Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies

Journal:	*Journal of Sports Sciences*
Manuscript ID:	RJSP-2014-0099.R1
Manuscript Type:	Special Issue
Keywords:	soccer, tracking system, match analysis, training load, calibration equations

Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies

Running title: Integrating different tracking systems in football

Buchheit, M.,[1,2] *Poon, T.K.,*[1] *Allen A.,*[3] *Modonutti, M.,* [1] *Gregson, W.,*[1,4] *and Di Salvo V.*[1,5]

[1]*Football Performance & Science Department, ASPIRE, Academy for Sports Excellence, Doha, Qatar*
[2]*Sport Sciences Departement, Myorobie Assocaition, Montvalezan, France*
[3]*Performance Sports New Zealand, Wellington, New Zealand*
[4]*Football Exchange, Research Institute for Sport & Exercise Sciences, Liverpool John Moores University, Liverpool, UK*
[5]*Department of Movement, Human and Health sciences, University of Rome "Foro Italico"*

Address for correspondence:
Martin Buchheit
Football Performance & Science Department
ASPIRE, Academy for Sports Excellence
P.O. Box 22287, Doha, Qatar.
Tel.: (+974) 4413 6103
Fax: (+974) 4413 6060
E-mail: martin.buchheit@aspire.qa

Submission type: Original Research
Abstract word count: 200
Text-only word count: 4000
Number of tables: 6
Number of figures: 3
Disclosures: nothing to disclose

Abstract

During the past decade substantial development of computer-aided tracking technology has occurred. Therefore, we aimed to provide calibration equations to allow the interchangeability of different tracking technologies used in soccer. Eighty-two highly-trained soccer players (U14-U17) were monitored during training and one game. Player activity was collected simultaneously with a semi-automatic multiple-camera (Prozone), local position measurement (LPM) technology (Inmotio) and two global positioning systems (GPSports and VX). Data were analyzed with respect to three different field dimensions (small, <30m^2 to full-pitch, match). Variables provided by the systems were compared, and calibration equations (linear regression models) between each system were calculated for each field dimension. Most metrics differed between the four systems with the magnitude of the differences dependant on both pitch size and the variable of interest. Trivial-to-small between-system differences in total distance were noted. However, high-intensity running distance (>14.4km/h) was slightly-to-moderately greater when tracked with Prozone, and accelerations, small-to-very largely greater with LPM. For most of the equations, the typical error of the estimate was of a moderate magnitude. Interchangeability of the different tracking systems is possible with the provided equations, but care is required given their moderate typical error of the estimate.

Key Words: soccer; tracking system; match analysis; training load; agreement; calibration equations.

Introduction

Monitoring players' physical activity during both matches and training is today a common practice in professional soccer (Carling, 2013). The detailed analysis of match activity demands allows individual physical performance profile to be determined, which can be used to define training orientations and/or design soccer-specific training drills (Di Salvo, et al., 2007). In parallel, the quantification of training sessions physical demands is an integral part of training load management and players monitoring (Casamichana, Castellano, Calleja-Gonzalez, San Roman, & Castagna, 2013, Scott, Lockie, Knight, Clark, & Janse de Jonge, 2013) permitting coaching staff to readjust training periodization on a day-by-day basis.

During the past decade, there has been a substantial development of computer-aided tracking technology for the examination of player activity during training and match-play. Sophisticated systems such as multiple camera semi-automatic systems, local position measurement (LPM) technology and global positioning system (GPS) technology are now capable of quickly recording and processing the data of all players' physical contributions throughout an entire match or training session (Carling, Bloomfield, Nelsen, & Reilly, 2008). Most of these systems/technologies have been shown to allow accurate examination of players' activity on the field, with the faster and/or less linear the movements, the poorer the accuracy and the reliability (Di Salvo, Collins, MacNeill, & Cardinale, 2006, Di Salvo, Gregson, Atkinson, Tordoff, & Drust, 2009, Frencken, Lemmink, & Delleman, 2010, Ogris, et al., 2012, Portas, Harley, Barnes, & Rush, 2010, Varley, Fairweather, & Aughey, 2012).

Despite their growing use in elite football, limited attempt to date has been made to comprehensively evaluate the agreement between these different systems/technologies (Harley, Lovell, Barnes, Portas, & Weston, 2011, Randers, et al., 2010). Results from these studies have shown that multiple camera semi-automatic systems tend to report slightly-to-moderately (Harley, et al., 2011) and moderately-to-largely (Randers, et al., 2010) greater distance covered at high intensity (i.e., >~18 km/h) than GPS technology. With the exception of one study comparing the accuracy of position detection between LPM and an image-based system (1 single camera) (Siegle, Stevens, & Lames, 2013), there is however, to our knowledge, no comparison between LPM technology and traditional tracking systems (i.e., multiple camera semi-automatic systems and GPS). Importantly also, the latter comparisons have been limited to match activities in a limited number of players (i.e., 6 (Harley, et al., 2011) and 13-20 (Randers, et al., 2010)), consequently solutions to use these systems interchangeably are still missing. Systems interchangeability is an important point for practitioners in professional clubs, who often use two different systems throughout the week, i.e., GPS technology during training sessions and multiple camera semi-automatic systems during matches. In order to have an integrated approach to monitoring training/competitive load over an entire week, it is therefore relevant to quantify the agreement between different technologies/systems, and compare their validity and reliability during the same running drills. Consequently, the first objective of this study was to provide calibration equations to allow the interchangeability of three different tracking technologies used in professional soccer (i.e., semi-automatic multiple-camera, Prozone; LPM, *Inmotio* and GPS technology, *GPSports* and *VX*) to measure on-field running movements in highly-trained young players. The second aim was to provide further information on the accuracy and reliability of the three technologies while comparing their measures with that obtained at the same time with timing gates used as gold-standards.

Methods

Participants. Eighty-two highly-trained soccer players representative of U14 to U17 teams from an elite soccer academy, and 14 athletes from the same academy track and field group (U16-U18) were studied. All soccer players participated on average in ~14 hours of combined soccer-specific training and competitive play per week (6-8 soccer training sessions, 1 strength training session, 1-2 conditioning sessions, 1 domestic match per week and 2 international club matches every 3 weeks). All players had a minimum of 3 years prior soccer-specific training. Physical characteristics and match running performance of the players have been reported elsewhere (Buchheit, Mendez-Villanueva, Simpson, & Bourdon, 2010). The track and field athletes had on average in ~14 hours of combined strength and speed training per week and were familiar with all the different sprints tests. These data arose as a condition of player monitoring in which player activities are routinely measured over the course of the competitive season. Therefore, usual appropriate ethics committee clearance was not required (Winter & Maughan, 2009). Nevertheless, to ensure team and player confidentiality, all physical performance data were anonymized before analysis, and the study followed the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Experimental overview. A summary of the different test protocols, variables measured and analyses conducted are presented in Table 1. The four systems examined/compared in the present study were representative of three different technology currently used in elite football: local position measurement technology (Inmotio Object tracking v2.6.9.545, 45 Hz, Amsterdam, The Netherlands), semi-automatic multiple-camera (Prozone 10 Hz, Leeds, UK), and GPS technology (GPSports, SPI Pro XII, 5 Hz, GPSport, chip version 2.6.4, Canberra, Australia) and VX, VX340a, chip version VXLogFW_V147-02_04Feb13, 4 Hz, Lower Hutt, New Zealand). Data were collected within 4 days in an open roof stadium (pitch: 105 x 70 m, natural grass), over 4 training sessions (2 with the U15 and 2 with the U17 team), a generic sprint session (no ball, with the track and field group) and a 20-min friendly match (combination of U14 and U15). During all sessions, players/athletes wore Inmotio, GPSports and VX sensors simultaneously (there is no need for a sensor for Prozone). The data collected with the 4 systems were processed within a month after the study by independent technicians/researchers, which were blinded to all other results.

To provide solutions to allow the interchangeability between the different systems, the variables provided by the 4 systems during the same training and matches were compared, and we provided calibration equations between each system (to predict the measures that could be expected with a given system from the data collected with another system).

In order to provide additional information on the validity of the 4 systems, we compared the measures provided by each of the 4 systems during standardized drills with running distance/times collected with timing gates (gold standard). Distance covered during different constant-velocity runs, peak velocity (V_{peak}) and peak acceleration (Acc_{peak}) were obtained from each system. The distance covered during the sprinting drill (Jennings, Cormack, Coutts, Boyd, & Aughey, 2010a, Waldron, Worsfold, Twist, & Lamb, 2011) was not examined. The following runs were examined (Figure 1): 1) Runs at 7.2 km/h, 14.4 km/h and 19.8 km/h on an oval 200-m course drawn on the soccer pitch (Gray, Jenkins, Andrews, Taaffe, & Glover, 2010), 2) 2 x 40-m sprints, 3) 2 x L-shaped sprints (10m+10m, 90° change of direction) and 4) 2 x ZigZag-shaped sprints (5m+5m+5m+5m, 90° changes of direction). It is worth noting, however, that the variables measured by the different systems are not all strictly comparable to the criterion measures. For instance, while V_{peak} could be measured by all four tracking systems (as the maximal velocity reached over a 0.8-s window) and the gates using the best 10-m split, Acc_{peak} (as the maximal changes in velocity reached over a 0.8-s window) could not be assessed with the gates; the criterion measure for acceleration was actually the average acceleration over the first 10 m. Similarly while we examined Acc_{peak} and V_{peak} during the change of direction speed drills, the criterion measure was the time over the course. At present, only Inmotio can easily provide estimated distance and time for set distances (which is a feature of the LPM method).

To examine the reliability of each system, all the above-mentioned drills were repeated twice within a few minutes.

The number of satellites for GPS was satisfactory during all sessions: range: 7-12, average 9.5 ± 2 for GPSports and 7-14, average 9.5 ± 2 and VX, respectively. The horizontal dilution of position (HDOP), which is a reflection of the geometrical arrangement of satellites and is related to both the accuracy and quality of the signal was not collected, which is a limitation. Data processing for GPS systems was achieved while using the software provided by each manufacturer/company (AMS R1-2012.9 for GPSports, 3.88.0.0 for VX, Inmotio Object tracking v2.6.9.545).

The Prozone system is configured for match situations which feature quantifiable team shape/formations as previously described (Di Salvo, et al., 2006, Di Salvo, et al., 2009). The standard system is not configured for monitoring training sessions and although best efforts were made to control the environment in this study, multiple player ID issues were experienced during data collection due to test participants leaving the playing field in an ad-hoc manner or entering covered areas with no visibility to the Prozone tracking sensors. In addition, for standard operational match situations, the Prozone system utilises a separate OB camera feed to continually validate trajectory player ID and cross reference with eventing data concerning player actions in matches. However this function was not available or employed for the duration of the data capture in this study, nor were any ball 'events' captured. Considerable care and attention was given in order to ensure that the Prozone data provided was as accurate as possible, however, it is unlikely to achieve the same quality standard as a normal match situation. To assess the level of reproducibility of this particular procedure, all data were processed twice by two independent Prozone teams. The differences and the CV between the two analyses were all non-substantial (standardized difference <0.2), the CV and ICC for the reliability analysis were comparable between the two sets of data analysis, and the comparisons with the other systems were similar (i.e., the magnitude of the between-system differences, the biases and the correlations coefficients were of similar magnitude). Therefore, only data from the first analysis were used for the study.

Finally, the Inmotio system was installed, calibrated and used by three qualified technicians from the Inmotio Company during all sessions. Players wore an LPM vest containing a transponder on the back and antennas on both shoulders, ensuring optimal line of fight to the base stations (Stevens, et al., 2014). A fourth qualified technician then processed the data as previously described (Stevens, et al., 2014) using the dedicated software (Inmotio Object tracking v2.6.9.545, 45 Hz, Amsterdam, The Netherlands).

Data Analysis

Data in text and figure are presented as mean with 90% confidence intervals. All data were first log-transformed to reduce bias arising from non-uniformity error. Only complete player-data sets were analysed (e.g., including data from all the 4 systems). Validity analysis included comparison of the different metrics with the reference measures (bias, but only when the metrics were similar, e.g., distance vs. distance or speed vs. speed), the typical error of the estimate (both in % and in standardized units) and the magnitude of the correlations between the systems. Data from all repeated trials (e.g., two runs at 19.8 km/h on the 200-m oval, two 40-m sprints, or two ZigZag-shaped sprints) were used to calculate for each system 1) the typical error of measurement, expressed as a CV (% absolute reliability) and 2) the intraclass coefficient correlation (ICC, relative reliability). To assess the agreement between the different systems, all analyses were performed with respect to three different field dimensions (i.e., small, <30 x 30 m; Medium, <50 x 50 m; and full pitch, match). Between-system standardized differences were calculated using pooled standard deviations. Uncertainty in each effect was expressed as 90% CL and as probabilities that the true effect was substantially positive and negative. These probabilities were used to make a qualitative probabilistic mechanistic inference about the true effect: if the probabilities of the effect being substantially positive and negative were both >5%, the effect was reported as unclear; the effect was otherwise clear and reported as the magnitude of the observed value. The scale was as follows: 25−75%, possible; 75−95%, likely; 95−99%, very likely; >99%, almost certain. Additionally, we provided calibration (linear) equations to predict the measures that could be expected with a given system from the data collected with another system. The typical error of the estimate (TEE, %) and the magnitude of the correlations between the systems were also calculated. When using the data from all training sessions/matches and speed zones pooled together, the TEEs for the calibration equations were consistently very large, irrespective of the model used (e.g., linear, log-linear, log-log). Different calibration equations were therefore provided for the same three pitch dimensions as for between-system differences, i.e., within the box area (<30 x 30m), half a pitch (<50 x 50m) and full pitch (match), and for different speed zones. This procedure substantially decreased the magnitude of the errors within each pitch dimension/speed zone. A linear model was retained for all calibration equations since using other models didn't decrease the TEEs. Threshold values for standardized differences, typical error and typical error of the estimate were >0.2 (small), >0.6 (moderate), >1.2 (large) and very large (>2) (Hopkins, Marshall, Batterham, & Hanin, 2009). The magnitude of the ICC was assessed using the following thresholds: >0.99, extremely high; 0.99-0.90, very high; 0.90-0.75, high; 0.75-0.50, moderate; 0.50-0.20, low; <0.20, very low (WG Hopkins, unpublished observations). Finally, the following criteria were adopted to interpret the magnitude of the correlation: ≤0.1, trivial; >0.1-0.3, small; >0.3-0.5, moderate; >0.5-0.7, large; >0.7-0.9, very large; and >0.9-1.0, almost perfect. If the 90% CI overlapped small positive and negative values, the magnitude was deemed unclear; otherwise that magnitude was deemed to be the observed magnitude (Hopkins, et al., 2009).

Results

There was no loss of data with Prozone and Inmotio. With the GPS systems, 4 (<2%, GPSsports) and 1 (<1%, VX) files could not be analyzed due to technical issues (no more battery and/or poor GPS signal).

Agreement.

Most of the metrics collected during either training sequences over different pitch sizes or during a friendly match on the full pitch differed between the 4 systems, and the magnitude of the between-system differences was related to both pitch size and the variable of interest. Overall, while there were small between-system differences in total distance, Prozone tended to report largely greater distance at high speed than the 3 other systems across all pitch sizes

(Figure 2). Table 2, 3 & 4 show selected calibration equations that can be used, at the team level, to predict the measures that could be expected with a given system from the data collected with another system. It is worth noting that the typical error of estimate is often of small-to-moderate magnitude.

Validity.

Distance ran at different speeds: When comparing the distance recorded during the three runs at three different running speeds with the reference distance (200 m), Inmotio tended to largely and moderately underestimate the distance ran at 7.2 km/h and 14.4 km/h (e.g., +47 m) respectively. Prozone and GPSports tended to moderately (e.g., +47 m at 14.4 km/h) and largely (e.g., +50 m at 14.4 km/h) overestimate the distance ran at all intensities, respectively (Figure 3). VX showed values within a small range of variation at 14.4 km/h (+37 m) but slightly underestimated and overestimated the distance ran at 7.2 km/h and 19.8 km/h, respectively. In overall, while VX showed the lower bias in overall, Prozone provided the most consistent distance output across all speed intensities (i.e., the magnitude of overestimation was constant).

Linear speed and acceleration: The data are presented in Table 5. The accuracy to measure V_{peak} was good for all systems. V_{peak} measures from all systems were very largely correlated with gate V_{peak}. There was no substantial bias for Inmotio and GPSports, while Prozone (moderate) and VX (small) tended to over- and underestimate V_{peak}, respectively. The accuracy to measure acceleration (based on 10-m sprint time) was poor-to-moderate for all four systems (at least moderate bias and all $r<0.90$).

Change of direction speed and acceleration: The accuracy to measure change of direction speed (based on L-shaped and Zig-Zag-shaped runs sprint time) was poor-to-moderate for all four systems (at least moderate bias and all $r<0.90$).

Reliability.

Linear speed and acceleration: The CV data are presented in Figure 4 and ICC results in Table 6. The CVs for V_{peak} during the 40-m sprint were good for all systems (1-3%). The CVs for

Acc_{peak} during the 40-m sprint were lower for Inmotio (5%) than for the 3 other systems (17-23%).

Change of direction speed and acceleration: The CVs for V_{peak} during L-shaped runs were moderate for all systems, and tended to be lower for Inmotio and VX. The CVs for V_{peak} during the Zig-Zag shaped run were moderate and similar for all systems. The CVs for Acc_{peak} during the both change of direction speed tests were large and similar for all systems. Finally, the CVs for the high-intensity run were small and comparable for all systems.

Discussion

In this study, we collected simultaneously for the time data from three different tracking technologies used in professional soccer (i.e., semi-automatic multiple-camera, *Prozone*; local position measurement, *Inmotio* and GPS technology, *GPSports* and *VX*). We were therefore able to provide for the first time some calibration equations to allow the interchangeability of the three systems to measure on-field running movements in highly-trained young players. We also provided further assessment of the validity and reliability of all technologies. The main results are as follow: 1) metrics collected during either training (different pitch sizes) or a friendly match (full pitch) differed between the 4 systems, and the magnitude of these differences was related to both pitch size and the variable of interest; 2) most of the calibration equations calculated were still associated with small-to-moderate typical error of the estimate; 3) all systems tended to either overestimate or underestimate the distance covered at different speed (linear runs), with the magnitude of error being speed- and technology-dependent; 4) while the validity of the four systems was good to measure speed over 30-40 m, the level of accuracy was only moderate for shorter distances with/without changes of direction; 5) the correlations between peak acceleration 10-m sprint times were moderate-to-low; and 6) the CV for intra-day test-retest were low for peak speed, but increased for measures of acceleration and changes of direction speed.

Agreement and interchangeability. In the present study, we found that during the same training sessions and the same match, the different metrics collected by the 4 systems differed, and that the magnitude of these differences was related the variable of interest (Figure 2). To date, the agreement between multiple camera semi-automatic systems and GPS had only been examined twice; distance covered at high-intensity (i.e., >~18 km/h) was slightly-to-moderately (Harley, et al., 2011) and moderately-to-largely (Randers, et al., 2010) greater with the camera systems than with GPS technology. Present results (Figure 4) are in alignment with these latter findings, but show in addition that the magnitude of the differences is also likely related to pitch size. These findings could actually be expected, since i) pitch size affects directly distance ran at high speed (Hill-Haas, Dawson, Impellizzeri, & Coutts, 2011) and ii) we observed speed-related differences in accuracy (see below, Table 5 and Figure 3). Taken together, these data suggest that pitch size should be considered when calculating calibration equations (Table 2, 3 and 4). While matches are obviously played over the entire pitch, this is rarely the case during training sessions. It is worth noting that while GPS and LPM technologies are well suited to both match and training monitoring, Prozone is designed to track players' activity during matches only. Therefore, despite the adjustments made specifically for the present study (see methods), data collection and proceeding with Prozone during the training are unlikely to have matched the standard applied to the match situation. Based on both statistical and practical considerations, we therefore provided three levels of equations, i.e., for playing sequences within the box area (<30 x 30m), half a pitch (<50 x 50m) and full pitch (match). This is to our knowledge the first time that such equations have been offered, so comparisons with the literature are difficult. In practice, sports scientists willing to convert match running distances collected with Prozone into GPS-expected distances can use the equation provided in Table 4 (i.e., GPS = 1.01 Prozone - 70 m). If we take the example of a player who covered 9500 m in total during a match, a total distance of 9525 m may have been measured if he had worn a one of the GPSports units examined in the present study. Similarly, during a training sequence organized over half a pitch (Table 2), a distance of 150 m >14.4 km.h^{-1} measured with Inmotio may be equivalent to 77 m when measured with a GPSports unit (GPS = 0.5 Inmotio) - 2 m). While we trust that the actual sample size (n = 163, 205 and 22 for small, medium and fill area, respectively) was large enough to obtain accurate equations, further cross-validations studies with external data collected during similar training/matches are required to examine the level of accuracy and precision of each equation. In fact, there were only a few equations with a non-substantial TEE (Table 2, 3 and 4); the conversion is therefore unfortunately not perfect and that small-to-moderate error will remain. The reasons for this source of error remains unknown, but it might be related, among others, to possible differences in transmitter movements on players' back, variations in postural sway affecting differently the different systems, and likely the difference in accuracy (Table 5 and Figure 3) and reliability (Table 6 and Figure 4) between different systems. While averaging repeated measures is an attractive option to reduce the noise related to matches- (Gregson, Drust, Atkinson, & Salvo, 2010) and training- (Hill-Haas, Rowsell, Coutts, & Dawson, 2008) variations in activity patterns, the TEE as calculated here is related to the inconsistent differences between the different measurement systems. Repeating the measure might therefore not affect the magnitude of the between-subject error. Another important question to practitioners is whether the TEE for a given equation is of importance clinically (Carling, 2013). The standardization of the error as presented in Table 2, 3 and 4 is a first approach to try to answer this question; all errors at least small or moderate might have practical implications in terms of actual workload on the field. It is, however, worth noting that whether small standardized differences should actually be considered as the smallest practical effect with respect to matches and training activity is still unknown. The relationship between match physical activity and success is unclear (Carling, 2013) and the impact that running performance during matches and training may have on players fatigue (Nedelec, et al., 2012, 2013) and fitness (Rollo, Impellizzeri, Zago, & Iaia, 2013) is likely mediated by a myriad of external factors (e.g., weekly periodization and content of the other sessions, the use or not recovery strategies), which make their relationship difficult to examine. Finally, the latest versions of the GPS systems (e.g., GPsports SPI HPU), which were not available at the time of the data recording, have been anecdotally shown to provide more accurate measures of acceleration (unpublished data). The effect of changes in hardware (e.g., new GPS units) and/or software (e.g., update) and the validity of the provided equations should be the matter of further research (Buchheit, et al., 2014). Finally, between-brand comparisons within a similar type of technology (e.g., GPS Catapult vs. Statsports vs. GPSports and VX, Inmotio vs. ClearSky Catapult, or Amisco vs. Prozone) should also be the matter of future research (Johnston, Watsford, Kelly, Pine, & Spurrs, 2013, Petersen, Pyne, Portus, & Dawson, 2009).

Validity. The validity of the four systems was examined at different levels, i.e., distance covered at set speeds, and maximal acceleration and velocity reached during standardized drills. With respect to distance covered during linear runs over the 200-m oval, Prozone and GPSsports tended to moderately and largely-to-very-largely overestimate the distance covered at all speeds, respectively, while the direction and magnitude of the error was speed-dependent with Inmotio and VX (Figure 2). While there is no comparable data for Prozone or VX in the literature, large overestimations (+1.5%) of distance covered over a 200-m course were also reported with 1-Hz GPSports units at 18 km/h (Gray, et al., 2010). In contrast, at the same speed on a 400- track, 5-Hz GPSports units were shown to moderately underestimate running distance by 1 to 3% (Petersen, et al., 2009). Differences within (Buchheit, et al., 2014, Johnston, et al., 2013, Varley, et al., 2012) and between (Johnston, et al., 2013, Petersen, et al., 2009) different GPS brands have already been reported and might be related to several factors within the GPS itself or differences in data flittering and analysis within the associated software (Buchheit, et al., 2014). The measurement error with GPS technology was also shown to be moderately correlated with the number of satellites to which each unit was connected (Gray, et al., 2010). In the present study, however, since we wanted the data to be analyzed in a real life scenario, where coaches and practitioners have to deal with the actual set of data collected (with a perfectly clear sky at least), the number of satellites was not considered in the analysis. Finally, the moderate-to-large underestimation of running distance with Inmotio is consistent with previous results with the same system (Frencken, et al., 2010, Stevens, et al., 2014); however, because of the large differences in running protocols (e.g., running speed, runs including changes of direction or not), results can't be directly compared.

With respect to the standardized sprinting drills, despite some small (VX) and moderate (Prozone) bias, the validity of the four systems was acceptable to measure V_{peak} over 30-40 m (all r > 0.80 and small TEE, Table 2). The validity was, however, only moderate for all systems when examining shorter distances with/without changes of direction (r: unclear to 0.7, moderate TEE, Table 2). It is, however, important to mention that in the present study we focused our analysis on the V_{peak} and Acc_{peak} detected by each system, rather than the average speed over the set distances (see methods). This is probably a reason for some of the poorer correlations coefficients and/or TEE compared with the literature. For example, Di Salvo reported almost perfect correlations for Prozone vs. timing gates average speed over different sprinting distances (Di Salvo, et al., 2006). In this latter study, the authors also reported summary correlation (including different sprint distances), which increases the magnitude of the correlation. With these data pooled and the limited number of data sets (i.e., n = 3 per distance (Di Salvo, et al., 2006)), examining the specific acceleration results (or at least time/speed over a short distance) is limited. A further analysis (Di Salvo, et al., 2009) of the same data (Di Salvo, et al., 2006) reported a mean bias (compared with timing gates) in average velocity for all runs pooled together of 0.4%. This lower bias compared with these reported in the present study (2.7 to 5.2 %, Table 5) may also be related to the simultaneous analysis of all runs pooled together, in comparison with the more practical within-run analysis of the present study. The low TEE (2-3%) for V_{peak} during the sprinting drills measured with Inmotio (Table 5) is similar to the data reported by Frencken et al. (Frencken, et al., 2010), who used average speed during both linear and change of direction sprints. The correlation coefficient for the majority of the drills were, however, lower that those reported in this latter study (Frencken, et al., 2010); this could be explained by the fact that instantaneous speed (as V_{peak} measured here) is far less reliable than average velocity (Ogris, et al., 2012). Acceleration measures collected with Inmotio have been recently compared with data obtained with high-speed cameras (i.e., Vicon system), which allow a continuous monitoring of players' movements (Stevens, et al., 2014). Results showed that the validity of Inmotio to measure accelerations was protocol-dependent. Peak acceleration was underestimated by 9% during runs involving 180º changes of direction, but overestimated up to 16% during movements with 90º changes of direction. Standard errors of the estimate ranged between 5 and 15% for average and peak accelerations, which tends to be greater than what we reported in the present study (i.e., <5%, Table 5). However, as discussed above for V_{peak}, care should be taken when comparing different methods of assessment, i.e., average acceleration with timing gates vs. peak acceleration with cameras. Finally, the GPS results are in agreement with recent studies showing a better accuracy of GPS for linear vs. change of direction speed (Gray, et al., 2010, Jennings, et al., 2010a, Vickery, et al., 2013). This is likely related to the fact that GPS might underestimate distance during non-linear movement due to an insufficient update rate and lean angle (standardized difference >3 to 5 when using 1-Hz GPS) (Gray, et al., 2010). The poor correlations and moderate TEE when comparing GPS-derived Acc_{peak} (in m.s^{-2}) and 10-m sprint time (Table 5) is consistent with the results from previous studies comparing GPS with timing gates (Waldron, et al., 2011) or speed radar (Varley, et al., 2012). Recent studies have also highlighted large between-unit variations in the different metrics, which is an indirect evidence of limited accuracy (Buchheit, et al., 2014, Jennings, Cormack, Coutts, Boyd, & Aughey, 2010b, Johnston, et al., 2013, Petersen, et al., 2009). The reasons for these sources of between-unit variations remain unclear.

Reliability. When examining the repeated high-speed (18 km/h) run, the CVs and ICC were small-to-moderate and low, respectively (Figure 4 and Table 6). To our knowledge, similar test-retest analyses have not been reported for Inmotio or Prozone, but the present GPS CVs (1.5% for GPSports, 3.3% for VX) are within the 1-3% range of these reported with 1- and 5-Hz devices (Petersen, et al., 2009). The greater CV for speed during the sprinting drills measured by Prozone in the present study (2 and 7% for V_{peak} and Acc_{peak}, respectively) compared with those in the Di Salvo study (~1%) (Di Salvo, et al., 2006) are likely related to their use of instantaneous velocity data (Ogris, et al., 2012). We could not locate any study on the reliability of the LPM technology. Present sprint data confirm that, irrespective of the measurement system, V_{peak} tends to be more reproducible than 10-m sprint time and change of direction speed. Present results also confirm that acceleration measures based on average speed (i.e, 10-m time, <2%) are more reliable than instantaneous velocity data (Ogris, et al., 2012): except for Inmotio (~5%), the CV for Acc_{peak} with Prozone and the GPS ranged between 16 and 25%. These data are in line with the large CV and low ICC reported for acceleration measures during short sprints and various racquet- and team sport-specific movement patterns (Vickery, et al., 2013). In previous studies where acceleration-derived variables showed lower CVs, the peak speed reached in the drills was used instead as a proxy of acceleration (<3% (Waldron, et al., 2011)), and/or instantaneous GPS data were compared with instantaneous speed radar data (<5% (Varley, et al., 2012)). Together with the limited validity of acceleration measures (see above), these data question the use of Acc_{peak} as provided routinely by the different systems (Buchheit, et al., 2014, Stevens, et al., 2014, Waldron, et al., 2011).

Conclusion

In practice, all systems are of interest to monitor players' activity in the field, and each has advantages and disadvantages (Table 7). The metrics collected during different training sessions on different pitch sizes or during a friendly match on the full pitch differ between the four systems, and the magnitude of the between-system differences is related to both pitch size and the variable of interest. These between-systems differences in running distance at set speeds can nevertheless be partially accounted for at the team level, using the calibration equations provided in this document. Because of the remaining small-to-moderate typical errors of the estimate in these equations, caution is still required when comparing/interpreting training/match activities measured by different systems. Whether the small-to-large between-system differences both in accuracy and running performance during matches/training have practical implications still remains to be determined. An important limitation of Inmotio and Prozone systems compared with GPS technology is that they can only be used on the football fields where they has been previously installed. Conversely, they can eventually be installed and used indoor and within stadiums where the use of GPS is limited (Table 7). The great interest of Prozone is that it can be used during matches, allow the collection of additional technical variables, and doesn't require any equipment to be worn by the players. Finally, GPS represent today the most time-efficient technology with respect to data management and reports. Between- GPS unit variability, within-unit chipset changes and software updates can have a substantial impact on the different metrics (Buchheit, et al., 2014). It is therefore paramount to consider these factors when monitoring players over the time.

References

Buchheit, M., Mendez-Villanueva, A., Simpson, B. M., & Bourdon, P. C. (2010). Match Running Performance and Fitness in Youth Soccer. *Int J Sports Med, 31*, 818-825.
Buchheit, M., Al Haddad, H., Simpson, B. M., Palazzi, D., Bourdon, P. C., Di Salvo, V., & Mendez-Villaneuva, A. (2014). Monitoring accelerations with GPS in football: time to slow down? *Int J Sports Physiol Perform, 9*, 442-445.
Carling, C., Bloomfield, J., Nelsen, L., & Reilly, T. (2008). The role of motion analysis in elite soccer: contemporary performance measurement techniques and work rate data. *Sports Med, 38*, 839-862.
Carling, C. (2013). Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? *Sports Med, 43*, 655-663.
Casamichana, D., Castellano, J., Calleja-Gonzalez, J., San Roman, J., & Castagna, C. (2013). Relationship between indicators of training load in soccer players. *J Strength Cond Res, 27*, 369-374.
Di Salvo, V., Collins, A., MacNeill, B., & Cardinale, M. (2006). Validation of Prozone ® :A new video-based performance analysis system *Int J Perf Anal Sport*, 108 - 119.
Di Salvo, V., Baron, R., Tschan, H., Calderon Montero, F. J., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *Int J Sports Med, 28*, 222-227.
Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2009). Analysis of high intensity activity in Premier League soccer. *Int J Sports Med, 30*, 205-212.
Frencken, W. G., Lemmink, K. A., & Delleman, N. J. (2010). Soccer-specific accuracy and validity of the local position measurement (LPM) system. *J Sci Med Sport, 13*, 641-645.
Gray, A. J., Jenkins, D., Andrews, M. H., Taaffe, D. R., & Glover, M. L. (2010). Validity and reliability of GPS for measuring distance travelled in field-based team sports. *J Sports Sci, 28*, 1319-1325.
Gregson, W., Drust, B., Atkinson, G., & Salvo, V. D. (2010). Match-to-match variability of high-speed activities in premier league soccer. *Int J Sports Med, 31*, 237-242.
Harley, J. A., Lovell, R. J., Barnes, C. A., Portas, M. D., & Weston, M. (2011). The interchangeability of global positioning system and semiautomated video-based performance data during elite soccer match play. *J Strength Cond Res, 25*, 2334-2336.
Hill-Haas, S., Rowsell, G., Coutts, A., & Dawson, B. (2008). The reproducibility of physiological responses and performance profiles of youth soccer players in small-sided games. *Int J Sports Physiol Perform, 3*, 393-396.
Hill-Haas, S. V., Dawson, B., Impellizzeri, F. M., & Coutts, A. J. (2011). Physiology of small-sided games training in football: a systematic review. *Sports Med, 41*, 199-220.
Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc, 41*, 3-13.
Jennings, D., Cormack, S., Coutts, A. J., Boyd, L., & Aughey, R. J. (2010a). The validity and reliability of GPS units for measuring distance in team sport specific running patterns. *Int J Sports Physiol Perform, 5*, 328-341.
Jennings, D., Cormack, S., Coutts, A. J., Boyd, L. J., & Aughey, R. J. (2010b). Variability of GPS units for measuring distance in team sport movements. *Int J Sports Physiol Perform, 5*, 565-569.
Johnston, R. J., Watsford, M. L., Kelly, S. J., Pine, M. J., & Spurrs, R. W. (2013). The Validity and reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. *J Strength Cond Res*, Nov 22. [Epub ahead of print].
Nedelec, M., McCall, A., Carling, C., Legall, F., Berthoin, S., & Dupont, G. (2012). Recovery in soccer: part I - post-match fatigue and time course of recovery. *Sports Med, 42*, 997-1015.

Nedelec, M., McCall, A., Carling, C., Legall, F., Berthoin, S., & Dupont, G. (2013). The influence of soccer playing actions on the recovery kinetics after a soccer match. *J Strength Cond Res,* [Published Ahead of Print].
Ogris, G., Leser, R., Horsak, B., Kornfeind, P., Heller, M., & Baca, A. (2012). Accuracy of the LPM tracking system considering dynamic position changes. *J Sports Sci*, *30*, 1503-1511.
Petersen, C., Pyne, D., Portus, M., & Dawson, B. (2009). Validity and reliability of GPS units to monitor cricket-specific movement patterns. *Int J Sports Physiol Perform, 4,* 381-393.
Portas, M. D., Harley, J. A., Barnes, C. A., & Rush, C. J. (2010). The validity and reliability of 1-Hz and 5-Hz global positioning systems for linear, multidirectional, and soccer-specific activities. *Int J Sports Physiol Perform, 5*, 448-458.
Randers, M., Mujika, I., Hewitt, A., Santisteban, J., Bischoff, R., Solano, R., Zubillaga, A., Peltola, E., Krustrup, P., & Mohr, M. (2010). Application of four different football match analysis systems: A comparative study. *J Sports Sci,* [published ahead of print].
Rollo, I., Impellizzeri, F. M., Zago, M., & Iaia, F. M. (2013). Effects of One Versus Two Games a Week on Physical and Subjective Scores of Sub-Elite Soccer Players. *Int J Sports Physiol Perform,* [Published ahaed of print].
Scott, B. R., Lockie, R. G., Knight, T. J., Clark, A. C., & Janse de Jonge, X. A. (2013). A comparison of methods to quantify the in-season training load of professional soccer players. *Int J Sports Physiol Perform*, *8*, 195-202.
Siegle, M., Stevens, T., & Lames, M. (2013). Design of an accuracy study for position detection in football. *J Sports Sci*, *31*, 166-172.
Stevens, T. G., de Ruiter, C. J., van Niel, C., van de Rhee, R., Beek, P. J., & Savelsbergh, G. J. (2014). Measuring Acceleration and Deceleration in Soccer-Specific Movements Using a Local Position Measurement (LPM) System. *Int J Sports Physiol Perform*, *9*, 446-456.
Varley, M. C., Fairweather, I. H., & Aughey, R. J. (2012). Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. *J Sports Sci*, *30*, 121-127.
Vickery, W. M., Dascombe, B. J., Baker, J. D., Higham, D. G., Spratford, W. A., & Duffield, R. (2013). Accuracy and reliability of GPS devices for measurement of sports-specific movement patterns related to cricket, tennis and field-based team sports. *J Strength Cond Res,* Oct 21. [Epub ahead of print].
Waldron, M., Worsfold, P., Twist, C., & Lamb, K. (2011). Concurrent validity and test-retest reliability of a global positioning system (GPS) and timing gates to assess sprint performance variables. *J Sports Sci*, *29*, 1613-1619.
Winter, E. M., & Maughan, R. J. (2009). Requirements for ethics approvals. *J Sports Sci, 27,* 985.

Figure Legends

Figure 1. Illustration of the different test protocols. Arrows show tests start and running direction. Black circles refer to timing gates locations. Grey circles refer to cones.

Figure 2. Mean total distance, distance ran above 14.4 $km.h^{-1}$, peak speed and the number of accelerations above 3 $m.s^{-2}$ (90% confidence intervals) measured by each tracking system during a friendly match played on the full pitch, and training sequences over medium or small pitch areas. The letters refer to substantial difference compared with the other system with 'p', 'g' and 'v' referring to Prozone, GPSport and VX, respectively. The number of letters indicates small, moderate, large or very large differences, respectively. n = 16, 205 and 163 for full (match), medium or small pitch areas, respectively. Only clear (at least possible) differences are reported.

Figure 3. Standardized bias (90% confidence intervals) in distance covered at 3 different set speeds for each tracking system. The number of * symbols indicate possible, likely, very likely and almost certain differences from the criterion measure, respectively. The grey area represents trivial difference (0.2 x pooled between-player SD). n = 14 for runs at 7.2 and 14.4 $km.h^{-1}$, and n = 28 for run at 19.8 $km.h^{-1}$.

Figure 4. Coefficients of variation (90% confidence intervals) for test-retest of different speed-related performance measures with each tracking system. Acc peak: peak acceleration reached during the test; Vpeak: peak velocity reached during the test; High-speed run refer to the distance covered within 16 s at 19.8 $km.h^{-1}$. Note that all measures were not available for all

systems, i.e., sprint times can only be computed with timing gates and Inmotio, and acceleration wasn't computed with timing gates (since based on a fixed time window with the 4 other systems). The number of '*' symbols stands for small and moderate typical error, respectively.

Table 1. Description of the different test protocols, variables measured and analyses conducted.

	Agreement between the systems	Validity	Reliability
Activity	Four training sessions and a friendly match	A generic sprint session (all tests repeated twice)	
		1. Runs on an oval 200-m course • 7.2 km/h • 14.4 km/h • 19.8 km/h 2. Standardized sprints • 40-m sprint • L-shaped sprint • ZigZag shaped sprint	
Measures	Distance into speed zones	1. Distance into speed zones during the runs 2. Peak speed and acceleration (4 systems) and sprint times (timing gates and Inmotio)	
Analysis	Standardized between-system differences and calibration equations with TEE and correlation coefficients	Standardized bias compared with timing gates, TEE and correlation coefficients	Standardized CV and ICC using data from the repeated trials within each system

TEE: typical error of the estimate; CV: coefficient of variation, ICC: intraclass correlation coefficient

Table 2. Calibration equations between the three different technologies during training within a small area (<20-30 m^2) for total distance covered and distance ran above 14.4 $km.h^{-1}$

		Inmotio	Prozone
TD	**Prozone**	P = (0.64 In) + 318 m TEE: 34.8% (31.5;39.0) ## r : 0.55 (0.45;0.64) **	
		In = (0.61 P + 261 m TEE: 45.5% (41.0;51.1) ## r : 0.55 (0.45;0.64) **	
	GPSports	G = (1.02 In) - 103 m TEE: 26.4% (23.9;29.4) # r : 0.94 (0.92;0.95) *****	G = (0.64 P) + 142 m TEE: 81.7% (72.9;93.0) ## r : 0.50 (0.40;0.59) **
		In = (0.82 G) + 207 m TEE: 16.4% (15.0;18.3) # r : 0.94 (0.92;0.95)*****	P = (0.54 G) + 437 m TEE: 36.4% (32.9;40.7) ## r : 0.50 (0.40;0.59) **
D> 14.4 $km.h^{-1}$	**Prozone**	P = (1.11 In) + 37 m TEE: 200.0% (183.6;220.7) ## r : 0.54 (0.44;0.62) **	
		In = (0.26 P) + 12 m TEE: 41.1% (37.6;54.2) ## r : 0.54 (0.44;0.62) **	
	GPSports	G = (0.50 In) - 2 m TEE: 61.6% (56.4;67.9) ## r : 0.68 (0.61;0.75) ***	G = (-1.4 P) + 754 m TEE: 381.1% (349.3;420.0) ## r : -0.35 (-0.46;-0.23) *
		In = (0.94 G) + 14 m TEE: 151.5% (138.9;167.0) ## r : 0.68 (0.61;0.75) ***	P = (-0.09 G) + 128 m TEE: 10.2% (9.3;11.2) ## r : -0.35 (-0.46;-0.23) *

TD: total distance, D> 14.4 $km.h^{-1}$: distance covered above 14.4 $km.h^{-1}$, In the equations, P, In and G refer to Prozone, Inmotio and GPSports, respectively. The number of '#' symbols stand for small, moderate, large and very large standardized typical error of the estimate, respectively. The number of '*' symbols refers to small, moderate, large, very large and almost perfect correlations between the two systems, respectively. n = 162

Table 3. Calibration equations between the three different technologies during training within a medium area (<50 m2) for total distance covered, distance ran above 14.4 km.h^{-1} and 19.8 km.h^{-1}. Given the good agreement between GPSports and VX measures, and the higher number of files available for GPSports, GPSports data were used as a generic measure for GPs technology.

		Inmotio	Prozone
TD	**Prozone**	P = (0.90 In) + 64 m TEE: 72.6% (65.7;81.3) ## r : 0.67 (0.60;0.73) ***	
		In = (0.73 P) + 348 m TEE: 43.3% (39.5;48.0) ## r : 0.67 (0.60;0.73) ***	
	GPSports	G = (0.97 In) - 67 m TEE: 14.3% (13.1;15.7) # r : 0.97 (0.96;0.97) *****	G = (0.73 P) + 257 m TEE: 51.1% (46.5;56.8) ## r : 0.63 (0.56;0.70) ***
		In = (0.98 G) + 122 m TEE: 12.8% (11.8;14.0) # r : 0.97 (0.96;0.97) *****	P = (0.90 G) + 155 m TEE: 76.5% (69.1;85.7) ## r : 0.63 (0.56;0.70) ***
D> 14.4 km.h^{-1}	**Prozone**	P = (0.93 In) + 48 m TEE: 68.0% (62.9;74.1) ## r : 0.70 (0.63;0.75) ****	
		In = (0.52 P) + 51 m TEE: 48.0% (37.7;44.4) ## r : 0.70 (0.63;0.75) ****	
	GPSports	G = (0.87 In) - 9 m TEE: 26.0% (24.1;28.4) # r : 0.92 (0.90;0.94) *****	G = (0.46 P) + 34 m TEE: 40.6% (37.6;44.3) ## r : 0.65 (0.58;0.71) ***
		In = (0.98 G) + 31 m TEE: 34.2% (31.6;37.3) # r : 0.92 (0.90;0.94) *****	P = (0.92 G) + 75 m TEE: 88.8% (82.2;96.8) ## r : 0.65 (0.58;0.71) ***
D> 19.8km.h^{-1}	**Prozone**	P = (0.84 In) + 51 m TEE: 109.8% (101.6;119.7) ## r : 0.60 (0.52;0.67) ***	
		In = 0.43 P + 13 m TEE: 54.5% (50.4;59.4) ## r : 0.60 (0.52;0.67) ***	
	GPSports	G = 0.74 In + 0 m TEE: 41.8% (38.7;45.6) ## r : 0.87 (0.84;0.89) ****	G = 0.32 P + 7 m TEE: 49.7% (45.9;54.1) ## r : 0.52 (0.43;0.60) ***
		In = 1.01 G + 9 m TEE: 66.1% (61.2;72.1) ## r : 0.87 (0.84;0.89) ****	P = 0.86 G + 28 m TEE: 158.2% (146.3;172.4) ## r : 0.52 (0.43;0.60) ***

TD: total distance, D> 14.4 km.h^{-1}: distance covered above 14.4 km.h^{-1}, D> 19.8 km.h^{-1}: distance covered above 19.8 km.h^{-1}. The number of '#' symbols stands for small, moderate, large and very large standardized typical error of the estimate, respectively. The number of '*' symbols refers to small, moderate, large, very large and almost perfect correlations between the two systems, respectively. n = 205.

Table 4. Calibration equations between the three different technologies during match play for total distance covered, distance ran above 14.4 km.h^{-1} and 19.8 km.h^{-1}

		Inmotio	Prozone
TD	**Prozone**	P = (1.03 In) + 14 m TEE: 3.0% (2.4;4.0) r : 0.99 (0.97;0.99) *****	
		In = (0.94 P) + 62 m TEE: 2.9% (2.3;4.0) r : 0.99 (0.97;0.99) *****	
	GPSports	G = (1.06 In) - 99 m TEE: 3.0% (2.4;4.0) # r : 0.97 (0.93;0.98) *****	G = (1.01 P) -70 m TEE: 5.4% (4.2;7.3) # r : 0.96 (0.92;0.98) *****
		In = (0.89 G) + 263 m TEE: 4.8% (3.8;6.6) # r : 0.97 (0.93;0.98) *****	P = (0.92 G) + 250 m TEE: 4.8% (3.8;6.6) r : 0.96 (0.92;0.98) *****
D> 14.4 km.h^{-1}	**Prozone**	P = (0.98 In) + 167 m TEE: 11.3% (8.9;15.7) # r : 0.91 (0.82;0.96) *****	
		In = (0.87 P) - 68 m TEE: 13.9% (11.0;14.4) # r : 0.91 (0.82;0.96) *****	
	GPSports	G = (0.81 In) + 77 m TEE: 23.2% (18.1;32.8) # r : 0.80 (0.62;0.90) ****	G = (0.75 P) -12 m TEE: 23.5% (14.8;33.2) ## r : 0.80 (0.61;0.90) ****
		In = (0.89 G) + 86 m TEE: 20.6% (16.1;29.0) # r : 0.80 (0.62;0.90) ****	P = (0.94 G) + 220 m TEE: 16.8% (13.2;23.5) ## r : 0.80 (0.61;0.90) ****
D> 19.8km.h^{-1}	**Prozone**	P = (1.05 In) + 77 m TEE: 17.0% (13.3;23.7) r : 0.91 (0.81;0.95) ******	
		In = (0.81 P) - 42 m TEE: 27.2% (21.2;38.6) r : 0.91 (0.81;0.95) ******	
	GPSports	G = (0.67 In) + 29 m TEE: 37.9% (29.2;54.7) r : 0.82 (0.65;0.91) *****	G = (0.58 P) - 7 m TEE: 42.3% (32.5;61.5) ## r : 0.77 (0.57;0.89) ****
		In = (0.96 G) + 23 m TEE: 38.9% (30.0;56.2) r : 0.82 (0.65;0.91) *****	P = (1.06 G) + 94 m TEE: 26.5% (20.6;37.6) ## r : 0.77 (0.57;0.89) ****

TD: total distance, D> 14.4 km.h^{-1}: distance covered above 14.4 km.h^{-1}, D> 19.8 km.h^{-1}: distance covered above 19.8 km.h^{-1}. The number of '#' symbols stands for small, moderate, large and very large standardized typical error of the estimate, respectively. The number of '*' symbols refers to small, moderate, large, very large and almost perfect correlations between the two systems, respectively. n = 22.

Table 5. Validity of the four tracking systems as compared with timing gates.

	Inmotio			Prozone		GPSport		VX	
Criterion (Gates)	Time	Acc_{Peak}	V_{peak}	Acc_{Peak}	V_{peak}	Acc_{Peak}	V_{peak}	Acc_{Peak}	V_{peak}
10-m time	B: -3.2% (-4.2;-3.2)## TEE: 3.1% (2.5;4.1)## r: 0.53 (0.23;0.74)***	TEE: 3.2% (2.6;4.3)## r: -0.49 (-0.71;-0.18)**	—	TEE: 3.6% (2.9;4.8)## r: -0.19 (-0.50;0.16)	—	TEE: 3.7% (3.0;4.9)## r: -0.06 (-0.39;0.29)	—	TEE: 3.7% (2.9;4.9)## r: 0.11 (-0.23;0.43)	—
30-m time	**B: -0.7% (-1.2;-0.2)** **TEE: 1.4% (1.2;1.9)#** **r: 0.93 (0.86;0.96)*******	—	—	—	—	—	—	—	—
V_{peak} (40-m sprint)	—	—	**B: 0.4 % (0.0;0.9)** **TEE: 1.2% (1.0;1.6)#** **r: 0.98 (0.95;0.99)******	—	B: 5.2 % (4.0;6.4)## TEE: 3.3% (2.6;4.4)# r: 0.82 (0.66;0.90)****	—	B: 0.3 % (-0.9;1.5) TEE: 3.3% (2.7;4.5)# r: 0.82 (0.65;0.91)****	—	B: 1.2% (0.1;2.4)# TEE: 3.4% (2.7;4.5)# r: 0.80 (0.64;0.90)****
L-shaped Time	B: -1.9% (-2.4;-1.4)## TEE: 1.6% (1.3;2.1)# r: 0.82 (0.67;0.91)****	TEE: 2.7% (2.2;3.5)## r: -0.29 (-0.56;0.05)	TEE: 1.9% (1.5;2.5)## r: -0.73 (-0.85;-0.53)****	TEE: 2.7% (2.2;3.5)## r: -0.29 (-0.57;0.04)	TEE: 2.7% (2.2;3.5)## r: -0.29 (-0.56;0.05)	TEE: 2.6% (2.1;3.5)## r: -0.35 (-0.61;-0.02)**	TEE: 2.1% (1.7;2.7)## r: -0.67 (-0.82;-0.43)**	TEE: 2.9% (2.3;3.8)## r: -0.16 (-0.48;0.20)	TEE: 2.6% (2.1;3.5)## r: -0.43 (-0.68;-0.10)**
Zizag-shaped Time	B: -0.1% (-0.9;-0.8) TEE: 2.6% (2.1;3.4)## r: 0.43 (0.11;0.64)**	TEE: 2.8% (2.2;3.8)## r: -0.20 (-0.51;0.16)	TEE: 2.8% (2.2;3.8)## r: 0.18 (-0.18;0.49)	TEE: 2.8% (2.3;3.8)## r: 0.05 (-0.30;0.38)	TEE: 2.8% (2.3;3.8)## r: -0.11 (-0.44;0.24)	TEE: 2.8% (2.3;3.8)## r: -0.12 (-0.45;0.23)	TEE: 2.9% (2.3;3.8)## r: 0.02 (-0.33;0.36)	TEE: 2.8% (2.3;3.8)## r: -0.09 (-0.43;0.26)	TEE: 2.8% (2.3;3.7)## r: 0.24 (-0.12;0.54)

V_{peak}: peak velocity reached during the test; Acc_{peak}: peak acceleration; B: mean bias (90% confidence limits); TEE: typical error of the estimate (90% confidence limits); r: correlation coefficient (90% confidence limits). The number of '#' symbols stand for small, moderate, large and very large standardized bias and TEE, respectively. For r values, the number of '*' symbols refers to small, moderate, large, very large and almost perfect correlations, respectively. In bold are the measures deemed to be clearly valid, i.e., non substantial bias, small TEE and r>0.90.

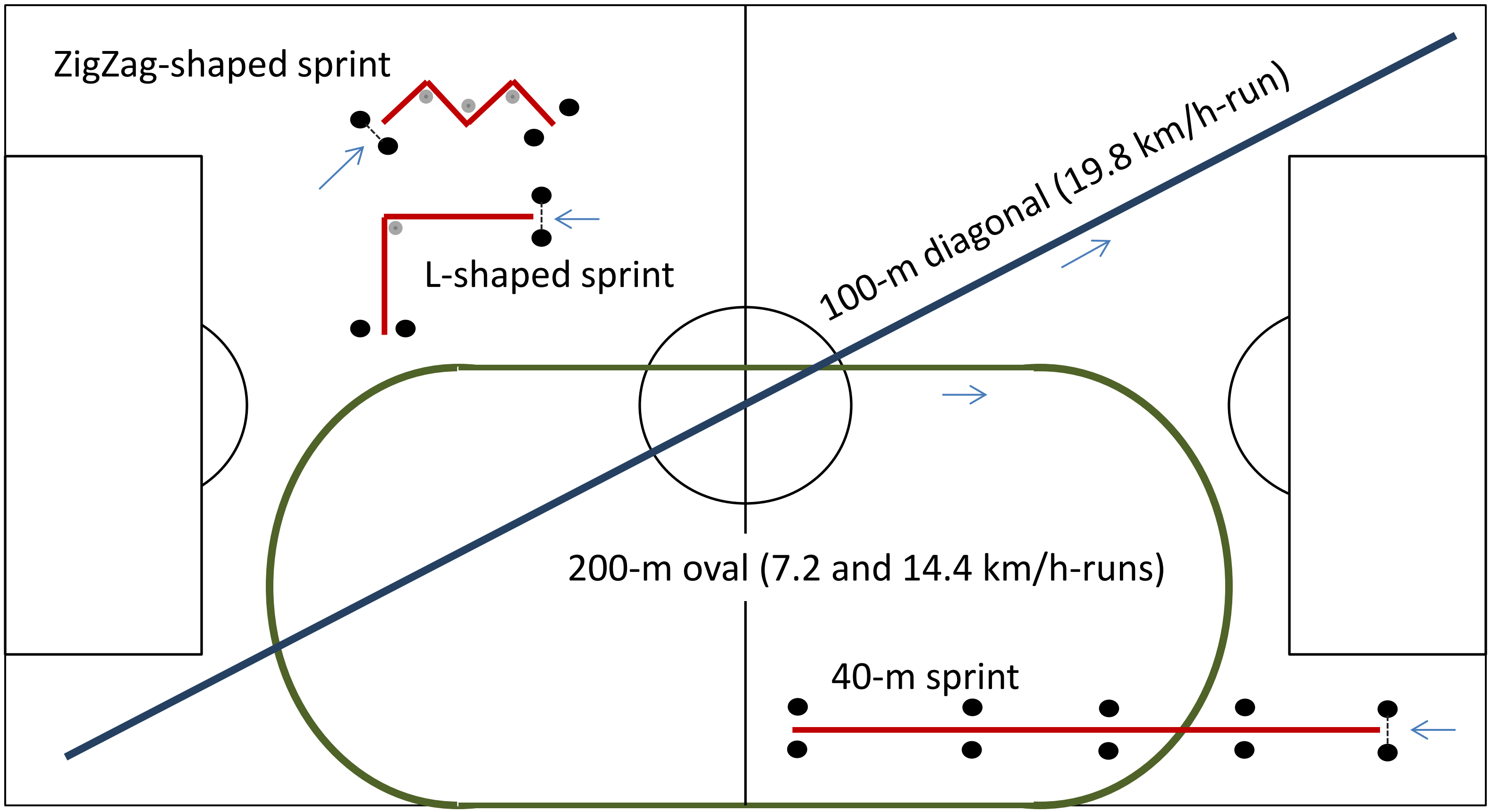
Table 6. Relative reliability for timing gates and the four tracking systems.

	Timing Gates	Inmotio	Prozone	GPSport	VX
10 m time	0.81 (0.48;0.93)***	0.70 (0.33;0.88)**			
20 m time	0.87 (0.65;0.96)***	0.80 (0.53;0.93)***			
30 m time	0.91 (0.73;0.97)****	0.91 (0.73;0.97)****			
40 m time	0.94 (0.83;0.98)****	0.94 (0.83;0.98)****			
Acc_{peak}		0.49 (0.02;0.78)*	-0.17 (-0.57;0.30)	-0.07 (-0.50;0.38)	0.04 (-0.38;0.50)
V_{peak}	0.97 (0.90;0.99)****	0.97 (0.90;0.99)****	0.88 (0.54;0.96)***	0.92 (0.86;0.92)****	0.97 (0.91;0.99)****
L-shaped time	0.60 (0.17;0.84)**	0.65 (0.30;0.85)**			
L-shaped Acc_{peak}		0.38 (-0.08;0.70)*	0.24 (-0.23;0.61)*	0.27 (-0.20;0.63)*	-0.05 (-0.50;0.42)
L-shaped V_{peak}		0.32 (-0.14;0.62)*	0.45 (0.01;0.74)*	0.07 (-0.38;0.50)	0.66 (0.29;0.86)**
Zig-Zag-shaped time	0.48 (-0.03;0.79)*	0.28 (-0.20;0.66)*			
Zig-Zag -shaped Acc_{peak}		0.21 (-0.27;0.61)*	0.43 (-0.02;0.73)*	0.36 (-0.12;0.70)*	-0.07 (-0.50;0.38)
Zig-Zag -shaped V_{peak}		-0.09 (-0.53;0.38)	0.30 (-0.16;0.65)*	0.61 (0.22;0.84)**	0.41 (-0.05;0.73)*
High-speed run distance		0.28 (-0.18;0.64)*	-0.47 (-0.76;0.04)*	0.20 (-0.26;0.59)*	-0.31 (-0.66;0.15)*

Intraclass correlation coefficient (90% confidence intervals) for test-retest of different speed-related performance measures with each tracking system. Acc_{peak}: peak acceleration reached during the test; V_{peak}: peak velocity reached during the test; High-speed run refer to the distance covered within 16 s at 19.8 km.h^{-1}. ICC rating: *:low, **:moderate, ***high, ****:very high.

Table 7. Practical considerations when using the four systems.

	Equipment	Possible Location(s)	Use	Data	Time to get reports	Database	Miscellaneous
Inmotio	Vest	Home training ground (outdoor and indoor)	Training + Friendly matches	Physical	~2-3 h	None	Sampling frequency related to player # Metrics likely affected by software/unit update
Prozone	None	Stadium(s) (outdoor and indoor)	Official matches	Physical + Match technical events	24 h	Extended	Players can't leave the pitch
GPSports	Vest	Anywhere (outdoor)	Training + Friendly matches	Physical	~30 min	Extended	Metrics affected by software/unit update
VX	Vest	Anywhere (outdoor)	Training + Friendly matches	Physical	~30 min	Extended	Metrics affected by software/unit update


ZigZag-shaped sprint
L-shaped sprint
100-m diagonal (19.8 km/h-run)
200-m oval (7.2 and 14.4 km/h-runs)
40-m sprint

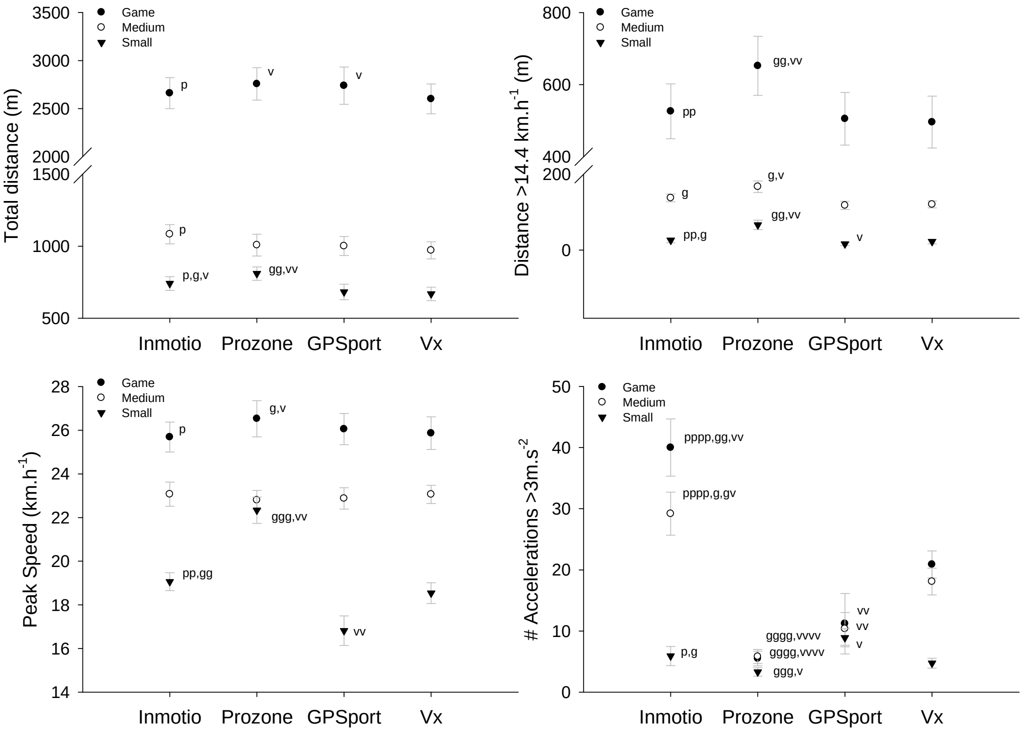


Figure 2. Mean total distance, distance ran above 14.4 km.h^{-1}, peak speed and the number of accelerations above 3 m.s^{-2} (90% confidence intervals) measured by each tracking system during a friendly match played on the full pitch, and training sequences over medium or small pitch areas. The letters refer to substantial difference compared with the other system with ‘p’, ‘g’ and ‘v’ referring to Prozone, GPSport and VX, respectively. The number of letters indicates small, moderate, large or very large differences, respectively. n = 16, 205 and 163 for full (match), medium or small pitch areas, respectively. Only clear (at least possible) differences are reported.

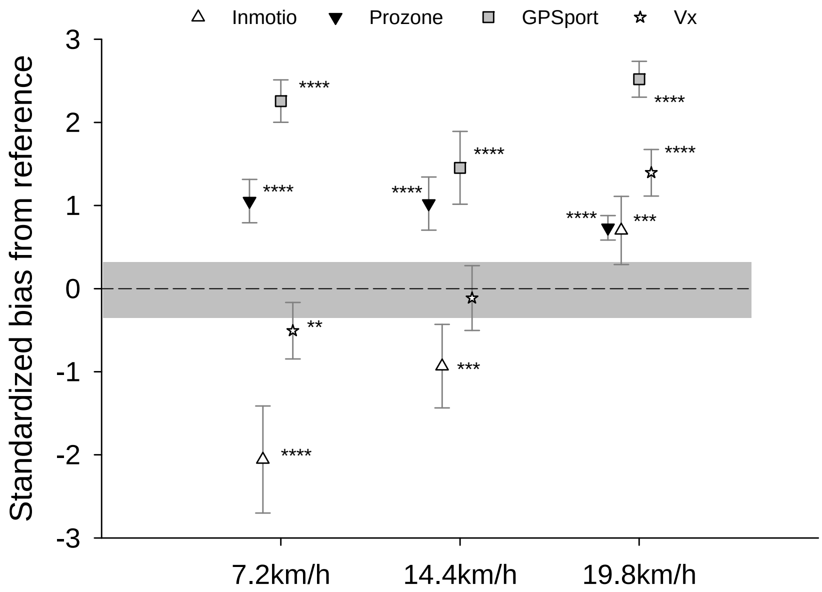


Figure 3. Standardized bias (90% confidence intervals) in distance covered at 3 different set speeds for each tracking system. The number of * symbols indicate possible, likely, very likely and almost certain differences from the criterion measure, respectively. The grey area represents trivial difference (0.2 x pooled between-player SD). n = 14 for runs at 7.2 and 14.4 $km.h^{-1}$, and n = 28 for run at 19.8 $km.h^{-1}$.

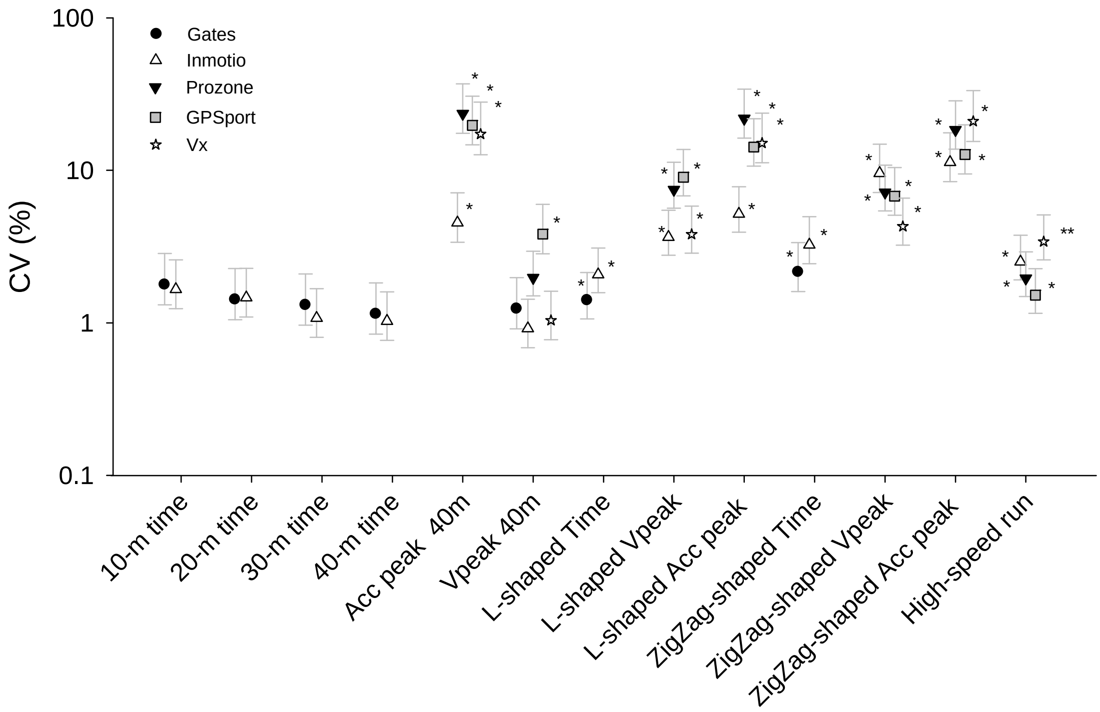


Figure 4. Coefficients of variation (90% confidence intervals) for test-retest of different speed-related performance measures with each tracking system. Acc peak: peak acceleration reached during the test; Vpeak: peak velocity reached during the test; High-speed run refer to the distance covered within 16 s at 19.8 km.h^{-1}. Note that all measures were not available for all systems, i.e., sprint times can only be computed with timing gates and Inmotio, and acceleration wasn't computed with timing gates (since based on a fixed time window with the 4 other systems). The number of '*' symbols stands for small and moderate typical error, respectively.