foot Ankle Int. 19, 26–31.

Mousavi, S.H., Hijmans, J.M., Moeini, F., Rajabi, R., Ferber, R., van der Worp, H., Zwerver, J., 2020. Validity and reliability of a smartphone motion analysis app for lower limb kinematics during treadmill running. Phys. Ther. Sport 43, 27–35.

Murray, L., Beaven, C.M., Hebert-Losier, K., 2018. Reliability of Overground Running Measures from 2D Video Analyses in a Field Environment. Sports (Basel) 7.

Neal, B.S., Lack, S.D., Barton, C.J., Birn-Jeffery, A., Miller, S., Morrissey, D., 2020. Is markerless, smart phone recorded two-dimensional video a clinically useful measure of relevant lower limb kinematics in runners with patellofemoral pain? A validity and reliability study. Phys. Ther. Sport 43, 36–42.

Noehren, B., Davis, I., Hamill, J., 2007. ASB clinical biomechanics award winner 2006 prospective study of the biomechanical factors associated with iliotibial band syndrome. Clin. Biomech. (Bristol, Avon) 22, 951–956.

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

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hrobjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71.

Peebles, A.T., Carroll, M.M., Socha, J.J., Schmitt, D., Queen, R.M., 2021. Validity of Using Automated Two-Dimensional Video Analysis to Measure Continuous Sagittal Plane Running Kinematics. Ann. Biomed. Eng. 49, 455–468.

Pfister, A., West, A.M., Bronner, S., Noah, J.A., 2014. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. J. Med. Eng. Technol. 38, 274–280.

Pipkin, A., Kotecki, K., Hetzel, S., Heiderscheit, B., 2016. Reliability of a Qualitative Video Analysis for Running. J. Orthop. Sports Phys. Ther. 46, 556–561.

Reinking, M.F., Dugan, L., Ripple, N., Schleper, K., Scholz, H., Spadino, J., Stahl, C., McPoil, T.G., 2018. Reliability of Two-Dimensional Video-Based Running Gait Analysis. Int. J. Sports Phys. Ther. 13, 453–461.

Vannatta, C.N., Heinert, B.L., Kernozek, T.W., 2020. Biomechanical risk factors for running-related injury differ by sample population: A systematic review and meta-analysis. Clin. Biomech. (Bristol, Avon) 75, 104991.

Weber, C.F., McClinton, S., 2020. Validity and Reliability of Video-Based Analysis of Upper Trunk Rotation during Running. Int. J. Sports Phys. Ther. 15, 910–919.

Willwacher, S., Kurz, M., Robbin, J., Thelen, M., Hamill, J., Kelly, L., Mai, P., 2022. Running-Related Biomechanical Risk Factors for Overuse Injuries in Distance Runners: A Systematic Review Considering Injury Specificity and the Potentials for Future Research. Sports Med. 52, 1863–1877.

Winter, D.A., 1982. Camera speeds for normal and pathological gait analyses. Med. Biol. Eng. Compu. 20, 408–412.

Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., D'Lima, D.D., Cristofolini, L., Witte, H., Schmid, O., Stokes, I., Standardization, Terminology Committee of the International Society of, B., 2002. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion–part I: ankle, hip, and spine. International Society of Biomechanics. J. Biomech. 35, 543-548.