Journal of Sports Sciences

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rjsp20

From lab to field: validity and reliability of inertial measurement unit-derived gait parameters during a standardised run

Elliot Webber, Cédric Leduc, Stacey Emmonds, Michael Eglon, Brian Hanley, Zafar Iqbal, Samrat Sheoran, Claire Chaisson & Dan Weaving

To cite this article: Elliot Webber, Cédric Leduc, Stacey Emmonds, Michael Eglon, Brian Hanley, Zafar Iqbal, Samrat Sheoran, Claire Chaisson & Dan Weaving (28 Sep 2024): From lab to field: validity and reliability of inertial measurement unit-derived gait parameters during a standardised run, Journal of Sports Sciences, DOI: 10.1080/02640414.2024.2408195

To link to this article: https://doi.org/10.1080/02640414.2024.2408195

Published online: 28 Sep 2024.

JOURNAL OF SPORTS SCIENCES
https://doi.org/10.1080/02640414.2024.2408195

Routledge
Taylor & Francis Group

SPORTS PERFORMANCE

OPEN ACCESS Check for updates

From lab to field: validity and reliability of inertial measurement unit-derived gait parameters during a standardised run

Elliot Webber[a,b], Cédric Leduc[a,b], Stacey Emmonds[a], Michael Eglon[b], Brian Hanley[a], Zafar Iqbal[b], Samrat Sheoran[a], Claire Chaisson[a] and Dan Weaving[c,d]

[a]Centre for Human Performance, Carnegie School of Sport, Leeds Beckett University, Leeds, UK; [b]Sport Science and Medicine Department, Crystal Palace FC, London, UK; [c]Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK; [d]Applied Sports Science and Exercise Testing Laboratory, The University of Newcastle, Ourimbah, NSW, Australia

ABSTRACT

The aim was to assess concurrent validity and test–retest reliability of spatiotemporal gait parameters from a thoracic-placed inertial measurement unit (IMU) in lab- (Phase One) and field-based (Phase Two) conditions. Spatiotemporal gait parameters were compared (target speeds 3, 5 and 7.5 $m{\cdot}s^{-1}$) between a 100 Hz IMU and an optical measurement system (OptoJump Next, 1000 hz) in 14 trained individuals (Phase One). Additionally, 29 English Premier League football players performed weekly 3×60 m runs (5 $m{\cdot}s^{-1}$; observations = 1227; Phase Two). Mixed effects modelling assessed the effect of speed on agreement between systems (Phase One) and test–retest reliability (Phase Two). IMU step time showed strong agreement (<0.3%) regardless of individual or running speed. Direction of mean biases up to 40 ms for contact and flight time depended on the running speed and individual. Step time, length and frequency were most reliable (coefficient of variation = 1.3-1.4%) but confounded by running speed. Step time, length and frequency derived from a thoracic-placed IMU can be used confidently. Contact time could be used if bias is corrected for each individual. To optimise test–retest reliability, a minimum running distance of 40 m is needed to ensure 10 constant-speed steps is gathered.

ARTICLE HISTORY
Received 28 January 2024
Accepted 17 September 2024

KEYWORDS
Accelerometer; monitoring; soccer; team sport

Introduction

Monitoring neuromuscular fatigue within elite football is a vital process as it can be present for up to 72 h after match-play, and thus potentially require consideration when prescribing subsequent training sessions (Girard et al., 2015; Rampinini et al., 2011). However, concerns have been raised surrounding the time efficiency and validity of common monitoring processes such as the countermovement jump and subjective questionnaires when applied in elite practical environments (Carling et al., 2018; Jeffries et al., 2020). Therefore, unobtrusive monitoring processes such as standardised running tests (SRT) has been proposed as more feasible and specific measures (Leduc et al., 2020). As elite football players habitually wear inertial measurement units (IMUs) on the thoracic spine, one such approach is a SRT in which players run at a constant speed between 50 to 60 m whilst concurrently wearing the device (Leduc et al., 2020).

IMUs contain a triaxial accelerometer that can assess the magnitude of instantaneous accelerations as an overall vector, or as individual vertical, mediolateral and anteroposterior vectors. Previous studies observing these vectors during SRTs have reported pre- and post-changes following team sport matches and training (Garrett et al., 2019; Leduc et al., 2020). It is thought that this is due to changes in spatiotemporal gait parameters such as step length and frequency (Leduc et al., 2020; Small et al., 2009). Although spatiotemporal gait parameters change when fatigued (Apte et al., 2021; Riazati et al., 2022), a direct association with accelerometer measures has not been provided. Additionally, accelerometer magnitude measures are conceptually ambiguous and limit the interpretation of the potential mechanisms of change, which potentially reduces their use in practice (Leduc et al., 2020). Considering that accelerometer measures are proposed as an indirect measure of spatiotemporal parameters, and a larger body of evidence has observed alterations to spatiotemporal gait parameters with fatigue (Apte et al., 2021), it is likely more beneficial to assess these measures directly rather than through global accelerometer measures. However, currently such variables are not provided by common manufacturers of IMUs used in elite sport (Fitzpatrick et al., 2021) and hence requires algorithm development for calculation.

A previous study has observed spatiotemporal parameters (i.e., step time) from a thoracic placed accelerometer to have good agreement ($R = 0.68–0.98$) with ground reaction force derived from an instrumented treadmill in a single team sport athlete (Buchheit et al., 2015). However, the external validity is questionable due to potential between-subject variability in gait kinematics (Phinyomark et al., 2015). For an algorithm to be used across multiple athletes, it must be robust to between-subject variability in gait strategy when deriving spatiotemporal parameters. Recently, Horsley et al. (2023) assessed the agreement between spatiotemporal parameters from a thoracic

CONTACT Elliot Webber e.webber@leedsbeckett.ac.uk Centre for Human Performance, Carnegie School of Sport, Leeds Beckett University, Leeds, UK
Supplemental data for this article can be accessed online https://doi.org/10.1080/02640414.2024.2408195.

accelerometer vs. lower back- and tibia-placed accelerometers. Despite absolute mean biases of less than 4% between placement sites for step time and step frequency, irrespective of running speed (3–8 m·s^{-1}), there was a lack of assessment against a criterion measure (e.g., OptoJump, video analysis). Therefore, to establish the suitability of these parameters for practical sporting environments, agreement against a criterion measure is also needed.

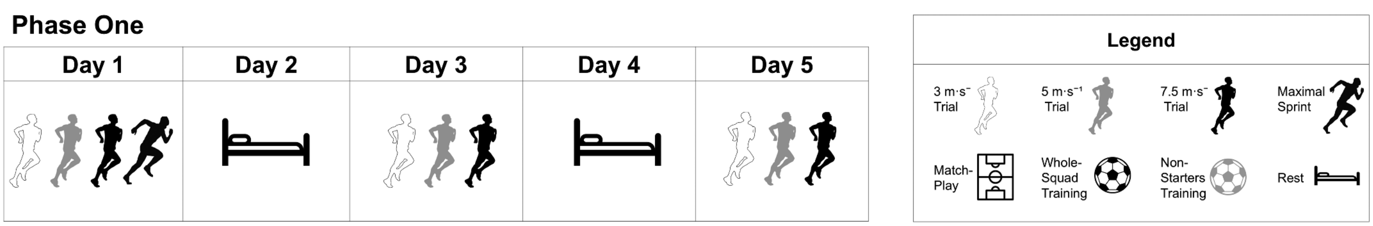
Regardless of the measure, the protocol used to identify alterations to gait also needs consideration. Currently, variation exists in running protocols in practice, including average running speeds (5.0 to 6.7 m·s^{-1}) and distances (50 to 60 m) (Leduc et al., 2020). Although both running speed (Novacheck, 1998) and the number of steps included (as per the central limit theorem) within analyses are likely to affect the observed average spatiotemporal parameters, no study has assessed how alterations to these affect parameter validity and reliability. Identifying a protocol that maximises these would allow for alignment of protocols in future studies, consequently leading to more accurate between-study comparisons, and provide useful guidelines for practitioners. Therefore, this study had two distinct phases. In phase one, the aim was to assess the concurrent validity of IMU gait parameters against a criterion measure. In phase two, the aim was to assess the test–retest reliability of further derived spatiotemporal gait parameters within an elite male footballing environment.

Methods

Participants and study design

Both experimental (Phase One) and observational (Phase Two) designs were conducted leading to two separate groups of participants (Figure 1). In Phase One, 14 trained male individuals (age: = 24.5 ± 2.4 yrs, height: = 1.81 ± 0.07 m, mass: = 85.1 ± 9.7 kg, maximum speed = 8.26 ± 0.68 m·s^{-1}) completed three testing visits (1 × familiarisation; 2 × data collection; 48 h between) comprising a warm-up and three sets of three 60 m runs in a single lane of an indoor athletics track. Concurrently, participants wore an IMU (STATSports Apex, STATSports, Newry, UK) whilst 40 m of the criterion OptoJump Next photocell system (Microgate, Bolzano, Italy) sampling at 1000 Hz was placed either side of the running lane. A final step count of 6429 was observed during this phase, with which the IMU data collected during these testing sessions was used to develop and validate an algorithm (Figure 2) to calculate temporal gait characteristics (i.e., step-, contact- and flight time and duty factor).

In Phase Two, 29 male players from a single English Premier League (EPL) football team (age: = 27.5 ± 5.4 yrs; height: = 1.84 ± 0.05 m; mass: = 84.1 ± 7.0 kg) participated. This number was limited by the number of players in the squad. Across the two pre-season and in-season periods, participants performed 42.3



Phase Two

	MD	MD+1	MD+2	MD-4	MD-3	MD-2	MD-1
Starters							
Total Number of Sessions per Player (n)	26.2 ± 17.0 (1 – 52)	0 ± 0 (0 - 0)	12.4 ± 9.6 (1 – 27)	19.0 ± 11.0 (2 – 34)	14.6 ± 9.0 (1 – 26)	24.9 ± 15.2 (2 – 46)	33.5 ± 23.0 (1 – 73)
Total Duration (minutes)	90.4 ± 12.6 (24 – 111)	0 ± 0 (0 - 0)	50.4 ± 18.8 (19 – 92)	82.4 ± 10.8 (42 – 118)	82.4 ± 12.5 (48 – 117)	69.9 ± 13.3 (27 – 101)	56.8 ± 13.0 (27 - 101)
Total Distance (m)	9490 ± 1233 (2430 – 12082)	0 ± 0 (0 - 0)	3486 ± 1276 (1304 – 6252)	5329 ± 731 (2597 – 7178)	6123 ± 1043 (2720 – 9056)	4255 ± 1003 (697 – 6764)	2980 ± 649 (980 – 5906)
HSR Distance (m)	673 ± 228 (182 – 1497)	0 ± 0 (0 - 0)	142 ± 103 (0 – 646)	218 ± 150 (0 – 931)	359 ± 200 (0 – 1013)	122 ± 90 (0 – 538)	55 ± 45 (0 – 537)
Sprint Distance (m)	143 ± 84 (5 – 510)	0 ± 0 (0 - 0)	5 ± 13 (0 – 86)	22 ± 30 (0 - 212)	52 ± 46 (0 – 242)	12 ± 19 (0 -183)	4 ± 9 (0 – 88)
Accelerations & Decelerations (n)	150 ± 40 (48 – 284)	0 ± 0 (0 - 0)	55 ± 27 (4 – 143)	128 ± 43 (22 – 327)	114 ± 39 (18 – 284)	87 ± 40 (2 – 268)	64 ± 25 (5 – 189)
Non-Starters							
Total Number of Sessions per Player (n)	7.8 ± 7.1 (1 – 30)	3.9 ± 3.4 (1 – 12)	6.7 ± 6.0 (1 - 23)	19.0 ± 11.0 (2 – 34)	14.6 ± 9.0 (1 – 26)	24.9 ± 15.2 (2 – 46)	8.8 ± 7.5 (1 – 29)
Total Duration (minutes)	23.1 ± 13.0 (3 – 87)	62.5 ± 7.3 (51 – 83)	62.8 ± 11.1 (41 – 90)	82.4 ± 10.8 (42 – 118)	82.4 ± 12.5 (48 – 117)	69.9 ± 13.3 (27 – 101)	58.4 ± 10.3 (37 – 82)
Total Distance (m)	2573 ± 1363 (264 – 8682)	4335 ± 651 (3334 – 7531)	4518 ± 831 (2967 – 7049)	5329 ± 731 (2597 – 7178)	6123 ± 1043 (2720 – 9056)	4255 ± 1003 (697 – 6764)	3172 ± 523 (1490 – 4526)
HSR Distance (m)	216 ± 127 (0 – 714)	285 ± 145 (17 – 630)	143 ± 69 (0 – 476)	218 ± 150 (0 – 931)	359 ± 200 (0 – 1013)	122 ± 90 (0 – 538)	70 ± 43 (0 – 176)
Sprint Distance (m)	43 ± 38 (0 – 242)	18 ± 29 (0 – 168)	3 ± 8 (0 – 47)	22 ± 30 (0 - 212)	52 ± 46 (0 – 242)	12 ± 19 (0 -183)	4 ± 7 (0 – 48)
Accelerations & Decelerations (n)	41 ± 22 (2 – 125)	117 ± 27 (56 – 223)	101 ± 32 (38 – 228)	128 ± 43 (22 – 327)	114 ± 39 (18 – 284)	87 ± 40 (2 – 268)	62 ± 23 (6 – 119)

Figure 1. Outline of the study design with descriptive statistics across a typical training week for the participants in phase two. Values are reported at mean ± standard deviation (range). MD = match-day. HSR = high speed running. High speed running and sprinting distances were deemed the distances covered above running speeds of 5.5 and 7.2 m·s^{-1}.

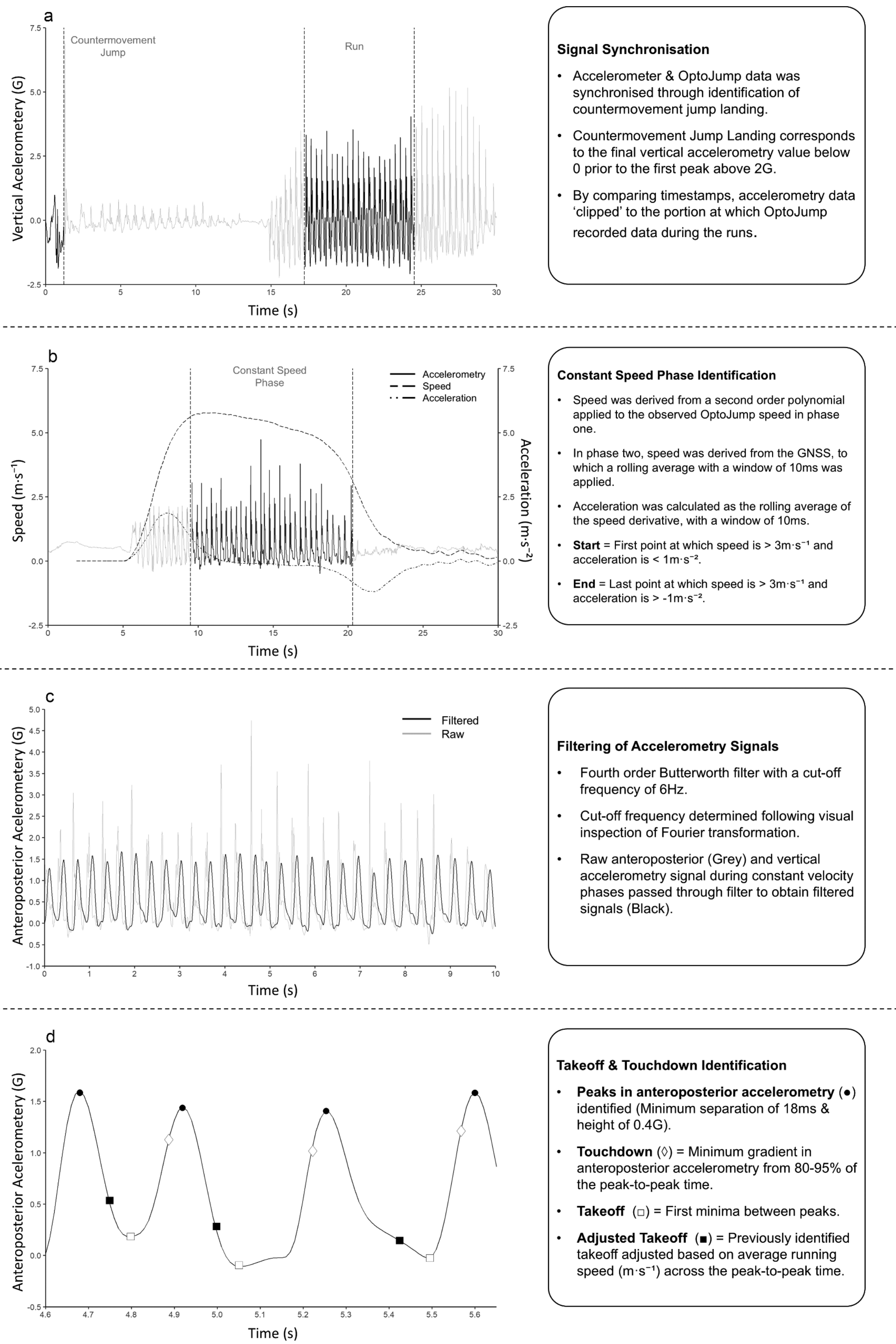


Figure 2. Outline of the algorithm development. A – a vertical accelerometry signal collected in phase 1 of the study highlighting the portion of the accelerometry signal occurring during the countermovement jump and run observed by OptoJump. b - anteroposterior accelerometry, running speed and acceleration collected during a single run during phase 2. c – the raw and filtered anteroposterior signal during the constant speed phase of a run. d – a representation of the identification of take-off and touchdown from the anteroposterior accelerometry signal.

± 23.7 SRTs during an on-field warm-up (Figure 1), typically 42 h (725 player-observations) or 76 h post-match (257 player-observations). Other trials occurred 1 day before matches (129 player-observations) or during pre-season (116 player-observations) for a final total of 1227 player-observations. Participants completed $3 \times 64 \pm 0.8$ m runs, each over 12 s, with 30 s passive rest. All testing was performed as part of the participants' regular monitoring regimen. Informed written consent was provided with ethical approval granted by the Leeds Beckett University Carnegie School of Sport Ethics Committee (Reference Number 112,885).

Protocols and data collection

Phase one – concurrent validity

During the trials (3 × 3×60 m run), running speed increased across the three sets ($3\,m{\cdot}s^{-1}$, $5\,m{\cdot}s^{-1}$, 7.5 $m{\cdot}s^{-1}$). Passive rest was provided between each repetition (30 s) and set (5 min). An audio cue controlled the speed, with participants aiming to pass cones every 15 m at the beep. Five minutes after the final set, participants completed a single 60 m maximal sprint to calculate each participant's maximal speed which was later used to express submaximal speeds as a relative percentage. The number of required steps was calculated using GPower (Erdfelder et al., 1996), with an alpha value of 0.05, power of 0.8, and correlation coefficient of 0.1. The subsequent requirement of 1077 steps was met ($n = 6429$) with a sample size of 14. Participants wore the IMU (100 Hz triaxial accelerometer) during trials between the scapulae within the manufacturer's vest and assigned the same device and vest throughout. Although intra-unit reliability has not been assessed for this specific accelerometer, similar devices have shown high reliability of instantaneous accelerometry derived measures (CV < 10%) (Crang et al., 2021). Additionally, 40 m of OptoJump Next photocell sensors (Microgate, Bolzano, Italy) were connected in series and placed either side of the running lane. Approximately 20 s before the start of each set, participants performed a submaximal countermovement jump between the OptoJump sensors (Figure 2(a)). This allowed for synchronization between the accelerometry signal and OptoJump.

OptoJump has demonstrated concurrent validity for spatiotemporal measures against video analysis and force plates (mean bias < 0.004 s) during overground race walking at speeds between 3.0 and 4.2 $m{\cdot}s^{-1}$ (Hanley & Tucker, 2019). OptoJump has additionally shown good inter- and intra-session reliability (CV = 5% to 7%) for contact- and step-time (Gomez Bernal et al., 2016). To minimise the bias of OptoJump, at least three light emitting diodes were required interruption to infer ground contact (Hanley & Tucker, 2019). During all runs, OptoJump continuously collected step-, contact-, and flight time and speed for each step. All accelerometry and OptoJump data were downloaded using SONRA and OptoJump Next software, respectively.

Gait characteristics algorithm

This algorithm involved the identification of runs and their constant speed phases (Figure 2(a,b)), correction & filtering of triaxial accelerometry data (Figure 2C), and touchdown and take-off identification (Figure 2(d)). The concurrent validity of IMU-derived temporal parameters were then compared with OptoJump Next-derived parameters.

Phase two – field based test-retest reliability

Over a period of two pre-seasons ($n = 14$ weeks) and in-season periods ($n = 48$ weeks) the participants performed three bouts (30 s passive recovery between bouts) of 60 m running on a grass pitch at a target speed of $5\,m{\cdot}s^{-1}$ ($n = 1227$ player observations), as per previous methods (Leduc et al., 2020). The number of participants in this phase was limited to 29 by the number of available players in the Running speed was controlled by a physical performance coach who ran alongside the participants with a stopwatch. Two familiarisation sessions were performed before data collection. Other than the inclusion of a countermovement jump, participants performed the same procedure as described for phase one. Following each training session, 100 Hz triaxial accelerometer and 10 Hz speed data were exported to a CSV file using SONRA software. Observations were not considered further if data were <8 connected GNSS satellites (16.8 ± 4.3) (Malone et al., 2017), resulting in the removal of three observations (final observations [n] = 1224).

Analysis

The algorithm to calculate step-, contact-, and flight time and duty factor was implemented on data collected in both phases of the study. This resulted in a total of 6429 steps extracted from 257 bouts of 60 m runs observed in Phase One, and 104,023 steps extracted from 3627 runs observed in Phase Two. In addition to step-, contact-, and flight-time and duty factor, during the second phase, step length, step frequency, running load index (RLI) metrics (Total, anteroposterior, mediolateral, and vertical), and vertical stiffness were also calculated on a step-by-step basis. Vertical stiffness, maximum vertical force, and vertical displacement were calculated in accordance with Morin et al. (2005). The definitions and equations for all metrics are provided in Supplementary Table S1. Additionally, during the second phase, following visual inspection of the step length histogram, steps were removed if their magnitude was greater than or equal to 2.36 m. Past this step length, the frequency of each step length was one or less, and therefore deemed erroneous. This resulted in the removal of 153 steps, and a final sample of 103,870 steps.

Statistical analysis – phase 1

For both IMU- and OptoJump-derived step characteristic variables, normality of step-, contact, flight time, and duty factor were assessed by visual inspection of Q-Q plots. All variables apart from flight time derived from OptoJump were deemed non-normal. It is likely that skewness was caused by disparities in the number of steps observed at each target speed ($3\,m{\cdot}s^{-1} = 2938$; $5\,m{\cdot}s^{-1} = 1884$; $7.5\,m{\cdot}s^{-1} = 1607$). To assess absolute agreement, Spearman correlations with bootstrapped 95% confidence intervals (95% CI) were computed and interpreted as per previous guidelines (Schoeber et al., 2018). To assess potential mean biases and non-parametric 95% limits of agreement (LoA) Bland-Altman plots were produced. The non-parametric LoAs were identified as the 2.5% and 97.5% quantiles of the difference scores. To identify any potential effect of absolute or relative speed on the absolute agreement between OptoJump and the IMU, a mixed-effects model with random intercept (τ_0) and slope (τ_1) was fitted using the *lmer* function in the *lme4* R package (Bates et al., 2015). The dependent variables were the absolute differences in step time, contact time, flight time, and duty factor between the IMU and OptoJump. Full model details are provided in Supplementary Table S1.

Statistical analysis – phase 2

The test–retest reliability of all metrics derived from the IMU was assessed iteratively on a step-by-step basis from the first three steps of the constant speed phase, up to 25 steps. Two general linear mixed-effect models were built for each metric per iteration. The first was an empty random effects model with participant as the random effect. Absolute speed and its quadratic term were added as fixed effects in the second. The *buildlme* function within the *buildmer* package (Voeten, 2023) optimised model fit via Aikake Information Criterion and backwards elimination of both fixed and random effects. This produced finalised general linear mixed effect models or linear regressions. A first-order autoregressive covariance structure with respect to days since the first observation was used within each model (Hecksteden et al., 2015). Model assumptions of normality and homoscedasticity of residuals, multicollinearity, and autocorrelation were assessed, and found unviolated. Coefficients of variation (CV) and smallest worthwhile changes were calculated at the group level by expressing the standard deviation (SD) of residual variation and 0.2 multiplied by the between-subject variation as percentages of the predicted score at the mean speed observed across all trials (5.76 $m{\cdot}s^{-1}$) respectively. Significant effects of fixed effects on each spatiotemporal gait parameter were determined if $p < 0.05$. The statistical power achieved by the models including speed was assessed via simulation-based analyses with the *mixedpower* function in the *mixedpower* R package (Kumle et al., 2021).

Results

Concurrent validity (phase one)

Table 1 provides descriptive statistics for all temporal variables for both IMU and OptoJump.

Figure 3 shows the correlation between measures for all temporal variables and Bland-Altman plots indicating overall mean bias and the change in mean biases across different absolute and relative speeds.

IMU step time showed the greatest agreement with OptoJump (*R* value [95% CI] = 0.88 [0.87 to 0.89]; mean bias [95% LoA] = −0.4 ms [−63.0 to 63.0 ms]). IMU contact time (mean bias [95% LoA] = −0.4 ms [−78.3 to 86.0 ms]) and flight time (−0.02 ms [−89.0 to 82.0 ms]) showed similar magnitudes of bias, but different degrees of consistency (contact time *R* value [95% CI] = 0.69 [0.68 to 0.71]; flight time = 0.17 [0.14 to 0.19]). Duty factor derived from the IMU also showed poorer consistency (*R* value [95% CI] = 0.10 [0.08 to 0.13]), and a mean bias of −0.58% (95% LoA = −10.4 to 15.0%).

Absolute and relative speed significantly affected ($p < 0.05$) the magnitude of mean biases for contact time, flight time and duty factor (Suplementary Table S2). At group level, biases for these temporal parameters were minimal at absolute and relative speeds of 4.5 $m{\cdot}s^{-1}$ and 53%, respectively (Figure 3).

Test–Retest Reliability (Phase Two)

Across all trials, the mean speed of the constant speed phases was 5.76 ± 0.29 $m{\cdot}s^{-1}$ (range = 3.93 to 6.56 $m{\cdot}s^{-1}$). Table 2 shows test–retest reliability results when calculated with and without the inclusion of speed across 25 steps. Across both methods, step length and step frequency produced a CV < 3%. However, the inclusion of speed as a fixed effect within the model approximately halved their respective CVs (step length = 2.98 to 1.48%; step frequency = 2.81 to 1.52%). Flight time (13.62 to 13.72%) consistently showed a poorer test–retest reliability than contact time (6.28 to 6.84%). For RLI metrics, RLI_{Total} (CV = 4.88 to 5.48%) and $RLI_{Vertical}$ (CV = 7.06 to 7.80%) were the most reliable. Apart from $RLI_{Mediolateral}$, vertical stiffness was the least reliable measure (CV = 14.68 to 15.57%).

Figure 4 provides a visual representation of the change in test–retest reliability of IMU derived metrics when the maximum number of steps ranging from 3 to 25 was used during their calculation. Lower variation in CV was observed post-10 steps (σ = 0.08 to 0.40%) than pre-10 steps (σ = 0.29 to 0.68%) for all metrics.

Discussion

The aims of this study were to quantify the concurrent validity of temporal gait parameters from a thoracic-placed IMU across running speeds and describe the test re-test reliability of further calculated spatiotemporal gait parameters in elite male football players. The findings showed that differences

Table 1. Descriptive statistics for all variables calculated from the algorithm and OptoJump at each target speed during phase one. Values are presented as mean ± SD. IMU = inertial measurement unit.

	IMU				OptoJump			
Variable	All Speeds	3 m · s^{-1}	5 m · s^{-1}	7.5 m · s^{-1}	All Speeds	3 m · s^{-1}	5 m · s^{-1}	7.5 m · s^{-1}
Actual Speed (m · s^{-1})	/	/	/	/	4.70 ± 1.70	3.07 ± 0.26	5.19 ± 0.42	7.13 ± 0.52
Step Time (ms)	334 ± 56	374 ± 36	330 ± 37	267 ± 35	334 ± 49	374 ± 25	330 ± 25	267 ± 21
Contact Time (ms)	216 ± 44	244 ± 38	203 ± 33	179 ± 30	220 ± 58	275 ± 28	197 ± 22	147 ± 16
Flight Time (ms)	118 ± 42	130 ± 37	126 ± 39	88 ± 40	114 ± 28	99 ± 28	133 ± 23	119 ± 18
Duty Factor (%)	32.4 ± 4.9	32.6 ± 4.4	31.0 ± 4.9	33.8 ± 5.5	32.4 ± 5.1	36.8 ± 3.4	29.8 ± 2.8	27.6 ± 2.5

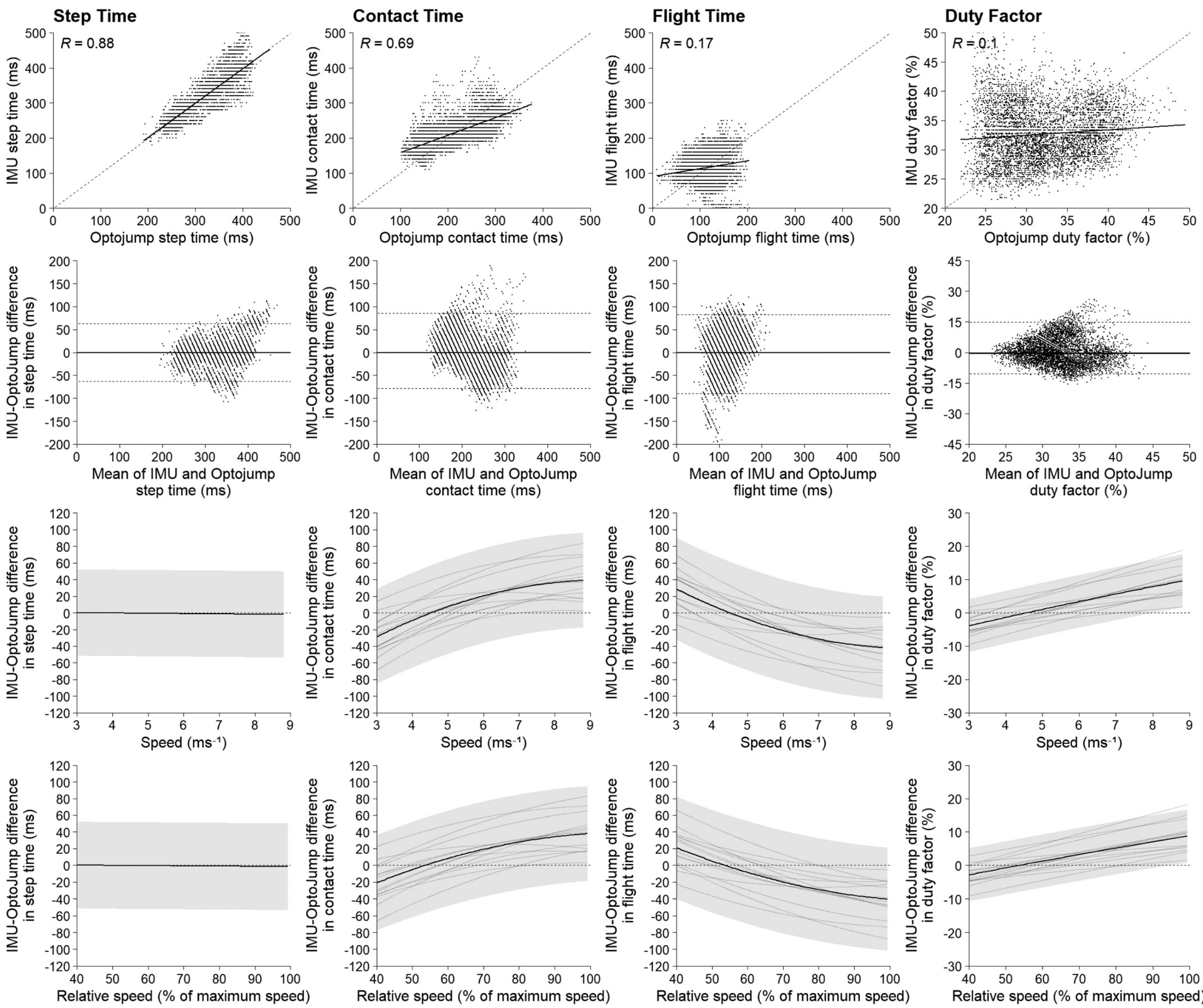


Figure 3. Absolute agreement between step characteristics (step time, contact time, flight time, duty factor) derived from the IMU and OptoJump (row one). Mean bias and 95% limits of agreement between IMU and Optojump derived step characteristics (row 2). Changes in the differences between the IMU and OptoJump derived step characteristic variables across absolute (row three) and relative speeds (row four). Graphs are presented at the group level (bold line) and at the individual level (faint lines).

Table 2. The coefficient of variation and smallest worthwhile changes of variables derived from the IMU when using different calculation methods in elite male football player during phase 2. SD = standard deviation. σ_e = SD of residual variance. CV = coefficient of variation. SWC = smallest worthwhile change. mR^2 = marginal R^2. cR^2 = conditional R^2. RLI = Running Load Index.

		Mixed Model without Speed					Mixed Model with Speed					Power Achieved	
Metric	Mean ± SD (Min-Max)	σ_e	CV (%)	SWC (%)	mR^2	cR^2	σ_e	CV (%)	SWC (%)	mR^2	cR^2	Speed	Speed^2
Step Length (m)	1.80 ± 0.9 (1.38 – 2.07)	0.05	2.98	0.78	0.000	0.633	0.03	1.48	0.74	0.310	0.905	1.00	1.00
Step Time (ms)	312 ± 14 (268 – 357)	8.71	2.79	0.73	0.000	0.638	4.81	1.54	0.74	0.266	0.893	0.48	0.98
Step Frequency (Hz)	3.21 ± 0.15 (2.80 – 3.73)	0.09	2.81	0.73	0.000	0.632	0.05	1.52	0.74	0.263	0.896	0.87	1.00
Contact Time (ms)	211 ± 21 (166 – 283)	14.4	6.84	1.38	0.000	0.512	13.3	6.28	1.36	0.082	0.584	0.31	0.58
Flight Time (ms)	101 ± 18 (50 – 147)	13.6	13.72	2.27	0.000	0.416	13.6	13.62	2.29	0.009	0.429	0.06	0.05
Duty Factor (%)	33.6 ± 2.7 (27.3 – 41.6)	2.11	6.25	1.04	0.000	0.421	2.10	6.24	1.04	0.003	0.422	0.17	0.19
Vertical Stiffness (kN·m^{-1})	12.5 ± 2.6 (6.0 – 21.8)	1.87	15.57	3.09	0.000	0.504	1.75	14.58	3.05	0.071	0.562	0.17	0.34
Maximum Force (kN)	1.92 ± 0.19 (1.35 – 2.46)	0.12	6.21	1.68	0.000	0.652	0.12	6.21	1.68	0.000	0.652	0.06	0.06
Change in Vertical Displacement (mm)	158 ± 24 (111 – 247)	16.1	10.23	2.23	0.000	0.550	14.0	8.84	2.21	0.122	0.662	0.37	0.65
RLI_{Total} (mG·m^{-1}·s^2)	432 ± 50 (311 – 561)	23.2	5.48	2.23	0.000	0.807	20.7	4.88	2.19	0.045	0.842	0.91	0.74
$RLI_{Anteroposterior}$ (mG·m^{-1}·s^2)	194 ± 34 (104 – 296)	18.7	9.88	3.01	0.000	0.701	18.3	9.64	3.02	0.015	0.717	0.23	0.14
$RLI_{Mediolateral}$ (mG·m^{-1}·s^2)	127 ± 38 (62 – 275)	18.8	16.06	5.81	0.000	0.768	18.8	16.06	5.81	0.000	0.768	0.11	0.11
$RLI_{Vertical}$ (mG·m^{-1}·s^2)	327 ± 47 (218 – 470)	25.1	7.80	2.51	0.000	0.723	22.7	7.06	2.47	0.057	0.769	0.88	0.71
Speed (m·s^{-1})	5.76 ± 0.29 (3.93 – 6.56)	0.28	4.87	0.23	0.000	0.096	/	/	/	/	/	/	/

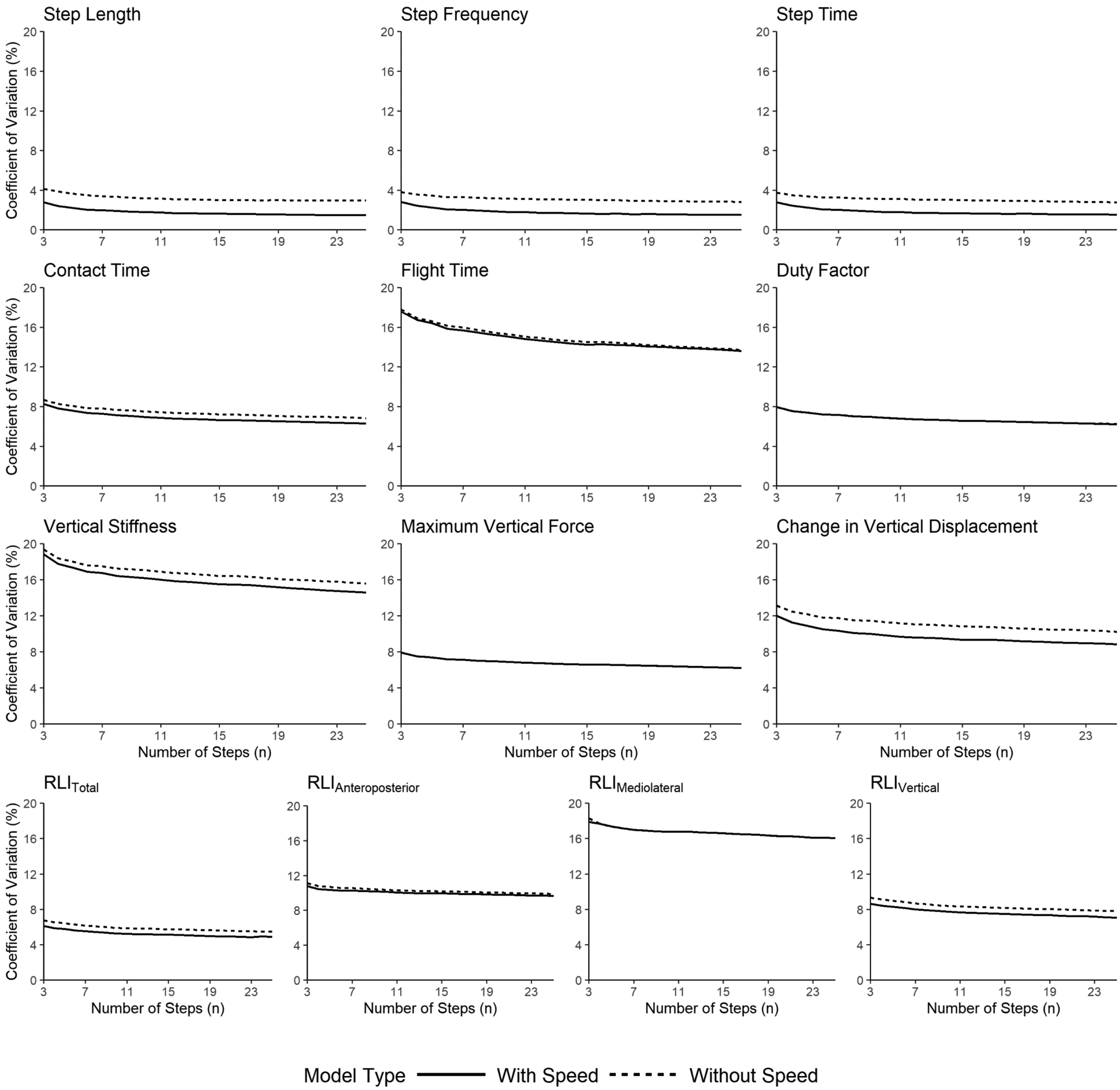


Figure 4. Coefficient of variation across the number steps involved within its calculation. CV = coefficient of variation. RLI = Running Load Index.

between step time from a thoracic IMU and OptoJump were negligible regardless of running speed or participant. Secondly, magnitudes of biases for contact time, flight time and duty factor can differ across participants and their running speeds. Thirdly, all metrics derived from step time (i.e., step length and frequency) showed excellent test–retest reliability in an elite sporting context (CV < 3%).

Phase one – concurrent validity

In the current study, step time calculated from a thoracic-worn IMU showed excellent agreement ($R = 0.88$) with OptoJump, and negligible mean bias across all speeds (<3 ms) regardless of the individual. This agrees with previous research where accelerometer-derived (foot and sacrum) step time showed good agreement ($R = 0.75$–0.95 and mean biases <3 ms) with criterion measures of motion capture and an instrumented treadmill (Falbriard et al., 2018; Lee et al., 2010). As the step time is typically calculated as the time from touchdown to touchdown, the rhythmic nature of running gait allows its accurate depiction if the conditions for touchdown identification remain consistent. Step time derived from an IMU placed between the scapulae can therefore be used as a viable measure in practice. Consequently, an SRT may provide practitioners a more direct insight into player's neuromuscular status (e.g., post-match fatigue, hamstring injury risk). Such variables, when compared with indirect accelerometer-based measures (i.e., RLI) would enhance practitioners' decision-making. Contact time derived from the IMU also showed a relatively high level of absolute agreement with OptoJump ($R = 0.69$). This agrees with previous observations assessing the agreement of the contact time derived from a foot-mounted accelerometer against OptoJump ($R = 0.66$) (García-Pinillos et al., 2021). Additionally, previous studies have typically observed similar underestimations in contact time of 11.5 and 48.7 ms at the foot and lumbar spine, respectively, at running speeds between 3.3 and 4.3 m·s^{-1} (Horsley et al., 2021). However, a novel finding from this study is the observation of overestimated contact times at previously unobserved higher speeds (>6 m·s^{-1}), with biases as high as 40 ms. Furthermore, the flight time showed similar magnitudes of bias across speed, but these biases acted in the opposite direction to the contact time (Horsley et al., 2021). Although, as there is a lack of literature investigating the validity and reliability of temporal gait parameters derived from OptoJump at running speeds above 6 m·s^{-1}, it is plausible that the observed biases are by means of OptoJump rather than the IMU. Despite this, as SRTs in practice are regularly performed at speeds slower than 6 m·s^{-1}, (5 to 5.5 m·s^{-1}) the observed biases of the IMU are applicable when implementing an SRT.

However, although lower biases at speeds regularly used during SRTs were observed compared to higher speeds (>7.5 m·s^{-1}), the presence of biases may still cause concern if looking to use an IMU during the SRT, especially when considering that the magnitude of bias differs at the individual level (Figure 3). As the contact and flight time biases are equal and opposite, practitioners could overcome this limitation by using a correction factor based on running speed. However, it is recommended that this is performed at the individual level because of between-subject variability in the running speed–bias relationship. To do this, practitioners could look to perform their own concurrent validity analysis as part of a pre-season testing battery using the methods described in the current study, or potentially with 1 m of OptoJump on a treadmill.

Phase two - test-retest reliability

The results of this study show that during an SRT, regardless of the number of steps at constant speed, step time, step length and step frequency show the greatest test–retest reliability. The reliability statistics for these, along with contact time and flight time, agreed with previous laboratory-based studies (Horsley et al., 2021), whilst RLI measures agreed with those observed in another field-based study (Leduc et al., 2020).

The CV of all measures tended to converge at approximately 10 steps (Figure 4). As the average speed of the run is above the typical threshold of high-speed running (5.5 m·s^{-1}), it is likely that practitioners would want to moderate this type of exposure solely for testing purposes. Therefore, if attempting to control for a similar speed as the current study and obtain 10 steps at constant speed, practitioners are advised to use a minimum of 40 m for the SRT to be completed in 8 s, with 20 m allowed for the constant speed phase, and 10 m for the acceleration and deceleration phases.

Despite aiming to minimise the between-trial variability in running speed during Phase Two, variation was still apparent (SD = 0.29 m·s^{-1}). This variation caused significant changes to temporal gait characteristics during the constant speed phase. For example, at the group level, an increase in running speed from 5 to 6 m·s^{-1}, led to a 0.19 m increase in step length (from 1.65 to 1.84 m). It is unknown whether this variation in speed was due only to measurement error or if neuromuscular fatigue also led to decreases in speed, as previously observed (Garrett et al., 2021). The dose–response effect between step length and speed, however, is not unexpected, as increased running speed is attributed to concurrent increases in step length and frequency, and decreased contact time (Novacheck, 1998). The effect of speed was greatest for step length, step frequency and contact time; however, the effect of speed on flight time was negligible, which contrasts with previous findings (Novacheck, 1998). This discrepancy is potentially due to a combination of the relatively high measurement error of flight time caused by the low sampling frequency of the accelerometer (100 Hz), and the low variation in speed across which flight time was observed. Another possibility is the low statistical power (Table 2) for the association between flight time and speed (Power = 0.06 & 0.05). This low power is due to the use of a smaller sample size than necessary to detect a potential significant association. This potential inaccuracy may have led to an inflated residual error value, and thus, reliability statistic. However, the magnitude of this for flight time and other metrics, also with poor statistical power (Table 2), is unknown, and requires investigation with a greater number of participants.

From a practical standpoint, if practitioners want to use such an approach, they need to consider any variations in speed because of its potential effect on spatiotemporal variables. To achieve such an optimal interpretation of changes over time, practitioners could consider a direct comparison between a "predicted" score calculated through statistical modelling against the observed score of the day. For an example of this approach, please see Lacome et al. (2018). Meaningful changes can then be interpreted by comparing the change score to the magnitude of the standard error (σ_e) as calculated in this study. Otherwise, practitioners will need to be careful when interpreting any observed changes as it could be the result of a change in speed. Therefore, they will have to monitor speed alongside those variables concurrently.

The reliability statistics calculated through the method used in this study is dependent upon appropriate confounding factors (e.g., speed) being identified and included within the analyses. It is likely that the exclusion of confounding factors (e.g., sleep, nutrition, surface stiffness) can inflate the calculated reliability statistic. As such factors were not monitored in the current study, it is unknown as to what extent the true measurement error was inflated. Therefore, exploring the magnitude of effects for these confounders alongside those already identified in the current study may be an area for future investigation. Another limitation of the current study is the sampling frequency of the IMU used to calculate spatiotemporal gait parameters. IMU in the current study, along with most other thoracic-placed IMUs available in practice, sample at 100 Hz. This means that temporal gait parameters can only be calculated to the closest 10 ms. As temporal gait parameters in the current study tended to be below 400 ms, it would be beneficial to use an IMU that samples at 1000 Hz to obtain optimal precision and sensitivity to changes. As technology develops (e.g., storage capacity), this could be feasible if IMUs placed within similar microtechnology have the capacity to sample at higher rates.

Conclusion

Step time and its derivatives, step length and step frequency showed agreement with a criterion and proved to be reliable measures when extracted from an IMU placed on the upper back whilst running at a constant speed. Contact time and flight time could also be considered by practitioners; however, because of potential inter-individual variation in the mean bias of these metrics, it is recommended that practitioners carry out their own analysis to allow for individualised corrections to the algorithm. It is also recommended that to enhance the reliability of their results, practitioners collect a minimum of 10 steps during the constant speed phase for analysis. These results provide an avenue for practitioners to monitor athlete gait in an unobtrusive way. However, currently, attempts to observe alterations to gait post-fatiguing exercise are limited. Future work should look to address whether longitudinal changes to spatiotemporal gait parameters in the days post-exercise are valid measures of altered neuromuscular status.

Acknowledgments

The authors would like to thank Joe Jackson and Johan Cassirame for providing additional OptoJump Next equipment, and the participants for their voluntary participation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Crystal Palace Football Club.

Data availability statement

Data supporting the findings from the first phase of the study are available from the corresponding author, EW, upon reasonable request. Due to legal constraints, supporting data is not available for phase two of this study.

References

Apte, S., Prigent, G., Stöggl, T., Martínez, A., Snyder, C., Gremeaux-Bader, V., & Aminian, K. (2021). Biomechanical response of the lower extremity to running-induced acute fatigue: A systematic review. *Frontiers in Physiology*, *12*. https://doi.org/10.3389/fphys.2021.646042

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Buchheit, M., Gray, A., & Morin, J.-B. (2015). Assessing stride variables and vertical stiffness with GPS-Embedded accelerometers: Preliminary insights for the monitoring of neuromuscular fatigue on the field. *Journal of Sports Science & Medicine*, *14*(4), 698–701.

Carling, C., Lacome, M., McCall, A., Dupont, G., Gall, F., Simpson, B., & Buchheit, M. (2018). Monitoring of post-match fatigue in professional soccer: Welcome to the real world. *SPORTS MEDICINE -AUCKLAND-*, *48*(12), 2695–2702. https://doi.org/10.1007/s40279-018-0935-z

Crang, Z. L., Duthie, G., Cole, M. H., Weakley, J., Hewitt, A., & Johnston, R. D. (2021). The validity and reliability of wearable microtechnology for intermittent team sports: A systematic review. *Sports Medicine*, *51*(3), 549–565. https://doi.org/10.1007/s40279-020-01399-1

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods Instruments & Computers*, *28*(1), 1–11. https://doi.org/10.3758/bf03203630

Falbriard, M., Mayer, F., Mariani, B., & Millet, G. P., & Aminian, K. (2018). Accurate estimation of running temporal parameters using foot-worn inertial sensors. *Frontiers in Physiology*, *9*(610). https://doi.org/10.3389/fphys.2018.00610

Fitzpatrick, J. F., Hicks, K. M., Russell, M., & Hayes, P. R. (2021). The reliability of potential fatigue-monitoring measures in elite youth soccer players. *The Journal of Strength & Conditioning Research*, *35*(12), 3448–3452. https://doi.org/10.1519/jsc.0000000000003317

García-Pinillos, F., Roche-Seruendo, L. E., Marcén-Cinca, N., Marco-Contreras, L. A., & Latorre-Román, P. A. (2021). Absolute reliability and concurrent validity of the stryd system for the assessment of running stride kinematics at different velocities. *The Journal of Strength & Conditioning Research*, *35*(1), 78–84. https://doi.org/10.1519/JSC.0000000000002595

Garrett, J., Akyildiz, Z., Leduc, C., van den Hoek, D., Manuel Clemente, F., & Ardigò, L. P. (2021). Peak running speed can be used to monitor neuromuscular fatigue from a standardized running test in team sport athletes. *Research in Sports Medicine (Print)*, *31*(4), 1–12. https://doi.org/10.1080/15438627.2021.1966012

Garrett, J., Graham, S. R., Eston, R. G., Burgess, D. J., Garrett, L. J., Jakeman, J., & Norton, K. (2019). A novel method of assessment for monitoring neuromuscular fatigue in Australian rules football players. *International Journal of Sports Physiology & Performance*, *14*(5), 598–605. https://doi.org/10.1123/ijspp.2018-0253

Girard, O., Nybo, L., Mohr, M., & Racinais, S. (2015). Plantar flexor neuromuscular adjustments following match-play football in hot and cool conditions. *Scandinavian Journal of Medicine & Science in Sports*, *25*(*Suppl* 1), 154–163. https://doi.org/10.1111/sms.12371

Gomez Bernal, A. B. S. M. S., Becerro de Bengoa-Vallejo, R. P. D. P. M., & Losa-Iglesias, M. E. P. (2016). Reliability of the OptoGait portable photoelectric cell system for the quantification of spatial-temporal parameters of gait in young adults. *Gait & posture*, *50*, 196–200. https://doi.org/10.1016/j.gaitpost.2016.08.035

Hanley, B., & Tucker, C. B. (2019). Reliability of the OptoJump next system for measuring temporal values in elite racewalking. *The Journal of Strength & Conditioning Research*, *33*(12), 3438–3443. https://doi.org/10.1519/JSC.0000000000003008

Hecksteden, A., Kraushaar, J., Scharhag-Rosenberger, F., Meyer, T., Theisen, D., & Senn, S. (2015). Individual response to exercise training - a statistical perspective. *Journal of Applied Physiology*, *118*(12), 1450–1459. https://doi.org/10.1152/japplphysiol.00714.2014

Horsley, B. J., Tofari, P. J., Halson, S. L., Kemp, J. G., Chalkley, D., Cole, M. H., Johnston, R., & Cormack, S. J. (2023). Validity and reliability of thoracic-mounted inertial measurement units to derive gait characteristics during running. *The Journal of Strength & Conditioning Research*, *38*(2), 274–282. https://doi.org/10.1519/JSC.0000000000004612

Horsley, B. J., Tofari, P. J., Halson, S. L., Kemp, J. G., Dickson, J., Maniar, N., & Cormack, S. J. (2021). Does site matter? Impact of inertial measurement unit placement on the validity and reliability of stride variables during running: A systematic review and meta-analysis. *Sports Medicine (Auckland*, *51*(7), 1449–1489. https://doi.org/10.1007/s40279-021-01443-8

Jeffries, A. C., Wallace, L., Coutts, A. J., McLaren, S. J., McCall, A., & Impellizzeri, F. M. (2020). Athlete-reported outcome measures for monitoring training responses: A systematic review of risk of bias and measurement property quality according to the COSMIN guidelines. *International Journal of Sports Physiology & Performance*, *15*(9), 1203–1215. https://doi.org/10.1123/ijspp.2020-0386

Kumle, L., Võ, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, *53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Lacome, M., Simpson, B., Broad, N., & Buchheit, M. (2018). Monitoring players' readiness using predicted heart-rate responses to soccer drills. *International Journal of Sports Physiology & Performance*, *13*(10), 1273–1280. https://doi.org/10.1123/ijspp.2018-0026

Leduc, C, & Buchheit M. (2020). The use of standardised runs (and associated data analysis) to monitor neuromuscular status in team sports players: A call to action. *Sport Performance & Science Reports*, *1*(117), 1–5.

Leduc, C., Tee, J., Lacome, M., Weakley, J., Cheradame, J., Ramirez, C., & Jones, B. (2020). Convergent validity, reliability, and sensitivity of a running test to monitor neuromuscular fatigue. *International Journal of Sports Physiology & Performance*, *15*(8), 1067–1073. https://doi.org/10.1123/ijspp.2019-0319

Lee, J. B., Mellifont, R. B., & Burkett, B. J. (2010). The use of a single inertial sensor to identify stride, step, and stance durations of running gait. *Journal of Science & Medicine in Sport*, *13*(2), 270–273. https://doi.org/10.1016/j.jsams.2009.01.005

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the Black Box: Applications and considerations for using GPS devices in sport. *International Journal of Sports Physiology & Performance*, *12*(s2), 2–18. https://doi.org/10.1123/ijspp.2016-0236

Morin, J. B., Dalleau, G., Kyröläinen, H., Jeannin, T., & Belli, A. (2005). A simple method for measuring stiffness during running. *Journal of Applied Biomechanics*, *21*(2), 167–180. https://doi.org/10.1123/jab.21.2.167

Novacheck, T. F. (1998). The biomechanics of running. *Gait & posture*, *7*(1), 77–95. https://doi.org/10.1016/S0966-6362(97)00038-6

Phinyomark, A., Osis, S., Hettinga, B. A., & Ferber, R. (2015). Kinematic gait patterns in healthy runners: A hierarchical cluster analysis. *Journal of Biomechanics*, *48*(14), 3897–3904. https://doi.org/10.1016/j.jbiomech.2015.09.025

Rampinini, E., Bosio, A., Ferraresi, I., Petruolo, A., Morelli, A., & Sassi, A. (2011). Match-related fatigue in soccer players. *Medicine & Science in Sports and Exercise*, *43*(11), 2161–2170. https://doi.org/10.1249/MSS.0b013e31821e9c5c

Riazati, S., Caplan, N., Matabuena, M., & Hayes, P. R. (2022). Gait and neuromuscular changes are evident in some masters club level runners 24-h after interval training run. *Frontiers in Sports and Active Living*, *4*, 830278. https://doi.org/10.3389/fspor.2022.830278

Schober, P., Boer, C., & Schwarte, L. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Small, K., McNaughton, L. R., Lovell, R., Greig, M., & Lohkamp, M. (2009). Soccer fatigue, sprinting and hamstring injury risk. *International Journal of Sports Medicine*, *30*(8), 573–578. https://doi.org/10.1055/s-0029-1202822

Voeten, C. (2023). Stepwise elimination and term reordering for mixed-effects regression. Version 2.11. https://cran.r-project.org/web/packages/buildmer