Durham Research Online

Deposited in DRO:

15 July 2021

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Oliva-Lozano, José M. and Muyor, José M. and Fortes, Víctor and McLaren, Shaun J. (2021) 'Decomposing the variability of match physical performance in professional soccer: Implications for monitoring individuals.', *European Journal of Sport Science*, 21 (11). pp. 1588-1596.

Further information on publisher's website:

<https://doi.org/10.1080/17461391.2020.1842513>

Publisher's copyright statement:

This is an Accepted Manuscript version of the following article, accepted for publication in *European Journal of Sport Science*. Oliva-Lozano, José M., Muyor, José M., Fortes, Víctor McLaren, Shaun J. (2021). Decomposing the variability of match physical performance in professional soccer: Implications for monitoring individuals. *European Journal of Sport Science* 21(11): 1588-1596.. It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

- 1 **DECOMPOSING THE VARIABILITY OF MATCH PHYSICAL PERFORMANCE IN**
- 2 **PROFESSIONAL SOCCER: IMPLICATIONS FOR MONITORING INDIVIDUALS**

ABSTRACT

The aims of this study were to establish sources of variability in match physical performance of professional soccer players and provide a method for monitoring individual between-match changes. Eleven players meeting the final inclusion criteria were monitored through an entire in-season competition phase ($n = 240$ individual match observations, median [range] match observations per player = 21 [15–31]). Ten-Hertz global positioning systems were used to measure match total distance (TD), total high-speed running distance ($\geq 21 \text{ km}\cdot\text{h}^{-1}$; HSRD), total accelerations (TAcc) and maximum running velocity ($V_{\max}$). Between-player, between-position, between-match and within-player variability were determined through linear mixed effects models. These data were then used to establish the practical significance of individual changes using a Minimum Effects Testing framework. All sources of variability were greater for HSRD (13–36%) when compared with all other metrics ($<6\%$). Using combined between-match and within-player variability along with the smallest worthwhile change ($0.2 \times$ between-player SD), between-match individual changes of $\pm 10\text{--}15\%$ in TD, TAcc and $V_{\max}$ were established as practically significant. For HSRD, these thresholds were considerably higher ($\geq 60\%$). In conclusion, the ability for soccer practitioners to identify meaningful changes in match physical performance can aid decision making around player management following competition. Our study provides a method to flag changes beyond the normal match-to-match variability and by a substantial magnitude. This may have implications for recovery but should be combined with other sources of data (internal load and response) and used only as an adjunct to practitioner domain knowledge, experience and expertise.

Keywords: Team Sport, Match Performance, Competition, Game Analysis.

26 INTRODUCTION

27 In recent years, load monitoring in sport has evolved through the increasing use of Global
28 Positioning Systems (GPS) [1–3]. Current GPS allow the measurement of velocity and
29 acceleration through doppler-derived methods, which allow the quantification of external training
30 and competition loads such as distance covered, the count of efforts and identification of peak
31 speed [4–7]. Importantly, while external loads are not the stimulus for functional training
32 adaptations, their measurement can provide a useful objective marker of physical performance
33 output, which plays a role in managing the overall training process [8]. This has become
34 increasingly important in modern professional soccer, where teams experience a large amount of
35 matches during the season and may often include two matches within a single week or three
36 matches within 7–10 days [6,9].

37 Managing the training process of soccer players presents as a daily challenge for sports
38 performance and medical practitioners. Information obtained from athlete monitoring during
39 competition can be used to guide decision making around the training schedule, such as the
40 programming of subsequent training or recovery activities. Soccer matches contribute to the largest
41 single session load of the training week [9] and their effects are typically considered when planning
42 the weekly microcycle. Players are usually given 24–48 hours of complete rest or active recovery
43 in anticipation of the physiological and neuromuscular responses to competition [10]. However,
44 the time-course of match-related recovery responses are extremely individual and may depend on
45 the specific match load [10]. Therefore, having the ability to determine whether match-play
46 demands are ‘unusual’ or abnormally high/low might prove to be a useful monitoring strategy to
47 inform practitioner decision making around the training schedule and individual management in
48 the days after competition.

The need to interpret changes in match physical performance is not a new dilemma and over the past decade there have been several investigations attempting to shed light on this issue [11–13]. Quantifying the match-to-match variability of physical performance can be a useful way to highlight changes that appear normal and those which seem higher or lower than usual. For instance, previous studies concluded that match-to-match variability was usually high (~14-53%) [11–15] for high-speed activities (e.g., high-speed running distance) but low for variables such as total distance covered (~3-5%) [12,14,15]. However, in isolation (i.e., without partitioning the usual sources of variability into their specific components), this approach does not identify the magnitude of a change. This can be achieved by scaling changes against so-called reference values for practical importance, such as the smallest worthwhile change. Probabilistic methods can then be used to determine the ‘practical significance’ of individual changes, taking into account both match-to-match variability and the smallest worthwhile change [16,17]. To our knowledge, this approach is yet to be applied in professional soccer.

The aims of our study were therefore twofold. We first aimed to provide a comprehensive breakdown of the variability in match physical performance of professional soccer players, including seasonal trends, variability between positions, players, and matches, and within players. Secondly, we aimed to apply a minimum effects testing framework for interpreting practically meaningful individual changes in soccer match physical performance.

MATERIALS AND METHODS

Study design

A cohort study was conducted in a professional soccer team from LaLiga123 during 2018-2019 season. The team played a total of 42 official matches (21 home matches and 21 away matches). Every match was played on natural grass fields with a 4-4-2 playing formation. The length of between-match microcycles varied from 3 to 9 days. The club authorized the data collection and this study was approved by the Bioethics Committee of the university.

Athletes

An initial sample of fourteen professional soccer players were monitored over a single in-season phase ($n = 515$ individual match observations). Goalkeepers were not included in the analysis, given their vastly different match activity-demands [4], and players were categorized as central defenders, full-backs, midfielders, wide-midfielders, and forwards. In an attempt to provide a representative seasonal profile of match external load, we elected to only include players who: a) completed the total duration of each match, b) played at least 5 matches during the season, and c) had at least one match observation in each trimester of the season (to avoid range effects). The final sample therefore included 240 individual match observations from 11 players (age: 28 ± 3 years; height: 180.9 ± 4.8 cm; body mass: 74.9 ± 4.1 kg; central defenders, $n = 3$ [67 individual match observations]; full-backs, $n = 2$ [42]; midfielders, $n = 2$ [48]; wide-midfielders, $n = 2$ [45]; forwards, $n = 2$ [38]. The median (range) number of match observations per player was 21 (15–31).

Procedures

Selection of Match External Load Metrics. Previous investigations employing dimension reduction techniques have suggested that 70% of the variance in soccer match external loads can be expressed by four principle components, each containing 2–4 variables [18]. Variables within a given component are thought to be similar (i.e., highly correlated and interchangeable), while components themselves provide different information (i.e., uncorrelated and unique). We therefore used a combination of both the previously reported component loadings (where a greater loading indicates a higher importance to the component) and our professional judgment to select four match external load indicators. These were total distance (TD), total high-speed running distance ($\geq 21 \text{ km}\cdot\text{h}^{-1}$; HSRD), total accelerations (TAcc) and maximum running velocity ($V_{\max}$).

Instruments. Every player was given a WIMU Pro device (RealTrack Systems, Almería, Spain) in order to collect physical performance variables. The device was placed on a vertical position in the back pocket of a specific vest (Rasán, Valencia, Spain). This device is considered as an electronic performance tracking system which mainly consists of 3D accelerometers, gyroscopes, and magnetometers as well as positioning sensors through 10 Hz Global Positioning System (GPS). These GPS have demonstrated good criterion validity (bias in mean velocity: 1.18-1.32 $\text{km}\cdot\text{h}^{-1}$; bias in distance: 2.32-4.32 m) and reliability (intraclass correlation coefficients: > 0.93) for the collection of physical performance variables in soccer [19].

All the devices were calibrated following manufacturer's instructions [20,21]. Prior to data collection, all the devices were first fully charged and placed on a Smart Station (RealTrack Systems, Almería, Spain). The station was placed on a flat surface without magnetic devices surrounding and then, the devices were turned on. Sixty seconds later, the start recording button was pressed and the calibration procedure was completed. Once this procedure finished, the

devices were given to the players, who were familiar with the tracking systems, 15 minutes before the start of the match. Finally, the session was analyzed at the end of the match using SPro software (RealTrack Systems, Almería, Spain).

Statistical Analysis

Our design located units of analysis (individual match observations) nested within clusters of units (Players), which were nested within larger clusters of clusters (Positions). To properly account for this hierarchical (correlated) nesting and to accurately quantify the variability in match physical performance, data were analysed using separate three-level linear mixed effect models. We used the MIXED procedure in SAS[®] Software (University Edition, SAS Institute Inc., Cary, NC, USA) to analyse both the original and log-transformed data, to quantify effect estimates in raw and percentage units. Model appropriateness was verified by examining plots of the studentized residual and predicted values, which were well behaved for both the raw and log-transformed data. To determine the linearized seasonal trend in each external load measure (the change in a given variable across the entire season), season week was re-scaled to range from -0.5 to 0.5 before being specified as a fixed effect (continuous covariate). Subsequently, separate random effects were added for Player ID, Position and Match Number. These effects were specified with a variance components covariance structure and estimated via Restricted Maximum Likelihood. Estimates—expressed as standard deviations (SD) and coefficients of variation (CV)—therefore included between-player, between-position, between-match and the remaining (residual) within-player variability, which is analogous to the typical error from a re-test reliability design. Uncertainty in all effects and ranges of values compatible with our data and statistical models were expressed as 90% confidence limits (CL). Boundary constraints were removed from covariance estimates to allow for negative variance. While negative variance (variability below zero) is illogical, it allows

for appropriate approximation of the uncertainty in the estimate when the true variability is close to zero.

Building on previous approaches [16,17], we used our variability estimates to provide a framework for practitioners to interpret individual changes in match physical performance indicators. This approach is founded on identifying changes that appear unusual and potentially meaningful. Here, an unusual change can be viewed as one beyond the normal match-to-match variability seen in any given player, after accounting for any seasonal trend and positional differences. We determined the observed match-to-match variability as the pooled (added) between-match SD and within-player typical error. These values were then multiplied by the square root of 2 and the appropriate value from the t distribution with the model degrees of freedom to establish 80% and 90% confidence limits, giving likely ranges for a normal or ‘usual’ individual change. A meaningful change can be viewed as one which exceeds a pre-established threshold of practical importance. Through lack of a known conceptual or empirical anchor between our match physical performance indicators and key match outcomes (e.g., result), we used a distribution-based approach to establish these boundaries [22]. Upon electing for a threshold equivalent to a small standardized change, reference values were given as 0.2 multiplied by the between-player SD or CV. Finally, we applied minimum effects tests (MET) [23] to establish probabilistic reference values for interpreting individual changes in match external loads, which combine both match-to-match variability and thresholds of practical importance. The t statistic for a change relative to practical importance ($\text{change} - \text{threshold} / [\text{match-to-match variability} \times \sqrt{2}]$) was converted to a probability via the one-tailed t -distribution. These probabilities were represented as continuous estimates across a plausible range of changes, as well as the changes associated with more conventional alpha levels (0.10 and 0.05).

RESULTS

The overall (mean) match external loads for TD, HSRD, TAcc and $V_{\max}$ were 10257 m, 699 m, 2976 (n) and $31.6 \text{ km}\cdot\text{h}^{-1}$. Observed external loads and the associated seasonal trends (individual and mean) are presented in Figure 1. There was a seasonal reduction in TD (mean slope: -220 m [90% confidence interval: -638 to 197], -2.1% [-6.1% to 1.9%]), and a seasonal increase in both HSRD (126 m [3 m to 249 m], 25% [3% to 52%]) and TAcc (112 [-16 to 241], 3.8% [-0.8% to 8.5%]). The magnitude and direction of the seasonal trend in $V_{\max}$ was inconclusive ($0.23 \text{ km}\cdot\text{h}^{-1}$ [$-0.49 \text{ km}\cdot\text{h}^{-1}$ to $0.95 \text{ km}\cdot\text{h}^{-1}$], 0.8% [-1.5% to 3.2%]).

Please insert Figure 1 near here

Estimates of between-player, between-position, between-match and within-player variability in each external load metric, expressed in raw (SD) and percentage (CV) units, are presented in Table 1, respectively. The observed match-to-match variability (combined between-match and within-player) for TD, HSRD, TAcc and $V_{\max}$ was 568 m (5.7%), 159 m (31%), 157 (5.6%) and $1.58 \text{ km}\cdot\text{h}^{-1}$ (5.2%). All sources of variability were greater for HSRD ($13\text{--}36\%$) when compared with all other external load metrics ($<6\%$). For each external metric, the greatest sources of variability were: between-player for TD (5.6%), between-position for HSRD (36%), between-match for TAcc (4.9%) and within-player for $V_{\max}$ 4.9% ; Table 1 & 2).

Please insert Table 1 near here

Reference values and methods for interpreting individual changes in match external loads are presented in Table 2 and Figure 2. Based on our thresholds for practical importance ($0.2 \times$ between-player SD or CV), between-match individual changes of $\pm 10\text{--}12\%$ (for a 80% CI/ $\alpha = 0.10$) and $\pm 13\text{--}15\%$ (for a 90% CI/ $\alpha = 0.05$) in TD, TAcc and $V_{\max}$ would be required to suggest practical

182 significance or substantially unusual. For HSRD, these thresholds were considerably higher ($\pm 58\%$
183 and $\pm 74\%$, respectively).

184 *****Please insert Table 2 near here*****

185 *****Please insert Figure 2 near here*****

186

DISCUSSION

The ability for soccer practitioners to identify meaningful changes in match physical performance has the potential to aid decision making around player management following competition. Through a detailed examination into the variability of match physical performance and using a minimum effects testing framework, our study is the first to apply a probabilistic method for monitoring individuals in professional soccer. Our main finding was that after accounting for seasonal trends and inter-position heterogeneity, between-match individual changes of $\pm\sim 10\text{--}15\%$ in representative GPS-derived measures of match physical performance (TD, TAcc and $V_{\max}$) can be considered practically significant within the current sample. This was with the exception HSRD, where thresholds were considerably higher ($\gtrsim 60\%$).

Our match physical performance values appear similar to those reported in other investigations and may therefore be considered representative of professional soccer competition [12,24]. Previous investigations have also observed that professional soccer players usually cover ~ 10 km per match [12,24], but only 4–6% of the TD is covered at high intensity [12]. In addition, soccer players are exposed to bouts of activity requiring running speeds in excess of $30 \text{ km}\cdot\text{h}^{-1}$ [12,25], as well as multiple changes of direction and duels, which significantly contribute to TAcc [18]. This supports the elite level population that participated in this study. We found HSRD and TAcc to increase throughout the season, whereas TD reduced and the trend in $V_{\max}$ was not clear. Previous studies have also found HSRD to increase from the start to the end of the season [14,15], but our observed decline in TD is not consistent with these studies [14,15], which reported a lower TD at the start of the season when compared to the middle and end. In this regard, there are several factors (e.g., ball possession, tactics, improved fitness, or adaptations to competitive league) that may explain these different trends [15]. For example, the team analyzed in our study adopted a

direct style of play with fast transitions and high defensive lines during the season, which may be another factor to explain the increase in HSRD and TAcc [26]. Also, although HSRD and TAcc are key performance indicators [18], a previous investigation observed that the teams with low-percentage ball possession covered more TD than teams with high-percentage ball possession [27]. Therefore, the seasonal trends need to be analyzed from a context-specific perspective in order to understand the potential practical applications of this type of analysis.

Regarding the estimates of between-player, between-position, between-match and within-player variability in each physical performance metric, a key finding was that sources of variability were greater for HSRD (13–36%) in comparison with all other external load metrics (<6%). These results are consistent with previous investigations which reported that the variability tended to increase with running intensity [11–15]. For instance, the match-to-match variability in TD from professional soccer players has been shown to be around 3.1% [14], 2.4% [15] and 5.3% [12] while the variability for HSRD was around 9.2% [14], 14.4% [15], 17.7% [13], 19.8% [11] and 53% [12]. Compared to previous investigations, however, a novel approach of our study was the ability to further partition the usual sources of variability into their specific components. We show, for example, that HSRD is far more variable between positions (36%) than it is between players within a given position (13%). Furthermore, we separate the observed match-to-match variability into true between-match variability (i.e., factors occurring at the level of a match) and within-player variability (i.e., factors occurring at the level of an individual). To our knowledge, this is the first attempt to do so in soccer. Our results indicate that the between-match and within-player variability of HSRD are similar, whereas the within-player variability of $V_{\max}$ was over 3 times that of the between-match variability. However, $V_{\max}$ did appear relatively stable, for all sources of variability (<5%), which is supported by previous research [12]. Our lower estimates (between match: 1.5%,

within-player: 4.9%) compared to Haddad and colleagues (observed match-to-match: 6–8%) are likely due to variance partitioning, but may also be influenced by different positional roles or playing styles [12].

The primary aim of our study was to break down the sources of variability in soccer match physical performance and use these data to shed light on monitoring individual between-match changes. We approached this twofold. First, between-match and within-player variability were combined to construct confidence intervals for an individual change in physical performance. This is similar to principles of the minimum detectable change (smallest detectable difference) and can be used to determine those which are beyond the observed match-to-match variability. For example, a change in whole match TD of +10.4% could be interpreted as higher than usual, using an alpha of 0.1 (80% CI). This can be a useful way to incorporate performance variability into decision making, but it does not consider the practical importance of a change. Building on the TD example, a whole match change of + 10.5% is greater than the smallest detectable difference for an 80%, but only by 0.1% which is unlikely to have any real-world relevance. Therefore, we additionally used effect size principles to estimate values of the smallest worthwhile change, given as 0.2 of the pure between-player variability (Table 1). Using a minimum effects testing framework, we then calculated the probability (p_{MET}) that a match change could be considered truly meaningful; that is, ‘unusual’ (beyond the normal match-to-match variability) by a substantial magnitude. A smaller p_{MET} indicates a greater practical significance and intuitively, p_{MET} of 0.10 and 0.05 correspond to individual changes where the respective 80% and 90% CI fall completely beyond the threshold of practical importance. For TD, this value was 11.5% for an 80% CI.

We have presented p_{MET} as both the value associated with fixed and conventional alpha levels (Table 2) and also as continuous estimates (Figure 2). The former is typical to making inference

from sample-based research studies but in the practice of monitoring individuals, this approach may be too conservative. The reality is that soccer practitioners are likely to interpret match physical performance data in a heuristic sense, so knowing if individual changes are trending towards practical significance might be of better use. The charts presented in Figure 2 could therefore be used to determine the actual practical significance, by drawing a vertical line up from the observed change (X-axis) and drawing a perpendicular horizontal line at the intercept of the Y-axis to determine the corresponding p_{MET} . These values can then be interpreted in the context of the decisions being made and the contextual constraints (e.g., associated time, resources and impact on players). For example, a lower p_{MET} for a positive change in HSRD and/or TAcc might be indicative of substantially more high-intensity match activity than normal. This might warrant the considerations for additional recovery, medical or nutritional intervention in the days following the match to ensure appropriate regeneration prior to the next competitive fixture.

We understand that it would be naïve to directly make an inference on fatigue or recovery from match physical performance alone, and this should be taken into consideration when making decisions using the aforementioned methods. External load is the means by which an internal load is induced; which is then the stimulus for consequential responses to the body's systems and functional performance (e.g., fatigue and recovery). Therefore, we do not advocate that changes in match physical performance are used to make direct and definitive inference on fatigue or recovery. Rather, these changes could imply a differential fatigue or recovery time-course response when compared to usual. For example, if practitioners detect changes in match physical performance that are substantially higher than usual (e.g., +10–15% for TD, TAcc or V_{max}) then an assessment of the subsequent 'response' might be warranted to inform athlete management decisions. This could include post-day subjective measures of fatigue or objective measures such

as heart rate variability or stiffness (e.g., leg/ whole body). Or, it may simply be used as an avenue to start a conversation with the player, where clinical and practical judgment can precede generic advice (e.g., sleep, nutrition) or further intervention. However, there are some general limitations, which need to be acknowledged and may also guide future investigations. First, only one professional soccer team was included in the analysis, so the sample size was limited to the most regular players of the league ($n = 13$) and these findings are representative of this sample. A multi-club study would therefore help extend our current findings. Additionally, recently implemented local positioning systems may provide more accurate data than GPS [1,19], particularly for match physical performance measures such as HSRD and $V_{\max}$ [5], and should therefore be considered both for interpreting our findings to data collected with these systems and future research into the variability of soccer match physical performance. Nonetheless, the changes in external load-based measures do not necessarily lead to a “worthwhile” change in the performance outcome. Finally, we have stressed the importance of including concepts such as the smallest worthwhile change to interpret differences in match physical performance, but this distribution-based approach is in itself limited. The extent to which group-based effect sizes principles can be used at the individual level can be questioned and this is likely a less-robust approach than using anchor-based methods, which are yet to be established for match physical performance in soccer.

Competitive fixtures present as the most demanding sessions within a professional soccer microcycle. However, the variability of match physical performance is often overlooked. Practitioners are responsible for quantifying weekly match demands to inform subsequent player management and devise an appropriate training schedule. Understanding the variability of match physical performance can aid this decision-making process by determining the practical significance of between-match changes. In our study applying this approach alongside a minimum

effects testing framework, we found that individual changes in representative GPS-derived measures of match physical performance (TD, TAcc and $V_{\max}$) of $\pm\sim 10\text{--}15\%$ can be considered practically significant. That is, beyond the normal match-to-match variability and by a magnitude greater than the smallest worthwhile change. This was with the exception of HSRD, where thresholds were considerably higher ($\gtrsim 60\%$). Soccer practitioners might therefore consider this approach to facilitate decision making. However, we maintain that other sources of data (internal load and the associated response) are needed to properly evaluate match demands and any athlete management decisions should be primarily based on domain knowledge (training principles and physiological or biomechanical theory) and experience.

311 **ACKNOWLEDGMENTS**

312 The authors would like to thank the club which authorized the data collection during the
313 competitive season. Also, the authors declare no conflict of interest.

314

REFERENCES

1. Linke, D.; Link, D.; Lames, M. Validation of electronic performance and tracking systems (EPTS) under field conditions. *PLoS One* **2018**, *13*, 1–19.
2. Cummins, C.; Orr, R.; O'Connor, H.; West, C. Global Positioning Systems and microtechnology sensors in team sports: a systematic review. *Sport. Med.* **2013**, *43*, 1025–1042.
3. Rojas-Valverde, D.; Gómez-Carmona, C.D.; Gutiérrez-Vargas, R.; Pino-Ortega, J. From big data mining to technical sport reports: the case of inertial measurement units. *BMJ Open Sport Exerc. Med.* **2019**, *5*, 1–3.
4. Oliva-Lozano, J.M.; Gómez-Carmona, C.D.; Pino-Ortega, J.; Moreno-Pérez, V.; Rodríguez-Pérez, M.A. Match and training high intensity activity-demands profile during a competitive mesocycle in youth elite soccer players. *J. Hum. Kinet.* **2020**, *Epub Ahead*, 1–11.
5. Buchheit, M.; Simpson, B.M. Player-tracking technology: half-full or half-empty glass? *Int. J. Sports Physiol. Perform.* **2017**, *12*, 35–41.
6. Palucci-Vieira, L.H.; Aquino, R.; Lago-Peñas, C.; Munhoz-Martins, G.H.; Puggina, E.F.; Barbieri, F.A.; Palucci Vieira, L.H.; Aquino, R.; Lago-Peñas, C.; Munhoz Martins, G.H.; et al. Running performance in Brazilian professional football players during a congested match schedule. *J. Strength Cond. Res.* **2018**, *32*, 313–325.
7. Oliva-Lozano, J.M.; Fortes, V.; Muyor, J.M. The first, second, and third most demanding passages of play in professional soccer: a longitudinal study. *Biol. Sport* **2020**, *38*, 165–

174.

8. Impellizzeri, F.M.; Marcora, S.M.; Coutts, A.J. Internal and external training load: 15 years on. *Int. J. Sports Physiol. Perform.* **2019**, *14*, 270–273.
9. Anderson, L.; Orme, P.; Di Michele, R.; Close, G.L.; Morgans, R.; Drust, B.; Morton, J.P. Quantification of training load during one-, two- and three-game week schedules in professional soccer players from the English Premier League: implications for carbohydrate periodisation. *J. Sports Sci.* **2016**, *34*, 1250–1259.
10. Nédélec, M.; McCall, A.; Carling, C.; Legall, F.; Berthoin, S.; Dupont, G. Recovery in soccer. *Sport. Med.* **2012**, *42*, 997–1015.
11. Carling, C.; Bradley, P.; McCall, A.; Dupont, G. Match-to-match variability in high-speed running activity in a professional soccer team. *J. Sports Sci.* **2016**, *34*, 2215–2223.
12. Haddad, H. Al; Méndez-Villanueva, A.; Torreño, N.; Munguia-Izquierdo, D.; Suárez-Arrones, L. Variability of GPS-derived running performance during official matches in elite professional soccer players. *J. Sports Med. Phys. Fitness* **2018**, *58*, 1439–1445.
13. Gregson, W.; Drust, B.; Atkinson, G.; Salvo, V. Match-to-match variability of high-speed activities in Premier League soccer. *Int. J. Sports Med.* **2010**, *31*, 237–242.
14. Mohr, M.; Krstrup, P.; Bangsbo, J. Match performance of high-standard soccer players with special reference to development of fatigue. *J. Sports Sci.* **2003**, *21*, 519–528.
15. Rampinini, E.; Coutts, A.; Castagna, C.; Sassi, R.; Impellizzeri, F. Variation in top level soccer match performance. *Int. J. Sports Med.* **2007**, *28*, 1018–1024.
16. McLaren, S.J.; Weston, M.; Smith, A.; Cramb, R.; Portas, M.D. Variability of physical

performance and player match loads in professional rugby union. *J. Sci. Med. Sport* **2016**, *19*, 493–497.

17. Weston, M.; Siegler, J.; Bahnert, A.; McBrien, J.; Lovell, R. The application of differential ratings of perceived exertion to Australian Football League matches. *J. Sci. Med. Sport* **2015**, *18*, 704–708.

18. Oliva-Lozano, J.M.; Rojas-Valverde, D.; Gómez-Carmona, C.D.; Fortes, V.; Pino-Ortega, J. Impact of contextual variables on the representative external load profile of Spanish professional soccer match-play: a full season study. *Eur. J. Sport Sci.* **2020**, *Epub Ahead*, 1–22.

19. Bastida-Castillo, A.; Gómez-Carmona, C.D.; De la Cruz, E.; Pino-Ortega, J.; Bastida Castillo, A.; Gómez Carmona, C.D.; la Cruz Sánchez, E.; Pino Ortega, J. Accuracy, intra- and inter-unit reliability, and comparison between GPS and UWB-based position-tracking systems used for time–motion analyses in soccer. *Eur. J. Sport Sci.* **2018**, *18*, 1–8.

20. Oliva-Lozano, J.M.; Rojas-Valverde, D.; Gómez-Carmona, C.D.; Fortes, V.; Pino, J. Worst case scenario match analysis and contextual variables in professional soccer players: a longitudinal study. *Biol. Sport* **2020**, *37*, 429–436.

21. Oliva-Lozano, J.M.; Fortes, V.; Krstrup, P.; Muyor, J.M. Acceleration and sprint profiles of professional male football players in relation to playing position. *PLoS One* **2020**, *15*, 1–12.

22. Cook, J.A.; Julious, S.A.; Sones, W.; Hampson, L. V; Hewitt, C.; Berlin, J.A.; Ashby, D.; Emsley, R.; Fergusson, D.A.; Walters, S.J.; et al. DELTA 2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a

randomised controlled trial. *BMJ* **2018**, 1–7.

23. Murphy, K.R.; Myers, B. Testing the hypothesis that treatments have negligible effects: Minimum-effect tests in the general linear model. *J. Appl. Psychol.* **1999**, *84*, 234–248.

24. Palucci-Vieira, L.H.; Aquino, R.; Moura, F.A.; Barros, R.; Arpini, V.M.; Oliveira, L.P.; Bedo, B.L.; Santiago, P.R.. Team dynamics, running, and skill-related performances of Brazilian U11 to professional soccer players during official matches. *J. Strength Cond. Res.* **2019**, *33*, 2202–2216.

25. Aquino, R.; Munhoz-Martins, G.H.; Palucci-Vieira, L.H.; Menezes, R.P. Influence of match location, quality of opponents, and match status on movement patterns in Brazilian professional football players. *J. Strength Cond. Res.* **2017**, *31*, 2155–2161.

26. Fernandez-Navarro, J.; Fradua, L.; Zubillaga, A.; McRobert, A.P. Influence of contextual variables on styles of play in soccer. *Int. J. Perform. Anal. Sport* **2018**, *18*, 423–436.

27. da Mota, G.R.; Thiengo, C.R.; Gimenes, S.V.; Bradley, P.S. The effects of ball possession status on physical and technical indicators during the 2014 FIFA World Cup Finals. *J. Sports Sci.* **2016**, *34*, 493–500.

FIGURES AND TABLES LEGEND

Figure 1. Seasonal trends in match physical performance. Data are presented as individual match observations (grey points), individual linear trends (grey lines) and the overall mean trend (thick blue line) with 90% confidence intervals (blue dotted line).

Figure 2. Probabilistic reference values for interpreting individual changes in match physical performance. Practical significance is given as the probability value from a minimum effects test (p_{MET} ; Y-axis). The combined between-match and within-player variability estimates (Table 1) multiplied by the square root of 2 are used as the error term. The threshold for a meaningful change is given as a small effect size (0.2 multiplied by the pure between-player SD; Table 1). A smaller p_{MET} indicates a greater likelihood that the change is meaningful, that is; ‘unusual’ or substantially beyond the normal match-to-match variability.

Table 1. Variability of match physical performance expressed in raw units and coefficients of variations (%).

Table 2. Reference values for interpreting individual changes in match physical performance.

Figure 1. Seasonal trends in match external loads. Data are presented as individual match observations (grey points), individual linear trends (grey lines) and the overall mean trend (thick blue line) with 90% confidence intervals (blue dotted line).

Figure 2. Probabilistic reference values for interpreting individual changes in match physical performance. Practical significance is given as the probability value from a minimum effects test (p_{MET} ; Y-axis). The combined between-match and within-player variability estimates (Table 1) multiplied by the square root of 2 are used as the error term. The threshold for a meaningful change is given as a small effect size (0.2 multiplied by the pure between-player SD; Table 1). A smaller p_{MET} indicates a greater likelihood that the change is meaningful, that is; ‘unusual’ or substantially beyond the normal match-to-match variability.

Table 1. Variability of match physical performance expressed in raw units and coefficients of variations (%).

	Metric	Variability			
		Between-player	Between-position	Between-match	Within-player
(SD; 90% CI)	TD (m)	546 (147 to 759)	376 (-416 to 675)	428 (322 to 513)	373 (344 to 408)
	HSRD (m)	72 (-24 to 104)	222 (-107 to 331)	101 (72 to 123)	124 (114 to 136)
	TAcc (#)	74 (12 to 104)	27 (-60 to 72)	137 (106 to 162)	76 (70 to 83)
	V _{max} (km·h ⁻¹)	0.90 (-0.30 to 1.31)	0.93 (-0.86 to 1.57)	0.44 (-0.23 to 0.66)	1.52 (1.40 to 1.66)
(CV; 90% CI)	TD (%)	5.6 (1.5 to 7.8)	3.8 (-4.2 to 6.9)	4.3 (3.2 to 5.1)	3.7 (3.4 to 4.1)
	HSRD (%)	12.7 (-3.9 to 18.9)	36 (-17 to 59)	19 (13 to 23)	23 (21 to 25)
	TAcc (%)	2.5 (0.4 to 3.5)	0.9 (-2.0 to 2.4)	4.9 (3.8 to 5.8)	2.6 (2.4 to 2.8)
	V _{max} (%)	2.8 (-1.0 to 4.2)	3.0 (-2.7 to 5.1)	1.5 (-0.6 to 2.2)	4.9 (4.5 to 5.4)

Table 2. Reference values for interpreting individual changes in match physical performance.

Metric	Unit	± Confidence limits for a true change*		Change (±) required to be practically significant*#	
		80%	90%	$\alpha = 0.10$	$\alpha = 0.05$
TD	m	1036	1333	1145	1442
	%	10.4	13.4	11.5	14.5
HSRD	m	291	374	305	389
	%	56	72	58	74
TAcc	#	286	368	301	383
	%	10.2	13.1	10.7	13.6
V _{max}	km·h ⁻¹	2.88	3.71	3.06	3.89
	%	9.4	12.1	10.0	12.7

#Based on the combined between-match and within-player variability (Table 1): TD = 568 m (5.7%), HSRD = 159 m (31%), TAcc = 157 (5.6%), V_{max} = 1.58 km·h⁻¹ (5.2%). See methods section for more details.

*The threshold for a practically important change is given as a small effect size: 0.2 multiplied by the pure between-player SD (Table 1). Changes are then estimate based on a minimum effects test against these thresholds at the given alpha level.