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

Horsley, Benjamin J., Tofari, Paul J., Halson, Shona L., Kemp, Justin G., Dickson, Jessica, Maniar, Nirav and Cormack, Stuart J.

Title: 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

Running heading: Impact of IMU Placement on Stride Variable Validity and Reliability

Authors: Benjamin J. Horsley[1], Paul J. Tofari[1], Shona L. Halson[1,2], Justin G. Kemp[1], Jessica Dickson[3], Nirav Maniar[1] and Stuart J. Cormack[1,2]

Affiliations: [1]School of Behavioural and Health Sciences, Australian Catholic University, Melbourne, Australia; [2]Sports Performance, Recovery, Injury and New Technologies (SPRINT) Research Centre, Australian Catholic University, Melbourne, Australia; [3]Library and Academic Research Services, Australian Catholic University, Melbourne, Australia

Corresponding author:

Benjamin J. Horsley

School of Behavioural and Health Sciences

Australian Catholic University, 115 Victoria Parade, Fitzroy, 3065, Australia

Email: benjamin.horsley@myacu.edu.au

Abstract word count: 441

Manuscript word count: 7180

Number of tables: 13

Number of figures: 5

Abstract

Background: Inertial measurement units (IMUs) are used for running gait analysis in a variety of sports. These sensors have been attached at various locations to capture stride data. However, it is unclear if different placement sites affect the derived outcome measures.

Objective: The aim of this systematic review and meta-analysis was to investigate the impact of placement on the validity and reliability of IMU-derived measures of running gait.

Methods: Online databases SPORTDiscus with Full Text, CINAHL Complete, MEDLINE (Ebscohost), EMBASE (Ovid) and Scopus were searched from the earliest record to 6 August 2020. Articles were included if they 1) used an IMU during running 2) reported spatiotemporal variables, peak ground reaction force (GRF) or vertical stiffness and 3) assessed validity or reliability. Meta-analyses were performed for a pooled validity estimate when 1) studies reported means and standard deviation for variables derived from the IMU and criterion 2) used the same IMU placement and 3) determined validity at a comparable running velocity ($\leq$1 $m \cdot s^{-1}$ difference).

Results: Thirty-nine articles were included, where placement varied between the foot, tibia, hip, sacrum, lumbar spine (LS), torso and thoracic spine (TS). Initial contact, toe-off, contact time (CT), flight time (FT), step time, stride time, swing time, step frequency (SF), step length (SL), stride length, peak vertical and resultant ground reaction force (GRF) and vertical stiffness were analysed. Four variables (CT, FT, SF and SL) were meta-analysed, where CT was compared between foot, tibia and LS placements and SF was compared between foot and LS. Foot placement data was meta-analysed for FT and SL. All data are mean difference (MD [95%CI]). No significant difference was observed for any site compared to the criterion for CT (foot: -11.47 ms [-45.68, 22.74], p=0.43; tibia: 22.34 ms [-18.59, 63.27], p=0.18; LS: -48.74 ms [-120.33, 22.85], p=0.12), FT (foot: 11.93 ms [-8.88, 32.74], p=0.13), SF (foot: 0.45 $step \cdot min^{-1}$ [-1.75, 2.66], p=0.47; LS: -3.45 $step \cdot min^{-1}$ [-16.28, 9.39], p=0.37) and SL (foot: 0.21 cm [-1.76, 2.18], p=0.69). Reliable derivations of CT (coefficient of variation [CV] <9.9%), FT (CV <11.6%) and SF (CV <4.4%) were shown using foot- and LS-worn IMUs, while the CV was <7.8% for foot-determined stride time, SL and stride length. Vertical GRF was reliable from the LS

(CV=4.2%) and TS (CV=3.3%) using a spring-mass model, while vertical stiffness was moderately (r=0.66) and nearly perfectly (r=0.98) correlated with criterion measures from the TS.

Conclusion: Placement of IMUs on the foot, tibia and LS are suitable to derive valid and reliable stride data, suggesting measurement site may not be a critical factor. However, evidence regarding the ability to accurately detect stride events from the TS is unclear and this warrants further investigation.

Key points

- Practitioners may attach inertial measurement units to the foot, tibia and lumbar spine to accurately and reliably derive stride variables during running.
- The computational method for gait event detection may be more critical to validity and reliability than the attachment location itself.
- These findings may open opportunities for practitioners to use inertial measurement units to analyse the gait patterns of athletes in a variety of running-based sports.

Declarations

Funding

This review received no funding.

Conflicts of interest

Benjamin Horsley, Paul Tofari, Shona Halson, Justin Kemp, Jessica Dickson, Nirav Maniar and Stuart Cormack declare that they have no conflicts of interest relevant to the content of this review.

Ethics approval

Not applicable.

Availability of data and material

The dataset and code used for meta-analysis are available from the corresponding author on request.

Authors' contributions

BJH, PJT, JD and SJC contributed to the development of the review and implementation of the search strategy. BJH carried out the meta-analysis with assistance from NM. BJH, PJT, NM and SJC collectively interpreted the results of the systematic review and meta-analysis, while BJH drafted the manuscript. All authors contributed to editing and revising the manuscript and approved the final version prior to submission.

Acknowledgments

We thank all those authors of reviewed papers who provided additional data that made some of the analyses reported in this manuscript possible.

1 Introduction

It is common practice to quantify the activities performed by athletes, or external load, to plan and monitor training and competition load [1]. Tracking technology, such as video-based systems, global positioning systems (GPS) and local positioning systems (LPS), measure athlete displacement and calculate velocity and acceleration [2-5]. However, due to its low sampling frequency (e.g. 10 Hz), GPS is limited in its ability to accurately capture changes in velocity or high-speed movements over short distances and when movements are nonlinear, such as changes of direction [6-8]. Although LPS (1000 Hz) sample at a higher rate than GPS, neither technology can account for non-locomotor activity, such as impacts or collisions [9-12]. To overcome some of the limitations of GPS and LPS, inertial measurement units (IMUs), comprising accelerometers, gyroscopes and magnetometers, can provide additional information on athlete activity profiles [13, 14].

Triaxial accelerometers measure acceleration in the anteroposterior, mediolateral and vertical axes and typically capture data between 100 and 1000 Hz [13, 15, 16]. Gyroscopes and magnetometers measure device orientation and direction, respectively [17]. Accelerometers have been used for quantifying daily physical activity and estimating energy expenditure [18-22] and their use is now common in athletes [11, 12, 23-25]. Accelerometer-derived metrics, such as PlayerLoad™, provide an indication of global external load from the summation of instantaneous rate of change of acceleration in the anteroposterior, mediolateral and vertical axes [15, 26]. However, PlayerLoad™ is a relatively gross measure that does not offer insight into discrete movements, such as stride variables. Instead, patterns in the signals of IMUs can be explored to identify foot contacts to calculate different stride variables, which may help in understanding the way in which athletes produce a given load [27, 28].

The detection of gait events, such as initial contact (IC) and toe-off (TO), is possible using accelerometer and gyroscope data [29-31]. Identifying these key events allows for the calculation of spatiotemporal parameters, including contact time, flight time, step and stride times, step frequency and step and stride lengths [30, 32-35]. The acceleration signal from IMUs may also be used to estimate ground reaction forces (GRFs) and vertical stiffness to describe the impact forces experienced by athletes and their ability to absorb force during running [36-40]. Deriving stride variables is important for evaluating an athlete's gait pattern and may help to inform injury mitigation and performance enhancement strategies [41]. However, device placement may influence the derived outcome measures and should be considered when using IMUs to capture stride data [42, 43].

Placement of IMUs for analysis of running gait can vary between the foot [30, 34, 44], distal and mid tibia [13, 31, 40], lumbosacral region [30, 32, 33] or thoracic spine [37, 39, 43]. Given accelerometers measure acceleration of the segment to which it is attached, there are some potential issues associated with placement on the upper body to measure accelerations occurring at the lower limb and derive valid and reliable stride data [42]. Attachment location is an important consideration due to signal attenuation, whereby acceleration magnitudes dissipate from the foot to the torso during ground contact in running [45-47]. Although securing IMUs to the foot may provide the most accurate derivations of stride variables [16, 34, 35], this site may not be practical in some sports (such as those that involve kicking), while other work has noted the potential for injury in contact sports using IMUs attached to the tibia [44]. Given IMUs have been utilised at various sites for the analysis of running gait in the literature, it is important to understand if IMU placement affects the derived outcome measures. This may help inform practitioners which attachment location is most appropriate for deriving valid and reliable stride data based on the constraints of the sport they work in. Therefore, the aim of this systematic review and meta-analysis is to report on the validity and reliability of inertial sensors to calculate spatiotemporal variables, GRF and vertical stiffness during running with respect to sensor placement.

2 Methods

2.1 Systematic Review Protocol

The protocol for this systematic review was registered on PROSPERO and can be accessed at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020160325. All procedures were performed in accordance with the PRISMA guidelines [48].

2.2 Eligibility Criteria

Articles were eligible for inclusion in this systematic review if they 1) were published in English 2) used an accelerometer, gyroscope or a combination of both technologies 3) had participants jog, run or sprint during data collection 4) reported at least one of the following outcome variables: IC, TO, contact time, flight time, step time, stride time, swing time, step frequency, step length, stride length, peak vertical or resultant GRF or vertical stiffness and 5) assessed validity or reliability.

2.3 Search Strategy

Keywords in the title and abstract of records, combined with relevant subject heading terms, such as Medical Subject Headings (MeSH), were systematically searched in SPORTDiscus with Full Text, CINAHL Complete, MEDLINE (Ebscohost), EMBASE (Ovid) and Scopus from the earliest record up until 6 August 2020. The following keyword search string was used in each electronic database (which is also detailed in Supplementary Information Appendix S1, Table S1):

(jog* OR run* OR sprint*) AND (acceleromet* OR "global positioning system" OR GPS OR gyroscope* OR IMU OR inertial* OR microtechnolog* OR “wearable sensor”) AND (acceleration* OR event* OR fatigue* OR force* OR GRF OR kinematic* OR kinetic* OR parameter* OR reliab* OR stance OR step* OR stiff* OR stride* OR strike* OR temporal OR valid*)

2.4 Study Selection

Search results were exported to reference management software EndNote X9.3.3 (Clarivate Analytics, Philadelphia, USA) where duplicates were removed. Two authors (BJH and PJT) then independently screened the title and abstract of each record in the Rayyan web-based systematic review tool (available at www.rayyan.qcri.org). The full text of potentially eligible articles was retrieved and one author (BJH) performed a final eligibility assessment, which was later checked by a second author (PJT). Discrepancies in article selection were resolved by a third author (SJC). The reference lists of all retrieved articles were also examined to determine any other articles that may be relevant to the review.

2.5 Data Extraction

Data relating to participant characteristics (age, body mass, height and activity level), sensor specifications (brand, model, range and sampling frequency), sensor location (foot, distal/mid tibia, hip, sacrum, lumbar and thoracic spine), criterion used for validity (brand, model and sampling frequency), running activity performed (number, duration or distance of runs, velocity), outcome variables analysed (temporal, spatial, GRF and vertical stiffness) and measures of validity and reliability were extracted from each included study. Definitions for the variables analysed in this review are presented in Table 1. Running velocity, temporal and spatial variables and GRF are reported in metres per second ($m{\cdot}s^{-1}$), milliseconds (ms), centimetres (cm) and Newtons (N), respectively. Where included studies did not report results in the aforementioned units, values were converted to enable better comparison between studies.

2.6 Assessment of Methodological Quality

The methodological quality of each included study was assessed using a modified assessment scale of Downs and Black [49]. Of the 27 criteria, the most relevant to the study designs included in this review were applied, which is consistent with other reviews [50, 51]. Each study was therefore assessed for quality of reporting (1-4, 6, 7 and 10), external validity (11 and 12) and internal validity bias (16, 18 and 20) based on 12 criteria. The criteria were evaluated as yes, no or unclear, with the score out of 12 determined from the number of items that were answered yes.

2.7 Data Analysis

The values of validity and reliability for each stride variable are presented in the tables below and included throughout the results sections.

2.7.1 Meta-analysis

Meta-analyses were performed when there were at least two studies that 1) reported means and standard deviation (SD) for stride variables calculated from IMUs and reference systems 2) used the same IMU attachment site and 3) assessed validity at a comparable running velocity (≤ 1 m·s^{-1} difference). Authors that did not include absolute mean ± SD values for the computed stride variables were contacted to gain the additional data. Raw outcome data was not obtained for 22 studies and were thereby ineligible for inclusion in any meta-analysis [31, 32, 34, 37, 38, 40, 43, 52-66]. Where there were multiple effects reported for different running velocities from a single study, data was aggregated so only a single effect was included in the meta-analysis [67]. However, when validity was assessed using IMUs from two different manufacturers [35] or criterion measures [29, 68] in a single study, effects were treated independently and both were included in the meta-analysis. Data pertaining to criterion validity was pooled from studies that used different reference measurement systems. Specifically, effects were pooled from studies that used motion capture [29], force plates [30, 69, 70], high-speed camera [16, 35, 68] and photocell systems [68, 71, 72]. This approach was used due to the limited number of studies with comparable methodologies and previous work demonstrating that optical timing and motion capture systems and force plate systems are all considered as criterion methods for gait analysis [73-75].

Where there was sufficient data to group effects based on eligibility criteria, meta-analyses were performed using random-effects models with the Meta statistical package in R software (version 3.6.3, R Foundation for Statistical Computing) to produce a pooled estimate of the mean difference (MD) in absolute units [76]. When studies could be pooled based on different IMU attachment sites for the same variable, subgroup analysis was performed to test whether placement differs in terms of their effects, with the significance level set at $p < 0.05$ [77].

The level of statistical heterogeneity was quantified by calculating the I^2 statistic [78]. Statistical heterogeneity was considered low ($I^2 < 25\%$), moderate ($I^2 = 25$-49%) and high ($I^2 > 50\%$) [78]. When 1^2 was high ($I^2 > 50\%$), leave-one-out analysis was performed to determine the studies that contributed most to heterogeneity and had a high influence on the overall effect [79]. Moderator analysis was also conducted to determine how much the criterion measure contributed to the observed variability of effect sizes between studies [80]. Where the criterion does not have a significant moderating effect, heterogeneity may be attributable to an unidentified source [80]. A meta-regression model was applied to the moderator analysis using the metareg function in R software [81]. Statistical significance was set at $p < 0.05$.

Effect sizes and their respective confidence intervals (CI), along with the overall MD for pooled effects, were visualised as forest plots [82]. In forest plots, studies are represented by a point estimate, bounded by a 95% CI for the effect [82]. The summary effect (MD) is symbolised by the polygon at the bottom of the plot [82]. The width of the polygon indicates the 95% CI. Studies that exhibit larger squares contribute more to the summary effect (MD) compared to studies with smaller squares [82].

Insert Table 1 about here

3 Results

3.1 Study Identification and Selection

A total of 4,654 records were identified through the database searches. An additional three articles were included through reference list searches. Following deduplication, title and abstract screening and a thorough full text screen of each record, 39 studies met the eligibility criteria and were included in the review [16, 29-40, 43, 52-66, 68-72, 83-87]. An outline of this process using the PRISMA flow diagram is presented in Figure 1.

3.2 Study Characteristics

A summary of the characteristics of each study is presented in Table 2. A total of 657 participants were included across 39 studies (mean ± SD 16.8 ± 10.2), where the populations sampled included healthy active adults (*n* = 15 studies), recreational/amateur (*n* = 12) and high-level runners (*n* = 5), team-sport athletes (*n* = 6), elite track and field athletes (*n* = 1) and triathletes (*n* = 1). Sensor placement varied between foot [16, 30, 34, 35, 52, 56, 59, 60, 65, 69, 71, 85], distal and mid tibia [29, 31, 40, 58, 60, 62, 69, 70, 84], hip [66], sacrum [32, 57], lumbar spine [30, 33, 38, 64, 68, 69, 72, 83], torso [53] and thoracic spine [36-39, 43, 54, 61, 86]. Two studies used multiple sensors and a combination of placements to derive stride variables [55, 87]. Validity was assessed using force plate systems (*n* = 17) [31, 33, 36, 38-40, 43, 54, 55, 57-59, 61, 62, 66, 69, 70], optical motion capture (*n* = 7) [29, 32, 52, 55, 64, 65, 85], instrumented treadmill (*n* = 7) [30, 34, 37, 53, 56, 60, 87], high-speed camera (*n* = 4) [16, 33, 35, 68], photocell systems (*n* = 3) [68, 71, 72], foot-mounted accelerometer (*n* = 1) [83], in-shoe piezo-electric force sensitive resistors (FSR) (*n* = 1) [63] and different stride time calculation methods (*n* = 1) [84] as criterions. Reliability was assessed in nine studies [16, 38, 40, 43, 59, 68, 71, 83, 86]. Contact time was the most commonly reported variable (*n* = 16) [16, 29, 30, 32-35, 37, 52, 53, 62, 68-71, 83], while six studies derived spatial data (step length and stride length) from accelerometers and gyroscopes [35, 52, 65, 71, 72, 85]. Eleven studies estimated peak vertical and resultant GRF [36, 38-40, 43, 55-57, 61, 66, 87], whereas three studies used accelerometers to derive vertical stiffness [37, 38, 86].

Insert Table 2 about here

3.3 Methodological Quality

Based on the number of criteria that were answered yes, the methodological quality of included studies ranged from 7 to 10 out of 12, with a mean score of 9 out of 12 (see Supplementary Information Appendix S1, Table S2). Out of the 39 studies, 24 did not include *p*-values alongside validity or reliability outcomes [29-34, 37-40, 53, 55, 56, 58, 59, 61, 63-66, 72, 83, 84, 86], two studies did not clearly report subject characteristics [33, 63], while another study did not provide a description of the running protocol used for assessing validity [60]. Five studies scored a yes for detailing the source population from which subjects were recruited [31, 60, 83, 86, 87], whereas this was unclear in the remaining studies.

3.4 Stride Variables

The results for each stride variable examined in this review are described in the following sections.

3.4.1 Initial contact

Validity outcomes for the detection of IC using IMUs secured to the foot, distal and mid tibia and lumbar spine are presented in Table 3. Mean relative differences (-16.0 to 3.3 ms) and estimation errors (-6.0 to 4.3 ms) were generally low for foot placement [30, 52, 69], while another study reported IC could be detected with a precision (median ± inter-quartile range [IQR]) of 2.0 ± 1.0 ms from a foot-mounted IMU [34]. Contrasting results were evident for placement on the tibia. Using only the angular velocity signal from a gyroscope, errors were as high as 64.2 ms compared to motion capture in one study [29], while another study detected IC from gyroscope data with an absolute mean error of 13.0 ± 6.0 ms to that of a force plate [70]. The mean relative difference (-38.0 ± 10.7 ms) was greater than that observed for the foot using tibial acceleration data [69], while other studies showed improved validity for determining IC from tibia-mounted IMUs compared to force plate measures (MD = -0.5 ± 0.3 ms, mean bias = -2.3 ± 4.7 ms, mean error = 1.68 ms) (see Table 3) [31, 58, 62]. In another study using tibia-mounted accelerometers, IC was detected with an accuracy of F_1 = 0.92-0.96 compared to those events determined from in-shoe piezo-electric FSRs [63]. The F_1 score is a measure of a test's accuracy, where an F_1 score of 1 reflects perfect precision and recall [88]. Detection of IC was slightly earlier (4.7 ms) at 3.3 $m \cdot s^{-1}$ from a lumbar spine-mounted IMU compared to the foot, but 2.4 ms slower at 4.1 $m \cdot s^{-1}$ [69]. The largest difference from force plate-identified IC was 53.0 ms for the lumbar spine [30].

3.4.2 Toe-off

Table 4 documents the validity statistics from studies that determined the accuracy of IMUs to detect TO. Between 2.2 and 4.1 m·s^{-1}, the mean relative difference and estimation errors for the detection of TO from foot-mounted IMUs ranged from -53.8 to 32.0 ms and -4.3 to 16.3 ms, respectively [30, 52, 69]. Errors up to -32.4 ms were shown using a gyroscope attached to the tibia [29], while another study using angular velocity data from the tibia showed TO was determined after force plate detection (absolute mean error > 23.0 ms) [70]. Smaller mean absolute and relative differences were observed for determining TO from tibial acceleration data (< 8.8 ms and < 1.0 ms, respectively) [69], while TO was detected with an accuracy of F_1 = 0.77-0.86 from accelerometers secured to the distal tibia when in-shoe piezo-electric FSRs were the criterion [63]. A time lag of 7.6 to 24.0 ms was present for the detection of TO from an IMU secured to the lumbar spine compared to values obtained from a force plate [30, 69].

Insert Tables 3 and 4 about here

3.4.3 Contact time

Validity outcomes reported from studies using placement on the foot, tibia, lumbar spine, torso and thoracic spine to derive contact time is presented in Table 5. The concurrent validity of an IMU fixed to the foot showed a deviation to high-speed camera measures between -3.3 and -0.1%, a mean bias between -5.6 and 0.4 ms and intraclass correlation coefficient (ICC) values as high as 0.97 for contact time across velocities of 4.3 ± 0.7 m·s^{-1}, 6.2 ± 0.7 m·s^{-1} and 8.0 ± 1.6 m·s^{-1} [16]. When a photocell system was the criterion, ICC values were as low as 0.1 at 5.6 m·s^{-1} using a foot placement [71]. Pearson correlation analysis showed a large agreement (r = 0.96) between a tibial accelerometer estimate of contact time and force plate [62], whereas contrasting results were evident for contact time calculated from gyroscope data (see Table 5) [29, 70]. True error and ICC outcomes were > 63.4 ms and < 0.32, respectively, compared to motion capture [29], whereas differences to force plate were smaller (> -12.0

ms) in another study using angular velocity data to determine contact time [70]. Compared to motion capture and force plate, small biases (0.8-1.1 ms) and estimation errors (5.0 ms) were shown for contact time when an IMU was placed on the sacrum and lumbar spine, respectively [32, 33]. However, significant differences ($p < 0.05$) were reported in another study using the lumbar spine when photocell (> -35.0%) and high-speed camera (> -31.0%) measures of contact time were used as the reference [68]. In a study comparing contact times derived from different accelerometer attachment sites, the lumbar spine showed a smaller difference from force plate-determined contact time (< 8.7%) to the values obtained from the tibia (< 17.3%) and foot (< 26.6%), with each site significantly correlated ($r > 0.74$, $p < 0.05$) with force plate (see Table 5) [69]. Similar results reported in a more recent study showing the mean lumbar spine-force plate difference (-29.0 ms) was less than that observed between foot-force plate (47.0 ms). In that study, accelerometers placed on the lumbar spine underestimated mean contact time compared to force plate, whereas foot acceleration overestimated by 18.0 ms [30]. Contact time derived from an accelerometer secured to the thoracic spine showed a mean bias of -10.4% and a nearly perfect correlation ($r = 0.98$) with an instrumented treadmill [37]. However, data from only one participant was analysed [37].

Three studies assessed the reliability of IMUs on the foot and lumbar spine to calculate contact time (see Table 6). The coefficient of variation (CV) was < 2.3% across velocities ranging between 2.2 and 5.6 $m \cdot s^{-1}$, while the standard error of measurement (SEM) was highest at 2.2 $m \cdot s^{-1}$ (5.0 ms) [71]. Good absolute and relative between-trial reliability was established using an accelerometer mounted on the lumbar spine (CV < 9.9%, ICC > 0.88) [68], while lower ICC values ranging from -0.24 to 0.67 were reported for inter-day reliability in another study using a lumbar spine-mounted accelerometer [83]. Greater SEM values were observed for lumbar spine-determined contact time (> 10.1 ms) compared to foot placement (< 5.0 ms) [71, 83].

Data collected between 3.3 and 4.3 $m \cdot s^{-1}$ was pooled to determine the effect of IMU placement on the accuracy of contact time compared to criterion measures (see Figure 2). There was a significant difference in the overall effect of different IMU attachment sites ($p = 0.02$). Contact time derived from the foot (MD [95% CI] -11.47 ms [-45.68, 22.74], $p = 0.43$), tibia (MD [95% CI] 22.34 ms [-18.59, 63.27], p = 0.18) and lumbar spine (MD [95% CI] -48.74 ms [-120.33, 22.85], p = 0.12) was not significantly different to the criterion. All subgroups were associated with high heterogeneity (I^2 > 54.1%). Leave-one-out analysis (see Supplementary Information Appendix S2) for foot and lumbar spine sites revealed that there was no single study influential enough to substantially change the overall heterogeneity (I^2 > 83.4%) or pooled MD. In contrast, heterogeneity could be explained for the tibia site by omitting one study [70] (I^2 = 0%), with the same study also having an influential effect on the overall result for tibia-determined contact time (MD [95% CI] 34.68 ms [11.16, 58.19], p = 0.02). Moderator analysis showed the type of criterion measure was not significantly associated with the observed variance in effect sizes (p = 0.15).

Insert Tables 5 and 6 about here

Insert Figure 2 about here

3.4.4 Flight time

Results from studies reporting the criterion validity of IMU-derived flight time are documented in Table 7. For placement at the foot, ICC values were as high as 0.81 at 5.6 $m \cdot s^{-1}$ and 0.86 at 3.3 $m \cdot s^{-1}$ compared to photocell and high-speed camera measures of flight time [35, 71]. Low estimation errors (< 8.2 ms) and median ± IQR bias (15.0 ± 12.0 ms) and precision (5.0 ± 3.0 ms) were reported for foot-determined flight time versus motion capture and instrumented treadmill values, respectively [34, 52]. There was a significant difference ($p < 0.05$) from high-speed camera and photocell system criterions when a lumbar spine placement was used to calculate flight time across a range of velocities (3.3-5.8 $m \cdot s^{-1}$; 41.0 to 103%) (see Table 7) [68], while the bias was -25.8% for thoracic spine-determined flight time in another study using an instrumented treadmill as the reference [37]. The observed difference for lumbar and thoracic spine sites was greater than that of a foot placement (< 15.1%) [35].

For reliability (see Table 6), the CV was as high as 11.6% at 2.2 m·s^{-1} for flight time derived from an IMU on the foot [71], while CV values were < 5.2% between trials using a lumbar spine-mounted accelerometer [68].

Outcome data between 3.3 and 4.2 m·s^{-1} was pooled from two studies [35, 71] to perform a meta-analysis assessing the effect of foot-determined flight time (I^2 = 59%; see Figure 3). Meta-analysis demonstrated that foot-determined flight time is not significantly different to reference measures (MD [95% CI] 11.93 ms [-8.88, 32.74], p = 0.13). Leave-one-out and moderator analyses were not performed due to only two studies in the meta-analysis.

Insert Table 7 about here

Insert Figure 3 about here

3.4.5 Step Time

Validity outcomes from two studies that calculated step time are presented in Table 8. Compared to values obtained from an instrumented treadmill, step time determined from a foot-worn IMU was shown to have perfect agreement and a median ± IQR precision of 3.0 ± 2.0 ms across velocities ranging from 2.8 to 5.6 m·s^{-1} [34]. The mean bias for step time calculated from a sacrum-worn accelerometer ranged from -1.3 to -0.4 ms across velocities ranging between 2.8 and 5.2 m·s^{-1}, showing a marginal underestimation of step time compared to measures derived from a motion capture system [32]. Sacrum-determined step time was most strongly correlated with motion capture at 2.8-3.3 m·s^{-1} (r = 0.93) [32].

3.4.6 Stride time

Validity outcomes for IMU-determined stride time are outlined in Table 8. Stride time was calculated from IMUs worn on the foot [52], tibia [29, 84], sacrum [32], lumbar spine [33] and thoracic spine [54]. There was no significant difference (p = 0.92) between foot-worn IMU and motion capture calculations of stride time, where the mean error ranged from -4.0 ± 24.0 ms at 2.2 m·s^{-1} to 0.3 ± 22.1 ms at 3.1 m·s^{-1} [52]. Comparison between different stride time calculation methods using tibial accelerometry showed ICC values were > 0.95 [84], while in another study using tibia-mounted IMUs, ICC values ranged between 0.55 and 0.83 using two motion capture methods (see Table 8) [29]. Stride time derived from the sacrum and lumbar spine showed low errors (standard error < 0.8 ms, mean estimation error < 5.0 ms) compared to motion capture, force plate and high-speed camera measures, respectively [32, 33]. However, when an accelerometer was attached to the thoracic spine, there was a significant bias of -26.0 ms (p = 0.00) compared to force plate stride time [54], which is greater than the bias reported for the sacrum (-1.0-1.2 ms) [32].

One study (see Table 6) established the reliability of accelerometer-derived stride time across different sampling frequencies [59]. The CV of stride time was < 3.5% for accelerometer signals between 100 and 1000 Hz [59].

Insert Table 8 about here

3.4.7 Swing Time

Only two studies, each using different attachment sites, reported the validity of IMUs to derive swing time (see Table 9). Swing time calculated from a foot-worn IMU was shown to have a median ± IQR bias of 15.0 ± 12.0 ms and a median ± IQR precision of 5.0 ± 2.0 ms compared to values obtained from an instrumented treadmill [34]. Swing time, derived from the angular velocity signal about the *y*-axis from a tibia-mounted gyroscope, showed poor to moderate agreement (ICC < 0.38) when two established motion capture methods were used as criterion measures [29, 89, 90].

Insert Table 9 about here

3.4.8 Step Frequency

Six studies quantified step frequency from foot-, tibia- and lumbar spine-worn IMUs, with reliability and validity values from each study presented in Tables 6 and 10, respectively. Foot-determined step frequency was nearly perfectly correlated (ICC > 0.95) with photocell and high-speed camera measures across a range of velocities (2.2 to 5.6 $m \cdot s^{-1}$) [35, 71]. Biases were small (< 4.5 $step \cdot min^{-1}$) and correlations exhibited close to perfect agreement ($r > 0.96$, $p < 0.001$) with an instrumented treadmill in one study that used IMUs from five different manufacturers on the foot, heel and distal tibia (see Table 10) [60]. However, the authors did not report running velocity during the trials [60]. The difference between step frequency derived from foot- and lumbar spine-worn IMUs and high-speed camera and photocell systems ranged between -0.9 and 0.8% [35, 68], while another study that directly compared values obtained from a lumbar spine-worn accelerometer to a foot-mounted accelerometer during the same run protocol deemed validity as "good” (ICC = 0.78-0.90) between 2.8 and 3.9 $m \cdot s^{-1}$ [83]. Maximal sprinting (6.8 ± 1.0 $m \cdot s^{-1}$) resulted in a bias ranging between -25.9 and -6.5 $step \cdot min^{-1}$ for step frequency derived from an IMU on the lumbar spine [72].

Reliability (see Table 6) was established for foot-determined step frequency, where the CV and SEM ranged between 1.1 to 2.0% and 1.7 to 2.8 $step \cdot min^{-1}$, respectively, across velocities (2.2 to 5.6 $m \cdot s^{-1}$) [71]. The ICC values representing the reliability of lumbar spine-determined step frequency were > 0.78 [68, 83].

Data collected between 3.3 and 4.2 $m \cdot s^{-1}$ was grouped to produce a pooled validity estimate for foot- and lumbar spine-determined step frequency (see Figure 4). There was no significant difference between foot and lumbar spine estimates of step frequency ($p = 0.20$). Derivations of step frequency from the foot (MD [95% CI] 0.45 $step \cdot min^{-1}$ [-1.75, 2.66], $p = 0.47$) and lumbar spine (MD [95% CI] -3.45 $step \cdot min^{-1}$ [-16.28, 9.39], $p = 0.37$) was shown to not be significantly different to the criterion. As there were only two studies in each subgroup, leave-one-out and moderator analyses were not performed.

Insert Table 10 about here

Insert Figure 4 about here

3.4.9 Step Length

The validity of foot-mounted IMUs to quantify step length during running at different velocities (2.2 to 5.6 m·s^{-1}) was investigated in three studies (see Table 11). Pearson's correlation and ICCs showed step length, calculated from Stryd™ and RunScribe™ devices, was nearly perfectly correlated ($r > 0.93$, $p < 0.001$) with photocell and high-speed camera measures across all velocities [35, 71]. One study used placement on the lumbar spine and showed that biases increased and ICC values decreased from jogging (bias = 8.1-12.2 cm; ICC = 0.90-0.94) to sprinting (bias = 11.5-28.4 cm; ICC = 0.79-0.85) compared to a photocell system [72].

One study assessed the reliability of step length derived from a foot-mounted IMU (see Table 6), which showed the CV ranged from 1.1 to 2.1% across all velocities (2.2 to 5.6 m·s^{-1}), while the SEM was highest at 5 m·s^{-1} (241.2 cm) [71].

Data collected between 3.3 and 4.2 m·s^{-1} was grouped to produce a pooled validity estimate for foot-determined step length. Results from the meta-analysis are presented in Figure 5 and show that IMUs worn on the foot produce step length values that are not significantly different to reference measures (MD [95% CI] 0.21 cm [-1.76, 2.18], $p = 0.69$). No moderator analysis was performed due to $I^2 = 0\%$.

3.4.10 Stride Length

Three studies determined the validity of foot-mounted IMUs to calculate stride length, where summary statistics from each study are documented in Table 11. Compared to motion capture, the mean error of IMU-derived stride length ranged from -0.5 to 46.0 cm [52, 65, 85]. The agreement between stride length determined from an IMU and motion capture system was improved during overground runs over

10 m (3.6 ± 0.3 m·s^{-1}; root mean square error [RMSE] = 8.3 cm) compared to running on a treadmill for 3 min at different velocities (2.2-3.1 m·s^{-1}; RMSE = 59.2-70.2 cm, r = 0.96, p < 0.001) [52, 85]. In a study comparing four different algorithms for computing stride length from IMU signals to a motion capture system, results showed that an algorithm based on foot trajectory performed best (mean error = 2.0 ± 14.1 cm, mean percentage error = 2.8%) than those based on stride time (mean error = 17.7 ± 57.3 cm, mean percentage error = 17.1%), foot acceleration (mean error = -0.5 ± 25.6 cm, mean percentage error = 7.9%) and deep learning (mean error = 2.5 ± 20.1 cm, mean percentage error = 5.9%) across a range of velocities up to 5.0 m·s^{-1} (see Table 11) [65].

The CV for within-subject variation of stride length across different sampling frequencies ranged from 4.9% at 1000 Hz to 7.8% at 100 Hz (see Table 6) [59].

Insert Table 11 about here

Insert Figure 5 about here

3.4.11 Ground reaction force

The outcomes for the 11 studies that investigated the validity of IMUs to estimate GRF are presented in Table 12. Two studies applied a neural network model to accelerometer data from the foot and thoracic spine to predict vertical and resultant GRF, respectively [56, 61]. The RMSE for vertical GRF determined from foot acceleration data was < 10.5 N compared to values obtained from an instrumented treadmill, while the mean signal cross-correlation was 0.99 when the entire vertical GRF waveform was evaluated [56]. A neural network method predicted resultant GRF from accelerometers worn on the thoracic spine with a mean coefficient of determination (r^2) value of 0.9 [61]. Attaching an accelerometer to the tibia [40] and hip [66] resulted in mean differences to force plate of 400.0 N and 106.4 N (~ 8.3%), respectively, for vertical GRF, whereas biases were smaller for the vertical (-34.1 N) and resultant (-29.7 N) components of peak force when an IMU was attached to the sacrum (see Table 12) [57]. One study that used a spring-mass model to calculate peak vertical force showed strong correlations between force plate-lumbar spine (*r* = 0.81) and force plate-thoracic spine (*r* = 0.79), while the CV was 9.2% and 9.6%, respectively [38]. When acceleration values were converted to Newtons by multiplying by body mass, larger measurement errors and weaker correlations were reported for both vertical (CV = 16.2%, *r* = 0.44, *p* < 0.01) and resultant GRF (CV = 16.4%) using a thoracic spine accelerometer [36, 43]. During slow (2 $m \cdot s^{-1}$) to moderate (5 $m \cdot s^{-1}$) speed running in another study, a single thoracic spine-mounted accelerometer was shown to be inadequate (RMSE > 509.2 N) for use with a mass-spring damper model to predict resultant GRF waveforms [39]. When multiple IMUs were used to estimate vertical GRF, the RMSE was 220.8 ± 45.7 N, while the root mean square deviation was 241.4 ± 59.6 N [55, 87].

The reliability of accelerometers to estimate vertical GRF was examined in four studies (see Table 6). For placement on the tibia, the SEM was 99.8 N (7.0%), whereas the minimal detectable change (MDC) was 276.7 N (19.3%) [40]. As with placement on the tibia (ICC = 0.88), lumbar spine (CV = 4.2%) and thoracic spine (CV = 3.3%) sites also showed reliable outcomes for vertical GRF derived from a spring-mass model during a continuous 2 min shuttle run [38]. However, when the same model was applied in another study using thoracic spine accelerometers, the authors classed the between-day typical error (TE; 0.8 N) and ICC (0.47) values as moderate [86]. Poor reliability was exhibited in a further study utilising accelerometers placed on the thoracic spine, whereby CV values were > 17.8% across velocities ranging between 3.3 and 6.7 $m \cdot s^{-1}$ [43].

Two studies reported mean ± SD values for thoracic spine-derived peak resultant GRF [36, 39]. However, as one study had an SD that was nearly as large as the mean [39], which suggests the data was not normally distributed and therefore not meeting the assumptions for a random-effects meta-analysis [91], these studies were not pooled.

3.4.12 Vertical stiffness

Three studies examined the reliability and validity of accelerometers placed at the lumbar and thoracic spine to calculate vertical stiffness (see Tables 6 and 13, respectively). A nearly perfect correlation (*r* =

0.98) between thoracic spine-determined vertical stiffness and that obtained from an instrumented treadmill was reported from a single participant in one study [37]. When a larger sample of participants were analysed in another study, correlations with force plate were not as strong between lumbar spine (r = 0.65) and thoracic spine (r = 0.66) estimates of vertical stiffness [38].

Inter-day reliability results were comparable between accelerometer placements, with a CV between 9.5 and 12.1% and ICC values 0.70-0.75 for both the lumbar and thoracic spine (see Table 6) [38, 86].

Insert Tables 12 and 13 about here

4 Discussion

This systematic review and meta-analysis summarises the validity and reliability of IMUs to derive spatiotemporal features of running gait and estimate peak GRF and vertical stiffness based on different attachment sites. Twelve variables were analysed across 39 studies, where the placement of IMUs varied between the foot, distal and mid tibia, hip, sacrum, lumbar spine, torso and thoracic spine. The results from reviewed studies and meta-analysis suggest it is possible to obtain valid and reliable stride data using IMUs attached at different sites. It appears that accuracy may depend more on the computational method used for identifying stride events (IC and TO) from inertial data rather than the attachment site itself.

Meta-analysis revealed that contact time and step frequency derived from IMUs placed at the foot, tibia and lumbar spine does not significantly differ to the criterion. However, some of these pooled analyses demonstrated high between-study heterogeneity (1^2 > 54.1%), which could not be explained by differing criterion methods, nor by omitting one study for the foot and lumbar spine subgroups. Subsequently, the source of heterogeneity remains unclear for these sites, but could be due to other methodological factors such as the type of sensor, sampling rate, or computational method for identifying stride events. These potential moderating variables could not be investigated further due to insufficient reporting of data within those studies. Although there were no influential studies for the foot and lumbar spine subgroups, the pooled MD for contact time determined from the tibia was distorted when one study [70] was omitted. Removal of this study from the meta-analysis resulted in an overall effect that was significantly different ($p = 0.02$) to the criterion, which would have suggested the tibia is not a suitable site to determine contact time had the study not been included. Other work reviewed here demonstrated valid results for contact time using IMUs secured to the distal tibia [62]. Although this study was not eligible for inclusion in the meta-analysis due to insufficient reporting of data, it is possible it may have supported our findings in the final meta-analysis, where no significant difference ($p = 0.18$) was observed between the tibia and criterion. Furthermore, IC and TO have been detected with good accuracy from tibial acceleration data [31, 62, 63], which suggests this site is a viable option for calculating temporal variables, such as contact time.

Subgroup analysis was not possible for flight time and step length due to a limited number of studies meeting eligibility criteria for inclusion. However, studies that used foot-worn IMUs to determine these metrics were meta-analysed and demonstrated that estimates of flight time and step length was not significantly different from criterion measures, which is similar to the results reported for contact time and step frequency. Collectively, the results from the four meta-analyses highlight the utility of using IMUs for gait analysis, where the findings reported here may open opportunities for practitioners to use placement on the foot, tibia or lumbar spine to capture spatiotemporal features of an athlete's stride in the field. However, there has been little work done (two reviewed studies) applying gait event detection methods to inertial data from the thoracic spine to investigate the validity of this site to derive temporal variables, with one study only reporting a single observation ($n = 1$) [37, 54]. It is therefore unclear whether placement on the thoracic spine is also suitable to derive temporal stride data.

Peak vertical or resultant GRFs during running have traditionally been measured from force platforms [92-94]. However, IMUs are more accessible to an athlete's normal training and competition environment than force platforms and may provide a useful tool for quantifying surrogate measures of force during running-based sports [95]. A variety of different approaches were used to estimate peak

GRFs in the studies reviewed here. Although meta-analysis was not possible, predictions of vertical GRF were shown to be most accurate when studies applied machine learning techniques or used multiple IMUs at different body segments [55, 56, 61, 87]. Given IMUs are commonly worn on the thoracic spine in sport, other studies investigated the validity of this site to predict GRFs from accelerometer data, with contrasting results. Acceleration data from the thoracic spine was inadequate to predict peak vertical and resultant GRF based on Newton's second law of motion (i.e. multiplying by body mass) [36, 43] and as input into a mass-spring-damper model [39]. Conversely, improved results were shown when peak vertical GRF was estimated from known contact time, flight time and body mass using a spring-mass model [38], while another study suggested accurate predictions of resultant GRFs from IMUs worn on the thoracic spine are possible by applying machine learning [61]. Based on the conflicting results from the studies reviewed here, it is unclear whether the accurate determination of peak vertical and resultant GRFs from accelerometer data at the thoracic spine is possible and warrants further investigation.

Two studies used estimations of peak vertical GRF to calculate vertical stiffness from IMUs worn the thoracic spine [37, 38]. Although the small biases and large to nearly perfect correlations in both studies appear promising for determining vertical stiffness using accelerometer data from this site, it is unclear whether placement on the thoracic spine is feasible for determining vertical stiffness when one study collected data from only one participant. Furthermore, calculating vertical stiffness using a spring-mass model approach, as per the method used in the two studies, is dependent on known contact time and flight time [96]. However, neither study provided a description of how IC and TO were determined mathematically from accelerometer data, nor how these events translated to accurate derivations of temporal variables [37, 38]. The ability of IMUs attached on the thoracic spine to correctly identify IC and TO events compared to a criterion should be explored more fully before practitioners can confidently use this site to 1) accurately calculate contact time and flight time and 2) use these metrics as inputs for estimating peak vertical GRF and vertical stiffness [23, 97].

Results from reviewed studies demonstrates that it is possible to obtain reliable derivations of contact time, flight time and step frequency from a foot or lumbar spine placement [16, 68, 71], while foot-worn IMUs can provide reproducible calculations of stride time, step length and stride length [59, 71]. Furthermore, placement on the tibia and lumbar and thoracic spine possessed excellent reliability for determining vertical GRF from accelerometer data [38, 40]. Collectively, these results indicate that IMUs possess good precision for calculating different stride variables [98]. Determining the sensitivity of IMU-derived stride variables by calculating the MDC or smallest worthwhile change (SWC) is also important so practitioners can determine whether changes in an athlete's gait pattern are real or due to error [99-101]. However, only two studies reported here determined the value (i.e. signal) that may constitute meaningful change for stride variables determined from IMUs [40, 86]. One study using tibia accelerometers calculated an MDC for peak vertical GRF that was higher than the SEM, suggesting that this metric may be sensitive to detect change when IMUs are secured to the tibia [40]. Conversely, the TE associated with thoracic spine-derived peak vertical GRF and vertical stiffness was greater than the SWC [86], which suggests this site is limited for detecting subtle changes in an athlete's gait pattern. No study determined the MDC or SWC for spatiotemporal variables, therefore future work may look to further our understanding of the signal-to-noise ratio of other stride metrics, such as from IMUs worn at various sites.

The use of IMUs in sport is increasingly being applied to gain additional insights (i.e. other than speed and distance) into the activity profiles of athletes. Practitioners can quantify proprietary designed metrics, such as PlayerLoad™ [10, 28, 102], estimate energy expenditure [103] and record the peak segmental acceleration values that occur during a variety of different team-sport movements [42, 104] using IMUs. There is an increasing body of evidence supporting the use of IMUs to capture characteristics of an athlete's stride, including spatiotemporal data [54], GRFs [36, 39, 40] and vertical stiffness [37, 86]. Capturing accurate stride variables appears possible across different sites using automated gait event detection techniques and may have practical application in profiling an athlete's stride in a variety of running-based sports. The use of IMUs may allow practitioners to perform gait analyses in the field to enhance their understanding of athlete movement strategy and monitor changes in stride variables that may occur with fatigue [28].

It is important to note that the meta-analyses in this review were impacted by a limited pool of eligible studies. It is likely that the results suffer from sparse data bias in instances where only two studies were meta-analysed due to relatively small sample sizes [105, 106]. Further research should include raw outcome data (mean ± SD values) alongside validity statistics in order to provide a complete summary of outcomes. Furthermore, the method adopted here treated three studies that used different IMUs or criterions as independent data sources [29, 35, 68]. It is possible that we may have observed a different finding had different IMUs or criterions not been treated independently within those studies. However, due to a limited number of studies, accounting for this dependency was not possible with the data available. Finally, data was only pooled within a velocity range of 3.3 to 4.3 $m \cdot s^{-1}$ due to eligibility criteria. As a result, meta-analyses here do not explain the effect of running velocity on validity, which may be an important distinction to make as previous work has shown that increased speed may lead to greater error in estimations of stride variables derived from IMUs [34, 43].

5 Conclusion

This review and meta-analysis demonstrated that valid and reliable derivations of stride metrics are possible from IMUs mounted on the foot, tibia and lumbar spine. This suggests that location may not be the most critical factor and that validity and reliability may be more dependent on the mathematic approach for detection of gait events. However, further work is warranted to explore the application of automated gait event detection algorithms on inertial data from the thoracic spine before practitioners can confidently use this site in the field to derive stride variables.

References

1. Vanrenterghem J, Nedergaard N, Robinson MA, Drust B. Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways. Sports Med. 2017;47(11):2135-42.
2. Spencer M, Lawrence S, Rechichi C, Bishop D, Dawson B, Goodman C. Time–motion analysis of elite field hockey, with special reference to repeated-sprint activity. J Sport Sci. 2004;22(9):843-50.
3. Aughey RJ. Applications of GPS technologies to field sports. Int J Sports Physiol Perform. 2011;6(3):295-310.
4. Buchheit M, Allen A, Poon TK, Modonutti M, Gregson W, Di Salvo V. Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. J Sport Sci. 2014;32(20):1844-57.
5. Stevens TG, de Ruiter CJ, van Niel C, van de Rhee R, Beek PJ, Savelsbergh GJ. Measuring acceleration and deceleration in soccer-specific movements using a local position measurement (LPM) system. Int J Sports Physiol Perform. 2014;9(3):446-56.
6. Duffield R, Reid M, Baker J, Spratford W. Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. J Sci Med Sport. 2010;13(5):523-5.
7. Jennings D, Cormack S, Coutts AJ, Boyd L, Aughey RJ. The validity and reliability of GPS units for measuring distance in team sport specific running patterns. Int J Sports Physiol Perform. 2010;5(3):328-41.
8. Varley MC, Fairweather IH, Aughey RJ. Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. J Sport Sci. 2012;30(2):121-7.
9. Gabbett T, Jenkins D, Abernethy B. Physical collisions and injury during professional rugby league skills training. J Sci Med Sport. 2010;13(6):578-83.
10. Boyd LJ, Ball K, Aughey RJ. Quantifying external load in Australian football matches and training using accelerometers. Int J Sports Physiol Perform. 2013;8(1):44-51.
11. Gastin PB, McLean O, Spittle M, Breed RV. Quantification of tackling demands in professional Australian football using integrated wearable athlete tracking technology. J Sci Med Sport. 2013;16(6):589-93.
12. Gastin PB, Mclean OC, Breed RV, Spittle M. Tackle and impact detection in elite Australian football using wearable microsensor technology. J Sport Sci. 2014;32(10):947-53.
13. Glassbrook DJ, Fuller JT, Alderson JA, Doyle TL. Foot accelerations are larger than tibia accelerations during sprinting when measured with inertial measurement units. J Sport Sci. 2019:1-8. doi:10.1080/02640414.2019.1692997.
14. Warman GE, Cole MH, Johnston RD, Chalkley D, Pepping G-J. Using microtechnology to quantify torso angle during match-play in field hockey. J Strength Cond Res. 2019;33(10):2648-54.
15. Boyd LJ, Ball K, Aughey RJ. The reliability of MinimaxX accelerometers for measuring physical activity in Australian football. Int J Sports Physiol Perform. 2011;6(3):311-21.
16. Ammann R, Taube W, Wyss T. Accuracy of PARTwear inertial sensor and Optojump optical measurement system for measuring ground contact time during running. J Strength Cond Res. 2016;30(7):2057-63.
17. Gabbett TJ. Quantifying the physical demands of collision sports: does microsensor technology measure what it claims to measure? J Strength Cond Res. 2013;27(8):2319-22.
18. Bouten CV, Westerterp KR, Verduin M, Janssen JD. Assessment of energy expenditure for physical activity using a triaxial accelerometer. Med Sci Sports Exerc. 1994;26(12):1516-23.
19. Bouten CV, Koekkoek KT, Verduin M, Kodde R, Janssen JD. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. IEEE Trans Biomed Eng. 1997;44(3):136-47.
20. Swartz AM, Strath SJ, Bassett DR, O'Brien WL, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. Med Sci Sports Exerc. 2000;32(9):450-6.
21. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40(1):181-8.
22. Skotte J, Korshøj M, Kristiansen J, Hanisch C, Holtermann A. Detection of physical activity types using triaxial accelerometers. J Phys Act Health. 2014;11(1):76-84.

23. McNamara DJ, Gabbett TJ, Chapman P, Naughton G, Farhart P. The validity of microsensors to automatically detect bowling events and counts in cricket fast bowlers. Int J Sports Physiol Perform. 2015;10(1):71-5.

24. Murray NB, Black GM, Whiteley RJ, Gahan P, Cole MH, Utting A et al. Automatic detection of pitching and throwing events in baseball with inertial measurement sensors. Int J Sports Physiol Perform. 2017;12(4):533-7.

25. Spangler R, Rantalainen T, Gastin PB, Wundersitz D. Inertial sensors are a valid tool to detect and consistently quanitfy jumping. Int J Sports Med. 2018;39(10):802-8. doi:10.1055/s-0044-100793.

26. Barrett S, Midgley A, Lovell R. PlayerLoad™: reliability, convergent validity, and influence of unit position during treadmill running. Int J Sports Physiol Perform. 2014;9(6):945-52.

27. Mooney MG, Cormack S, O'Brien BJ, Morgan WM, McGuigan M. Impact of neuromuscular fatigue on match exercise intensity and performance in elite Australian football. J Strength Cond Res. 2013;27(1):166-73.

28. Cormack SJ, Mooney MG, Morgan W, McGuigan MR. Influence of neuromuscular fatigue on accelerometer load in elite Australian football players. Int J Sports Physiol Perform. 2013;8(4):373-8.

29. McGrath D, Greene BR, O'Donovan KJ, Caulfield B. Gyroscope-based assessment of temporal gait parameters during treadmill walking and running. Sports Eng. 2012;15(4):207-13.

30. Benson LC, Clermont CA, Watari R, Exley T, Ferber R. Automated accelerometer-based gait event detection during multiple running conditions. Sensors. 2019;19(7):1483. doi:10.3390/s19071483.

31. Aubol KG, Milner CE. Foot contact identification using a single triaxial accelerometer during running. J Biomech. 2020;105:109768. doi:10.1016/j.jbiomech.2020.109768.

32. Lee JB, Mellifont RB, Burkett BJ. The use of a single inertial sensor to identify stride, step, and stance durations of running gait. J Sci Med Sport. 2010;13(2):270-3.

33. Bergamini E, Picerno P, Pillet H, Natta F, Thoreux P, Camomilla V. Estimation of temporal parameters during sprint running using a trunk-mounted inertial measurement unit. J Biomech. 2012;45(6):1123-6.

34. Falbriard M, Meyer F, Mariani B, Millet GP, Aminian K. Accurate estimation of running temporal parameters using foot-worn inertial sensors. Front Physiol. 2018;9. doi:10.3389/fphys.2018.00610.

35. García-Pinillos F, Latorre-Román PÁ, Soto-Hermoso VM, Párraga-Montilla JA, Pantoja-Vallejo A, Ramírez-Campillo R et al. Agreement between the spatiotemporal gait parameters from two different wearable devices and high-speed video analysis. PLoS ONE. 2019;14(9). doi:10.1371/journal.pone.0222872.

36. Wundersitz DWT, Netto KJ, Aisbett B, Gastin PB. Validity of an upper-body-mounted accelerometer to measure peak vertical and resultant force during running and change-of-direction tasks. Sports Biomech. 2013;12(4):403-12.

37. Buchheit M, Gray A, Morin JB. Assessing stride variables and vertical stiffness with GPS-embedded accelerometers: preliminary insights for the monitoring of neuromuscular fatigue on the field. J Sports Sci Med. 2015;14:698-701.

38. Eggers TM, Massard TI, Clothier PJ, Lovell R. Measuring vertical stiffness in sport with accelerometers: exercise caution! J Strength Cond Res. 2018;32(7):1919-22.

39. Nedergaard NJ, Verheul J, Drust B, Etchells T, Lisboa P, Robinson MA et al. The feasibility of predicting ground reaction forces during running from a trunk accelerometry driven mass-spring-damper model. PeerJ. 2018;6. doi:10.7717/peerj.6105.

40. Raper DP, Witchalls J, Philips EJ, Knight E, Drew MK, Waddington G. Use of a tibial accelerometer to measure ground reaction force in running: a reliability and validity comparison with force plates. J Sci Med Sport. 2018;21(1):84-8.

41. Dugan SA, Bhat KP. Biomechanics and analysis of running gait. Phys Med Rehabil Clin N Am. 2005;16(3):603-21.

42. Nedergaard NJ, Robinson MA, Eusterwiemann E, Drust B, Lisboa PJ, Vanrenterghem J. The relationship between whole-body external loading and body-worn accelerometry during team-sport movements. Int J Sports Physiol Perform. 2017;12(1):18-26.

43. Edwards S, White S, Humphreys S, Robergs R, O'Dwyer N. Caution using data from triaxial accelerometers housed in player tracking units during running. J Sport Sci. 2019;37(7):810-8.

44. Glassbrook DJ, Fuller JT, Alderson JA, Doyle TL. Measurement of lower-limb asymmetry in professional rugby league: a technical note describing the use of inertial measurement units. PeerJ. 2020;8:e9366. doi:10.7717/peerj.9366.
45. Lafortune MA, Lake MJ, Hennig EM. Differential shock transmission response of the human body to impact severity and lower limb posture. J Biomech. 1996;29(12):1531-7.
46. Derrick TR, Hamill J, Caldwell GE. Energy absorption of impacts during running at various stride lengths. Med Sci Sports Exerc. 1998;30(1):128-35. doi:10.1097/00005768-199801000-00018.
47. Lucas-Cuevas AG, Encarnación-Martínez A, Camacho-García A, Llana-Belloch S, Pérez-Soriano P. The location of the tibial accelerometer does influence impact acceleration parameters during running. J Sport Sci. 2017;35(17):1734-8. doi:10.1080/02640414.2016.1235792.
48. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9.
49. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52(6):377-84.
50. Cummins C, Orr R, O'Connor H, West C. Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. Sports Med. 2013;43(10):1025-42.
51. Glassbrook DJ, Doyle TL, Alderson JA, Fuller JT. The demands of professional rugby league match-play: a meta-analysis. Sports Med. 2019;5:1-20. doi:10.1186/s40798-019-0197-9.
52. Chew D-K, Ngoh KJ, Gouwanda D, Gopalai AA. Estimating running spatial and temporal parameters using an inertial sensor. Sports Eng. 2018;21(2):115-22.
53. Watari R, Hettinga B, Osis S, Ferber R. Validation of a torso-mounted accelerometer for measures of vertical oscillation and ground contact time during treadmill running. J Appl Biomech. 2016;32(3):306-10. doi:10.1123/jab.2015-0200.
54. Kenneally-Dabrowski CJ, Serpell BG, Spratford W. Are accelerometers a valid tool for measuring overground sprinting symmetry? Int J Sports Sci Coach. 2018;13(2):270-7.
55. Dorschky E, Nitschke M, Seifer A-K, van den Bogert AJ, Eskofier BM. Estimation of gait kinematics and kinetics from inertial sensor data using optimal control of musculoskeletal models. J Biomech. 2019;95:109278. doi:10.1016/j.jbiomech.2019.07.022.
56. Ngoh KJ, Gouwanda D, Gopalai AA, Chong YZ. Estimation of vertical ground reaction force during running using neural network model and uniaxial accelerometer. J Biomech. 2018;76:269-73. doi:10.1016/j.jbiomech.2018.06.006.
57. Gurchiek RD, McGinnis RS, Needle AR, McBride JM, van Werkhoven H. The use of a single inertial sensor to estimate 3-dimensional ground reaction force during accelerative running tasks. J Biomech. 2017;61:263-8. doi:10.1016/j.jbiomech.2017.07.035.
58. Mitschke C, Heß T, Milani TL. Which method detects foot strike in rearfoot and forefoot runners accurately when using an inertial measurement unit? Appl Sci. 2017;7(9):959. doi:10.3390/app7090959.
59. Mitschke C, Zaumseil F, Milani TL. The influence of inertial sensor sampling frequency on the accuracy of measurement parameters in rearfoot running. Comput Methods Biomech Biomed Engin. 2017;20(14):1502-11. doi:10.1080/10255842.2017.1382482.
60. Pairot de Fontenay B, Roy J, Dubois B, Bouyer L, Esculier J. Validating commercial wearable sensors for running gait parameters estimation. IEEE Sens J. 2020;20(14):7783-91. doi:10.1109/JSEN.2020.2982568.
61. Pogson M, Verheul J, Robinson MA, Vanrenterghem J, Lisboa P. A neural network method to predict task- and step-specific ground reaction force magnitudes from trunk accelerations during running activities. Med Eng Phys. 2020;78:82-9. doi:10.1016/j.medengphy.2020.02.002.
62. Sinclair J, Hobbs SJ, Protheroe L, Edmundson CJ, Greenhalgh A. Determination of gait events using an externally mounted shank accelerometer. J Appl Biomech. 2013;29(1):118-22.
63. Tan HX, Aung NN, Tian J, Chua MCH, Yang YO. Time series classification using a modified LSTM approach from accelerometer-based data: a comparative study for gait cycle detection. Gait Posture. 2019;74:128-34. doi:10.1016/j.gaitpost.2019.09.007.
64. Winter SC, Lee JB, Leadbetter RI, Gordon SJ. Validation of a single inertial sensor for measuring running kinematics overground during a prolonged run. J Fit Res. 2016;5(1):14-23.

65. Zrenner M, Gradl S, Jensen U, Ullrich M, Eskofier BM. Comparison of different algorithms for calculating velocity and stride length in running using inertial measurement units. Sensors. 2018;18(12):4194. doi:10.3390/s18124194.
66. Neugebauer JM, Collins KH, Hawkins DA. Ground reaction force estimates from ActiGraph GT3X+ hip accelerations. PLoS ONE. 2014;9(6):e99023. doi:10.1371/journal.pone.0099023.
67. Hall JP, Barton C, Jones PR, Morrissey D. The biomechanical differences between barefoot and shod distance running: a systematic review and preliminary meta-analysis. Sports Med. 2013;43(12):1335-53.
68. Gindre C, Lussiana T, Hebert-Losier K, Morin J-B. Reliability and validity of the Myotest for measuring running stride kinematics. J Sport Sci. 2016;34(7):664-70.
69. Mo S, Chow DH. Accuracy of three methods in gait event detection during overground running. Gait Posture. 2018;59:93-8. doi:10.1016/j.gaitpost.2017.10.009.
70. Fadillioglu C, Stetter BJ, Ringhof S, Krafft FC, Sell S, Stein T. Automated gait event detection for a variety of locomotion tasks using a novel gyroscope-based algorithm. Gait Posture. 2020;81:102-8. doi:10.1016/j.gaitpost.2020.06.019.
71. García-Pinillos F, Roche-Seruendo LE, Marcén-Cinca N, Marco-Contreras LA, Latorre-Román PA. Absolute reliability and concurrent validity of the Stryd system for the assessment of running stride kinematics at different velocities. J Strength Cond Res. 2018:1-8. doi:10.1519/jsc.0000000000002595.
72. Machulik C, Hamacher D, Lindlein K, Zech A, Hollander K. Validation of an inertial measurement unit based magnetic timing gate system during running and sprinting. German Journal of Sports Medicine. 2020;71(3):69-75. doi:10.5960/dzsm.2020.426.
73. Simon SR. Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. J Biomech. 2004;37(12):1869-80.
74. Lienhard K, Schneider D, Maffiuletti NA. Validity of the Optogait photoelectric system for the assessment of spatiotemporal gait parameters. Med Eng Phys. 2013;35(4):500-4.
75. Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. Sensors. 2014;14(2):3362-94.
76. Schwarzer G. meta: an R package for meta-analysis. R News. 2007;7(3):40-5.
77. Borenstein M, Higgins JP. Meta-analysis and subgroups. Prev Sci. 2013;14(2):134-43.
78. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J. 2003;327(7414):557-60.
79. Viechtbauer W, Cheung MWL. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010;1(2):112-25.
80. Quintana DS. From pre-registration to publication: a non-technical primer for conducting a meta-analysis to synthesize correlational data. Front Physiol. 2015;6. doi:10.3389/fpsyg.2015.01549.
81. Schwarzer G, Carpenter JR, Rücker G. Meta-Analysis with R. 1 ed. Use R! Springer International Publishing Switzerland; 2015.
82. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. Br Med J. 2001;322(7300):1479-80.
83. Gouttebarge V, Wolfard R, Griek N, de Ruiter CJ, Boschman JS, van Dieën JH. Reproducibility and validity of the Myotest for measuring step frequency and ground contact time in recreational runners. J Hum Kinet. 2015;45:19-26.
84. Norris M, Kenny IC, Anderson R. Comparison of accelerometry stride time calculation methods. J Biomech. 2016;49(13):3031-4.
85. Brahms C, Zhao Y, Gerhard D, Barden JM. Stride length determination during overground running using a single foot-mounted inertial measurement unit. J Biomech. 2018;71:302-5.
86. Buchheit M, Lacome M, Cholley Y, Simpson BM. Neuromuscular responses to conditioned soccer sessions assessed via GPS-embedded accelerometers: insights into tactical periodization. Int J Sports Physiol Perform. 2018;13:577-83.
87. Wouda FJ, Giuberti M, Bellusci G, Maartens E, Reenalda J, van Beijnum B-JF et al. Estimation of vertical ground reaction forces and sagittal knee kinematics during running using three inertial sensors. Front Physiol. 2018;9:218. doi:10.3389/fphys.2018.00218.
88. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J Mach Learn Technol. 2011;2(1):37-63.

89. Hreljac A, Marshall RN. Algorithms to determine event timing during normal walking using kinematic data. J Biomech. 2000;33(6):783-6.
90. Zeni Jr. J, Richards J, Higginson J. Two simple methods for determining gait events during treadmill and overground walking using kinematic data. Gait Posture. 2008;27(4):710-4.
91. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods. 2010;1(2):97-111. doi:10.1002/jrsm.12.
92. Cavanagh PR, Lafortune MA. Ground reaction forces in distance running. J Biomech. 1980;13(5):397-406. doi:10.1016/0021-9290(80)90033-0.
93. Keller TS, Weisberger A, Ray J, Hasan SS, Shiavi RG, Spengler DM. Relationship between vertical ground reaction force and speed during walking, slow jogging, and running. Clin Biomech. 1996;11(5):253-9.
94. Hunter JP, Marshall RN, McNair PJ. Relationships between ground reaction force impulse and kinematics of sprint-running acceleration. J Appl Biomech. 2005;21(1):31-43.
95. Verheul J, Nedergaard NJ, Vanrenterghem J, Robinson MA. Measuring biomechanical loads in team sports – from lab to field. Sci Med Footb. 2020:1-7.
96. Morin J-B, Dalleau G, Kyröläinen H, Jeannin T, Belli A. A simple method for measuring stiffness during running. J Appl Biomech. 2005;21(2):167-80.
97. Ruddy JD, Cormack SJ, Whiteley R, Williams MD, Timmins RG, Opar DA. Modeling the risk of team sport injuries: a narrative review of different statistical approaches. Front Physiol. 2019;10. doi:10.3389%2Ffphys.2019.00829.
98. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1-15.
99. Cormack SJ, Newton RU, McGuigan MR, Doyle TL. Reliability of measures obtained during single and repeated countermovement jumps. Int J Sports Physiol Perform. 2008;3(2):131-44.
100. Thorborg K, Petersen J, Magnusson SP, Hölmich P. Clinical assessment of hip strength using a hand-held dynamometer is reliable. Scand J Med Sci Sports 2010;20(3):493-501.
101. Garrett J, Graham SR, Eston RG, Burgess DJ, Garrett LJ, Jakeman J et al. A novel method of assessment for monitoring neuromuscular fatigue in Australian rules football players. Int J Sports Physiol Perform. 2018;14(5):598-605.
102. Rowell AE, Aughey RJ, Clubb J, Cormack SJ. A standardized small sided game can be used to monitor neuromuscular fatigue in professional A-League football players. Front Physiol. 2018;9. doi:10.3389/fphys.2018.01011.
103. Walker EJ, McAinch AJ, Sweeting A, Aughey RJ. Inertial sensors to estimate the energy expenditure of team-sport athletes. J Sci Med Sport. 2016;19(2):177-81.
104. Wundersitz DW, Gastin PB, Richter C, Robertson SJ, Netto KJ. Validity of a trunk-mounted accelerometer to assess peak accelerations during walking, jogging and running. Eur J Sport Sci. 2015;15(5):382-90.
105. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. Br Med J. 2016;352:i1981. doi:10.1136/bmj.i1981.
106. Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. PLoS ONE. 2018;13(9):e0204056. doi:10.1371/journal.pone.0204056.

Table 1 Definitions of stride variables.

Variable	Definition
Initial contact	The time instant when the foot initiates contact with the ground [34].
Toe-off	The time instant when the foot ends contact with the ground [34].
Contact time	Time between initial contact to toe-off of each foot [32, 37]
Flight time	Time between toe-off and initial contact of the contralateral foot [37].
Step time	Time between initial contacts of the contralateral foot [32].
Stride time	Time between initial contacts of the same foot [32, 54].
Swing time	Time between toe-off to initial contact of the same foot [29].
Step frequency	Number of ground contact events per minute [35].
Step length	Length or distance between initial contacts of the contralateral foot [35].
Stride length	Length or distance between initial contacts of the same foot [65].
Ground reaction force	The force the ground exerts on the body during foot-ground contact [92].
Vertical stiffness	The quotient of maximum ground reaction force and centre of mass displacement [96].

Table 2 Study characteristics for the 39 studies included in the review.

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Ammann et al. (2016) [16]	7 male and 5 female high-level running athletes (31 ± 6 y, 68.6 ± 11.6 kg, 1.70 ± 0.08 m)	IMU (PARTwear, HuCE-microLab, University of Applied Sciences, Biel, Switzerland)	Triaxial accelerometer (± 16 g), triaxial gyroscope and triaxial magnetometer sampled at 1000 Hz	Foot (fixed to the lace of the shoe)	High-speed camera (Camera Marathon Ultra CL600, Videal AG, Niederönz, Switzerland) sampled at 1000 Hz	3 x 40 m runs at 4.3 ± 0.7, 6.2 ± 0.7 and 8.0 ± 0.5 $m{\cdot}s^{-1}$	Contact time	10/12
Aubol & Milner (2020) [31]	9 male and 10 female recreational runners (26.2 ± 3.8 y, 71.5 ± 7.1 kg, 1.78 ± 0.06 m)	Accelerometer (Model 356A45, PCB Piezotronics, Depew, NY)	Triaxial accelerometer (± 16 g) sampled at 201.03 Hz)	Distal tibia	Force plate (AMTI, Inc., Watertown, MA) sampled at 1000 Hz	10 x 17 m runs at 3.0 ± 0.2 $m{\cdot}s^{-1}$	Initial contact	10/12
Benson et al. (2019) [30]	8 male and 4 female recreational runners (26.2 ± 3.8 y, 71.5 ± 7.1 kg, 1.78 ± 0.06 m)	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Triaxial accelerometer (± 16 g) sampled at 201.03 Hz)	Dorsal foot and lumbar spine	Instrumented split-belt treadmill (Bertec Inc., Columbus, OH) sampled at 1000 Hz	90 s runs at 2.7, 3.3 and 3.6 $m{\cdot}s^{-1}$	Initial contact, toe-off and contact time	9/12
Bergamini et al. (2012) [33]	5 elite track and field athletes	IMU (FreeSense, Sensorize, Italy)	Triaxial accelerometer (± 6 g) and triaxial gyroscope (± 500°s^{-1}) sampled at 200 Hz	Lumbar spine (L1 level)	Six adjacent force platforms (Z20740AA, Kistler, Switzerland) sampled at 200 Hz and high-speed camera (Casio Exilim EX-F1, Japan) sampled at 300 Hz	3 x 60 m maximal sprints	Contact time and stride time	8/12
Brahms et al. (2018) [85]	7 male and 4 female healthy adults (22.3 ± 1.5 y, 76.04 ± 3.19 kg, 175.2 ± 23.1 cm)	IMU (Xsens, Enschede, the Netherlands)	Triaxial accelerometer, triaxial gyroscope and triaxial magnetometer sampled at 100 Hz	Mid-foot	6-camera 3D motion capture system sampled at 100 Hz	20 x 10 m runs at 2.7 to 4.4 $m{\cdot}s^{-1}$	Stride length	10/12
Buchheit et al. (2015) [37]	1 team-sport athlete (36 y, 80 kg, 182 cm)	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Triaxial accelerometer (± 16 g) sampled at 100 Hz	Thoracic spine (T2 level)	Instrumented treadmill (ADAL3D-WR, MD, HEF Tecmachine, Andrézieux-Boutheon, France) sampled at 1000 Hz	2 x 3 runs at 2.8 $m{\cdot}s^{-1}$; 6 runs at 4.7 $m{\cdot}s^{-1}$; 6 runs at 6.7 $m{\cdot}s^{-1}$	Contact time, flight time and vertical stiffness	9/12
Buchheit et al. (2018) [86]	18 elite academy soccer athletes (17 ± 2 y)	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Triaxial accelerometer (± 16 g) sampled at 100 Hz	Thoracic spine	N/A	4 x ~60 m runs at 6.1-6.7 $m{\cdot}s^{-1}$)	Vertical stiffness	10/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Chew et al. (2018) [52]	10 healthy males (25.5 ± 3.8 y, 65.5 ± 15.2 kg, 174.4 ± 19.5 cm)	IMU (Opal, APDM Inc.)	Triaxial accelerometer (± 6 g) and triaxial gyroscope (± 2000°s^{-1}) sampled at 128 Hz	Foot (fixed to the shoe)	Optical motion capture system (Qualisys, Qualisys AB) sampled at 128 Hz	3 min runs at 2.2, 2.5, 2.8 and 3.1 m·s^{-1}	Initial contact, toe-off, contact time, flight time, stride time and stride length	10/12
Dorschky et al. (2019) [55]	10 healthy male subjects (27.1 ± 2.6 y, 76.9 ± 8.6 kg, 1.82 ± 0.05 m)	IMU (Portabiles GmbH, Erlangen, DE)	Triaxial accelerometer (± 16 g) and triaxial gyroscope (± 200°s^{-1}) sampled at 1000 Hz	Midfoot, lateral tibia, left and right lateral thigh and lumbar spine	Optical motion capture system (Vicon MX, Oxford, UK) sampled at 200 Hz and one force plate (Kistler Instruments Corp, Winterhur, CH) sampled at 1000 Hz	Runs over a force platform at 3.0-3.3 m·s^{-1}, 3.9-4.1 m·s^{-1} and 4.7-4.9 m·s^{-1}	Vertical ground reaction force	9/12
Edwards et al. (2019) [43]	10 male rugby union athletes (21 ± 2 y, 81.8 ± 11.1 kg, 1.81 ± 0.50 m)	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Triaxial accelerometer (± 16 g) sampled at 100 Hz	Thoracic spine (T1-T6 vertebrae)	Two floor-embedded force platforms (Type 9281CA and 9821EA, Kistler, Winterhur, Switzerland) sampled at 1200 Hz	Ten runs over a force platform at 3.3, 5.0 and 6.7 m·s^{-1}	Vertical ground reaction force	10/12
Eggers et al. (2018) [38]	10 male and 7 female healthy adults (18-40 y, 70.4 ± 9.7 kg, 1.73 ± 0.06 m)	Accelerometer (wGT3X-BT, ActiGraph, Pensacola, FL, USA)	Triaxial accelerometer (± 8 g) sampled at 100 Hz	Lumbar spine (L2) and thoracic spine	Four 600 x 400 mm force plates (model BP400600-1000, Advanced Mechanical Technology, Inc., Watertown, MA, USA) sampled at 2000 Hz	2 min continuous shuttle runs over 20 m at 3.3 m·s^{-1}	Vertical ground reaction force and vertical stiffness	9/12
Fadillioglu et al. (2020) [70]	13 male healthy adults (26.1 ± 2.9 y, 78.4 ± 5.9 kg, 178.7 ± 5.5 cm)	Gyroscope (ADXRS652, Analog Devices Inc., Norwood, MA, USA)	Uniaxial gyrosope (± 250°s^{-1}) sampled at 1500 Hz	Tibia	Two floor-embedded force plates (BP600900, Advanced Mechanical Technology, Inc., Watertown, MA, USA)	3 trials of moderate and fast running (velocity not reported)	Initial contact, toe off and contact time	10/12
Falbriard et al. (2018) [34]	28 male and 13 female healthy adults (29 ± 6 y, 70 ± 10 kg, 174 ± 8 cm)	IMU (Physilog 4, Gait Up, Switzerland)	Triaxial accelerometer (± 16 g) and triaxial gyroscope (± 2000°s^{-1}) sampled at 500 Hz	Dorsal foot	Instrumented treadmill (T-170-FMT, Arsalis, Belgium) sampled at 1000 Hz	30 s runs ranging between 2.8 m·s^{-1} and 5.6 m·s^{-1}	Initial contact, toe-off, contact time, flight time, swing time and step time	9/12
Garcia-Pinillos et al. (2018) [71]	18 male recreational endurance runners (34 ± 7 y, 70.5 ± 6.2 kg, 1.76 ± 0.05 m)	IMU (Stryd^{TM}, Stryd Powermeter, Stryd Inc., Boulder, CO, USA)	Triaxial accelerometer and triaxial gyroscope	Foot	OptoGait system (Optogait; Microgate, Bolzano, Italy)	3 min runs ranging between 2.2 m·s^{-1} and 5.6 m·s^{-1} (0.3 m·s^{-1} increments)	Contact time, flight time, step frequency and step length	10/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Garcia-Pinillos et al. (2019) [35]	44 male and 5 female amateur endurance runners (26 ± 8 y, 71 ± 10 kg, 1.74 ± 0.07 m)	IMU (Stryd™ [Stryd Powermeter, Stryd Inc. Boulder CO, USA]; RunScribe™ [Scribe Lab. Inc. San Francisco CA, USA])	Triaxial accelerometer and triaxial gyroscope Triaxial accelerometer, triaxial gyroscope and triaxial magnetometer sampled at 500 Hz	Foot	High-speed camera (Imaging Source DFK 33UX174, The Imaging Source Europe GmbH; Germany) sampled at 1000 Hz	3 min self-selected comfortable running velocity (3.25 ± 0.36 $m{\cdot}s^{-1}$)	Contact time, flight time, step frequency and step length	10/12
Gindre et al. (2016) [68]	20 male runners (31.6 ± 9.2 y, 72.5 ± 9.8 kg, 178 ± 5.4 cm)	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Triaxial accelerometer sampled at 500 Hz	Lumbar spine (level with naval)	Optojump Next® (Microgate, Bolzano, Italy) sampled at 1000 Hz and high-speed video camera (Casio High Speed EXILIM EX-FH25®, CASIO Europe GmbH, Norderstedt, Germany) sampled at 300 Hz	2 x 60 m runs at 3.3, 4.2, 5 and 5.8 $m{\cdot}s^{-1}$	Contact time, flight time and step frequency	10/12
Gouttebarge et al. (2015) [83]	11 male and 3 female recreational runners (45 ± 14 y, 77 ± 11 kg, 181 ± 7 cm)	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Triaxial accelerometer sampled at 200-500 Hz	Lumbar spine	Foot-mounted accelerometer (± 6 g, MMA7361L, Freescale Semiconductor, Austin, Texas, USA) sampled at 1000 Hz	3 x 400 m runs at 2.8, 3.3 and 3.9 $m{\cdot}s^{-1}$	Contact time and step frequency	10/12
Gurchiek et al. (2017) [57]	12 male and 3 female subjects (23.2 ± 2.1 y, 75.5 ± 12.6 kg, 1.8 ± 0.1 m)	IMU (Yost Data Logger 3-Space Sensor, YEI Technology, Portsmouth, OH)	Triaxial accelerometer (± 24 g) and triaxial gyroscope (± 2000°s^{-1}) sampled at 450 Hz	Sacrum	Force plate (AMTI, Watertown, MA, USA) sampled at 1000 Hz	Six linear standing sprint starts	Vertical and resultant ground reaction force	10/12
Kenneally-Dabrowski et al. (2018) [54]	13 male professional rugby union athletes (23.8 ± 2.4 y, 102.5 ± 12.2 kg, 186.6 ± 8.4 cm)	Accelerometer (GPSports, Canberra, Australia)	Triaxial accelerometer (± 16 g) sampled at 100 Hz	Thoracic spine	Eight 600 x 900 mm force plates (Kistler, Amherst, MA, USA) sampled at 1000 Hz	3 x 40 m maximal sprints (8.64 ± 0.5)	Stride time	10/12
Lee et al. (2010) [32]	6 male and 4 female national standard runners (30.3 ± 7.9 y, 67.7 ± 9.5 kg, 174.3 ± 5.7 cm)	Accelerometer (KXM52 – 1050 Kionix, NY, USA)	Triaxial accelerometer sampled at 100 Hz	Sacrum (S1)	Optical motion capture system (Proreflex MCU, Qualisys Medical AB, Gothenburg, Sweden) sampled at 500 Hz	3 x 5 min runs at 2.8-3.3 $m{\cdot}s^{-1}$, 3.6-4.2 $m{\cdot}s^{-1}$ and 4.4-5.3 $m{\cdot}s^{-1}$	Contact time, step time and stride time	9/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Machulik et al. (2020) [72]	18 male and 10 female runners (28.2 ± 3.8 y, 70.6 ± 10.7 kg, 175.5 ± 9.5 cm)	IMU (Humotion SmarTracks Integrated)	Triaxial accelerometer (± 16 g, 400 Hz), triaxial gyroscope (± 2000°s^{-1}, 400 Hz) and triaxial magnetometer (100 Hz)	Lumbar spine	Optojump Next® (Microgate, Bolzano, Italy) sampled at 1000 Hz	3 x 60 m runs jogging (3.8 ± 0.7 m·s^{-1}) and sprinting 6.8 ± 1.0 m·s^{-1})	Step frequency and step length	9/12
McGrath et al. (2012) [29]	4 male and 1 female healthy adults (26-32 y)	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Triaxial gyroscope sampled at 102.4 Hz	Tibia	Cartesian Optoelectronic Dynamic Anthropometer (CODA) motion analysis system (Charnwood Dynamics Ltd, Leicestershire, UK) sampled at 200 Hz	2 x 20 s runs at 2.2 m·s^{-1} and 3.3 m·s^{-1}	Initial contact, toe-off, contact time, swing time and stride time	9/12
Mitschke et al. (2017) [59]	21 male subjects (28.9 ± 10.8 y, 74.4 ± 7.1 kg, 177.0 ± 5.2 cm)	Accelerometer (ADXL278, Analog Devices Inc., Norwood, MA, USA)	Biaxial accelerometer (± 687 m·s^{-2}) sampled at 1000 Hz	Heel	One 60 x 90 cm force platform (Kistler, 9287 BA) sampled at 1000 Hz	30 x 15 m runs at 3.5 ± 0.1 m·s^{-1}	Initial contact	9/12
Mitschke et al. (2017) [58]	12 recreational rearfoot strike runners (24.8 ± 4.5 y, 72.3 ± 7.8 kg, 176.0 ± 5.4 cm) and 11 recreational forefoot strike runners (26.3 ± 3.2 y, 74.5 ± 7.5 kg, 177.0 ± 3.6 cm)	IMU (ICM-20601, InvenSense, San Jose, CA, USA)	Triaxial accelerometer (± 353 m·s^{-2}), triaxial gyroscope (± 4000°s^{-1}) sampled at 3570 Hz	Tibia (medial aspect)	One 0.6 x 0.9 m force plate (9287 BA, Kistler, Winterthur, Switzerland) sampled at 3570 Hz	5 x 15 m runs at a self-selected velocity (3.26 ± 0.4 m·s^{-1})	Initial contact	9/12
Mo & Chow (2018) [69]	7 male and 4 female healthy adults (25.5 ± 4.2 y, 58.8 ± 5.3 kg, 168.3 ± 9.1 cm)	IMU (MyoMOTION MR3, Noraxon, USA)	Triaxial accelerometer (± 16 g) sampled at 200 Hz	Dorsal foot, tibia and lumbar spine (L5-S1)	Three force platforms (Bertec, FP4060-07, USA) sampled at 2000 Hz	10 x 10 m runs at 3.1 ± 0.1 m·s^{-1} and 4.1 ± 1.2 m·s^{-1}	Initial contact, toe-off and contact time	10/12
Nedergaard et al. (2018) [39]	20 healthy male athletes (22 ± 4 y, 76 ± 11 kg, 178 ± 8 cm)	Accelerometer (MinimaxX S4, Catapult Innovations, Scoresby, Australia)	Triaxial accelerometer (± 13 g) sampled at 100 Hz	Thoracic spine	One 0.9 x 0.6 m^2 Kistler force platform (9287C, Kistler Instruments Ltd., Winterthur, Switzerland) sampled at 3000 Hz	Four runs over a force platform at 2, 3, 4 and 5 m·s^{-1}	Resultant ground reaction force	9/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Neugebauer et al. (2014) [66]	19 male and 20 female subjects (21.2 ± 1.3 y, 67.6 ± 11.5 kg, 1.73 ± 0.12 m)	Accelerometer (GT3X+ AM, ActiGraph, Pensacola, FL, USA)	Triaxial accelerometer (± 6 g) sampled at 100 Hz	Hip	Force plate (Kistler Corporation, Model 9281B, Amherst, NY, USA) sampled at 1000 Hz	8-10 x 15 m runs ranging between 2.2 $m{\cdot}s^{-1}$ and 4.1 $m{\cdot}s^{-1}$ (0.3 $m{\cdot}s^{-1}$ increments)	Vertical ground reaction force	9/12
Ngoh et al. (2018) [56]	7 healthy male subjects (21.3 ± 0.5 y, 63 ± 6.1 kg, 174.9 ± 6.6 cm)	IMU (Opal, APDM Inc.)	Triaxial accelerometer (± 6 g), triaxial gyroscope (± 200°s^{-1}) and triaxial magnetometer (± 6 Gauss)	Dorsal foot	Instrumented treadmill (Mercury, H/P Cosmos Sports and Medical GmbH)	1 min runs at 2.2, 2.5 and 2.8 $m{\cdot}s^{-1}$	Vertical ground reaction force	9/12
Norris et al. (2016) [84]	1 male and 5 female recreational runners (33.5 ± 5.8 y, 71.1 ± 12.2 kg, 1.66 ± 0.08 m)	Accelerometer (Shimmer 2r, Shimmer Inc., Dublin, Ireland)	Triaxial accelerometer (± 16 g) sampled at 204.8 Hz	Distal tibia	Four varying accelerometer-derived stride time calculation methods	Running at a self-selected velocity during a half-marathon training programme	Stride time	9/12
Pairot de Fontenay et al. (2020) [60]	19 male and 13 female healthy adults (27.0 ± 5.5 y, 69.1 ± 11.4 kg, 174.4 ± 8.5 cm)	IMU (MilestonePod [Milestone Sports, Long Beach, CA, USA]; Zoi [Runteq, Tampere, Finland]; RunScribe™ [Montara, CA, USA]; Moov Now™ [Moov, San Mateo, CA, USA]; TgForce, Kelsec Systems Inc., Montreal, Canada])	Not reported	Dorsal foot and distal tibia	Instrumented treadmill (Bertec, Columbus, OH, USA) sampled at 1000 Hz	Not reported	Step frequency	10/12
Pogson et al. (2020) [61]	10 male and 5 female team-sport athletes (23 ± 1 y, 74 ± 9 kg, 1.74 ± 0.08 m)	Accelerometer (MinimaxX S5, Catapult Innovations, Scoresby, Australia)	Triaxial accelerometer (± 16 g) sampled at 100 Hz	Thoracic spine	In-ground force platform (9287B, Kistler Holding AG, Winterthur, Switzerland) sampled at 3000 Hz	Straight overground accelerated, decelerated and constant speed running between 2 $m{\cdot}s^{-1}$ and 8 $m{\cdot}s^{-1}$ (1 $m{\cdot}s^{-1}$ increments)	Resultant ground reaction force	9/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Raper et al. (2018) [40]	4 male (27.00 ± 5.48 y, 66.00 ± 5.29 kg, 177.50 ± 4.65 cm) and 6 female (26.83 ± 3.06 y, 54.83 ± 3.19 kg, 164.50 ± 2.88 cm) professional triathletes	Accelerometer (ViPerform v5, DorsaVi, Melbourne, Australia)	Triaxial accelerometer	Tibia (medial border)	Eight piezoelectric force plates (Kistler Instrument Group, Amherst, New York, United States of America) sampled at 1000 Hz	10 x 50 m runs at 5.2 ± 0.6 $m{\cdot}s^{-1}$	Vertical ground reaction force	9/12
Sinclair et al. (2013) [62]	11 male and 5 female healthy adults (29.4 ± 5.7 y, 67.8 ± 10.7 kg, 1.73 ± 4.87 m)	Accelerometer (Biometrics ACL 300, UK)	Triaxial accelerometer sampled at 1000 Hz	Distal tibia	Force platform (Kistler Ltd; Model 9281CA, Kistler Instruments Ltd., Alton, Hampshire) sampled at 1000 Hz	10 runs at 4 $m{\cdot}s^{-1}$	Initial contact, toe-off and contact time	10/12
Tan et al. (2019) [63]	20 healthy subjects	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Triaxial accelerometer (± 8 g) sampled at 128 Hz	Distal tibia	In-shoe piezo-electric force sensitive resistors	Treadmill running overground running and outdoor running	Initial contact and toe-off	7/12
Watari et al. (2016) [53]	14 male and 8 female semi-elite runners (28.2 ± 10.1 y, 65.4 ± 8.1 kg, 1.73 ± 0.75 m)	Accelerometer (Forerunner 620, Garmin International Inc., Olathe, KS)	Not reported	Torso (near xiphoid process of the sternum)	Instrumented treadmill (Bertec, Columbus, OH) sampled at 1000 Hz	60 s runs at 2.7, 3.0, 3.3, 3.6 and 3.9 $m{\cdot}s^{-1}$	Contact time	9/12
Winter et al. (2016) [64]	6 male and 4 female recreational runners (27.5 ± 9.5 y, 69.5 ± 11.8 kg, 175.8 ± 8.1 cm)	Accelerometer (ADXL202, Analog Devices Inc., Norwood, MA, USA)	Triaxial accelerometer (± 8 g) sampled at 100 Hz	Lumbar spine	12-camera motion analysis system (NEXUS v1.8, Vicon Motion Systems Ltd. UK) sampled at 100 Hz	5 x 50 m runs overground	Initial contact and toe-off	9/12
Wouda et al. (2018) [87]	8 experienced male runners (25.1 ± 5.2 y, 77.7 ± 9.4 kg, 183.7 ± 4.5 cm)	IMU (Xsens, Enschede, the Netherlands)	Triaxial accelerometer, triaxial gyroscope and triaxial magnetometer sampled at 240 Hz	Lower legs and pelvis	S-Mill instrumented treadmill (ForceLink, Culemborg, the Netherlands) sampled at 1000 Hz	3 min runs at 2.8, 3.3 and 3.9 $m{\cdot}s^{-1}$	Vertical ground reaction force	11/12
Wundersitz et al. (2013) [36]	12 male and 5 female team-sport athletes (21 ± 2 y, 78.2 ± 11.6 kg, 1.82 ± 0.08 m)	Accelerometer (SPI Pro, ASP00725, GPSports, Canberra, Australia)	Triaxial accelerometer (± 8 g) sampled at 100 Hz	Thoracic spine (T2)	In-ground force plate (BP600900, Advanced Mechanical Technology Inc., Watertown, MA, USA) sampled at 100 Hz	5 x 10 m runs in a straight-line (5.4 ± 0.4 $m{\cdot}s^{-1}$) and 5 x 10 m angled runs at 45º (4.8 ± 0.4 $m{\cdot}s^{-1}$), 90º (4.1 ± 0.3 $m{\cdot}s^{-1}$) and 180º (3.5 ± 0.3 $m{\cdot}s^{-1}$)	Vertical and resultant ground reaction force	10/12

Table 2 Study characteristics for the 39 studies included in the review (continued).

Study	Sample (age, mass, height)	Sensor	Hardware	Sensor placement	Criterion	Activity	Variable(s)	Methodological quality
Zrenner et al. (2018) [65]	21 male and 6 female amateur runners (24.9 ± 2.4 y, 178. 6 ± 8.0 cm)	IMU (miPod sensor)	Triaxial accelerometer (± 16 g) and triaxial gyroscope (± 2000°s^{-1}) sampled at 200 Hz	Foot (midsole)	Motion capture system (Vicon Motion Systems Inc., Oxford, UK) sampled at 200 Hz	10 runs at 2-3 m·s^{-1}, 10 runs at 3-4 m·s^{-1}, 15 runs at 4-5 m·s^{-1}, 15 runs at 5-6 m·s^{-1}	Stride length	9/12

Abbreviations: °s^{-1}, degrees per second; cm, centimetres; g, gravitational acceleration; Hz, hertz; IMU, inertial measurement unit; kg, kilograms; m, metres; m^2, metres squared; m·s^{-1}, metres per second; mm, millimetres; s, seconds; y, years.

Table 3 Validity summary statistics for initial contact.

Study	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Statistic 1	Statistic 2
Chew et al. (2018) [52]	IMU (Opal, APDM Inc.)	Motion capture system	Foot		ME ± SD (ms)	RMSE (ms)
				2.2	-2.6 ± 12.8	4.7
				2.5	-6.0 ± 14.1	5.3
				2.8	4.3 ± 17.9	8.3
				3.1	3.0 ± 14.1	4.7
Falbriard et al. (2018) [34]	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot		Median bias ± IQR (ms)	Median precision ± IQR (ms)
				2.8-5.6	11.0 ± 10.0	2.0 ± 1.0
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Foot		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1	-7.3 ± 3.3	5.2 ± 3.4
				4.1 ± 1.2	3.3 ± 4.7	4.2 ± 4.7
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Foot		MD (ms)	95% LoA (ms)
				3.3	-16.0	-58.0, 27.0
Mitschke et al. (2017) [59]	Accelerometer (ADXL278, Analog Devices Inc., Norwood, MA, USA)	Force plate	Heel		MD (ms)	
				3.5 ± 0.1	0.7 ± 2.6	
Sinclair et al. (2013) [62]	Accelerometer (Biometrics ACL 300, UK)	Force plate	Tibia		ME (95% CI) (ms)	AE (95% CI) (ms)
				4.0	1.7 (-2.9, 6.3)	5.5 (1.9, 9.0)
Tan et al. (2019) [63]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	In-shoe piezo-electric force sensitive resistors	Tibia		F_1 score	
				Not reported	0.92-0.96	
McGrath et al. (2012) [29]	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Hreljac and Marshal (2000)	Tibia		True error (ms)	% error
				2.2	33.4	0.8
				3.3	24.1	0.5
	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Zeni et al., (2008)	Tibia		True error (ms)	% error
				2.2	64.2	1.5
				3.3	61.7	1.4
Mitschke et al. (2017) [58]	IMU (ICM-20601, InvenSense, San Jose, CA, UDA) Sinclair et al., (2013)	Force plate	Tibia		MD (ms)	
				3.26 ± 0.4	11.5 ± 4.2	
	IMU (ICM-20601, InvenSense, San Jose, CA, UDA) Mercer et al., (2003)	Force plate	Tibia		MD (ms)	
				3.26 ± 0.4	-1.1 ± 10.7	
	IMU (ICM-20601, InvenSense, San Jose, CA, UDA) Maiwald et al., (2015)	Force plate	Tibia		MD (ms)	
				3.26 ± 0.4	-0.5 ± 0.3	
	IMU (ICM-20601, InvenSense, San Jose, CA, UDA) Sabatini et al., (2005)	Force plate	Tibia		MD (ms)	
				3.26 ± 0.4	-5.1 ± 3.0	
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Tibia		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1	-38.0 ± 10.7	19.5 ± 6.5
				4.1 ± 1.2	-16.7 ± 11.9	17.4 ± 11.0
Aubol & Milner (2020) [31]	Accelerometer (Model 356A45, PCB Piezotronics, Depew, NY)	Force plate	Tibia		Mean bias (ms)	95% LoA (ms)
				3.0 ± 0.2	-2.3 ± 4.7	-6.8, 11.5
Fadillioglu et al. (2020) [70]	Gyroscope (ADXRS652, Analog Devices Inc., Norwood, MA, USA)	Force plate	Tibia		AME ± SD (ms)	RAME ± SD (%)
				Moderate	10.0 ± 4.0	3.4 ± 1.4
				Fast	13.0 ± 6.0	5.5 ± 2.7

Table 3 Validity summary statistics for initial contact (continued).

Study	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Statistic 1	Statistic 2
Winter et al. (2016) [64]	Accelerometer (ADXL202, Analog Devices Inc., Norwood, MA, USA)	Motion capture system	Lumbar spine		TEE (ms)	Pearson's *r*
				Self-paced	0.8	0.99
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Lumbar spine		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1	-2.6 ± 4.9	9.0 ± 2.0
				4.1 ± 1.2	5.7 ± 5.0	6.2 ± 4.6
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Lumbar spine		MD (ms)	95% LoA (ms)
				3.3	53.0	24.0, 82.0

Abbreviations: AE, absolute error; AME, absolute mean error; CI, confidence interval; F_1 score, weighted average of precision and recall; IMU, inertial measurement unit; IQR, inter-quartile range; LoA, limits of agreement; MAD, mean absolute difference; MD, mean difference; ME, mean error; MRD, mean relative difference; ms, milliseconds; m·s^{-1}, metres per second; RAME, relative absolute mean error; RMSE, root mean square error; SD, standard deviation; TEE, typical error of the estimate. Negative values represent a time lead in the detection of initial contact by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Falbriard et al. (2018) as validity outcomes were reported from pooled speeds.
Values converted to milliseconds where required.

Table 4 Validity summary statistics for toe-off.

Study	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1} \pm SD$	Statistic 1	Statistic 2
Chew et al. (2018) [52]	IMU (Opal, APDM Inc.)	Motion capture system	Foot		ME ± SD (ms)	RMSE (ms)
				2.2 2.5 2.8 3.1	3.3 ± 20.9 16.3 ± 16.7 -4.3 ± 15.0 2.6 ± 19.5	9.0 11.1 7.6 11.0
Falbriard et al. (2018) [34]	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot		Median bias ± IQR (ms)	Median precision ± IQR (ms)
				2.8-5.6	-4.0 ± 7.0	4.0 ± 2.0
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Foot		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1 4.1 ± 1.2	-32.0 ± 14.1 -53.8 ± 8.1	25.0 ± 7.5 27.6 ± 7.6
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Foot		MD (ms)	95% LoA (ms)
				3.3	32.0	-84.0, 148.0
Sinclair et al. (2013) [62]	Accelerometer (Biometrics ACL 300, UK)	Force plate	Tibia		ME (95% CI) (ms)	AE (95% CI) (ms)
				4.0	-3.6 (-5.4, 1.8)	5.0 (3.5, 8.5)
Tan et al. (2019) [63]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	In-shoe piezo-electric force sensitive resistors	Tibia		F_1 score	
				Not reported	0.77-0.81	
McGrath et al. (2012) [29]	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Hreljac and Marshal (2000)	Tibia		True error (ms)	% error
				2.2 3.3	-32.4 -28.8	0.7 0.8
	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Zeni et al., (2008)	Tibia		True error (ms)	% error
				2.2 3.3	-15.1 -24.2	0.7 0.7
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Tibia		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1 4.1 ± 1.2	0.0 ± 4.1 1.0 ± 7.8	5.1 ± 2.1 8.8 ± 3.7
Fadillioglu et al. (2020) [70]	Gyroscope (ADXRS652, Analog Devices Inc., Norwood, MA, USA)	Force plate	Tibia		AME ± SD (ms)	RAME ± SD (%)
				Moderate Fast	26.0 ± 20.0 23.0 ± 23.0	8.0 ± 4.8 9.4 ± 8.8
Winter et al. (2016) [64]	Accelerometer (ADXL202, Analog Devices Inc., Norwood, MA, USA)	Motion capture system	Lumbar spine		TEE (ms)	Pearson's *r*
				Self-paced	0.8	0.99
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Lumbar spine		MRD ± SD (ms)	MAD ± SD (ms)
				3.1 ± 0.1 4.1 ± 1.2	7.6 ± 9.9 9.4 ± 12.7	15.2 ± 5.0 20.3 ± 8.2
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Lumbar spine		MD (ms)	95% LoA (ms)
				3.3	24.0	-15.0, 63.0

Abbreviations: AE, absolute error; AME, absolute mean error; CI, confidence interval; F_1 score, weighted average of precision and recall; IMU, inertial measurement unit; IQR, inter-quartile range; LoA, limits of agreement; MAD, mean absolute difference; MD, mean difference; ME, mean error; MRD, mean relative difference; ms, milliseconds; $m \cdot s^{-1}$, metres per second; RAME, relative absolute mean error; RMSE, root mean square error; SD, standard deviation; TEE, typical error of the estimate. Negative values represent a time lead in the detection of toe-off by the IMU compared to the criterion.

Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Falbriard et al. (2018) as validity outcomes were reported from pooled speeds.

Values converted to milliseconds where required.

Table 5 Validity summary statistics for contact time.

Study	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Ammann et al. (2016) [16]	IMU (PARTwear, HuCE-microLab, University of Applied Sciences, Biel, Switzerland)	High-speed camera	Foot				ICC (95% CI)	Systematic bias (ms)	%D		
				4.3 ± 0.7	185.5 ± 21.7	194.6 ± 34.3	0.97 (0.92, 0.99)	-5.6*	-3.3 ± 5.0		
				6.2 ± 0.7	145.5 ± 20.9	147.4 ± 20.3	0.96 (0.92, 0.98)	-0.7	-0.8 ± 6.2		
				8.0 ± 1.6	118.3 ± 11.6	117.5 ± 9.0	0.81 (0.65, 0.89)	0.4	-0.1 ± 6.7		
Chew et al. (2018) [52]	IMU (Opal, APDM Inc.)	Motion capture system	Foot				ME ± SD (ms)	RMSE (ms)			
				2.2	Not reported	Not reported	-6.1 ± 6.2	7.8			
				2.5			-8.2 ± 5.3	9.2			
				2.8			-8.1 ± 3.1	9.1			
				3.1			-8.1 ± 2.5	10.0			
Falbriard et al. (2018) [34]	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot				Median bias ± IQR (ms)	Median precision ± IQR (ms)			
				2.8-5.6	Not reported	Not reported	-15.0 ± 12.0	5.0 ± 3.0			
Garcia-Pinillos et al. (2018) [71]	IMU (Stryd^{TM}, Stryd Powermeter, Stryd Inc., Boulder, CO, USA)	Photocell system	Foot				ICC	Pearson's *r*			
				2.2-5.6	311.5 ± 11.5 to 175.5 ± 3.3	340.0 ± 28.0 to 175.0 ± 6.0	0.06-0.46	0.08-0.66**			
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Foot				MRD ± SD (ms)	Pearson's *r*	%D	MAD ± SD (ms)	
				3.1 ± 0.1	228.0 ± 23.0	253.0 ± 10.0	-24.7 ± 14.8	0.88*	11.9 ± 4.8	29.3 ± 11.5	
				4.1 ± 1.2	159.0 ± 13.0	215.0 ± 7.0	-56.0 ± 9.6	0.74*	26.6 ± 4.3	34.2 ± 10.4	
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Foot				MD (ms)	95% LoA (ms)			
				3.3	320.1 ± 41.5	270.6 ± 25.4	47.0	-59.0, 154.0			
Garcia-Pinillos et al. (2019) [35]	IMU (RunScribe^{TM}, Scribe Lab. Inc. San Francisco CA, USA)	High-speed camera	Foot				ICC (95% CI)	Pearson's *r*	MD (%)	MD (ms)	Systematic bias ± RE (ms)
				3.3 ± 0.4	261.0 ± 28.0	267.0 ± 28.0	0.90 (0.80, 0.94)	0.83	2.3**	-6.0	-6.0 ± 16.0
	IMU (Stryd^{TM}, Stryd Powermeter, Stryd Inc. Boulder CO, USA)	High-speed camera	Foot				ICC (95% CI)	Pearson's *r*	MD (%)	MD (ms)	Systematic bias ± RE (ms)
				3.3 ± 0.4	253.0 ± 22.0	267.0 ± 28.0	0.81 (0.29, 0.93)	0.82	5.2***	-14.0	-15.0 ± 16.0
Sinclair et al. (2013) [62]	Accelerometer (Biometrics ACL 300, UK)	Force plate	Tibia				ME (95% CI) (ms)	Pearson's *r*	AE (95% CI) (ms)		
				4.0	185.30	190.46	-5.2 (0.9, 10.2)	0.96	11.5 (8.1, 14.9)		
McGrath et al. (2012) [29]	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system [Hreljac and Marshal (2000)]	Tibia				ICC	True error (ms)	% error		
				2.2	390.0 ± 30.0	440.0 ± 20.0	0.32	-66.4	15.2		
				3.3	450.0 ± 51.0	390.0 ± 60.0	0.30	-63.4	16.7		

Table 5 Validity summary statistics for contact time (continued).

Study	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1}$ ± SD	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
McGrath et al. (2012) [29]	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system [Zeni et al., (2008)]	Tibia				ICC	True error (ms)	% error		
				2.2	390.0 ± 30.0	460.0 ± 10.0	0.26	-79.1	19.8		
				3.3	450.0 ± 51.0	420.0 ± 10.0	0.29	-90.2	22.4		
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Tibia				MRD ± SD (ms)	Pearson's *r*	%D	MAD ± SD (ms)	
				3.1 ± 0.1	291.0 ± 15.0	253.0 ± 10.0	38.0 ± 9.4	0.78*	15.6 ± 3.0	39.4 ± 8.0	
				4.1 ± 1.2	248.0 ± 39.0	215.0 ± 7.0	32.9 ± 34.1	0.74*	17.3 ± 14.1	30.9 ± 18.9	
Fadillioglu et al. (2020) [70]	Gyroscope (ADXRS652, Analog Devices Inc., Norwood, MA, USA)	Force plate	Tibia				MD (ms)	95% LoA (ms)			
				2.6 ± 0.4	297.0 ± 43.0	309.0 ± 45.0	-12.0	-83.0, 59.0			
				3.9 ± 0.6	225.0 ± 42.0	232.0 ± 26.0	-7.0	-85.0, 71.0			
Lee et al. (2010) [32]	Accelerometer (KXM52 – 1050 Kionix, NY, USA)	Motion capture system	Sacrum				Mean bias (ms)	Pearson's *r*	SE (ms)	95% LoA (ms)	
				2.8-3.3	Not reported	Not reported	1.1	0.91	0.9	-25.0, 22.0	
				3.6-4.2			2.2	0.94	0.7	-20.0, 16.0	
				4.4-5.2			0.8	0.90	0.9	-24.0, 23.0	
Bergamini et al. (2012) [33]	IMU (FreeSense, Sensorize, Italy)	Force plate	Lumbar spine				ME (ms)	LoA (ms)			
				Maximal sprint	122.9 ± 10.9	123.3 ± 13.1	5.0	25.0			
	IMU (FreeSense, Sensorize, Italy)	High-speed camera	Lumbar spine				ME (ms)	LoA (ms)			
				Maximal sprint	105.2 ± 4.5	103.6 ± 7.7	5.0	25.0			
Gouttebarge et al. (2015) [83]	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Foot-mounted accelerometer	Lumbar spine				ICC (95% CI)				
				2.8	172.0 ± 15.0	297.1 ± 20.0	0.49 (-0.03, 0.80)				
				3.3	159.1 ± 17.0	278.4 ± 25.0	0.50 (-0.02, 081)				
				3.9	144.2 ± 16.0	251.3 ± 24.0	0.48 (-0.07, 0.81)				
Gindre et al. (2016) [68]	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Photocell system	Lumbar spine				ICC	CV%	MD (%)		
				3.3	166.0 ± 15.0	268.0 ± 17.0	0.63	17.7	-38.0*		
				4.2	154.0 ± 15.0	237.0 ± 15.0	0.67	16.2	-35.0*		
				5.0	135.0 ± 16.0	208.0 ± 13.0	0.75	16.4	-35.0*		
				5.8	116.0 ± 16.0	182.0 ± 16.0	0.82	17.9	-36.0*		
	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	High-speed camera	Lumbar spine				ICC	CV%	MD (%)		
				3.3	166.0 ± 15.0	252.0 ± 17.0	0.72	15.9	-34.0*		
				4.2	154.0 ± 15.0	223.0 ± 13.0	0.47	14.1	-31.0*		
				5.0	135.0 ± 16.0	198.0 ± 12.0	0.63	14.7	-32.0*		
				5.8	116.0 ± 16.0	173.0 ± 12.0	0.74	16.0	-33.0*		

Table 5 Validity summary statistics for contact time (continued).

Study	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1} \pm SD$	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Mo & Chow (2018) [69]	IMU (MyoMOTION MR3, Noraxon, USA)	Force plate	Lumbar spine				MRD ± SD (ms)	Pearson's *r*	%D	MAD ± SD (ms)	
				3.1 ± 0.1	263.0 ± 15.0	253.0 ± 10.0	10.3 ± 8.9	0.83*	6.3 ± 1.8	15.9 ± 4.7	
				4.1 ± 1.2	220.0 ± 18.0	215.0 ± 7.0	4.6 ± 12.1	0.89*	8.7 ± 3.7	18.7 ± 7.5	
Benson et al. (2019) [30]	Accelerometer (Shimmer3®, Shimmer Inc., Dublin, Ireland)	Force plate	Lumbar spine				MD (ms)	95% LoA (ms)			
				3.3	241.8 ± 30.2	270.6 ± 25.4	-29.0	-69.0, 10.0			
Watari et al. (2016) [53]	Accelerometer (Forerunner 620, Garmin International Inc., Olathe, KS)	Instrumented treadmill	Torso				Mean bias (ms)	CCC			
				2.7	Not reported	Not reported	-17.0	0.69			
				3.0			-10.1	0.77			
				3.3			-5.8	0.87			
				3.6			-2.6	0.83			
				3.9			-1.4	0.84			
Buchheit et al. (2015) [37]	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Instrumented treadmill	Scapula				Mean bias (90% CI) (%)	Pearson's *r* (90% CI)	CV% (90% CI)		
				2.8-7.5	Not reported	Not reported	-10.4 (-12.3, -9.8)	0.98 (0.97, 0.99)	3.9 (3.4, 4.6)		

Abbreviations: %D, percentage difference; AE, absolute error; CCC, concordance correlation coefficient; CI, confidence interval; CV, coefficient of variation; ICC, intraclass correlation coefficient; IMU, inertial measurement unit; IQR, inter-quartile range; LoA, limits of agreement; MAD, mean absolute difference; MD, mean difference; ME, mean error; MRD, mean relative difference; ms, milliseconds; $m \cdot s^{-1}$, metres per second; RE, random error; RMSE, root mean square error; SD, standard deviation; SE, standard error.
Negative values represent an underestimation of contact time calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Falbriard et al. (2018) and Buchheit et al. (2015) as validity outcomes were reported from pooled speeds.
Values converted to milliseconds where required.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

Table 6 Reliability summary statistics for each analysed stride variable.

Study	Variable	Site	Running velocity m·s^{-1} ± SD	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Ammann et al. (2016) [16]	Contact time	Foot		CV%	ICC			
			4.3-8.0	2.9-3.8	0.91-0.96			
Garcia-Pinillos et al. (2018) [71]	Contact time	Foot		CV%	SEM (ms)			
			2.2-5.6	1.2-2.3	1.0-5.0			
Gouttebarge et al. (2015) [83]	Contact time	Lumbar spine		SEM (ms)	ICC (95% CI)			
			2.8		-0.24 (-0.69, 0.32)			
			3.3	14.8	0.35 (-0.23, 0.74)			
			3.9	10.1	0.67 (0.22, 0.88)			
Gindre et al. (2016) [68]	Contact time	Lumbar spine		CV%	ICC			
			3.3	6.5	0.99			
			4.2	6.7	0.88			
			5.0	8.3	0.95			
			5.8	9.9	0.97			
Garcia-Pinillos et al. (2018) [71]	Flight time	Foot		CV%	SEM (ms)			
			2.2-5.6	3.7-11.6	3.0-8.0			
Gindre et al. (2016) [68]	Flight time	Lumbar spine		CV%	ICC			
			3.3	4.6	0.94			
			4.2	4.8	0.95			
			5.0	5.2	0.98			
			5.8	5.2	0.98			
Mitschke et al. (2017) [59]	Stride time	Heel		CV%				
			3.5 ± 0.1	2.6-3.5				
Garcia-Pinillos et al. (2018) [71]	Step frequency	Foot		CV%	SEM (step·min^{-1})			
			2.2-5.6	1.1-2.0	1.7-2.8			
Gouttebarge et al. (2015) [83]	Step frequency	Lumbar spine		SEM (step·min^{-1})	ICC (95% CI)			
			2.8	3.5	0.82 (0.52, 0.94)			
			3.3	4.1	0.78 (0.44, 0.92)			
			3.9	3.0	0.92 (0.77, 0.97)			
Gindre et al. (2016) [68]	Step frequency	Lumbar spine		CV%	ICC			
			3.3	4.4	0.94			
			4.2	3.9	0.82			
			5.0	4.1	0.84			
			5.8	4.1	0.94			
Garcia-Pinillos et al. (2018) [71]	Step length	Foot		CV%	SEM (cm)			
			2.2-5.6	1.1-2.1	107.1-241.2			
Mitschke et al. (2017) [59]	Stride length	Heel		CV%				
			3.5 ± 0.1	4.9-7.8				
Raper et al. (2018) [40]	vGRF	Tibia		SEM (95% CI) (N)	ICC (95% CI)	SEM (95% CI) (%)	MDC (95% CI) (N)	MDC (95% CI) (%)
			5.2 ± 0.6	99.8 (82.1, 119.1)	0.88 (0.83, 0.92)	7.0 (5.7, 8.3)	276.7 (227.3, 330.1)	19.3 (15.9, 23.0)
Eggers et al. (2018) [38]	vGRF	Lumbar spine		CV% (90% CI)	ICC (90% CI)	TEE (90% CI) (N)		
			3.3	4.2 (3.3, 6.1)	0.93 (0.84, 0.97)	0.3 (0.2, 0.4)		
Buchheit et al. (2018) [86]	vGRF	Thoracic spine		CV% (90% CI)	ICC (90% CI)	TE (90% CI) (N)	SWC (%)	
			6.1-6.7	17.1 (13.6, 25.1)	0.47 (0.12, 0.72)	0.8 (0.6, 1.1)	5.0	
Eggers et al. (2018) [38]	vGRF	Thoracic spine		CV% (90% CI)	ICC (90% CI)	TEE (90% CI) (N)		
			3.3	3.3 (2.5, 4.7)	0.95 (0.89, 0.98)	0.3 (0.2, 0.4)		
Edwards et al. (2019) [43]	vGRF	Thoracic spine		CV%	ICC	TE (N)		
			3.3	17.8	0.47	2.6		
			5.0	18.6	0.50	2.9		
			6.7	21.8	0.31	2.9		
Eggers et al. (2018) [38]	Vertical stiffness	Lumbar spine		CV% (90% CI)	ICC (90% CI)	TEE (90% CI) (kN·m^{-1})		
			3.3	12.1 (9.3, 17.6)	0.70 (0.41, 0.86)	0.7 (0.5, 1.0)		

Table 6 Reliability summary statistics for each analysed stride variable.

Study	Variable	Site	Running velocity m·s^{-1} ± SD	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Buchheit et al. (2018) [86]	Vertical stiffness	Thoracic spine		CV% (90% CI)	ICC (90% CI)	TE (90% CI) (kN·m^{-1})	SWC (%)	
			6.1-6.7	11.0 (8.6, 15.6)	0.75 (0.52, 0.88)	0.5 (0.7, 1.2)	4.0	
Eggers et al. (2018) [38]	Vertical stiffness	Thoracic spine		CV% (90% CI)	ICC (90% CI)	TEE (90% CI) (kN·m^{-1})		
			3.3	9.5 (7.3, 13.7)	0.71 (0.44, 0.87)	0.7 (0.5, 1.0)		

Abbreviations: CI, confidence interval; cm, centimetres; CV, coefficient of variation; ICC, intraclass correlation coefficient; kN·m^{-1}; kilo Newtons per metre; MDC, minimal detectable change; ms, milliseconds; m·s^{-1}, metres per second; N, Newtons; SD, standard deviation; SEM, standard error of measurement; step·min^{-1}; steps per minute; SWC, smallest worthwhile change; TE, typical error; TEE, typical error of the estimate. Running velocity reported with or without ± SD, depending on the method used in each study. Values presented for Ammann et al. (2016) and Garcia-Pinillos et al. (2018) represents reliability assessed at a range of speeds.
Values converted to milliseconds, centimetres or Newtons where required.

Table 7 Validity summary statistics for flight time.

Study	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Chew et al. (2018) [52]	IMU (Opal, APDM Inc.)	Motion capture system	Foot				ME ± SD (ms)	RMSE (ms)			
				2.2	Not reported	Not reported	6.1 ± 6.2	7.8			
				2.5			8.2 ± 5.3	9.2			
				2.8			8.1 ± 3.1	10.1			
				3.1			8.1 ± 2.5	10.0			
Falbriard et al. (2018) [34]	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot				Median bias ± IQR (ms)	Median precision ± IQR (ms)			
				2.8-5.6	Not reported	Not reported	15.0 ± 12.0	5.0 ± 3.0			
Garcia-Pinillos et al. (2018) [71]	IMU (Stryd™, Stryd Powermeter, Stryd Inc., Boulder, CO, USA)	Photocell system	Foot				ICC	Pearson's *r*			
				2.2-5.6	62.0 ± 16.9 to 137.6 ± 6.5	36.5 ± 25.4 to 133.7 ± 8.4	0.56-0.81	0.60^{**}-0.83^{*}			
Garcia-Pinillos et al. (2019) [35]	IMU (RunScribe™, Scribe Lab. Inc. San Francisco CA, USA)	High-speed camera	Foot				ICC (95% CI)	MD (%)	Pearson's *r*	MD (ms)	Systematic bias ± RE (ms)
				3.3 ± 0.4	96.0 ± 26.0	93.0 ± 25.0	0.86 (0.75, 0.92)	3.2	0.75^{***}	3.0	3.0 ± 17.0
	IMU (Stryd™, Stryd Powermeter, Stryd Inc. Boulder CO, USA)	High-speed camera	Foot				ICC (95% CI)	MD (%)	Pearson's *r*	MD (ms)	Systematic bias ± RE (ms)
				3.3 ± 0.4	107.0 ± 23.0	93.0 ± 25.0	0.81 (0.18, 0.93)	15.1^{***}	0.81^{***}	14.0	15.0 ± 15.0
Gindre et al. (2016) [68]	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Photocell system	Lumbar spine				ICC	MD (%)	CV%		
				3.3	205.0 ± 13.0	101.0 ± 20.0	0.67	103^{*}	24.7		
				4.2	204.0 ± 14.0	119.0 ± 20.0	0.72	71.0^{*}	19.8		
				5.0	205.0 ± 15.0	131.0 ± 18.0	0.78	57.0^{*}	17.1		
				5.8	202.0 ± 15.0	135.0 ± 17.0	0.82	50.0^{*}	15.3		
	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	High-speed camera	Lumbar spine				ICC	MD (%)	CV%		
				3.3	205.0 ± 13.0	117.0 ± 17.0	0.69	75.0^{*}	20.5		
				4.2	204.0 ± 14.0	133.0 ± 18.0	0.66	52.0^{*}	16.3		
				5.0	205.0 ± 15.0	143.0 ± 20.0	0.66	43.0^{*}	14.8		
				5.8	202.0 ± 15.0	144.0 ± 18.0	0.77	41.0^{*}	13.7		
Buchheit et al. (2015) [37]	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Instrumented treadmill	Scapula				Mean bias (90% CI) (%)	CV% (90% CI)	Pearson's *r* (90% CI)		
				2.8-7.5	Not reported	Not reported	-25.8 (-18.8, -27.7)	15.7 (13.5, 18.9)	0.68 (0.55, 0.78)		

Abbreviations: CI, confidence interval; CV, coefficient of variation; ICC, intraclass correlation coefficient; IMU, inertial measurement unit; IQR, inter-quartile range; MD, mean difference; ME, mean error; ms, milliseconds; m·s^{-1}, metres per second; RE, random error; RMSE, root mean square error; SD, standard deviation.
Negative values represent an underestimation of flight time calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Falbriard et al. (2018) and Buchheit et al. (2015) as validity outcomes were reported from pooled speeds.
Values converted to milliseconds where required.
$^{*}p < 0.05$.
$^{**}p < 0.01$.
$^{***}p < 0.001$.

Table 8 Validity summary statistics for step time and stride time.

Study	Variable	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4
Falbriard et al. [34] (2018)	Step time	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot				Median bias ± IQR (ms)	Median precision ± IQR (ms)		
					2.8-5.6	Not reported	Not reported	0 ± 0	3.0 ± 2.0		
Lee et al. (2010) [32]	Step time	Accelerometer (KXM52 – 1050 Kionix, NY, USA)	Motion capture system	Sacrum				Mean bias (ms)	95% LoA (ms)	SE (ms)	Pearson's *r*
					2.8-3.3	Not reported	Not reported	-0.7	-20.0, 18.0	0.7	0.93
					3.6-4.2			-1.3	-21.0, 18.0	0.8	0.78
					4.4-5.2			-0.4	-19.0, 19.0	0.8	0.76
Chew et al. (2018) [52]	Stride time	IMU (Opal, APDM Inc.)	Motion capture system	Foot				ME ± SD (ms)	RMSE (ms)		
					2.2	Not reported	Not reported	-4.0 ± 24.0	17.6		
					2.5			-3.2 ± 22.7	17.3		
					2.8			-1.0 ± 25.6	24.8		
					3.1			0.3 ± 22.1	21.4		
Norris et al. (2016) [84]	Stride time	Accelerometer (Shimmer 2r, Shimmer Inc., Dublin, Ireland)	Stride time calculation method [Mercer et al., (2003)]	Tibia				SE (ms)	CV%	ICC	
					Self-paced	740.0 ± 90.0	740.0 ± 100.0	8.0	1.1	0.99	
		Accelerometer (Shimmer 2r, Shimmer Inc., Dublin, Ireland)	Stride time calculation method [Mizrahi et al., (2000)]	Tibia				SE (ms)	CV%	ICC	
					Self-paced	740.0 ± 90.0	740.0 ± 100.0	7.0	0.9	0.99	
		Accelerometer (Shimmer 2r, Shimmer Inc., Dublin, Ireland)	Stride time calculation method [Purcell et al., (2006)]	Tibia				SE (ms)	CV%	ICC	
					Self-paced	740.0 ± 90.0	740.0 ± 100.0	10.0	1.3	0.99	
McGrath et al. (2012) [29]	Stride time	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system [Hreljac and Marshal (2000)]	Tibia				True error (ms)	% error	ICC	
					2.2	810.0 ± 10.0	810.0 ± 10.0	0.1	1.5	0.55	
					3.3	770.0 ± 20.0	780.0 ± 10.0	0.2	1.2	0.83	
		IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system [Zeni et al., (2008)]	Tibia				True error (ms)	% error	ICC	
					2.2	810.0 ± 10.0	810.0 ± 10.0	0.29	1.27	0.57	
					3.3	770.0 ± 20.0	780.0 ± 10.0	0.29	1.26	0.69	
Lee et al. (2010) [32]	Stride time	Accelerometer (KXM52 – 1050 Kionix, NY, USA)	Motion capture system	Sacrum				Mean bias (ms)	95% LoA (ms)	SE (ms)	Pearson's *r*
					2.8-3.3	Not reported	Not reported	-0.5	-16.0, 15.0	0.6	0.98
					3.6-4.2			1.2	-17.0, 20.0	0.7	0.95
					4.4-5.2			-1.0	-21.0, 19.0	0.8	0.92

Table 8 Validity summary statistics for step time and stride time (continued).

Study	Variable	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1} \pm SD$	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3	Statistic 4
Bergamini et al. (2012) [33]	Stride time	IMU (FreeSense, Sensorize, Italy)	Force plate	Lumbar spine				ME (ms)	LoA (ms)		
					Maximal sprint	485.0 ± 42.2	483.8 ± 41.4	5.0	25.0		
		IMU (FreeSense, Sensorize, Italy)	High-speed camera	Lumbar spine				ME (ms)	LoA (ms)		
					Maximal sprint	453.8 ± 14.2	453.7 ± 16.2	5.0	25.0		
Kenneally-Dabrowski et al. (2018) [54]	Stride time	Accelerometer (GPSports, Canberra, Australia)	Force plate	Thoracic spine				Mean bias (ms)	95% LoA (ms)	Spearman's *r*	
					8.64 ± 0.5	Not reported	Not reported	-26.0*	-91.0, 39.0	-0.18	

Abbreviations: CV, coefficient of variation; ICC, intraclass correlation coefficient; IMU, inertial measurement unit; IQR, inter-quartile range; LoA, limits of agreement; ME, mean error; ms, milliseconds; $m \cdot s^{-1}$, metres per second; RMSE, root mean square error; SD, standard deviation; SE, standard error.
Negative values represent an underestimation of step time and stride time calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Falbriard et al. (2018) as validity outcomes were reported from pooled speeds.
Values converted to milliseconds where required.
*$p < 0.05$.

Table 9 Validity summary statistics for swing time.

Study	Sensor	Criterion	Site	Running velocity m·s⁻¹ ± SD	Sensor mean ± SD ms	Criterion mean ± SD ms	Statistic 1	Statistic 2	Statistic 3
Falbriard et al. (2018) [34]	IMU (Physilog 4, Gait Up, Switzerland)	Instrumented treadmill	Foot				Median bias ± IQR (ms)	Median precision ± IQR (ms)	
				2.8-5.6	Not reported	Not reported	15.0 ± 12.0	5.0 ± 3.0	
McGrath et al. (2012) [29]	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Hreljac and Marshal (2000)	Tibia				True error (ms)	% error	ICC
				2.2	460.0 ± 330.0	360.0 ± 10.0	65.9	18.7	0.38
				3.3	450.0 ± 20.0	390 ± 10.0	54.8	16.6	0.32
	IMU (Shimmer, Shimmer Inc., Dublin, Ireland)	Motion capture system Zeni et al., (2008)	Tibia				True error (ms)	% error	ICC
				2.2	460.0 ± 330.0	340.0 ± 10.0	78.8	26.8	0.32
				3.3	450.0 ± 20.0	360 ± 10.0	90.0	26.4	0.28

Abbreviations: ICC, intraclass correlation coefficient; IMU, inertial measurement unit; IQR, inter-quartile range; ms, milliseconds; $m \cdot s^{-1}$, metres per second; SD, standard deviation.
A velocity range is presented for Falbriard et al. (2018) as validity outcomes were reported from pooled speeds.
Values converted to milliseconds where required.

Table 10 Validity summary statistics for step frequency.

Study	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1} \pm SD$	Sensor mean ± SD $step \cdot min^{-1}$	Criterion mean ± SD $step \cdot min^{-1}$	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Garcia-Pinillos et al. (2018) [71]	IMU (Stryd™, Stryd Powermeter, Stryd Inc., Boulder, CO, USA)	Photocell system	Foot				ICC	Pearson's *r*			
				2.2-5.6	160.9 ± 6.8 to 191.8 ± 5.4	159.6 ± 6.3 to 193.2 ± 5.9	0.96-0.99	$0.96\text{-}0.99^{***}$			
Garcia-Pinillos et al. (2019) [35]	IMU (RunScribe™, Scribe Lab. Inc. San Francisco, CA, USA)	High-speed camera	Foot				ICC (95% CI)	Pearson's *r*	MD ($step \cdot min^{-1}$)	MD (%)	Systematic bias ± RE ($step \cdot min^{-1}$)
				3.3 ± 0.4	168.1 ± 7.4	166.8 ± 7.7	0.96 (0.92, 0.98)	0.95^{***}	1.3	0.8	1.3 ± 2.5
	IMU (Stryd™, Stryd Powermeter, Stryd Inc. Boulder, CO, USA)	High-speed camera	Foot				ICC (95% CI)	Pearson's *r*	MD ($step \cdot min^{-1}$)	MD (%)	Systematic bias ± RE ($step \cdot min^{-1}$)
				3.3 ± 0.4	166.7 ± 7.3	166.8 ± 7.7	0.97 (0.94, 0.98)	0.93^{***}	-0.1	0.1	-0.1 ± 2.8
Pairot de Fontenay et al. (2020) [60]	IMU (MilestonePod, Milestone Sports, Long Beach, CA, USA)	Instrumented treadmill	Foot				Mean bias ($step \cdot min^{-1}$)	Pearson's *r*	95% LoA ($step \cdot min^{-1}$)		
				Not reported	Not reported	Not reported	1.6	0.99^{***}	± 1.4		
	IMU (Zoi, Runteq, Tampere, Finland)	Instrumented treadmill	Foot				Mean bias ($step \cdot min^{-1}$)	Pearson's *r*	95% LoA ($step \cdot min^{-1}$)		
				Not reported	Not reported	Not reported	0.9	0.99^{***}	± 1.3		
	IMU (RunScribe™, Montara, CA, USA)	Instrumented treadmill	Heel				Mean bias ($step \cdot min^{-1}$)	Pearson's *r*	95% LoA ($step \cdot min^{-1}$)		
				Not reported	Not reported	Not reported	1.1	0.99^{***}	± 0.9		
Pairot de Fontenay et al. (2020) [60]	IMU (Moov Now™, Moov, San Mateo, CA, USA)	Instrumented treadmill	Tibia				Mean bias ($step \cdot min^{-1}$)	Pearson's *r*	95% LoA ($step \cdot min^{-1}$)		
				Not reported	Not reported	Not reported	2.3	0.98^{***}	± 2.0		
	IMU (TgForce, Kelsec Systems Inc., Montreal, Canada)	Instrumented treadmill	Tibia				Mean bias ($step \cdot min^{-1}$)	Pearson's *r*	95% LoA ($step \cdot min^{-1}$)		
				Not reported	Not reported	Not reported	4.5	0.96^{***}	± 6.1		
Gouttebarge et al. (2015) [83]	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Foot-mounted accelerometer	Lumbar spine				ICC (95% CI)				
				2.8 3.3 3.9	164.3 ± 7.0 168.9 ± 8.0 175.9 ± 10.0	165.6 ± 8.0 169.4 ± 8.0 175.7 ± 13.0	0.89 (0.69, 0.96) 0.78 (0.45, 0.96) 0.90 (0.72, 0.97)				

Table 10 Validity summary statistics for step frequency (continued).

Study	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD step·min^{-1}	Criterion mean ± SD step·min^{-1}	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5
Gindre et al. (2016) [68]	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	Photocell system	Lumbar spine				ICC	MD (%)	CV%		
				3.3	163.0 ± 10.0	163.0 ± 9.0	0.86	0.1	4.1		
				4.2	168.0 ± 9.0	169.0 ± 9.0	0.94	-0.5	3.9		
				5.0	177.0 ± 10.0	178.0 ± 10.0	0.93	-0.3	3.8		
				5.8	188.0 ± 11.0	190.0 ± 12.0	0.87	-0.8	4.6		
	Accelerometer (Myotest®, Myotest SA, Sion, Switzerland)	High-speed camera	Lumbar spine				ICC	MD (%)	CV%		
				3.3	163.0 ± 10.0	163.0 ± 9.0	0.89	0.2	3.9		
				4.2	168.0 ± 9.0	168.0 ± 9.0	0.95	-0.2	3.7		
				5.0	177.0 ± 10.0	176.0 ± 11.0	0.84	0.3	3.8		
				5.8	188.0 ± 11.0	190.0 ± 12.0	0.86	-0.9	4.6		
Machulik et al. (2020) [72]	IMU (Humotion SmarTracks Integrated System)	Photocell system	Lumbar spine				ICC (95% CI)	Systematic bias (step·min^{-1})	95% LoA (step·min^{-1})		
				3.8 ± 0.7	159.6 ± 7.8	168.6 ± 7.8	0.75-0.89 (0.48, 0.95)	-11.9 to -5.2	-20.8, 1.7		
				6.8 ± 1.0	206.4 ± 15.9	228.0 ± 19.2	0.90-0.94 (0.79, 0.97)	-25.9 to -6.5	-47.76, 6.1		

Abbreviations: CI, confidence interval; CV, coefficient of variation; ICC, intraclass correlation coefficient; IMU, inertial measurement unit; LoA, limits of agreement; MD, mean difference; m·s^{-1}, metres per second; RE, random error; SD, standard deviation; step·min^{-1}, steps per minute.
Negative values represent an underestimation of step frequency calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study.
***$p < 0.001$.

Table 11 Validity summary statistics for step length and stride length.

Study	Variable	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD cm	Criterion mean ± SD cm	Statistic 1	Statistic 2	Statistic 3	Statistic 4	Statistic 5	Statistic 6	Statistic 7
Garcia-Pinillos et al. (2018) [71]	Step length	IMU (Stryd™, Stryd Powermeter, Stryd Inc., Boulder, CO, USA)	Photocell system	Foot	2.2-5.6	83.0 ± 3.6 to 173.9 ± 84.4	83.8 ± 3.6 to 172.9 ± 5.1	ICC 0.93-0.99	Pearson's *r* 0.93-0.99***					
Garcia-Pinillos et al. (2019) [35]	Step length	IMU (RunScribe™, Scribe Lab. Inc. San Francisco CA, USA)	High-speed camera	Foot	3.3 ± 0.4	116.3 ± 12.1	116.9 ± 12.5	ICC (95% CI) 0.97 (0.95, 0.98)	Pearson's *r* 0.96***	MD (cm) -0.6	MD (%) 0.5	Systematic bias ± RE (cm) -0.6 ± 4.3		
		IMU (Stryd™, Stryd Powermeter, Stryd Inc. Boulder CO, USA)	High-speed camera	Foot	3.3 ± 0.4	118.05 ± 13.47	116.89 ± 12.50	ICC (95% CI) 0.98 (0.96, 0.99)	Pearson's *r* 0.94***	MD (cm) 1.2	MD (%) 1.0	Systematic bias ± RE (cm) 1.2 ± 3.9		
Machulik et al. (2020) [72]	Step length	IMU (Humotion SmarTracks Integrated)	Photocell system	Lumbar spine	3.8 ± 0.7 6.8 ± 1.0	141.0 ± 26.0 196.0 ± 36.0	131.0 ± 20.0 173.0 ± 21.0	ICC (95% CI) 0.90-0.94 (0.79, 0.98) 0.79-0.85 (0.58, 0.93)	Systematic bias (cm) 8.1-12.2 11.5-28.4	95% LoA (cm) -14.2, 30.4 -7.1, 62.3				
Brahms et al. (2018) [85]	Stride length	IMU (Xsens, MTw)	Motion capture system	Foot	3.6 ± 0.3	259.2 ± 27.6	262.3 ± 27.2	ICC (95% CI) 0.96 (0.93, 0.97)	Pearson's *r* 0.96***	MD (cm) -3.2	ME (cm) 5.0	% error 2.0	95% LoA (cm) -18.3, 11.8	RMSE (cm) 8.3
Chew et al. (2018) [52]	Stride length	IMU (Opal, APDM Inc.)	Motion capture system	Foot	2.2 2.5 2.8 3.1	Not reported	Not reported	ME ± SD (cm) 32.3 ± 48.2 14.1 ± 46.0 44.0 ± 56.7 46.0 ± 62.6	RMSE (cm) 62.4 70.2 63.8 59.2					
Zrenner et al. (2018) [65]	Stride length	IMU (miPod sensor): stride time-based algorithm	Motion capture system	Foot	2.0-6.0	Not reported	Not reported	ME ± SD (cm) 17.7 ± 57.3	MAE (cm) 45.2	% error 17.1				
		IMU (miPod sensor): acceleration-based algorithm	Motion capture system	Foot	2.0-6.0	Not reported	Not reported	ME ± SD (cm) -0.5 ± 25.6	MAE (cm) 19.9	% error 7.9				
		IMU (miPod sensor): foot trajectory-based algorithm	Motion capture system	Foot	2.0-6.0	Not reported	Not reported	ME ± SD (cm) 2.0 ± 14.1	MAE (cm) 7.6	% error 2.8				
		IMU (miPod sensor): deep learning-based algorithm	Motion capture system	Foot	2.0-6.0	Not reported	Not reported	ME ± SD (cm) 2.5 ± 20.1	MAE (cm) 15.3	% error 5.9				

Abbreviations: CI, confidence interval; cm, centimetres; ICC, intraclass correlation coefficient; IMU, inertial measurement unit; LoA, limits of agreement; MAE, mean absolute error; MD, mean difference; ME, mean error; m·s^{-1}, metres per second; RE, random error; RMSE, root mean square error; SD, standard deviation.
Negative values represent an underestimation of step length and stride length calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Zrenner et al. (2018) as validity outcomes were reported from pooled speeds.
Values converted to centimetres where required.
***$p < 0.001$.

Table 12 Validity summary statistics for ground reaction force.

Study	Variable	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1} \pm SD$	Sensor mean ± SD N	Criterion mean ± SD N	Statistic 1	Statistic 2	Statistic 3	Statistic 4
Ngoh et al. (2018) [56]	vGRF	IMU (Opal, APDM Inc.)	Instrumented treadmill	Foot				RMSE ± SD (N)	Signal cross-correlation		
					2.2 2.5 2.8	Not reported	Not reported	10.5 ± 6.2 9.3 ± 4.4 10.5 ± 5.6	0.99 0.99 0.99		
Raper et al. (2018) [40]		Accelerometer (ViPerform v5, DorsaVi, Melbourne, Australia)	Force plate	Tibia				MD (N)			
					5.2 ± 0.6	Not reported	Not reported	400.0			
Neugebauer et al. (2014) [66]		Accelerometer (GT3X+ AM, ActiGraph, Pensacola, FL, USA)	Force plate	Hip				MD (N)	% MD ± SD	Mean bias ± SD (N)	95% LoA (N)
					2.2-4.1	Not reported	Not reported	106.4	8.3 ± 3.7	-50.5 ± 130.4	-311.3, 210.3
Gurchiek et al. (2017) [57]		IMU (Yost Data Logger 3-Space Sensor, YEI Technology, Portsmouth, OH)	Force plate	Sacrum				RMSE (N)	Pearson's *r*	Systematic bias (N)	95% LoA (N)
					Not reported	Not reported	Not reported	77.1	0.88**	-34.1	-171.8, 103.7
Eggers et al. (2018) [38]		Accelerometer (wGT3X-BT, ActiGraph, Pensacola, FL, USA)	Force plate	Lumbar spine				CV% (90% CI)	Pearson's *r* (90% CI)	TEE (90% CI) (N)	
					3.3	Not reported	Not reported	9.2 (7.6, 11.7)	0.81 (0.69, 0.89)	0.71 (0.51, 1.05)	
Wundersitz et al. (2013) [36]		Accelerometer (SPI Pro, ASP00725, GPSports Pty. Ltd., Canberra, Australia)	Force plate	Scapula				CV%	Spearman's *r*		
					5.4 ± 0.5	1582.0 ± 408.0	1731.0 ± 245.0	16.2	0.12		
Eggers et al. (2018) [38]		Accelerometer (wGT3X-BT, ActiGraph, Pensacola, FL, USA)	Force plate	Scapula				CV% (90% CI)	Pearson's *r* (90% CI)	TEE (90% CI) (N)	
					3.3	Not reported	Not reported	9.6 (8.0, 12.3)	0.79 (0.54, 1.0)	0.76 (0.54, 1.14)	
Edwards et al. (2019) [43]		Accelerometer (SPI HPU, GPSports Pty. Ltd., Canberra, Australia)	Force plate	Scapula				Pearson's *r*			
					3.3-6.7	Not reported	Not reported	0.44**			
Wouda et al. (2018) [87]		IMU (Xsens, Enschede, the Netherlands)	Instrumented treadmill	Lower legs and pelvis				RMSE ± SD (N)	Pearson's *r*		
					3.3	2338.8 ± 256.4	2261.1 ± 101.0	220.8 ± 45.7	0.96		
Dorschky et al. (2019) [55]		IMU (Portabiles GmbH, Erlangen, DE)	Force plate	Foot, tibia, thighs and lumbar spine				RMSD ± SD (N)	Pearson's *r*	rRMSD ± SD (%)	
					3.0-4.9	Not reported	Not reported	241.4 ± 59.6	0.94	12.8 ± 3.6	

Table 12 Validity summary statistics for ground reaction force (continued).

Study	Variable	Sensor	Criterion	Site	Running velocity m·s^{-1} ± SD	Sensor mean ± SD N	Criterion mean ± SD N	Statistic 1	Statistic 2	Statistic 3	Statistic 4
Gurchiek et al. (2017) [57]	rGRF	IMU (Yost Data Logger 3-Space Sensor, YEI Technology, Portsmouth, OH)	Force plate	Sacrum				RMSE (N)	Pearson's r	Systematic bias (N)	95% LoA (N)
					Not reported	Not reported	Not reported	73.6	0.90**	-29.7	-163.9, 104.4
Wundersitz et al. (2013) [36]		Accelerometer (SPI Pro, ASP00725, GPSports Pty. Ltd., Canberra, Australia)	Force plate	Thoracic spine				CV%	Spearman's r		
					5.4 ± 0.5	2194.0 ± 317.0	1755.0 ± 253.0	16.4	0.31		
Nedergaard et al. (2018) [39]		Accelerometer (MinimaxX S4, Catapult Innovations, Scoresby, Australia)	Force plate	Thoracic spine				RMSE (N)			
					2.0	3124.4 ± 1422.7	1714.6 ± 162.6	509.2			
					3.0	4769.0 ± 3979.4	1896.2 ± 149.7	509.2			
					4.0	17562.8 ± 30118.0	2068.0 ± 118.6	706.8			
					5.0	6818.0 ± 5999.4	2084.7 ± 186.2	972.8			
Pogson et al. (2020) [61]		Accelerometer (MinimaxX S5, Catapult Innovations, Scoresby, Australia)	Force plate	Thoracic spine				r^2			
					2.0-8.0	Not reported	Not reported	0.9			

Abbreviations: CI, confidence interval; CV, coefficient of variation; IMU, inertial measurement unit; LoA, limits of agreement; MD, mean difference; m·s^{-1}, metres per second; N, Newtons; r^2, coefficient of determination; rGRF, resultant ground reaction force; RMSD, root mean square deviation; RMSE, root mean square error; rRMSD, relative root mean square deviation; SD, standard deviation; TEE, typical error of the estimate; vGRF, vertical ground reaction force.
Negative values represent an underestimation of ground reaction force calculated by the IMU compared to the criterion.
Velocity reported with or without ± SD, depending on the method used in each study. A velocity range is presented for Neugebauer et al. (2014), Edwards et al. (2019), Dorschky et al. (2019) and Pogson et al. (2020) as validity outcomes were reported from pooled speeds.
Values converted to Newtons where required.
**$p < 0.01$.

Table 13 Validity summary statistics for vertical stiffness.

Study	Sensor	Criterion	Site	Running velocity $m \cdot s^{-1}$	Sensor mean ± SD $kN \cdot m^{-1}$	Criterion mean ± SD $kN \cdot m^{-1}$	Statistic 1 (90% CI)	Statistic 2 (90% CI)	Statistic 3 (90% CI)
Eggers et al. (2018) [38]	Accelerometer (wGT3X-BT, ActiGraph, Pensacola, FL, USA)	Force plate	Lumbar spine				CV%	Pearson's *r*	TEE ($kN \cdot m^{-1}$)
				3.3	26.0 ± 5.0	24.9 ± 3.7	12.9 (10.7, 16.5)	0.65 (0.44, 0.79)	1.2 (0.8, 2.0)
Buchheit et al. (2015) [37]	Accelerometer (SPI HPU, GPSports, Canberra, Australia)	Instrumented treadmill	Thoracic spine				CV%	Pearson's *r*	Mean bias (%)
				2.8-7.5	Not reported	Not reported	6.3 (5.5, 7.5)	0.98 (0.97, 0.99)	-13.3 (-14.6, -11.9)
Eggers et al. (2018) [38]	Accelerometer (wGT3X-BT, ActiGraph, Pensacola, FL, USA)	Force plate	Thoracic spine				CV%	Pearson's *r*	TEE ($kN \cdot m^{-1}$)
				3.3	24.4 ± 3.8	24.9 ± 3.7	12.8 (10.6, 16.3)	0.66 (0.46, 0.79)	1.2 (0.8, 2.0)

Abbreviations: CI, confidence interval; CV, coefficient of variation; $kN \cdot m^{-1}$, kilo Newtons per metre; $m \cdot s^{-1}$, metres per second; SD, standard deviation; TEE, typical error of the estimate.
Negative values represent an underestimation of vertical stiffness calculated by the IMU compared to the criterion.
A velocity range is presented for Buchheit et al. (2015) as validity outcomes were reported from pooled speeds.

Fig. 1 Flow chart of study selection process

Fig. 2 Forest plot displaying the effect of contact time (ms) calculated from IMUs worn on the foot, tibia and lumbar spine between 3.3 and 4.3 m·s^{-1}. Data are presented as means and SD of IMU- and criterion-derived contact time. Data from the RunScribe™ sensor is shown in Garcia-Pinillos et al. (2019a), while data from the Stryd™ device is shown in Garcia-Pinillos et al. (2019b). Two different motion capture algorithms were used as criterions for McGrath et al. (2012a and b). Gindre et al. (2016a and b) is represented by high-speed camera and photocell system criterions, respectively. CI, confidence interval; df, degrees of freedom; IMU, inertial measurement unit; IV, instrumental variable; over, overestimation; SD, standard deviation; under, underestimation

Fig. 3 Forest plot displaying the effect of flight time (ms) calculated from IMUs worn on the foot between 3.3 and 4.2 m·s^{-1}. Data are presented as means and SD of IMU- and criterion-derived flight time. Data from the RunScribe™ sensor is shown in Garcia-Pinillos et al. (2019a), while data from the Stryd™ device is shown in Garcia-Pinillos et al. (2019b). CI, confidence interval; df, degrees of freedom; IMU, inertial measurement unit; IV, instrumental variable; over, overestimation; SD, standard deviation; under, underestimation

Fig. 4 Forest plot displaying the effect of step frequency (step·min^{-1}) calculated from IMUs worn on the foot and lumbar spine between 3.3 and 4.2 m·s^{-1}. Data are presented as means and SD of IMU- and criterion-derived step frequency. Data from the RunScribe™ sensor is shown in Garcia-Pinillos et al. (2019a), while data from the Stryd™ device is shown in Garcia-Pinillos et al. (2019b). Gindre et al. (2016a and b) is represented by high-speed camera and photocell system criterions, respectively, where the authors reported the same values for each. CI, confidence interval; df, degrees of freedom; IMU, inertial measurement unit; IV, instrumental variable; over, overestimation; SD, standard deviation; under, underestimation

Fig. 5 Forest plot displaying the effect of step length (cm) calculated from IMUs worn on the foot between 3.3 and 4.2 m·s^{-1}. Data are presented as means and SD of IMU- and criterion-derived flight time. Data from the RunScribe™ sensor is shown in Garcia-Pinillos et al. (2019a), while data from the Stryd™ device is shown in Garcia-Pinillos et al. (2019b). CI, confidence interval; df, degrees of freedom; IMU, inertial measurement unit; IV, instrumental variable; over, overestimation; SD, standard deviation; under, underestimation