BJMB
Brazilian Journal of Motor Behavior

Research Article

Intra and inter-rater reliability of manual two-dimensional kinematics analysis of running

ELIANE C. GUADAGNIN[1] | JEAN S. CARVALHO[2] | VICTOR L. COSTA[2] | CAROLINA S. MARTINS[2] | FRANCESCA C. SONDA[3] | FELIPE P. CARPES[2]

[1] Instituto Brasil de Tecnologias da Saúde (IBTS), Rio de Janeiro, RJ, Brazil
[2] Applied Neuromechanics Group, Laboratory of Neuromechanics, Universidade Federal do Pampa, Uruguaiana, RS, Brazil
[3] Biomechanics and Kinesiology Research Group, Exercise Research Laboratory, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

Correspondence to: Eliane C. Guadagnin, Ph.D.
Postal address: Instituto Brasil de Tecnologias da Saúde; Visconde de Pirajá 407, room 905, Ipanema; Rio de Janeiro, RJ, Brazil, ZIP 22.410-003
email: elianecguadagnin@hotmail.com
https://doi.org/10.20338/bjmb.v19i1.468

HIGHLIGHTS
• Foot strike and angles presented good to excellent intra-rater reliability.
• The inter-rater ICC was good to excellent for foot strike and hip angle.
• The inter-rater ICC was weak to excellent for knee and ankle angle in Analysis 1.
• The inter-rater reliability was better for Analysis 2 than the first analysis.

ABBREVIATIONS

CV	Coefficient of variation
ICC	Intraclass correlation coefficient
MDC	Minimum detectable change
SD	Standard deviation
SEM	Standard error of measurement
2D	Two-dimensional
3D	Three-dimensional

PUBLICATION DATA
Received 16 02 2025
Accepted 31 05 2025
Published 10 07 2025

BACKGROUND: Two-dimensional (2D) kinematics analysis of running is often rater-dependent, but it provides a low-cost option, allowing the implementation of kinematics analysis in the field and clinical settings.
AIM: To determine the intra- and inter-rater reliability of manual 2D running kinematics assessment.
METHODS: Ten male runners (45±9 years) ran on a treadmill, while sagittal plane movements were recorded using a digital camera. Three raters manually analyzed the same videos twice (Analysis 1 and Analysis 2; one week apart) using Kinovea to manually identify the foot strikes and define hip, knee, and ankle angles.
RESULTS: Intra-rater reliability was good to excellent for all variables (ICC: 0.766–1.000), with SEM values up to 1.6° for the angles and MDC ranging from 0.6° to 4.3° for the angles. Inter-rater ICC ranged from weak to excellent (ICC: 0.544-1.00), being worse for knee and ankle angles. SEM reached 2.6°, and MDC values ranged up to 7.2° for the angles. Reliability indices improved in the second analysis for inter-rater agreement.
INTERPRETATION: The manual analysis of 2D running kinematics using Kinovea shows, in general, excellent intra-rater and moderate to excellent inter-rater reliability. Knee and ankle angles showed the poorest results, highlighting the importance of training in their kinematic analysis. Overall, reliability improved in the second analysis, which emphasizes the importance of training for achieving better outcomes.

KEYWORDS: Angular kinematics | Biomechanics | Motion capture | Lower extremity

INTRODUCTION

Running is a popular sport, often practiced to improve general fitness [1]. However, this activity is also associated with a high occurrence of injuries [2]. Biomechanical analysis of running can help improve performance and prevent injuries by identifying potential overloading factors, which can help guide more targeted training strategies.

Although a three-dimensional (3D) biomechanics assessment is considered the gold standard for accurate multi-planar and dimensional kinematics analysis [3], it requires expensive equipment and is often limited to major cities and research laboratories. In addition, it requires trained personnel, considerable time for data processing, and a large space for installation [4]. These factors limit the application of 3D motion capture systems in clinical settings [3] and create a challenge for implementing biomechanical analysis of runners in the field, sports clubs, and physical therapy clinics, especially in economically developing countries.

As an alternative, two-dimensional (2D) kinematics assessments are simpler and cheaper to conduct, as they can be performed using commercial video cameras and free software, such as Kinovea. 2D video analysis is portable, time-effective and cost-effective [3], making it a good alternative to subjective analysis [5] and observational scales [4]. Although there are important differences between these two approaches, we need to consider that despite the limitations, a 2D analysis can be of fundamental importance for amateur athletes, for example. Furthermore, it can be a first step when screening for important indicators of performance or injury risks.

A previous systematic review found that the intra- and inter-rater reliability of 2D analysis is generally good to excellent [6]. However, its validity compared to 3D analysis is typically poor to moderate for most angular parameters during running [6], differences that may stem from the out-of-plane movements of segments, which are common in the running motion pattern. This complexity can make it difficult to accurately identify the true angle, especially for the ankle [7]. Furthermore, camera placement is a matter of concern that impacts more a 2D than a 3D analysis. A proper protocol and adequate camera positioning can reduce the influence of these technical issues. Considering the clinical application of the 2D analyses, a confounding factor that deserves attention is the analyst influence, like when using video recordings to perform a 2D assessment with Kinovea. This influence pertains to the correct identification of events (e.g., foot strikes), which are commonly identified visually, and the accurate determination of joint angles, which are manually drawn using software. Therefore, intra- and inter-rater reliability is crucial to ensure high-quality movement assessments.

However, knowledge about the reliability of 2D kinematic analysis of running is limited. Few studies have examined intra- and inter-rater reliability using Kinovea. In one of them, gait analysis was performed with the participants visiting the lab twice within one week [4]. The authors reported a good intraclass correlation coefficient (ICC) for intra-rater assessment (0.859-0.886) and an excellent ICC (>0.90) for the inter-rater analysis [4]. This procedure considers both data acquisition and analysis influences, as different videos were used for the intra-rater analysis. This approach does not allow us to determine which specific aspect is influencing the results. Analyzing the same video twice by the same rater could provide more information about intra-rater data analysis and reduce the influence of data collection factors (e.g., marker placement, camera positioning, intra-individual variability). When the same rater assessed the same video twice, differences of approximately 3-6º in knee and hip angles at foot strike during running were found. In contrast, differences of about 6-8º were observed in the inter-rater analysis. However, the authors did not report ICC, Standard Error of Measurement (SEM), or Minimum Detectable Change (MDC) values [8], which could give information regarding the random error associated with the measurements and the minimum threshold required to consider a change as real. The determination of angular parameters depends on the correct identification of the critical events for the running technique. One of the events is the foot strike instant, which, in a previous study, varied from 5.6 to 11.6 frames between intra-rater and inter-rater assessments, with the differences being smaller for the intra-rater assessment [9]. Furthermore, foot strike patterns have been consistently associated with biomechanical loading patterns that influence the risk of common running-related injuries, such as patellofemoral pain syndrome [10]. Additionally, these variables are directly linked to running efficiency and performance, as they affect stride length, ground contact time, and energy transfer during the gait cycle [11].

Kinematic analysis using Kinovea demonstrates good to excellent intra- and inter-rater ICC for other movements such as wrist range of motion [12], shoulder range of motion [13], and the knee valgus in the lateral step down test [14]. These various studies, when combined, suggest good reliability in using 2D motion capture to analyze sagittal running kinematics. However, no previous study has combined the analysis of foot strike and angles to determine the reliability of analyzing the same videos on different days and by different examiners. Therefore, here we determine the intra-analyst (analysis of the same video twice) and inter-analyst (comparison between three analysts) reliability of 2D kinematics of running, including foot strike identification and lower extremity angular kinematics in the sagittal plane, using the Kinovea software. Our hypothesis is that the reliability of 2D kinematics of running using Kinovea would be greater for intra-rater than inter-rater assessment.

METHODS

Experimental design

The participants were recruited from the local community, visited the laboratory one day for anthropometric measurements (body mass and height), and performed a treadmill running test. During this test, the motion was captured at the sagittal plane. The 2D kinematics analysis considered the intra- and inter-rater reliability assessment. The analyses were part of a broader project addressing the effects of compression garments on running [15]. Here, the data from the pre-intervention condition were analyzed.

Participants

Runners from the local community were invited to participate in the study. To be included, they should be aged from 30 to 60 years, have at least one year of running experience, a regular weekly training volume of around 30 km, participate in running competitions, and not have a history of musculoskeletal injuries in the lower limbs in the past year. Ten amateur competitive runners were included (age 45 ± 9 years old, body mass 69 ± 7 kg, height 166 ± 4 cm, Body Mass Index 25 ± 2 kg/m², 5 km personal best 20 ± 2 min, running experience of 14 ± 13 years). The local institutional ethics committee approved this research (CAAE 76274117.4.0000.5323), and all participants signed an informed consent form. All procedures followed the Declaration of Helsinki.

Running protocol

The running protocol was conducted using a motorized treadmill (Gait Trainer 3, Biodex Inc., USA) with an inclination set at 1%. The protocol started with a 5-min warm-up walking at 1.4 m/s followed by a 3-min run at low speed (2.2 m/s). After the warm-up, the speed was increased to represent 90% of the performance time in the individual best 10 km personal record, reported by participants previously [16]. Then participants ran for 5 minutes. The average speed for all the participants was 3.5 ± 0.2 m/s. All participants ran wearing their running shoes. All participants were recommended to refrain from intense physical exercise 48 hour prior to the test, and any exercise 24 hours before the test.

Kinematic assessment

The kinematic data were obtained using 2D video analysis in the first kilometer completed after the warm-up and at the target speed. The videos were recorded using a digital camera (Panasonic, Japan, HDC-HS80, 1920x1080 full HD, shutter speed 1/125) placed aside (90 cm height and 3 m away) the treadmill with the lens perpendicular to the plane of movement. Ten consecutive strides were registered.

Spherical reflexive markers of 14 mm diameter were placed at four anatomic reference points from the left body side [5th metatarsal (foot), lateral malleolus (ankle), lateral femoral condyle (knee), and greater trochanter of the femur (hip)] to determine the kinematic parameters from the left lower limb. Images were sampled at 30 Hz, and the data were further deinterlaced to 60 Hz using a free software for kinematics analysis of movement (Kinovea, 0.9.3), used for running kinematics assessments [8].

Ten-foot strikes were visually identified from the videos, using the magnifier tool with a magnification factor of 2.5x, and registered in a spreadsheet. After this, the sagittal plane angles at the foot strike were determined and registered in the same spreadsheet, always in the same order (hip-knee-ankle). The absolute hip angle to the horizontal was drawn using the angle to the horizontal tool from Kinovea. The knee and ankle relative angles were drawn using the angle tool from Kinovea. Finally, the ankle angle was determined by drawing a line from the 5th metatarsal to the lateral malleolus and a line from the lateral malleolus to the lateral femoral condyle. The knee angle was determined by a line drawn from the lateral malleolus to the lateral femoral condyle and a line from the lateral femoral condyle to the femur's greater trochanter. The absolute hip angle was drawn from a horizontal line created by the tool and a line drawn from the femur's greater trochanter to the lateral femoral condyle. Figure 1 shows how the angles were determined and the reference values for each angle.

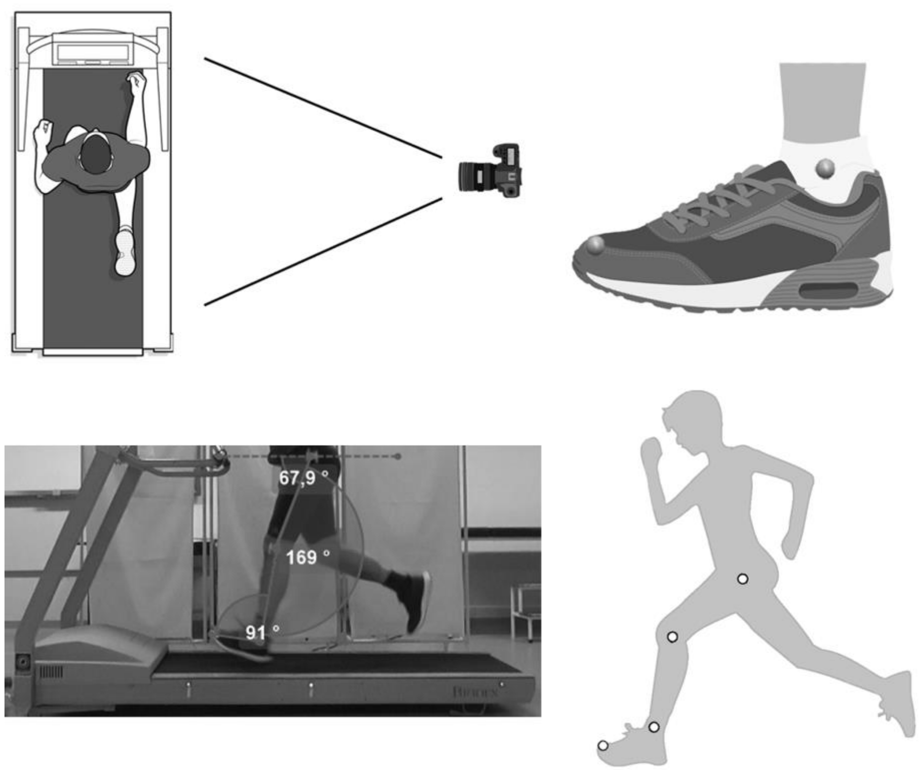

Figure 1. Camera placement, reference markers, and angle determination for the analysis.

Intra-rater and inter-rater reliability

Three raters (R1, R2 and R3) with different levels of experience in the kinematic assessment of running (raters 1 and 3: about one year; rater 2: about two years) conducted all analyses, using the same methodology as described above. A step-by-step guide was utilized to follow the same sequence of steps in the analyses, and all of the raters used the same version of the software Kinovea. Before the final analysis, the three raters carried out a 2D-kinematic analysis training for a total of 20 strides to guarantee that all of them were following the same steps.

After this training, each rater performed the kinematic analyses of the same videos two times, with about one week of interval in between. All raters used a different file for the second analysis and did not have access to the results from the first analysis while performing the second.

The intra-rater analysis represents the reliability of the same rater when analyzing the data in two different moments. These results are presented separately for each rater.

The inter-rater analyses are presented in pairs for the three raters (i.e., Rater 1 vs. Rater 2, Rater 1 vs. Rater 3, and Rater 2 vs. Rater 3). These results are presented separately for Analysis 1 and for Analysis 2 (conducted about one week after the first).

Statistical Analysis

The statistical analysis was performed with the software SPSS 23.0 and the Jamovi software [The jamovi project (2023). jamovi (Version 2.3) (Computer Software); Retrieved from https://www.jamovi.org]. Mean and standard deviation (SD) were computed in the descriptive analysis. The data normality and homogeneity were verified using the Shapiro-Wilk and Levene tests, respectively. Values analyzed by the same rater on different days were used to obtain the intra-rater reliability. Values of different raters were used to obtain the inter-rater reliability.

The intraclass correlation coefficient (ICC), the standard error of the measurement (SEM), the minimum detectable change (MDC), and the coefficient of variation (CV) were determined to verify the intra- and inter-rater reliability. The ICC was classified as weak (ICC<0.75), good (0.75>ICC<0.90), and excellent (ICC>0.90) [17]. Intra- and inter-rater reliability were assessed using a two-way mixed effects model with fixed effects for raters, ICC (3,1) model, once a fixed set of three raters performed repeated assessments on the same data. The SEM was estimated using the following equation: *SEM = SD diff**√(1-ICC). SD diff is the standard deviation of differences between measurements and was determined by the equation:

$$SD_{diff} = \sqrt{\frac{(n_1-1)\,SD^2_1+(n_2-1)\,SD^2_2}{(n_1+n_2-2)}}$$

The MDC was estimated based on a 95% confidence interval, where MDC = 1.96*SEM*√2. The CV was determined by the standard deviation to the mean ratio. A significance level of 5% was adopted for all analyses.

Bland-Altman analysis was also conducted using the following analyses: a one-sample t-test was performed to test if the differences between raters and analyses differed from zero; a linear regression was applied to verify the presence of systematic errors of measurement; and the differences and mean between analyses and raters and the limits of agreement were determined and graphically represented.

RESULTS

The intra-rater ICC was good or excellent for all angles (range of 0.766 - 0.993) and was excellent for the foot strike detection (range of 0.996-1.000) for the three raters. SEM values ranged from 0.00 to 2.2 frames for the foot strike and 0.22 to 1.6º for the angles. The MDC ranged from 0.0 to 6.0 frames for the foot strike instant and from 0.6 to 4.3° for the angles. The CV values were increased for the foot strike instant, ranging from 1.5% to 4.8% for the joint angles, with the knee angle showing the lowest values (Table 1).

Bland-Altman analyses (Supplementary File) of the intra-rater assessment revealed that the differences between the first and second analyses were significantly different from zero only for the ankle angle in the assessments performed by Rater 1 (Table S1). For all other raters and measurements, the differences did not significantly differ from zero (Table S1). None of the analyses showed systematic measurement errors (regression analyses; Table S1). Figure S1 illustrates that most of the values fell within the limits of agreement.

Considering the inter-rater analyses, the ICC values for the foot strike ranged from 0.998 to 1.000 in all analyses, being classified as excellent, with SEM values ranging from 0.0 to 1.5 frames. MDC ranged from 0.0 to 4.2 frames and the CV from 27.0 to 27.6% (Table 2).

ICC was good or excellent for the absolute hip angle in Analysis 1 and excellent for Analysis 2, with SEM values ranging from 0.3° to 0.9° in Analysis 1 and corresponding to 0.1° in Analysis 2. The MDC ranged from 0.9 to 2.6° in the first analysis and from 0.2 to 0.3° in the second analysis. The CV values ranged from 3.5% to 3.9% (Table 2).

When considering the knee angle, the ICC ranged from 0.544 to 0.948 (weak to excellent) in the first analysis (Analysis 1), and in Analysis 2, the ICC ranged from 0.820 to 0.866 (good). SEM values ranged from 0.7 to 1.8º in the first analysis and from 1.1 to 1.2° in the second. MDC ranged from 2.0 to 5.1° and from 3.0 to 3.3 in the first and second analyses, respectively. The CV values ranged from 1.4% to 1.9% (Table 2).

For the ankle angle, the SEM values ranged from 1.5 to 2.6º in Analysis 1 and from 0.7 to 2.0 in Analysis 2. In the first analysis, the ICC ranged from 0.579 to 0.915 (moderate to excellent) and in Analysis 2 from 0.780 to 0.952 (good to excellent). The MDC varied from 4.1 to 7.2° in the first analysis and from 2.1 to 5.5 in the second. The CV values ranged from 3.3% to 5.0% (Table 2).

Bland-Altman analyses (Supplementary File, Table S2 and Figures S2 to S5) of the inter-rater assessments revealed that, in Analysis 1, the differences between raters were significantly different from zero for the ankle angle (R1 vs R2; R2 vs R3; $p < 0.05$) as well as for the foot strike instant and knee angle (R2 vs R3; $p<0.05$). For the other parameters, the differences were not significantly different from zero (Table S2). In Analysis 2, the results showed improvement, with differences observed only for the ankle angle when comparing Rater 1 with Rater 2 and Rater 2 with Rater 3 ($p<0.001$). Systematic measurement errors were found only for the ankle angle (R1 vs R2; $p=0.037$) in the first analysis, and for the ankle angle (R1 vs R2 and R2 vs R3; $p<0.05$) as well as the hip angle (R2 vs R3; $p=0.035$) in the second analysis. As shown in Figures S1 to S5, most values fell within the limits of agreement.

Table 1. Intra-rater assessment.

	Analysis 1 (mean ± SD)	Analysis 2 (mean ± SD)	ICC	95%CI	p	SEM	MDC	CV (%)
				Foot strike instant (frames)				
Rater 1	122.7 ± 34.3	122.6 ± 34.4	0.996	0.983-0.999	<0.001	2.2	6.0	27.9
Rater 2	123.5 ± 33.4	123.5 ± 33.4	1.000	1.000-1.000	<0.001	0.0	0.0	27.0
Rater 3	123.6 ± 33.3	123.4 ± 33.6	1.000	1.000-1.000	<0.001	0.0	0.0	27.1
				Hip (º)				
Rater 1	67.9 ± 2.9	67.6 ± 2.5	0.924	0.692-0.981	<0.001	0.8	2.1	3.9
Rater 2	67.6 ± 2.6	67.2 ± 2.2	0.975	0.901-0.994	<0.001	0.4	1.1	3.6
Rater 3	67.7 ± 2.6	67.6 ± 2.6	0.993	0.971-0.998	<0.001	0.2	0.6	3.8
				Knee (º)				
Rater 1	161.4 ± 2.3	160.9 ± 2.9	0.869	0.471-0.967	0.003	0.9	2.6	1.5
Rater 2	162.8 ± 3.1	162.2 ± 2.6	0.978	0.911-0.995	<0.001	0.4	1.2	1.8
Rater 3	160.0 ± 3.1	161.4 ± 3.3	0.766	0.570-0.942	0.021	1.6	4.3	1.8
				Ankle (º)				
Rater 1	97.8 ± 5.2	95.4 ± 4.3	0.923	0.690-0.981	<0.001	1.3	3.7	4.8
Rater 2	90.5 ± 2.8	89.9 ± 2.4	0.955	0.818-0.989	<0.001	0.6	1.5	2.8
Rater 3	96.6 ± 4.9	96.2 ± 4.2	0.935	0.740-0.984	<0.001	1.2	3.2	4.6

ICC: Intraclass Correlation Coefficient; CI: Confidence Interval; SEM: Standard Error of the Measurement; MDC: Minimum Detectable Change; CV: Coefficient of Variation.

Table 2. Inter-rater assessment.

	Rater 1 (mean ± SD)	Rater 2 (mean ± SD)	Rater 3 (mean ± SD)	ICC	95%CI	p	SEM	MDC	CV (%)
Foot strike instant (frames)									
					Analysis 1				
R1 *vs* R2	122.7 ± 34.3	123.5 ± 33.4	--	0.999	0.992-1.000	<0.001	1.1	3.0	27.5
R1 *vs* R3	122.7 ± 34.3	--	123.6 ± 33.3	0.998	0.992-0.999	<0.001	1.5	4.2	27.4
R2 *vs* R3	--	123.5 ± 33.4	123.6 ± 33.3	1.000	1.000-1.000	<0.001	0.0	0.0	27.0
					Analysis 2				
R1 *vs* R2	122.6 ± 34.4	123.5 ± 33.4	--	0.998	0.991-0.999	<0.001	1.5	4.2	27.5
R1 *vs* R3	122.6 ± 34.4	--	123.4 ± 33.6	0.998	0.991-0.999	<0.001	1.5	4.2	27.6
R2 *vs* R3	--	123.5 ± 33.4	123.4 ± 33.6	0.999	1.000-1.000	<0.001	1.1	2.9	27.1
Hip joint (°)									
					Analysis 1				
R1 *vs* R2	67.9 ± 2.9	67.6 ± 2.6	--	0.892	0.565-0.973	<0.001	0.9	2.6	3.9
R1 *vs* R3	67.9 ± 2.9	--	67.7 ± 2.6	0.924	0.694-0.981	<0.001	0.8	2.1	3.9
R2 *vs* R3	--	67.6 ± 2.6	67.7 ± 2.6	0.983	0.932-0.966	<0.001	0.3	0.9	3.8
					Analysis 2				
R1 *vs* R2	67.6 ± 2.5	67.2 ± 2.2	--	0.998	0.991-0.999	<0.001	0.1	0.3	3.5
R1 *vs* R3	67.6 ± 2.5	--	67.6 ± 2.6	0.999	0.997-1.000	<0.001	0.1	0.2	3.8
R2 *vs* R3	--	67.2 ± 2.2	67.6 ± 2.6	0.998	0.991-0.999	<0.001	0.1	0.3	3.5
Knee joint (°)									
					Analysis 1				
R1 *vs* R2	161.4 ± 2.3	162.8 ± 3.1	--	0.780	0.114-0.945	0.017	1.3	3.5	1.5
R1 *vs* R3	161.4 ± 2.3	--	160.0 ± 3.1	0.544	-0.836-0.887	0.129	1.8	5.1	1.4
R2 *vs* R3	--	162.8 ± 3.1	160.0 ± 3.1	0.948	0.792-0.987	<0.001	0.7	2.0	1.9
					Analysis 2				
R1 *vs* R2	160.9 ± 2.9	162.2 ± 2.6	--	0.820	0.274-0.955	0.009	1.2	3.3	1.6
R1 *vs* R3	160.9 ± 2.9	--	161.4 ± 3.3	0.859	0.432-0.965	0.004	1.2	3.3	1.8
R2 *vs* R3	--	162.2 ± 2.6	161.4 ± 3.3	0.866	0.461-0.967	0.003	1.1	3.0	1.7
Ankle joint (°)									
					Analysis 1				
R1 *vs* R2	97.8 ± 5.2	90.5 ± 2.8	--	0.691	-0.245-0.923	0.048	2.3	6.5	3.9
R1 *vs* R3	97.8 ± 5.2	--	96.6 ± 4.9	0.915	0.657-0.979	0.001	1.5	4.1	5.0
R2 *vs* R3	--	90.5 ± 2.8	96.6 ± 4.9	0.579	-0.695-0.895	0.107	2.6	7.2	3.6
					Analysis 2				
R1 *vs* R2	95.4 ± 4.3	89.9 ± 2.4	--	0.853	0.407-0.963	0.004	1.3	3.7	3.5
R1 *vs* R3	95.4 ± 4.3	--	96.2 ± 4.2	0.780	0.113-0.945	0.017	2.0	5.5	4.3
R2 *vs* R3	--	89.9 ± 2.4	96.2 ± 4.2	0.952	0.806-0.988	<0.001	0.7	2.1	3.3

R: Rater; Pooled: Analysis of all analysts pooled; ICC: Intraclass Correlation Coefficient; CI: Confidence Interval; SEM: Standard Error of the Measurement; MDC: Minimum Detectable Change; CV: Coefficient of Variation.

DISCUSSION

Our primary goal was to determine the intra-rater and inter-rater reliability of 2D kinematics of running, including foot strike identification and angular kinematics in the sagittal plane, using the Kinovea software. We tried to replicate a common condition of runners' field assessment when limited resources for a 3D analysis are available with this approach. Moreover, different raters can be enrolled in the athletes' assessment. We confirmed our hypothesis regarding greater reliability for intra-rater than inter-rater assessments of 2D running kinematics using Kinovea.

Different from a previous study [4], the intra-rater analysis was performed with the same video twice. In doing so, we could assess specifically the kinematic manual data analysis effect once the data collecting setup and the videos were the same, thus not influencing the results. Aspects like the camera and markers' placement can also affect the analysis, but this was not the case here. An analysis of the same running video was performed in a previous study [8]. Although ICC, SEM, and MDC values were not presented, they found a difference of about 3-6º for knee and hip angles [8], which is larger than the values found here, ranging from 0.1 to 1.4º between the first and the second analysis.

The intra-rater results demonstrated that the analysis is reliable in raters with different levels of experience. Good to excellent ICC values were found for the knee angle, and excellent ICC values were found for the other parameters, with the highest SEM value of 2.2º. The rater with more significant experience (~2 years) with the Kinovea presented, in general, greater ICC values and lower SEM values compared to the other two raters (~1 year of experience). MDC values ranged from 0.6 to 4.3 for the angles that represent less than 5% of the angle values and suggest that the minimum difference necessary to be considered a real change in the individual/participant is low.

Inter-rater ICC was greater for the foot strike instant and the hip angle than the knee and ankle angles. This could be related to the fact that the foot strike identification is more reliable and more comfortable to identify due to less motion of the markers. We also consider that the hip angle, determined as an absolute measurement, while the knee and ankle angles were relative, was a factor of influence. By considering the absolute angle, the hip angle relied only on the identification of two markers. It might have influenced the ICC and SEM values, which were greater and lower for the hip angle, respectively, compared to the knee and ankle angles. These results were consistent with the Bland-Altman analyses, which showed satisfactory outcomes overall (most of the data fell within the limits of agreement), although the ankle angle presented the poorest results.

The inter-rater reliability was better for the second analysis (Analysis 2) than for the first analysis. We found SEM values smaller for inter-rater reliability, which indicates a small, significant number of errors. Most of the ICC classified as weak or good in the first analysis were considered, respectively, as good and excellent in the second. These aspects indicate that even though the intra-rater analysis is reliable, performing the analysis of the same video twice or having more training can improve reliability. It is also important to highlight that we had three raters, and all the pairwise analyses between them showed similar results.

When using video records to perform a 2D assessment using Kinovea, the video rater has a determinant influence on the correct identification of the foot strikes and the manually drawn angles. From our results, it seems that this type of analysis can be performed reliably by the same rater and by different raters, contributing to a high-quality movement assessment. It was previously demonstrated that a 2D analysis could present similar results as the 3D [5,7,18], despite its known limitations. These limitations include the out-of-plane movements performed by the segments, common in the running pattern of movement, and influence the 2D angle determination [7]. Our results strengthen the applicability of a 2D analysis using a commercial camera and the free software Kinovea. This is especially important once a 2D analysis is cheaper and simpler to conduct than the 3D and is also portable and time, and cost-effective [3], and can be utilized when limited resources are available in different clinical and health centers.

Similar to previous studies [9,12], the intra-rater analyses were better than the inter-rater analyses. Although training and a discussion of the analysis were performed before starting the data analysis, and although all raters followed the same steps on data analysis, different raters still influenced the results. We attribute these differences to the manual actions required in the analysis, despite most of the parameters presenting good to excellent ICC for the inter-rater analyses, especially in the second analysis. Additionally, each rater independently identified the foot strike instant and determined the angles at this instant. Even if they have not always identified the foot strike at the same frame, high ICC was found for most of the angles measured, which is also an important aspect that favors the systematic steps performed during all the analyses.

Our study has limitations. We tried to replicate the condition of clinicians and opted for a low-cost setup for data collection. Therefore, our camera was limited to 60 Hz for the video recording. A higher-speed camera may contribute to better results to providing more realistic data. Our participants were typical runners, not presenting significant alterations in the running technique and no movement constraints. In the case of a clinical context considering runners with altered motion patterns, primarily increased non-sagittal movements, further attention should be paid.

CONCLUSION

The manual analysis of 2D sagittal running kinematics shows good to excellent intra-rater and weak to excellent inter-rater reliability. Foot strike identification and absolute hip angle, manually determined, were more reliable than knee and ankle relative angles. Inter-rater analysis improves in the second analysis of the same video, with weak or good ICC becoming good or excellent in the second analysis. We conclude that the analyses are reliable and allow the comparison of results from the same rater and the same video on different days and the comparison between three raters on different days of video-analysis.

REFERENCES

1. Fields KB, Sykes JC, Walker KM, Jackson JC. Prevention of running injuries. *Curr Sports Med Rep.* 2010;9:176-182. doi: 10.1249/JSR.0b013e3181de7ec5.
2. Van Der Worp MP, Ten Haaf DS, Van Cingel R, De Wijer A, Nijhuis-Van Der Sanden MW, Staal JB. Injuries in runners; a systematic review on risk factors and sex differences. *PLoS One.* 2015;10:e0114937. doi: 10.1371/journal.pone.0114937.
3. Schurr SA, Marshall AN, Resch JE, Saliba SA. Two-Dimensional Video Analysis Is Comparable to 3d Motion Capture in Lower Extremity Movement Assessment. *Int J Sports Phys Ther*. 2017;12:163-172. PMID: 28515970
4. Fernandez-Gonzalez P, Koutsou A, Cuesta-Gomez A, Carratala-Tejada M, Miangolarra-Page JC, Molina-Rueda F. Reliability of Kinovea((R)) Software and Agreement with a Three-Dimensional Motion System for Gait Analysis in Healthy Subjects. *Sensors (Basel).* 2020;20. doi: 10.3390/s20113154.
5. Zult T, Allsop J, Tabernero J, Pardhan S. A low-cost 2-D video system can accurately and reliably assess adaptive gait kinematics in healthy and low vision subjects. *Sci Rep.* 2019;9:18385. doi: 10.1038/s41598-019-54913-5.
6. Leporace G, Metsavaht L, Gonzalez FF, Arcanjo De Jesus F, Machado M, Celina Guadagnin E, et al. Validity and reliability of two-dimensional video-based assessment to measure joint angles during running: A systematic review and meta-analysis. *J Biomech.* 2023;157:111747. doi: 10.1016/j.jbiomech.2023.111747.
7. Davis RB, Õunpuu S. A Comparison of 2D and 3D Techniques for the Determination of Normal Pediatric Running Kinematics. In: *Proceedings of the 1991 IEEE Seventeenth Annual Northeast Bioengineering Conference* 1991.
8. Damsted C, Nielsen RO, Larsen LH. Reliability of video-based quantification of the knee- and hip angle at foot strike during running. *Int J Sports Phys Ther.* 2015;10:147-154.
9. Damsted C, Larsen LH, Nielsen RO. Reliability of video-based identification of footstrike pattern and video time frame at initial contact in recreational runners. *Gait Posture.* 2015;42:32-35. doi: 10.1016/j.gaitpost.2015.01.029.
10. Wei Z, Hou X, Qi Y, Wang L. Influence of foot strike patterns and cadences on patellofemoral joint stress in male runners with patellofemoral pain. *Phys Ther Sport*. 2024;65:1-6. doi: 10.1016/j.ptsp.2023.10.006.
11. Chen H, Sun D, Zhang Q, Liu Q, Bíró I, Gu Y. Biomechanical differences between habitual forefoot and rearfoot strike running: A systematic review. *Advanced Exercise and Health Science.* 2024;1:241-247. doi: 10.1016/j.aehs.2024.11.004.
12. El-Raheem RMA, Kamel RM, Ali MF. Reliability of Using Kinovea Program in Measuring Dominant Wrist Joint Range of Motion. *Trends in Applied Sciences Research.* 2015;10:224-230.
13. Elrahim RMA, Embaby EA, Ali MF, Kamel RM. Inter-rater and intra-rater reliability of Kinovea software for measurement of shoulder range of motion. *Bulletin of Faculty of Physical Therapy*. 2016;21.
14. Ribeiro DB, Rodrigues GM, Bertoncello D. Intra and inter-rater reliability in dynamic valgus in soccer players. *Rev Bras Med Esporte.* 2020;26.
15. Carvalho J, Kunzler MR, Priego-Quesada JI, Aparicio I, Perez-Soriano P, Machado AS, et al. Effects of 24 h Compression Interventions with Different Garments on Recovery Markers during Running. *Life (Basel).* 2021;11. doi: 10.3390/life11090905.
16. Brophy-Williams N, Driller MW, Kitic CM, Fell JW, Halson SL. Effect of Compression Socks Worn Between Repeated Maximal Running Bouts. *Int J Sports Physiol Perform.* 2017;12:621-627. doi: 10.1123/ijspp.2016-0162.
17. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15:155-163. doi: 10.1016/j.jcm.2016.02.012.
18. Hanley B, Tucker CB, Bissas A. Differences between motion capture and video analysis systems in calculating knee angles in elite-standard race walking. *J Sports Sci.* 2018;36:1250-1255. doi: 10.1080/02640414.2017.1372928

Citation: Guadagnin EC, Carvalho JS, Costa VL, Martins CS, Sonda FC, Carpes FP. (2025). Intra and inter-rater reliability of manual two-dimensional kinematics analysis of running. *Brazilian Journal of Motor Behavior*, 19(1):e468.

Funding: The Brazilian National Council for Scientific and Technological Development (CNPq – Brazil) supports FPC. ECG was funded by the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 for ECG, JSC and FCS.
Competing interests: The authors have declared that no competing interests exist.
DOI: https://doi.org/10.20338/bjmb.v19i1.468