

Original Research

Reliability of the Main 2D Kinematic Variables of Running Evaluated Categorically in Amateur Runners, Aged 18 to 55 Years

Annie V Stappung, PT¹ , Paulina M Espinoza, PT^{1,2}, Matilda J Letelier, PT¹, Felipe H Palma, MSc¹ ^a

¹ Laboratorio Integrativo de Biomecánica y Fisiología del Esfuerzo (LIBFE), Universidad de Los Andes, Chile, ² Motion Center, Chile

Keywords: running, concordance, reliability, kinematics, biomechanics

<https://doi.org/10.26603/001c.142061>

International Journal of Sports Physical Therapy

Vol. 20, Issue 8, 2025

Background

Running offers numerous health benefits, yet the prevalence of running-related injuries remains perfect. Evaluating running biomechanics through kinematic variables has gained attention due to their association with injuries. Currently, 3D video analysis is considered the gold standard, however, its complexity and cost limit widespread use. Consequently, 2D kinematic analysis using standard cameras has become popular, although research on its reliability is lacking.

Hypothesis/Purpose: The purpose of this study was to evaluate the reliability of the main kinematic variables of running, assessed categorically using 2D video analysis, in amateur runners aged 18 to 55 years.

Study Design

A cross-sectional analytical observational study.

Methods

Thirty-three amateur runners aged 18-55 underwent evaluation. Reflective markers were placed on key anatomical landmarks of the dominant lower limb to enhance motion tracking in 2D video recordings from sagittal and posterior frontal planes. The videos were analyzed by three independent raters: a consensus group (four persons), a separate physical therapist, and an athletic trainer. Pairwise inter-rater agreement was assessed among all three raters using the Kappa coefficient. Nine categorical kinematic variables related to running biomechanics were assessed.

Results

The variables that presented the greater concordance were the foot progression angle (K 0.847, 0.654, 0.792), heel whip (K 0.780, 0.847, 0.835), knee window (K 0.847, 0.847, 0.835) and overstriding (K 0.920, 0.780, 0.857). The variables with the lowest concordance were pelvic drop (K 0.257, 0.047, 0.597) and hindfoot eversion (K 0.604, 0.236, 0.604). Perfect concordance was obtained in seven of the nine variables between the consensus group and the physical therapist.

Conclusion

2D kinematic analysis exhibited promising reliability for specific running variables, suggesting its potential in clinical settings. However, its effectiveness depends on both the evaluated variables and the expertise of the evaluators.

^a Corresponding Author:

Felipe H Palma

Laboratorio Integrativo de Biomecánica y Fisiología del Esfuerzo (LIBFE)

Escuela de Kinesiología, Universidad de los Andes, Santiago, 7620086, Chile

Telephone number: +569 40132267; e-mail: fpalma@uandes.cl

Level of Evidence

Level 3

INTRODUCTION

Running has become a widely popular activity due to its low economic cost, minimal equipment requirements, and well-established health benefits, such as improved cardiovascular and neurological functions, weight control, and reduced risk of heart disease.¹⁻⁶ However, this increasing participation has been accompanied by a significant rise in running-related injuries (RRI). Reported musculoskeletal injury rates range from 30% to 79% per 100 at-risk runners, with incidence estimates varying from 7 to 59 injuries per 1000 running hours.^{7,8}

Injuries in running can be broadly classified as either acute, such as muscle tears, strains, sprains, or superficial skin injuries, or overuse injuries, with the latter accounting for approximately 80% of RRI.⁶⁻¹¹ Common overuse injuries include patellofemoral syndrome, iliotibial band friction syndrome, medial tibial stress syndrome, Achilles tendinopathy, and plantar fasciitis; these conditions typically arise as a consequence of excessive repetitive loading.^{10,12-14} Although intrinsic factors like sex, race, and skeletal structure have been widely investigated as contributors to injury risk, biomechanical factors, particularly those related to running technique, remain relatively underexplored despite being estimated to contribute up to 40% of RRI.¹⁴

In the context of running biomechanics, researchers have examined a variety of kinematic variables to assess their relationship with injury occurrence. Depending on the plane of motion and the lower extremity joints evaluated, researchers have identified several variations associated with specific injuries.¹⁵ Variables such as foot strike pattern, foot progression angle, pelvic drop, hindfoot eversion, tibia angle in response to loading, peak knee flexion during the stance phase, overstriding, vertical displacement of the center of mass, trunk inclination, and knee window have all been reported to have potential links with RRI.¹⁶

Kinematic evaluation involves assessing both the displacement and angular movement of different body segments during running, typically through measurements of position, velocity, and acceleration in both linear and angular forms.¹⁷ Although 3D motion capture is considered the gold standard in biomechanical research, its complexity and high cost have led to an increased use of 2D motion analysis with conventional video cameras. While 2D analysis has demonstrated reliability and validity for evaluating lower limb movements in the sagittal and frontal planes, it does face limitations when capturing movements in multiple dynamic planes.¹⁷⁻¹⁹

The application of 2D video analysis has enabled evaluations in larger groups of athletes; however, its practical use in clinical settings remains challenging. The necessary equipment, such as cameras, tripods, high-contrast markers, and controlled environments, along with the potential influence of artifacts and software constraints, increases both the complexity and cost of the process. Although per-

fect reliability ($K = 0.81-1$) has been demonstrated under certain conditions, the validity of 2D analysis compared to 3D motion capture can vary significantly.^{18,19} One proposed solution to these limitations is the Automated Video Analysis for Dynamic Systems program, which facilitates the manual digitization of high-contrast markers in video recordings, thereby aiding in the quantification of continuous kinematic data.²⁰ Another approach to streamline the kinematic analysis involves converting quantitative measurement into qualitative or categorical assessments, which could offer a quick and practical tool for clinicians. The challenge with this method lies in ensuring the reproducibility and reliability of the qualitative categorization of kinematic variables. It is critical to identify which variables can be consistently evaluated across different evaluators and testing sessions. In this regard, previous work by Souza et al., has made initial progress by categorizing some kinematic variables as the foot strike pattern.¹⁶

Building on these insights, the purpose of this study was to evaluate the reliability of the main kinematic variables of running, assessed categorically using 2D video analysis, in amateur runners aged 18 to 55 years. Establishing that these variables can be consistently and reliably categorized is crucial for enabling physical therapists and athletic trainers to accurately assess running techniques and tailor individualized training programs. Furthermore, demonstrating robust inter-rater reliability in this evaluation approach could serve as the foundation for future advances, including the development of automated video analysis systems that streamline and enhance clinical assessment of running biomechanics.

METHODS

PARTICIPANTS

Thirty-three recreational runners (both genders; age 18-55 years; body mass index 18.5-30.0 kg/m²) from Santiago, Chile, were recruited via social media and local university outreach. Recreational runners were defined as individuals who had been running for more than three months, without competitive purposes, with a running frequency equal to or greater than three times per week and a maximum volume of 40 kilometers. Participants with musculoskeletal lower limb injuries in the prior six months or surgery in the prior two years were excluded.

The study was conducted following the principles of the Declaration of Helsinki and was approved by the ethics committee of Universidad de los Andes, Chile (CEC2021020). Participants were instructed about the assessment, its purpose, and the procedures to be performed. Before the test, all participants signed an informed consent form, outlining the assessment activities, as well as the benefits and risks they might face, with the latter minimized by the researchers. The information provided by the participants was processed under strict confidentiality to

Table 1. Variables evaluated in the posterior view of the frontal plane, according to the classification by Souza et al.¹⁶

Variable	Categorization	Definition
Foot progression Angle	Normal	5° to 10° Toe-out
	Moderate	Neutral foot alignment
	Severe	5° to 10° Toe-in
Heel Whip	Internal or external rotation < 5°	Toe-out or toe-in < 5° at initial swing
	Internal or external rotation > 5°	Toe-out or toe-in > 5° at initial swing
Knee Window	Presence of space	Alignment of the hip, knee, and ankle
	Absence of space	Excessive hip adduction, hip internal rotation, and knee valgus
Pelvic Drop	Absence	Alignment between the posterior superior iliac spines
	Moderate presence < 10°	Fall of the vertical < 10°
	Severe presence > 10°	Fall of the vertical > 10°
Heel Eversion	Normal	Vertical relationship of the hindfoot with the Achilles tendon
	Increased angle	Increase in eversion

Table 2. Variables evaluated in the sagittal plane, according to the classification by Souza et al.¹⁶

Variable	Categorization	Definition
Foot Strike Patterns	Forefoot	Foot strike occurs in the anterior third of the foot, with the ball and toes engaging first.
	Midfoot	Foot strike happens in the middle third, with a balanced engagement of both the forefoot and rearfoot.
	Rearfoot	Foot strike occurs in the posterior third, with the heel making contact first.
Tibia Angle at Loading Response	Tibia extension	Lateral knee joint marker is posterior to the lateral malleolus marker
	Vertical tibia	Lateral knee and malleolus markers are directly vertical
	Tibia flexion	Lateral knee marker is anterior to the lateral malleolus
Knee Flexión During Stance	Normal 45°	Difference in angle from initial contact to peak knee flexion
	Less than 45°	
Overstriding	Aligned with the pelvis	The foot is landing under the center of mass
	Anterior to the pelvis	The foot is landing in front of the center of mass (overstride)

prevent any moral harm. Additionally, volunteers were free to withdraw from the research at any time, and their decision was respected.

VARIABLES

See Tables 1 and 2.

MEASUREMENT PROTOCOL

Before recording, proper camera installation was carried out to ensure video recording with adequate resolution and image sharpness.²¹ Two Apple® phones, iPhone 11 and iPhone 13 models were used, with the following characteristics: 1080p HD resolution at a capture frequency of 60 frames per second (FPS). The cameras recorded in the horizontal plane, without the use of lenses such as wide-angle

or fisheye, to avoid peripheral distortion, as the focal length is shorter. They were positioned with a stable tripod along with a level on the tripod to orient the cameras in space and avoid manual adjustment.²¹ Cameras were positioned at an approximate distance of two meters from the runner, with the final height and exact distance being fine-tuned to each volunteer's anthropometry, following the protocol by Roche for both planes.²¹ Both cameras were synchronized to obtain simultaneous recordings in both planes and enhance the analysis of the variables. This was done using intermittent red light, which did not obstruct the camera lens, blinking three times before turning it off. To maximize image quality, natural or LED light was favored. Two ring lights were used just behind the camera, allowing light to bounce as much as possible onto the camera and avoiding flickering effects.²¹

Before the evaluation, the volunteer's anthropometric variables (body mass, height, and body mass index) were recorded. For body mass, a digital scale (Eufy®) was used, and height was measured using a measuring tape. Additionally, the dominant limb was recorded, determined by a step-up and step-down test on a 30 cm high step.²² Each volunteer then performed a 10-minute warm-up at a self-selected running speed, below the test speed, to adapt to the treadmill conditions for comfortable execution,²³ making it as similar as possible to normal conditions. After the warm-up, 3x3 cm markers with distinctive colored tape were placed on the dominant lower extremity at the following anatomical points: posterior superior iliac spines, greater trochanter, lateral femoral epicondyle, mid-thigh, mid-tibia, lateral malleolus, heel, and forefoot superiorly. This was done to improve perspective during video analysis.

Subsequently, recording began and continued for two minutes while the volunteer ran at a self-selected speed. Following the run, the volunteer walked on the treadmill for three minutes to gradually return to a resting state. Afterwards, evaluators conducted a series of lower limb stretches, including those targeting the quadriceps, hamstrings, and gastrocnemius.

The final video for each participant, synchronized from both camera planes, contained 20 running cycles for analysis. To facilitate detailed observation, evaluators were permitted to reduce the video playback speed by 10-25% using iMovie.

The kinematic analysis was performed through purely visual inspection, without the use of any specialized motion analysis software. Three independent rating entities were established for this process:

- A consensus group (reference standard): this entity comprised a physical therapist with eight years of experience in biomechanics and three fifth-year physical therapy students with specialized training. Through collaborative review and discussion based on established literature, this group reached a definitive categorization for each variable ([Table 1](#), [Table 2](#)). Their final decisions served as the reference standard for this study.¹⁶
- An independent physical therapist: a specialist in sports with at least three years of experience working with amateur runners.
- An independent athletic trainer: also, a specialist in sports with a minimum of three years of experience with amateur runners.

The 33 videos were compiled into a folder and shared with the two independent evaluators via AirDrop. Each analyzed the nine kinematic variables without knowledge of the ratings from the other evaluator or the consensus group. This blinded evaluation process was implemented to mitigate a priori bias. Subsequently, pairwise inter-rater reliability was calculated to compare the ratings among all three entities (consensus group vs. physical therapist, consensus group vs. athletic trainer, physical therapist vs. athletic trainer).

SAMPLE SIZE

The sample size was calculated using the Liao study, employing a 95% confidence level (alpha 0.05) and 80% power (beta of 0.2), resulting in a minimum required sample size of 32 individuals.²⁴

STATISTICAL ANALYSIS

After analyzing the recordings, the agreement between the results obtained by the consensus group and those provided by external evaluators was determined. Concordance referred to how closely the observations on the measured variables matched each other and, therefore, did not assess the validity or certainty of the variables.²⁵ Since all mentioned variables were measured categorically, the Cohen's kappa coefficient and weighted kappa were calculated based on the number of categories to analyze inter-rater agreement using STATA 14/IC software (STATA CORP. LTA, USA). This coefficient is interpreted at different levels depending on its magnitude, including weak (< 0.40), moderate (0.40-0.60), strong (0.61-0.80), or perfect (0.81-1) categories.^{25,26}

RESULTS

A total of 33 recreational runners were included in the final analysis. The demographic and anthropometric characteristics of the sample, including age, sex, and body mass index, are detailed in [Table 3](#).

The pairwise inter-rater reliability for the nine kinematic variables is presented in [Table 4](#) and summarized visually in [Figure 1](#). The analysis revealed substantial to almost perfect agreement for several variables. Specifically, the variables with the highest concordance across the rater pairs were the knee window (K = 0.847, 0.847, 0.835), overstriding (K = 0.920, 0.780, 0.857), foot progression angle (K = 0.847, 0.654, 0.792), and heel whip (K = 0.780, 0.847, 0.835). In contrast, the variables demonstrating the weakest and most inconsistent agreement were pelvic drop (K 0.257, 0.047, 0.597) and rearfoot eversion (K 0.604, 0.236, 0.604).

A direct comparison between specific rater pairs, presented in [Table 5](#), highlights that the agreement between the consensus group and the independent physical therapist was particularly high, achieving strong to perfect concordance for seven of the nine variables.

DISCUSSION

Running-related injuries are common among amateur runners, and while 3D motion capture is the gold standard for biomechanical analysis, its cost and complexity limit clinical use. In contrast, 2D video analysis offers a more practical alternative, yet its reliability for categorically evaluating key kinematic variables remains unclear. Therefore, this study aimed to address that gap by assessing the inter-rater reliability of this approach. The primary finding was that the analysis of running kinematic variables on a treadmill through 2D video is a reliable tool for implementation in

Table 3. Characterization of the volunteers (n=33).

	Minimum	Maximum	Mean	Standard Deviation
Age (years)	18.00	54.00	28.00	9.09
Body mass (kg)	44.80	89.10	70.50	11.50
Height (m)	1.52	1.88	1.69	0.09
Body mass index (kg/m ²)	19.10	29.50	24.40	2.80
	Absolute Frequency (n)		Relative Frequency (%)	Total N (%)
Gender	Male	24	73	33 (100)
	Female	9	27	
Dominant limb	Right	17	52	33 (100)
	Left	16	48	

Table 4. Kappa Coefficient, confidence interval and percentage of agreement among evaluators.

Variable	K [95% CI] 1-2	K [95% CI] 1-3	K [95% CI] 2-3	% A 1-2	% A 1-3	% A 2-3
Frontal Plane (posterior view)						
Foot progression Angle	0.847 [0.642 – 1.000]	0.654 [0.382 – 0.925]	0.792 [0.573 – 1.000]	93.9	84.8	90.9
Heel Whip	0.780 [0.549 – 1.000]	0.634 [0.348 – 0.919]	0.835 [0.614 – 1.000]	90.9	84.8	93.9
Knee Window	0.847 [0.646 – 1.000]	0.847 [0.646 – 1.000]	0.835 [0.614 – 1.000]	93.9	93.9	93.9
Pelvic drop	0.257 [-0.020 – 0.536]	0.047 [-0.160 – 0.255]	0.597 [0.318 – 0.876]	83.3	74.2	81.8
Heel eversion	0.604 [0.292 – 0.916]	0.236 [-0.123 – 0.595]	0.604 [0.292 – 0.916]	84.8	69.7	84.8
Sagittal Plane						
Foot Strike Patterns	1.000 [1.000 – 1.000]	0.552 [0.277 – 0.826]	0.552 [0.277 – 0.826]	100.0	86.4	86.4
Knee Flexion During Stance	0.778 [0.541 – 1.000]	0.153 [-0.171 – 0.479]	0.355 [0.038 – 0.672]	90.9	60.6	69.7
Tibia Angle at Loading Response	0.758 [0.502 – 1.000]	0.450 [0.151 – 0.748]	0.588 [0.301 – 0.875]	93.9	84.4	87.9
Overstriding	0.920 [0.768 – 1.000]	0.780 [0.549 – 1.000]	0.857 [0.667 – 1.000]	97.0	90.9	93.9

K=Kappa Coefficient; CI= Confidence interval; 1=Consensus; 2=Physical Therapist; 3=Athletic Trainer; %A=Agreement Percentage.

the clinical setting, aligning with other studies that identify it as a useful instrument in certain contexts.^{3,7,18,27} However, according to the current findings, its reliability is affected by the variables assessed and the preparation and experience of the evaluators.

An analysis of the data did not support the hypothesis that one observational plane is inherently more reliable than the other. Instead, the findings indicate that reliability is a function of the specific kinematic variable in question. This is demonstrated by the mixed results within each plane (Table 5). At first glance, the sagittal plane presented challenges, as only overstriding ($K = 0.920, 0.780, 0.857$) achieved high reliability across all rater comparisons. This seems to contrast with literature suggesting sagittal plane kinematics are easier to assess.²⁸ However, this interpretation is incomplete. When the analysis is limited to the expert consensus group and the experienced physical ther-

apist, all four sagittal plane variables achieved strong-to-perfect agreement. This highlights that rater experience may be a more critical factor than the plane of view. Similarly, in the frontal plane, variables such as knee window showed high reliability, while others like pelvic drop were consistently unreliable regardless of the rater. Therefore, the evidence suggests that clinicians should focus on the known reliability of individual variables rather than assuming an entire observational plane is reliable.

Focusing on the successful assessments, four variables demonstrated consistently high inter-rater agreement: knee window, heel whip, foot progression angle, and overstriding. Among these, the knee window stood out, achieving perfect agreement ($K = 0.847, 0.847, 0.835$) across all three rater comparisons. These variables likely performed well because they represent distinct, easily observable

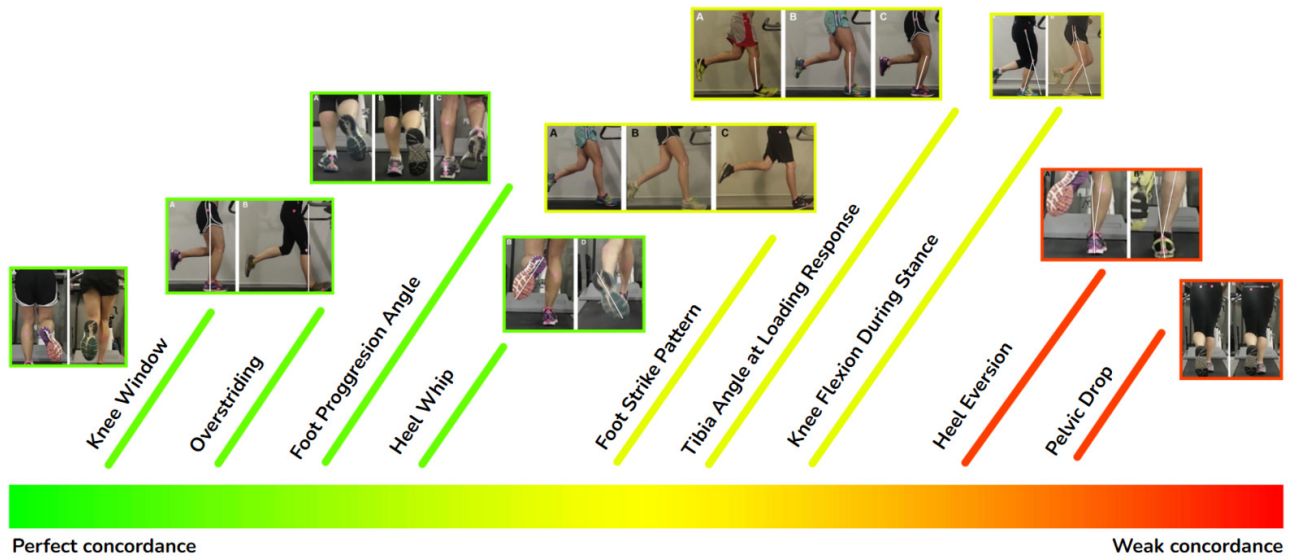


Figure 1. Illustration of the order of variables according to their concordance. Adapted from: Souza RB. An Evidence-Based Videotaped Running Biomechanics Analysis. Physical Medicine and Rehabilitation Clinics of North America. 2016.

Table 5. Level of concordance between evaluators according to each category.

Variables	CL 1-2	CL 1-3	CL 2-3
Frontal Plane (posterior view)			
Foot progression Angle	Perfect	Strong	Strong
Heel Whip	Strong	Strong	Perfect
Knee Window	Perfect	Perfect	Perfect
Pelvic Drop	Weak	Weak	Moderate
Heel eversion	Moderate	Weak	Moderate
Sagittal Plane			
Foot Strike Patterns	Perfect	Moderate	Moderate
Knee Flexion During Stance	Strong	Weak	Weak
Tibia Angle at Loading Response	Strong	Moderate	Moderate
Overstriding	Perfect	Strong	Perfect

CL=Concordance Level: weak (< 0,40), moderate (0,40-0,60), strong (0,61-0,80) or perfect (0,8- 1). 1=Consensus; 2=Physical Therapist; 3=Athletic Trainer.

events with clear visual cues, making them suitable for categorical, even dichotomous, judgment.

In contrast, variables such as foot strike, peak knee flexion during stance, and tibia angle at loading response demonstrated weak and inconsistent inter-rater concordance. The poor performance of these variables can be attributed to several factors. Firstly, assessing them often requires identifying a single precise frame within a rapid movement, a task that is inherently challenging in standard 2D video. Secondly, a polytomous scale (multiple options) may introduce more subjective ambiguity than a simple binary choice. The lower agreement for these variables underscores the limitations of visual 2D analysis for capturing subtle, transient kinematic events.

A central theme emerging from these contrasting results is the profound impact of the measurement scale. The data

suggests that using a dichotomous (yes/no) scale is a powerful strategy for increasing reliability in 2D analysis. Even when a variable had quantitative undertones (like the degree of heel whip), the binary classification proved effective. This practical observation is supported by statistical theory; dichotomizing variables, particularly around the median, creates a robust classification that is invariant to data transformations and less sensitive to asymmetric distributions, which in turn enhances the stability of agreement coefficients like Kappa.^{27,29}

The analysis of rearfoot eversion yielded poor and inconsistent reliability (K 0.604, 0.236, 0.604), suggesting its assessment via 2D video is not advisable for clinical use. This low agreement is likely attributable to the inherent difficulty of capturing a complex, tri-planar motion from a single frontal plane view, a limitation previously high-

lighted by Souza et al.¹⁶ The challenge is compounded by the foot's natural abduction in the transverse plane (12° to 18°), which can obscure a true posterior view and introduce significant observational bias.³⁰

Pelvic drop exhibited the lowest concordance of all variables measured (K 0.257, 0.047, 0.597), confirming its unsuitability for clinical evaluation with this method. This poor performance directly illustrated a central theme of this study: the negative impact of complex rating scales. The assessment of pelvic drop requires a multi-category judgment combined with an estimation of movement magnitude, in contrast to the simple, dichotomous variables that proved more reliable. This finding reinforces the conclusion that combining multi-category scales with quantitative estimations severely hinders observational consistency in 2D analysis.²⁹

Rater expertise emerged as a critical factor influencing measurement reliability. The consensus group and the experienced physical therapist achieved strong-to-perfect agreement on seven of the nine variables, whereas comparisons involving the athletic trainer only reached this level for four. This disparity suggests that reliability is closely linked to academic training in biomechanics and clinical experience in movement analysis, which was a competency profile that distinguished the physical therapist and the consensus group. Consequently, the findings support a two-tiered classification for the clinical applicability of these variables. A core group (foot progression angle, heel whip, knee window, and overstriding) proved robust enough for reliable assessment regardless of the rater's background. In contrast, the reliability of other key variables, namely foot strike pattern, peak knee flexion during stance, and tibia angle at loading response, was highly dependent on evaluator expertise. This finding for foot strike is especially noteworthy, as it contradicts previous research suggesting it should be a universally reliable measure.³ In the present study, these three variables consistently yielded weak to moderate concordance in evaluations involving the athletic trainer, while achieving strong agreement exclusively between the expert raters (Table 5). Therefore, while some variables are broadly applicable, the clinical utility of this latter group is contingent upon assessment by practitioners with demonstrable expertise in 2D kinematic analysis.

The findings of this study identify a core set of variables that demonstrated high inter-rater concordance, suggesting their suitability for reliable clinical categorization. While this high concordance serves as a strong indicator of inter-rater reliability, it is critical to acknowledge that reliability does not inherently establish the method's validity. In other words, consistent agreement between raters does not guarantee that the 2D assessment accurately reflects the true biomechanical event. Therefore, to progress from a reliable method to a validated clinical tool, the essential next step is a direct comparison of this 2D analysis approach with the gold-standard 3D motion capture. Such a study is required to confirm the method's accuracy and support its future development for widespread clinical application.

LIMITATIONS

Despite the rigorous methodology employed, this study has several limitations. 1) Minor inaccuracies in camera synchronization may have occurred. Although a reflective light marker was used to facilitate frame-by-frame alignment, perfect synchronization between cameras cannot be fully guaranteed. 2) The posterior view was susceptible to perspective errors, as the natural medial-lateral displacement of a runner on a treadmill can affect the precision of ankle and foot kinematic measurements.²¹ 3) The treadmill's sidebar periodically obstructed the view of the greater trochanter marker, potentially compromising the assessment of the overstriding variable. Future research should utilize equipment that allows for an uninterrupted sagittal view. 4) Marker visibility was occasionally compromised by participant clothing that lacked sufficient contrast. While participants were instructed to wear dark attire, non-adherence in some cases suggests that stricter clothing protocols are needed to optimize marker detection. 5) Despite detailed protocols, variability may have arisen from subjective differences among raters in identifying the precise video frame for analysis. 6) Finally, although some studies mentioned not finding significant differences when studying kinematic variables on a treadmill versus pavement,³¹ the present study does not replicate the usual external conditions of overground running. However, this method remains a standard approach for conducting assessments in a controlled environment.²⁷

FUTURE DIRECTIONS

To make this type of assessment accessible to various types of evaluators without compromising reliability, as mentioned earlier, future interventions should include pre-training for less experienced evaluators to ensure better results. This approach is consistent with the results of Reinking et al., where minimal differences were observed between experienced and inexperienced evaluators in the analysis of kinematic variables running in 2D video, and all evaluators underwent a three-hour training session before conducting the video analysis.³² Following this same line, it would be suggestive for future studies to examine whether differences in the analysis of variables between evaluators are modified with the addition of training, aiming for greater concordance between less experienced and experienced evaluators. Additionally, it is important to mention that the creation of a standardized measurement instrument is recommended for replication by different research centers, thereby obtaining comparable results across different studies. Additionally, recording with cameras with an image capture rate greater than 60 FPS is suggested to improve the observation of the exact frame in which the variable is analyzed, especially for peak knee flexion in the stance phase and tibia angle in response to load. Furthermore, it is suggested that the variables to be evaluated be purely qualitative and dichotomous, not requiring assessment of exact angles of movement as mentioned earlier and not having more than one category. For example, in the

case of this study, it would be prudent to modify the categorization of the pelvic drop variable to define the absence or presence of this pattern²⁷ or provide greater clarity through illustrative images of the alternatives.

CONCLUSION

The current findings support using categorical assessments of running kinematic variables via 2D video analysis for some variables assessed in running gait analysis. However, to transform this method into a reliable tool that enables valid comparisons across studies, standardization of measurement conditions is essential. The current results reveal significant variability among evaluator groups, underscoring that reliability depends on both the specific variables assessed and the expertise of the evaluators. Notably, variables such as foot progression angle, heel whip, knee window, and overstriding achieved perfect agreement, suggesting their potential as robust indicators in this evaluation context.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

ACKNOWLEDGMENTS

We would like to express our gratitude to Badir Zara for offering the facilities and supporting the participant assessments, and to Rodrigo Guzmán Venegas and Óscar Valencia Cayupán for their valuable methodological guidance and suggestions.

Submitted: December 31, 2024 CDT. Accepted: June 22, 2025 CDT. Published: August 01, 2025 CDT.

© The Author(s)



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc/4.0> and legal code at <https://creativecommons.org/licenses/by-nc/4.0/legalcode> for more information.

REFERENCES

1. Schache AG, Dorn TW, Pandy MG. Lower-Limb Muscle Function in Human Running. In: Kumar V, Schmiedeler J, Sreenivasan SV, Su HJ, eds. *Advances in Mechanisms, Robotics and Design Education and Research*. Springer International Publishing; 2013:323-327. doi:[10.1007/978-3-319-00398-6_25](https://doi.org/10.1007/978-3-319-00398-6_25)
2. Levangie PK, Norkin CC. *Joint Structure and Function: A Comprehensive Analysis*. McGraw-Hill Education; 2011. <http://fadavispt.mhmedical.com/content.aspx?aid=1132749857>
3. de Oliveira FCL, Fredette A, Echeverría SO, Batcho CS, Roy JS. Validity and reliability of 2-dimensional video-based assessment to analyze foot strike pattern and step rate during running: A systematic review. *Sports Health*. 2019;11(5):409-415. doi:[10.1177/1941738119844795](https://doi.org/10.1177/1941738119844795)
4. Maykut JN, Taylor-Haas JA, Paterno MV, DiCesare CA, Ford KR, Maykut J. Concurrent validity and reliability of 2D kinematic analysis of frontal plane motion during running. *Int J Sports Phys Ther*. 2015;10(2):136.
5. Ceyssens L, Vanelderden R, Barton C, Malliaras P, Dingenen B. Biomechanical risk factors associated with running-related injuries: A systematic review. *Sports Med*. 2019;49(7):1095-1115. doi:[10.1007/s40279-019-01110-z](https://doi.org/10.1007/s40279-019-01110-z)
6. Mora CAA. Revisión de los beneficios de la intensidad y modalidades de ejercicio físico sobre el estrés psicológico. *Pensar en Movimiento: Revista de Ciencias del Ejercicio y la Salud*. 2018;16(1):30335. doi:[10.15517/pensarmov.v16i1.30335](https://doi.org/10.15517/pensarmov.v16i1.30335)
7. Dingenen B, Staes F, Vanelderden R, Ceyssens L, Malliaras P, Barton CJ, et al. Subclassification of recreational runners with a running-related injury based on running kinematics evaluated with marker-based two-dimensional video analysis. *Phys Ther Sport*. 2020;44:99-106. doi:[10.1016/j.ptsp.2020.04.032](https://doi.org/10.1016/j.ptsp.2020.04.032)
8. Buist I, Bredeweg SW, Bessem B, van Mechelen W, Lemmink KAPM, Diercks RL. Incidence and risk factors of running-related injuries during preparation for a 4-mile recreational running event. *Br J Sports Med*. 2010;44(8):598. doi:[10.1136/bjsm.2007.044677](https://doi.org/10.1136/bjsm.2007.044677)
9. Assumpção CDO, Lima LCR, Oliveira FBD, Greco CC, Denadai BS. Exercise-induced muscle damage and running economy in humans. *Sci World J*. 2013;2013. doi:[10.1155/2013/189149](https://doi.org/10.1155/2013/189149)
10. Hespanhol Junior LC, Pena Costa LO, Lopes AD. Previous injuries and some training characteristics predict running-related injuries in recreational runners: a prospective cohort study. *J Physiother*. 2013;59(4):263-269. doi:[10.1016/S1836-9553\(13\)70203-0](https://doi.org/10.1016/S1836-9553(13)70203-0)
11. Scott A, Docking S, Vicenzino B, Alfredson H, Zwerver J, Lundgreen K, et al. Sports and exercise-related tendinopathies: a review of selected topical issues by participants of the second International Scientific Tendinopathy Symposium (ISTS) Vancouver 2012. *Br J Sports Med*. 2013;47(9):536. doi:[10.1136/bjsports-2013-092329](https://doi.org/10.1136/bjsports-2013-092329)
12. Bredeweg SW, Zijlstra S, Bessem B, Buist I. The effectiveness of a preconditioning programme on preventing running-related injuries in novice runners: a randomised controlled trial. *Br J Sports Med*. 2012;46(12):865. doi:[10.1136/bjsports-2012-091397](https://doi.org/10.1136/bjsports-2012-091397)
13. Cosca DD, Navazio F. Common Problems in Endurance Athletes. 2007. <http://familydoctor.org/online/famdocen/home/healthy/physical/injuries/479.html>
14. Van Der Worp MP, Ten Haaf DSM, Van Cingel R, De Wijer A, Nijhuis-Van Der Sanden MWG, Bart Staal J. Injuries in runners; a systematic review on risk factors and sex differences. *PLoS One*. 2015;10(2). doi:[10.1371/journal.pone.0114937](https://doi.org/10.1371/journal.pone.0114937)
15. Nicola TL, Jewison DJ. The anatomy and biomechanics of running. *Clin Sports Med*. 2012;31(2):187-201. doi:[10.1016/j.csm.2011.10.001](https://doi.org/10.1016/j.csm.2011.10.001)
16. Souza RB. An evidence-based videotaped running biomechanics analysis. *Phys Med Rehabil Clin N Am*. 2016;27(1):217-236. doi:[10.1016/j.pmr.2015.08.006](https://doi.org/10.1016/j.pmr.2015.08.006)
17. Simon SR. Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. *J Biomech*. 2004;37(12):1869-1880. doi:[10.1016/j.jbiomech.2004.02.047](https://doi.org/10.1016/j.jbiomech.2004.02.047)
18. Dingenen B, Barton C, Janssen T, Benoit A, Malliaras P. Test-retest reliability of two-dimensional video analysis during running. *Phys Ther Sport*. 2018;33:40-47. doi:[10.1016/j.ptsp.2018.06.009](https://doi.org/10.1016/j.ptsp.2018.06.009)
19. Martinez C, Garbett S, Hiromasa K, Jackson R, Miya E, Miya M, et al. Comparison of 2-D and 3-D analysis of running kinematics and actual versus predicted running kinetics. *Int J Sports Phys Ther*. 2022;17(4):566-573. doi:[10.26603/001c.34432](https://doi.org/10.26603/001c.34432)

20. Peebles AT, Carroll MM, Socha JJ, Schmitt D, Queen RM. Validity of using automated two-dimensional video analysis to measure continuous sagittal plane running kinematics. *Ann Biomed Eng.* 2021;49(1):455-468. doi:[10.1007/s10439-020-02569-y](https://doi.org/10.1007/s10439-020-02569-y)
21. Roche Seruendo LE. Methodological approach for running gait 2D video analysis. *Revista internacional de deportes colectivos.* 2017;30:69-92.
22. Sadeghi H, Allard P, Prince F, Labelle H. Symmetry and limb dominance in able-bodied gait: a review. *Gait Posture.* 2000;12(1):34-45. doi:[10.1016/S0966-6362\(00\)00070-9](https://doi.org/10.1016/S0966-6362(00)00070-9)
23. Lussiana T, Gindre C. Feel your stride and find your preferred running speed. *Biol Open.* 2016;5(1):45-48. doi:[10.1242/bio.014886](https://doi.org/10.1242/bio.014886)
24. Liao JJZ. Sample size calculation for an agreement study. *Pharm Stat.* 2010;9(2):125-132. doi:[10.1002/pst.382](https://doi.org/10.1002/pst.382)
25. Cortés-Reyes É, Rubio-Romero JA, Gaitán-Duarte H. Métodos estadísticos de evaluación de la concordancia y la reproducibilidad de pruebas diagnósticas. *Rev Colomb Obstet Ginecol.* 2010;61(3):247-255. doi:[10.18597/rcog.271](https://doi.org/10.18597/rcog.271)
26. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med.* 2012;22(3):276-282. doi:[10.11613/BM.2012.031](https://doi.org/10.11613/BM.2012.031)
27. Hensley CP, Kontos D, Feldman C, Wafford QE, Wright A, Chang AH. Reliability and validity of 2-dimensional video analysis for a running task: A systematic review. *Phys Ther Sport.* 2022;58:16-33. doi:[10.1016/j.ptsp.2022.08.001](https://doi.org/10.1016/j.ptsp.2022.08.001)
28. Wille CM, Lenhart RL, Wang S, Thelen DG, Heiderscheit BC. Ability of sagittal kinematic variables to estimate ground reaction forces and joint kinetics in running. *J Orthop Sports Phys Ther.* 2014;44(10):825-830. doi:[10.2519/jospt.2014.5367](https://doi.org/10.2519/jospt.2014.5367)
29. Santa Escobar CD. Efectos en la estimación de modelos de estructuras de covarianza con variables dicotomizadas en diferentes puntos de corte. Published online 2018. <https://repositorio.unal.edu.co/handle/unal/63518>
30. Magee DJ. *Orthopedic Physical Assessment.* 6th ed. Elsevier Health Sciences; 2013. https://books.google.cl/books?id=_wcyAgAAQBAJ
31. Riley PO, Paolini G, Della Croce U, Paylo KW, Kerrigan DC. A kinematic and kinetic comparison of overground and treadmill walking in healthy subjects. *Gait Posture.* 2007;26(1):17-24. doi:[10.1016/j.gaitpost.2006.07.003](https://doi.org/10.1016/j.gaitpost.2006.07.003)
32. Reinking MF, Dugan L, Ripple N, Schleper K, Scholz H, Spadino J, et al. Reliability of two-dimensional video-based running gait analysis. *Int J Sports Phys Ther.* 2018;13(3):453-461. doi:[10.26603/ijsp.20180453](https://doi.org/10.26603/ijsp.20180453)