

# Validity and reproducibility of match-derived ratios of selected external and internal load parameters in soccer players: A simple way to monitor physical fitness?

**AUTHORS:** Jan Schimpchen<sup>1,2</sup>, Paulo Freitas Correia<sup>1</sup>, Tim Meyer<sup>2</sup>

<sup>1</sup> Sport Lisboa e Benfica, Human Performance Department, Lisbon, Portugal

<sup>2</sup> Institute of Sport and Preventive Medicine, Saarland University, Saarbrücken, Germany

**ABSTRACT:** The study aim was to assess whether match-derived external-to-internal load ratios are a valid and reliable tool to measure physical fitness. Sixteen elite youth soccer players ( $17 \pm 1$  years) performed two maximal fitness tests. Subsequently, players participated in three intra-squad soccer matches in three consecutive weeks. Three GPS-based parameters of external load (total distance, PlayerLoad, high-intensity distance) were divided by three heart rate-based parameters of internal load (iTRIMP, Banister TRIMP, average percentage heart rate reserve) for the ratio calculations. Validity was established by comparing the ratios with results of the fitness tests, while between-athlete and within-athlete reliability were quantified. Most integrated load ratios were moderately-to-largely correlated with the various fitness parameters. Overall, a ratio consisting of PlayerLoad and average percentage heart rate reserve demonstrated the most consistent correlations with maximum treadmill speed ( $r = 0.69$ ,  $P = 0.003$ ) and the speeds associated with 4 mmol/L of blood lactate ( $r = 0.56$ ,  $P = 0.024$ ) and 80% of heart rate reserve ( $r = 0.54$ ,  $P = 0.031$ ). Most of the ratios displayed acceptable levels of reproducibility (intraclass correlation coefficient  $> 0.8$  and coefficient of variation  $< 10\%$ ), with the minimal detectable change of all ratios ranging between 7.1 and 37.8%. Given their associations with physical fitness and non-invasive nature, certain external-to-internal load ratios may be used to monitor physical fitness in soccer players. However, the ratios may not be sensitive enough to detect small yet practically relevant alterations in player fitness.

**CITATION:** Schimpchen J, Correia PF, Meyer T. Validity and reproducibility of match-derived ratios of selected external and internal load parameters in soccer players: A simple way to monitor physical fitness? *Biol Sport.* 2023;40(4):1039–1046.

Received: 2022-10-22; Reviewed: 2022-11-14; Re-submitted: 2023-01-18; Accepted: 2023-01-20; Published: 2023-03-06.

Corresponding author:

**Jan Schimpchen**

Sport Lisboa e Benfica  
Quinta da Trindade  
2840-600 Seixal, Portugal  
E-mail: [jan.schimpchen@web.de](mailto:jan.schimpchen@web.de)

**ORCID:**

Jan Schimpchen  
0000-0001-8339-0638

Paulo Freitas Correia  
0000-0002-1238-2566

Tim Meyer  
0000-0003-3425-4546

**Key words:**

Heart rate  
GPS  
Physical performance  
Lactate  
Football

## INTRODUCTION

The development and in-season maintenance of high levels of physical fitness is an important consideration for practitioners tasked with maximising the performance potential of soccer players. Ideally, accurate information on fluctuations in player fitness can be obtained regularly to tailor training interventions to the individual player needs [1]. These insights are typically derived from (maximal) laboratory- or field-based fitness tests. However, due to their exhaustive nature and the difficulty of finding suitable time points within an oftentimes congested playing schedule, it is unrealistic to expect frequent testing opportunities at the higher levels of professional soccer [2]. Given these limitations, it would seem desirable to establish alternative testing protocols which are valid, reliable, minimally invasive and require little to no extra time or effort by the players [3, 4].

Within elite soccer, training and match physical demands are commonly recorded using a host of different internal and external load monitoring tools [2]. It has been proposed that the relationship between identified markers of internal and external load may provide suitable information regarding how a player is coping with

their training programme [5]. As such, under standardised and non-fatigued conditions, improved fitness would be assumed for players experiencing a lower internal load relative to a standardised external task, which may be attributed to various central and peripheral adaptations resulting in improved exercise economy, whole-body peak oxygen uptake, lactate threshold and oxygen uptake kinetics [6]. Vice versa, players with a higher internal load relative to the same external task would be suspected to have experienced negative fitness adaptations [5]. Similar conclusions may be deduced from relative changes in internal load in response to non-standardised external stimuli. Several studies have demonstrated significant correlations between external-to-internal load ratios during sport-specific, non-standardised activity and parameters of aerobic performance in soccer [7, 8], rugby [9] and hurling players [10]. Given that markers of external load have been shown to be highly variable in competition [11] and may be impacted by contextual factors [12], the integration of external and internal load into a ratio represents a rather elegant way to assess athlete fitness since the parameters necessary for calculation are typically

recorded as part of the monitoring routine of elite soccer teams [13], avoiding the need for any additional ‘touch points’ with and burden on the players [14].

Akubat *et al.* [7, 8] investigated different combinations of integrated external-internal load ratios obtained during a soccer match-specific running simulation. The authors used five different GPS- or accelerometer-derived parameters of external load (total distance, high-intensity distance, PlayerLoad, mean metabolic power, distance covered at high metabolic power) and individualised TRIMP (iTRIMP) as the only parameter of internal load. Since heart rate (HR)-based parameters are easily and precisely measured, observing the cardio-circulatory response to exercise is frequently the preferred method of monitoring internal load in a team-sport context. In a recent study, iTRIMP was found to be the training load measure with the strongest relationship to changes in aerobic fitness over a 6-week pre-season period in elite youth soccer players [15], and therefore its inclusion in the integrated external-internal load ratios seems to be justified. However, since the calculation of iTRIMP is based on each player’s lactate profile to generate an individual exponential weighting factor for the TRIMP calculation [16], it is very context-dependent whether the necessary testing procedures can be carried out (potentially multiple times per season). It would appear worthwhile to investigate whether other more easily obtainable parameters of cardio-circulatory load might be a feasible alternative to iTRIMP in the calculation of the integrated ratios. Furthermore, since the two studies assessing the relationship between different external-internal load ratios and aerobic fitness in soccer players were based on running protocols looking to replicate typical match running demands [7, 8], it remains to be confirmed whether ratios based on actual match-derived markers of external and internal load are equally related to parameters of cardio-circulatory fitness. If so, this would potentially allow for continuous fitness monitoring based on match data alone.

When interpreting any test score, it is important to be aware of and appreciate the uncertainty associated with the measure [17].

Preferably, a test chosen for implementation should be responsive to small but relevant changes over time. As such, if match-derived external-to-internal load ratios are to be used for monitoring changes in player fitness, the ratios should be able to differentiate signal (real change) from noise (random measurement error). In this context, the minimal detectable change (MDC) and the smallest worthwhile change (SWC) [18–20] are two important test statistics which can help quantify signal (SWC) in comparison to noise (MDC). MDC refers to the smallest change that can be detected with an acceptable probability level, while SWC relates to the smallest change that can be considered important.

Therefore, the aims of this study were twofold: (1) to investigate whether different combinations of match-derived external-internal load ratios were associated with markers of aerobic fitness; and (2) to quantify the reliability of these measures across multiple time points.

## MATERIALS AND METHODS

### Participants

Twenty-one youth outfield players ( $17 \pm 1$  years,  $179 \pm 8$  cm,  $72 \pm 7$  kg) belonging to an elite Portuguese soccer team were recruited to participate in the study [21]. An a priori sample size calculation (G\*Power, version 3.1.9.7, Heinrich Heine University Düsseldorf, Germany) indicated that 21 participants were required to have a sufficient power ( $> 0.8$ ) based on the alpha selected (5%) and a large effect size ( $d = 0.5$ ). This effect size was based on previous studies on this topic [7, 8]. Players were eligible to participate in the study if they had been free of injury up to four weeks prior to the beginning of the study and accustomed to the regular training routine of the team. The players and their parents were informed about the study procedures and provided their written informed consent. Performance data were anonymised to ensure complete player confidentiality. The investigation was conducted in accordance with the Declaration of Helsinki and ethical approval was received

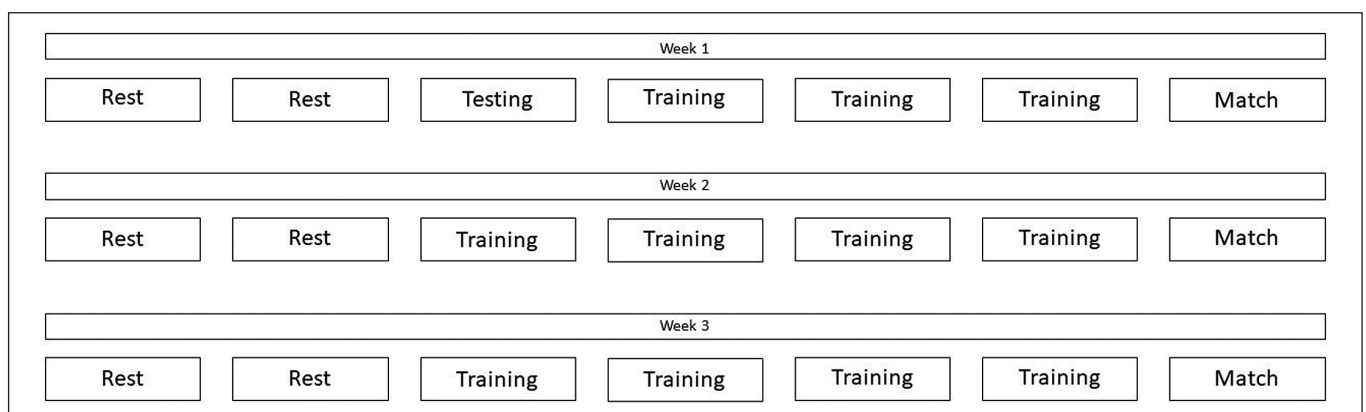


FIG. 1. Overview of study design.

from the local ethics committee (Saarland University, approval number 20–26).

### *Design and Data Collection*

Due to local COVID-19 regulations during the study period, the participating team was not involved in official competition. However, the team had been training for 14 weeks straight before the beginning of the investigation, with the training week consisting of four training sessions from Monday to Thursday and an intra-squad, full-pitch 90 min match each Friday, followed by two days of rest. This weekly structure was largely maintained for our study purposes (Fig. 1).

This observational study began with an evaluation of the players' cardiocirculatory fitness. Throughout the morning (8 am–12 pm) and 90 minutes after their last meal, all players performed a stepwise incremental treadmill running test with determination of capillary blood lactate and HR responses. The test was initiated at a running velocity of  $8 \text{ km} \cdot \text{h}^{-1}$ , with a step duration of 3 min, a  $2 \text{ km} \cdot \text{h}^{-1}$  velocity increment for each successive step until volitional exhaustion occurred, with steps separated by a 1 min passive recovery period for the collection of capillary blood samples from the participants' earlobes (Lactate Pro 2, Arkray KDK, Kyoto, Japan). HR was continuously monitored (Firstbeat Sports Sensor, Firstbeat Technologies Oy, Jyväskylä, Finland) at a sampling frequency of 0.2 Hz and exercise HR for a given velocity was defined as the mean HR over the last 30 s for each step. The treadmill gradient was set at 1% to reflect the energetic cost associated with outdoor running [22]. After the test, the data were fed into specialised software (Ergonizer, Kai Röcker, Freiburg, Germany) and the following parameters were calculated for further analysis: (1) running velocity at volitional exhaustion ( $V_{\text{max}}$ ) (if subjects did not complete the final stage,  $V_{\text{max}}$  was calculated as: (running velocity of last completed stage) +  $(x/180 \cdot 2 \text{ km} \cdot \text{h}^{-1})$ , where  $x$  equals the duration of the final stage in s), (2) running velocity associated with a fixed lactate concentration of 4 mmol/L ( $V_4 \text{ mmol}$ ), and (3) the physical working capacity (running velocity) at 80% of the participants' individual HR reserve ( $V_{80\% \text{HR}_{\text{Res}}}$ ). During the late afternoon of day 1 (4 pm–6 pm) and after at least 6 h of passive recovery from the previous assessment, the players performed a Yo-Yo Intermittent Recovery level 1 test (Yo-Yo IR1). The test was executed on an artificial outdoor turf, at an ambient temperature of 15°C. The Yo-Yo IR1 has been shown to possess high levels of construct validity and may therefore be considered a good indicator of intermittent team sport-specific physical fitness [23]. The running distance associated with the last completed stage before failing to maintain the indicated running velocity was recorded as the final test result.

All training sessions throughout the study period were planned and led by the team's coaching staff. Before the beginning of the investigation, team coaches were instructed to maintain their normal training periodisation leading up to the matches and, as such, expose the players to the regular training volumes and intensities associated with their training approach. Overall training load throughout the study protocol was kept consistent to maintain the players'

physical condition. In addition, players continued to perform two gym-based training sessions per week as per their usual training routine. These sessions took place four and two days prior to the match and consisted of full-body, multi-joint strength and power exercises (first weekly session) and exercises focused on lumbo-pelvic control, landing mechanics and balance (second weekly session).

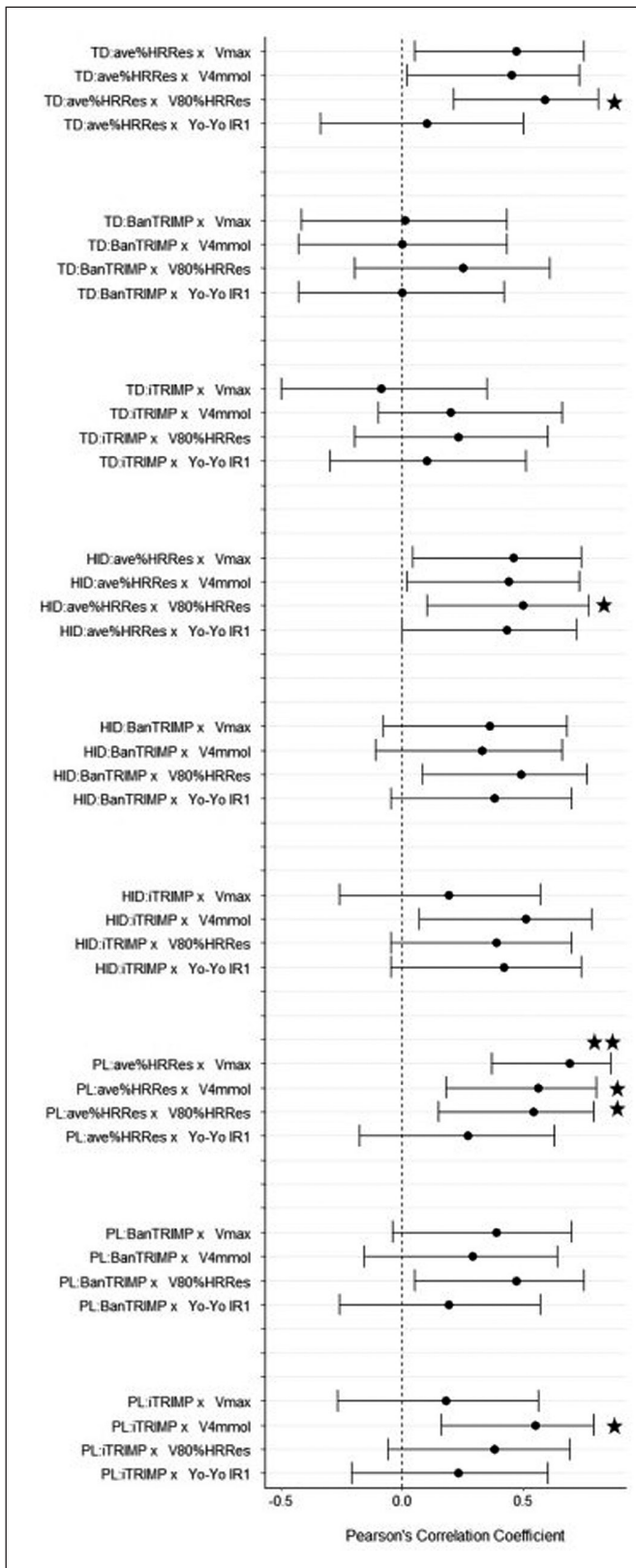
During all matches, player movement was recorded using portable GPS devices (Catapult Vector, Melbourne, Australia). Previous studies have found 10 Hz GPS systems to be the most valid and reliable across linear and team sport simulated running. Accuracy measures of total distance and high-intensity distance displayed a coefficient of variation (CV) of between 1.9 and 10.5% compared to criterion measures [24]. After the match, all data were downloaded using the respective proprietary software (Catapult Openfield, version 3.2.0) and adequate positional data quality was verified. In accordance with previous research [25], GPS connection to at least 6 satellites and a horizontal dilution of precision (HDOP)  $\leq 2$  were required for data to be included. Throughout the final dataset, the number of connected satellites and HDOP were  $14.7 \pm 0.6$  and  $0.7 \pm 0.1$ , respectively. The players' total distance (TD) covered, the distance covered at running speeds  $> 18 \text{ km} \cdot \text{h}^{-1}$  (HID) and PlayerLoad (PL – a widely used external load metric derived from three-dimensional measures of the instantaneous rate of change of acceleration) [26] were collected as markers of external load over the course of the three matches. These three parameters were deemed relevant since they are commonly used to quantify training and match demands. HR was continuously recorded throughout the matches and the resulting cardiorespiratory intensity and load were calculated in three different ways: (1) as the average percentage of  $\text{HR}_{\text{Res}}$  ( $\text{ave}\% \text{HR}_{\text{Res}}$ ), (2) as a Banister TRIMP score (BanTRIMP) [27], and (3) as an iTRIMP score [16]. Consequently, nine different external-to-internal load ratios were calculated and investigated.<sup>1</sup>

Resting HR for each player was determined one morning following two days of rest, in a quiet room with players remaining in a supine position for 5 min. Resting HR was assumed as the lowest 5 s value. Maximal HR was assumed as the highest 5 s value throughout all testing, training and match activities during the study period.

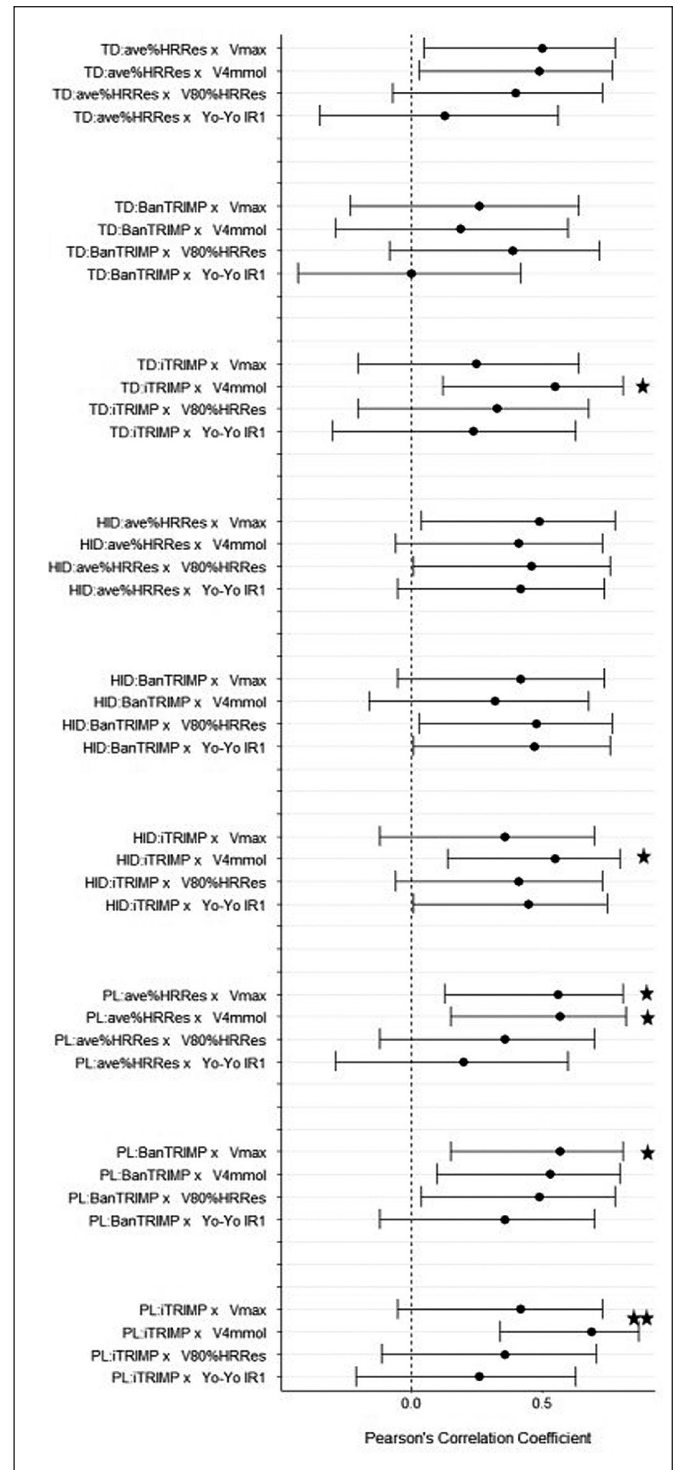
### *Statistical Analysis*

Unless stated otherwise, all data are presented as mean  $\pm$  standard deviation (SD). The assumption of normality was verified using the Shapiro-Wilk test. Pearson's correlation coefficients were used to quantify the relationship between external-to-internal load ratios recorded during the first match and fitness markers. The magnitude of the correlation was interpreted as follows:  $< 0.1$ , trivial;  $0.1\text{--}0.29$ , small;  $0.3\text{--}0.49$ , moderate;  $0.5\text{--}0.69$ , large;  $0.7\text{--}0.89$ , very large;

<sup>1</sup> It must be pointed out that  $\text{ave}\% \text{HR}_{\text{Res}}$  represents internal intensity rather than internal load. As such, ratios involving this parameter should be more correctly classified as load-intensity ratios.



**FIG. 2.** Relationship between the different external-to-internal load ratios derived during the first half of match play and selected markers of physical fitness.  
 Note: TD – total distance; HID – high-intensity distance; PL – PlayerLoad; Vmax – velocity at volitional exhaustion; V4 mmol – velocity associated with 4 mmol/L of blood lactate; V80%HR<sub>Res</sub> – velocity at 80% of heart rate reserve; Yo-Yo IR1 – Yo-Yo intermittent recovery level 1 test performance; \* – P ≤ 0.05; \*\* – P ≤ 0.01.



**FIG. 3.** Relationship between the different external-to-internal load ratios derived during the second half of match play and selected markers of physical fitness.  
 Note: TD – total distance; HID – high-intensity distance; PL – PlayerLoad; Vmax – velocity at volitional exhaustion; V4 mmol – velocity associated with 4 mmol/L of blood lactate; V80%HR<sub>Res</sub> – velocity at 80% of heart rate reserve; Yo-Yo IR1 – Yo-Yo intermittent recovery level 1 test performance; \* – P ≤ 0.05; \*\* – P ≤ 0.01.

0.9–0.99, nearly perfect; and 1, perfect [28]. Between-athlete and within-athlete reliability were assessed using intraclass correlation coefficients (ICC) and CV, respectively. ICC estimates and their 90% CI were calculated based on a mean-rating ( $k = 3$ ), absolute agreement, 2-way random-effects model. The following thresholds for interpretation of ICC results were used:  $< 0.5$ , low;  $0.5–0.74$ , moderate;  $0.75–0.89$ , high;  $0.9–0.98$ , very high; and  $> 0.98$ , extremely high. CV was calculated by dividing the within-subject SD by the mean multiplied by 100. Furthermore, MDC at 90% CI level was calculated as  $1.645 \times SEM \times \sqrt{2}$  [18] and SWC was calculated using a distribution-based method as  $0.2 \times$  between-subject standard deviation [19]. Standard error of measurement (SEM) was estimated as the square root of the mean square error from the ANOVA [29]. Ratio reliability and association with fitness markers were analysed separately for first and second halves of the matches. A player entering a match at halftime was considered the same as playing the first half of the match and only players competing in at least a full half were included for analysis. Only players competing in all three matches were considered for the reliability analysis. An  $\alpha$ -level of  $P \leq 0.05$  was accepted as statistically significant. All statistical analyses were conducted using specialised Excel spreadsheets [30].

**RESULTS**

During the initial incremental treadmill test, players achieved a  $V_{max}$  of  $17.7 \pm 1.3 \text{ km} \cdot \text{h}^{-1}$ , at a HR of  $195 \pm 7 \text{ beats} \cdot \text{min}^{-1}$ , representing  $98.2 \pm 1.7\%$  of their maximal HR ( $HR_{max}$ ), with a maximal blood lactate concentration of  $8.4 \pm 2.5 \text{ mmol/l}$ .  $V_4 \text{ mmol}$  was associated with a running velocity of  $15.1 \pm 1.4 \text{ km} \cdot \text{h}^{-1}$ , at  $93.3 \pm 2.5\%$  of  $HR_{max}$ .  $V_{80\%HR_{Res}}$  occurred at  $12.1 \pm 1.3 \text{ km} \cdot \text{h}^{-1}$  and an average blood lactate concentration of  $2.3 \pm 0.4 \text{ mmol/l}$ . The last completed stage of the Yo-Yo IR1 test corresponded to a running distance of  $2181 \pm 484 \text{ m}$ , at  $96.6 \pm 2\%$  of  $HR_{max}$ .

*Relationship between load ratios and fitness markers*

Two players dropped out of the investigation during the first week of training, one due to injury and the other due to illness. Furthermore, three players were excluded from the analysis due to incomplete HR recordings during the first match. 14 of the remaining 16 players completed the entire first match, while two players only completed the first half. The relationships between the different external-to-internal load ratios and fitness markers are presented for the first and second half in Figures 2 and 3, respectively.

**TABLE 1.** Between-athlete and within-athlete reproducibility of the different external-to-internal load ratios

External: Internal Ratio	Half	ICC (90% CI)	CV (90% CI)	SEM	MDC	SWC	SEM %	MDC %	SWC %
TD:iTRIMP	1 <sup>st</sup>	0.90 (0.77; 0.96)	7.6 (6.0; 10.8)	4.5	10.5	2.4	8.1	18.8	4.3
	2 <sup>nd</sup>	0.90 (0.8; 0.98)	8.7 (6.4; 14.5)	5.6	13.0	3.0	10.0	22.3	5.2
PL:iTRIMP	1 <sup>st</sup>	0.78 (0.56; 0.91)	9.6 (7.6; 13.5)	0.5	1.2	0.2	9.2	21.3	3.8
	2 <sup>nd</sup>	0.88 (0.65; 0.97)	10.7 (7.9; 18.0)	0.6	1.5	0.3	10.7	24.8	4.9
HID:iTRIMP	1 <sup>st</sup>	0.88 (0.73; 0.95)	15.6 (12.3; 21.8)	0.9	2.2	0.5	16.3	37.8	8.0
	2 <sup>nd</sup>	0.91 (0.73; 0.98)	15.5 (11.5; 25.9)	0.9	2.1	0.5	14.9	34.7	8.1
TD:BanTRIMP	1 <sup>st</sup>	0.83 (0.64; 0.93)	4.8 (3.8; 6.7)	2.5	5.8	1.1	5.0	11.7	2.1
	2 <sup>nd</sup>	0.75 (0.38; 0.93)	6.7 (5.0; 11.3)	3.7	8.6	1.1	7.4	17.1	2.3
PL:BanTRIMP	1 <sup>st</sup>	0.70 (0.43; 0.87)	6.9 (5.5; 9.8)	0.3	0.8	0.1	6.6	15.4	2.4
	2 <sup>nd</sup>	0.79 (0.45; 0.94)	7.9 (5.7; 13.2)	0.4	1.0	0.1	8.2	19.0	2.8
HID:BanTRIMP	1 <sup>st</sup>	0.80 (0.59; 0.92)	14.5 (11.5; 20.4)	0.8	1.8	0.3	15.2	35.4	6.0
	2 <sup>nd</sup>	0.79 (0.45; 0.94)	14.7 (10.9; 24.6)	0.7	1.5	0.3	12.6	29.3	5.3
TD:ave%HR <sub>Res</sub>	1 <sup>st</sup>	0.83 (0.65; 0.93)	2.8 (2.2; 4.0)	2.1	4.9	0.9	3.1	7.1	1.3
	2 <sup>nd</sup>	0.89 (0.67; 0.97)	3.3 (2.5; 5.6)	2.6	6.1	1.1	3.8	8.9	1.6
PL:ave%HR <sub>Res</sub>	1 <sup>st</sup>	0.87 (0.71; 0.95)	5.6 (4.4; 8.0)	0.4	0.9	0.2	5.6	13.1	2.8
	2 <sup>nd</sup>	0.95 (0.85; 0.99)	4.2 (3.1; 7.0)	0.4	0.8	0.2	4.9	11.4	2.9
HID:ave%HR <sub>Res</sub>	1 <sup>st</sup>	0.75 (0.50; 0.90)	15.0 (11.8; 21.1)	1.1	2.6	0.4	16.1	37.6	5.5
	2 <sup>nd</sup>	0.65 (0.22; 0.9)	16.2 (12.1; 27.2)	0.9	2.1	0.4	12.9	29.9	4.9

Note: ICC – intraclass correlation coefficient; CI – confidence interval; CV – coefficient of variation; SEM – standard error of measurement; MDC – minimum detectable change; SWC – smallest worthwhile change; TD – total distance; PL – PlayerLoad™; HID – high-intensity distance; iTRIMP – individualized TRIMP; BanTRIMP – Banister TRIMP; ave%HR<sub>Res</sub> – average percentage of heart rate reserve

*Reproducibility of load ratios*

The three matches were played on the same regulation size, natural grass outdoor pitch under very similar environmental conditions (humidity:  $92 \pm 9\%$ , temperature:  $14 \pm 3^\circ\text{C}$ , wind speed:  $17 \pm 6 \text{ km} \cdot \text{h}^{-1}$ ). Both teams played in a 4-3-3 formation with players maintaining their tactical position through all games. Three players missed a game due to being called up to train with an older team within the club's academy structure and one player had to be excluded from the analysis due to injury during the final match. Of the remaining 12 players, 7 completed the full 90 minutes in all three consecutive matches. Throughout the three matches, those 7 players on average covered a TD of  $11,130 \pm 909 \text{ m}$ , with a HID of  $1,067 \pm 236 \text{ m}$ , a PL of  $1,184 \pm 183 \text{ AU}$ , at an average of  $80 \pm 4\%$  of  $\text{HR}_{\text{Res}}$ , with a BanTRIMP score of  $226 \pm 23 \text{ AU}$ , and an iTRIMP score of  $214 \pm 40 \text{ AU}$ . Between-athlete and within-athlete reproducibility of the different integrated external-to-internal load ratios is presented in Table 1.

**DISCUSSION**

The present study sought to establish the association of different combinations of match-derived external-to-internal load ratios with various submaximal and maximal fitness parameters and to quantify their reproducibility across three time points. The main findings were that (i) while most combinations of integrated load ratios were moderately-to-largely correlated with the various fitness parameters at a single time point, these results must be interpreted cautiously due to the rather wide confidence intervals; (ii) the combination of PL as an external load parameter and  $\text{ave}\% \text{HR}_{\text{Res}}$  as an internal load parameter demonstrated the most consistent associations with different fitness markers in comparison to other load ratios; (iii)  $\text{ave}\% \text{HR}_{\text{Res}}$  can be used as a simpler internal load alternative to iTRIMP for the ratio calculations; (iv) combinations involving HID showed reduced within-athlete reproducibility compared to combinations involving TD and PL as external load parameters; and (v) none of the ratios may be sensitive enough to detect small but relevant changes in athlete fitness.

In comparison to previous studies [7, 8], the relationships between different combinations of external-to-internal load ratios and markers of physical fitness found in the current study were less strong and less consistent. This may be partially explained by the use of actual match-derived external and internal load data as opposed to data derived from soccer match running simulations. While the results of simulations of physical match demand may be associated with isolated markers of physical fitness and may reproduce certain aspects of match running performance [31], these protocols miss important physical actions and their impact on internal load, specifically in relation to short but intense accelerations and decelerations, changes of directions, dives, jumps or tackles [32]. Furthermore, the cognitive stressors associated with ball movement activities and the tactical reading of the game cannot be adequately replicated by match running simulations. In fact, it has been suggested that team

sport athletes may experience mental fatigue due to prolonged and demanding match-related cognitive activity and that mental fatigue may induce reductions in physical performance during intermittent team sport activity [33]. It is important to point out that the observed reduction in physical performance in those studies was most likely due to an increased perception of effort rather than cardiovascular or metabolic mechanisms. Therefore, it seems fair to speculate that match-derived external-to-internal load ratios may be influenced not only by physical, but also by mental requirements. As such, comparably weaker associations between these load ratios and laboratory-derived markers of physical fitness are not necessarily surprising. This should not, however, automatically be interpreted as a potential weakness of the load ratios. While isolated fitness tests can more accurately determine changes in the physiological profile of a player, integrated load ratios may still be able to more holistically capture a player's ability to withstand competition-specific demands.

Apart from match-derived load ratios, another promising approach may be the integration of external and internal load data derived from standardised training drills. Training drills can be designed to possess fewer degrees of freedom compared to match play and may therefore correlate better with markers of fitness tests. In fact, Taylor *et al.* [9] demonstrated that integrated load ratios derived from a sprint interval training drill (6  $\times$  6-second sprint with 54 seconds recovery) and from a small-sided game (6 vs. 6, 10 minutes on a 39  $\times$  51 m pitch) showed large-to-very large associations with submaximal fitness markers and acceptable reliability in academy rugby union players. While similar results might be expected for soccer players and, therefore, potentially represent an interesting alternative to match-derived data, we want to highlight that no training drill can totally replicate all physical, technical, tactical, and cognitive demands that players experience during actual match play. As such, a certain trade-off between standardization and specificity should be expected.

Previous studies have already investigated different external load parameters to be included in the integrated load ratio calculations [7–10]. However, iTRIMP was the sole parameter of internal load to be examined. The results of the present study indicate for the first time that  $\text{ave}\% \text{HR}_{\text{Res}}$  can be used as a less complex alternative to iTRIMP, facilitating an easier use of external-to-internal load ratios as a marker of physical fitness. While these findings seemingly favour a simpler approach with regards to the internal load parameter, a reduction of the chaotic nature of professional soccer matches into a single locomotive parameter might be overly simplistic. Lacombe *et al.* [34] argue, for instance, that a more individualised approach taking into account player-specific combinations of external load variables may be more suitable to detect alterations in player-specific fitness. While their approach largely correlates with the results of a submaximal running test, the magnitude in correlation is not substantially larger than the best performing ratio in the present study. Therefore, it appears that more sophisticated approaches with regards to both external and internal load calculations do not

necessarily outperform integrated ratios consisting of more easily derived parameters. Also, the attractiveness of easy-to-use indicators with high face value must be emphasised.

The present study was the first to quantify the between-athlete reliability of integrated load ratios, generally demonstrating high-to-very high reproducibility. With regards to within-athlete reliability, one prior study examined several external-to-internal load ratios during three different training drills (continuous shuttle run, sprint interval training, small-sided game) in academy rugby union players [9] and found the CV of the ratios to range between 7.1 and 36.3%. Higher variability was found for ratios involving high-speed running as an external load parameter, which was in line with the present study. Given that match-to-match variability in high-speed running has been shown to be high among professional soccer players [11], it seems likely that this variability carries over into the integrated load ratios. As such, it is not advised to use these specific ratios as indicators of aerobic fitness. In contrast, ratios involving TD and PL as external load parameters showed reduced variability. One possible explanation may be that TD and PL are more volume-based parameters and, therefore, are not as affected by match-to-match variability as more intensity-based parameters such as HID. Nevertheless, an analysis of MDC and SWC revealed that none of the ratios may be sensitive enough to detect small but practically meaningful changes in athlete fitness. This may not be entirely surprising given that the ratios entail the variability specific to both the external [11, 35] and the internal load [36] parameters. As such, for an integrated load ratio consisting of PL and  $\text{ave\%HR}_{\text{Res}}$ , only changes greater than 11–13% can be confidently assumed to be greater than the normal day-to-day variability, while changes greater than 3% may already be practically meaningful. It needs to be pointed out, however, that due to the way MDC and SWC are calculated and the probabilities of finding the true value associated with both approaches (90% confidence for MDC vs. 16% confidence for SWC), readers should not compare SWC and MDC. Instead, the values provided by the present study should be critically evaluated and interpreted specifically to a given context.

Previous research on match-derived external-to-internal load ratios has shown that there may be substantial within-athlete ratio differences between the first and second halves of a match [37], indicating that – as the match goes on – external output decreases disproportionately in relation to internal load due to accumulated fatigue. Taking this finding into account, first and second halves were analysed separately to investigate any potential differences with regards to ratio validity and reproducibility. The results showed, however, that there were no clear and consistent between-half differences, indicating that fatigue does not seem to impact these ratio properties and that periods as short as a single half can be used to calculate and interpret changes in integrated load ratios.

There are limitations to this study that readers should be aware of. Given that the ratios are based on match-derived external and internal load data, they can only be calculated for players participating in the match but not for the rest of the squad. While it is fair to argue that the fitness of starters and substitutes is more relevant to team success than the fitness of fringe players, an alternative perspective would be that changes in fitness in those players might be more likely due to a lack of regular playing time and therefore require closer attention. Furthermore, the present study established load ratio validity at a single time point. While the between-athlete comparison is an important element of a given test, analyses of within-athlete changes over time are likely the more pertinent issue for athlete monitoring purposes and should be evaluated in the future. Thirdly, both maximal fitness assessments were performed on the same day. While a passive rest period of at least 6 h between the two tests was ensured, it is possible that results of the Yo-Yo IR1 test may have been negatively influenced by the earlier treadmill test. Fourthly, it has been pointed out previously that there may be statistical problems associated with the use of ratios. In certain circumstances, ratios might not be scaling consistently across a range of measured values of the denominator, indicating that the ratio would provide biased estimates at the lower and higher ends of the range [38]. Finally, the relatively wide confidence intervals indicate that a larger sample size is required to be more confident about the results of this study.

## CONCLUSIONS

In conclusion, the findings of this study indicate that certain match-derived external-to-internal load ratios are largely correlated with submaximal and maximal markers of cardiorespiratory fitness, with a combination of PL and  $\text{ave\%HR}_{\text{Res}}$  demonstrating the most consistent associations. Practitioners looking to use integrated ratios to monitor changes in fitness in soccer players need to be aware of the uncertainty associated with the ratios, however. In the case of PL: $\text{ave\%HR}_{\text{Res}}$ , only changes in excess of 11.4–13.1% can be confidently assumed to represent a true change. Hence, the ratios may not be sensitive enough to detect small yet practically relevant alterations in player fitness.

## Acknowledgements

The authors would like to extend their gratitude to all players and the team's coaching and supporting staff for their participation and collaboration in the study.

## Disclosure of Interest

The authors report no conflict of interest.

## REFERENCES

1. Anderson L, Orme P, Di Michele R, Close GL, Milsom J, Morgans R, Drust B, Morton JP. Quantification of seasonal-long physical load in soccer players with different starting status from the English premier league: Implications for maintaining squad physical fitness. *Int J Sports Physiol Perform.* 2016 Nov; 11(8):1038–46.
2. Akenhead R, Nassis GP. Training Load and Player Monitoring in High-Level Football: Current Practice and Perceptions. *Int J Sports Physiol Perform.* 2016 Jul; 11(5):587–93.
3. Shushan T, McLaren SJ, Buchheit M, Scott TJ, Barrett S, Lovell R. Submaximal Fitness Tests in Team Sports: A Theoretical Framework for Evaluating Physiological State. *Sports Med.* 2022 Nov; 52(11):2605–26.
4. Schimpchen J, Correia PF, Meyer T. Minimally Invasive Ways to Monitor Changes in Cardiac Fitness in Running-based Sports: A Systematic Review. *Int J Sports Med.* 2022 Nov 4; a-1925–7468.
5. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and External Training Load: 15 Years On. *Int J Sports Physiol Perform.* 2019; 14(2):270–3.
6. Stone NM, Kilding AE. Aerobic Conditioning for Team Sport Athletes. *Sports Med.* 2009 Aug; 39(8):615–42.
7. Akubat I, Barrett S, Abt G. Integrating the internal and external training loads in soccer. *Int J Sports Physiol Perform.* 2014; 9(3):457–62.
8. Akubat I, Barrett S, Sagarra M, Abt G. The Validity of External: Internal Training Load Ratios in Rested and Fatigued Soccer Players. *Sports.* 2018; 6(2):44.
9. Taylor RJ, Sanders D, Myers T, Akubat I. Reliability and Validity of Integrated External and Internal Load Ratios as Measures of Fitness in Academy Rugby Union Players. *J Strength Cond Res.* 2020 Jun; 34(6):1723–30.
10. Malone S, Doran D, Akubat I, Collins K. The Integration of Internal and External Training Load Metrics in Hurling. *J Hum Kinet.* 2016; 53(1):211–21.
11. Gregson W, Drust B, Atkinson G, Salvo VD. Match-to-match variability of high-speed activities in premier league soccer. *Int J Sports Med.* 2010; 31(4):237–42.
12. Rampinini E, Coutts A, Castagna C, Sassi R, Impellizzeri F. Variation in Top Level Soccer Match Performance. *Int J Sports Med.* 2007 Dec; 28(12):1018–24.
13. Lima-Alves A, Claudino JG, Boussoia D, Rangel Couto C, Teixeira-Coelho F, Pimenta EM. The relationship between internal and external loads as a tool to monitor physical fitness status of team sport athletes: a systematic review. *Biol Sport.* 2022; 39(3):629–38.
14. Drust B. Applied science and soccer: a personal perspective on the past, present and future of a discipline. *Sport Perform Sci Rep.* 2019;
15. Ellis M, Penny R, Wright B, Noon M, Myers T, Akubat I. The dose–response relationship between training-load measures and aerobic fitness in elite academy soccer players. *Sci Med Footb.* 2021; 5(2):128–36.
16. Manzi V, Iellamo F, Impellizzeri F, D’Ottavio S, Castagna C. Relation between individualized training impulses and performance in distance runners. *Med Sci Sports Exerc.* 2009 Nov; 41(11):2090–6.
17. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med.* 2000; 30(1):1–15.
18. Beckerman H, Roebroek ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek ALM. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res.* 2001; 10:571–8.
19. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes.* 2006; 4(May 2014).
20. Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc.* 1999; 31(3):472–85.
21. McKay AKA, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, Sheppard J, Burke LM. Defining Training and Performance Caliber: A Participant Classification Framework. *Int J Sports Physiol Perform.* 2022 Feb 1; 17(2):317–31.
22. Jones AM, Doust JH. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *J Sports Sci.* 1996 Aug; 14(4):321–7.
23. Bangsbo J, Iaia FM, Krstrup P. The YoYo intermittent recovery test: A useful tool for evaluation of physical performance in intermittent sports. *Sports Med.* 2008; 38(1):37–51.
24. Scott MTU, Scott TJ, Kelly VG. The validity and reliability of global positioning systems in team sport: A brief review. *J Strength Cond Res.* 2016; 30(5):1470–90.
25. 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(6):704–8.
26. Barrett S, Midgley A, Lovell R. PlayerLoad: Reliability, Convergent Validity, and Influence of Unit Position During Treadmill Running. *Int J Sports Physiol Perform.* 2014; 9(6):945–52.
27. Banister EW. Modeling elite athletic performance. In: MacDougall JD, Wenger HA, Green HJ, editors. *Physiological testing of elite athletes. Champaign (Ill.): Human Kinetics; 1991.* p. 403–24.
28. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Med Sci Sports Exerc.* 2009; 41(1):3–12.
29. Weir JP. Quantifying the test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005; 19(1):231–40.
30. Hopkins WG. Spreadsheets for Analysis of Validity and Reliability. *Sportscience.* 2015; 19:36–45.
31. Williams JD, Abt G, Kilding AE. Ball-Sport Endurance and Sprint Test (BEAST90): validity and reliability of a 90-minute soccer performance test. *J Strength Cond Res.* 2010; 24(12):3209–18.
32. Bloomfield J, Polman RCJ, Donoghue PGO. Physical Demands of Different Positions in FA Premier League Soccer. *J Sports Sci Med.* 2007; 6(1):63–70.
33. Smith MR, Marcora SM, Coutts AJ. Mental fatigue impairs intermittent running performance. *Med Sci Sports Exerc.* 2015; 47(8):1682–90.
34. Lacombe M, Simpson B, Broad N, Buchheit M. Monitoring Players’ Readiness Using Predicted Heart-Rate Responses to Soccer Drills. *Int J Sports Physiol Perform.* 2018 Nov; 13(10):1273–80.
35. Thornton HR, Nelson A, Delaney JA, Serpiello F, Duthie GM. Inter-Unit Reliability and Effect of Data Processing Methods of Global Positioning Systems. *Int J Sports Physiol Perform.* 2019; 14(4):432–8.
36. Lamberts RP, Lambert MI. Day-to-day variation in heart rate at different levels of submaximal exertion: implications for monitoring training. *J Strength Cond Res.* 2009 May; 23(3):1005–10.
37. Torreño N, Munguía-Izquierdo D, Coutts A, Sáez de Villarreal E, Asian-Clemente J, Suarez-Arrones L. Relationship Between External and Internal Loads of Professional Soccer Players During Full Matches in Official Games Using Global Positioning Systems and Heart-Rate Technology. *Int J Sports Physiol Perform.* 2016; 11(7):940–6.
38. Atkinson G, Batterham AM. The use of ratios and percentage changes in sports medicine: Time for a rethink? *Int J Sports Med.* 2012; 33(7):505–6.