Can small-sided games assess the training-induced aerobic adaptations in elite football players?

Running title: Small-sided games and training status

Andrea RIBOLI [1] *, Alexandre DELLAL[2, 3], Fabio ESPOSITO [1], Giuseppe CORATELLA [1]

[1] *Department of Biomedical Sciences for Health, Università degli Studi di Milano, Italy.*

[2] *Inter-university Laboratory of Human Movement Biology, University Claude Bernard Lyon 1, University of Lyon, Lyon, France.*

[3] *FIFA Medical Centre of Excellence, Centre Orthopédique Santy, Sport Science and Research Department, Lyon, France*

**Corresponding author:*

Andrea Riboli

Department of Biomedical Sciences for Health, Università degli Studi di Milano, Italy.

E-mail: riboliandrea@outlook.com

Abstract

BACKGROUND: To investigate whether or not SSGs could be used to evaluate the aerobic fitness status and the longitudinal training-induced adaptations in football players. Additionally, the capacity of SSGs to recreate the official match demands was investigated.

METHODS: Twenty-five elite football players were monitored. Total distance (TD), high-speed running, very high-speed running, sprint and accelerations plus decelerations distance were measured during 20 SSGs formats and 25 official-matches; in SSGs, average heart rate was also collected. During submaximal Yo-Yo test, heart rate at peak exercise, heart rate post-60s recovery and rate of perceived exertion were collected. Coefficient of variation, interclass correlation-coefficient and correlation-coefficient analysis were used to calculate validity, reliability, construct validity and, internal and external responsiveness of SSGs demands.

RESULTS: In SSGs, a *small* variability (~6.0%) with *moderate* reliability (~0.542 to ~0.663) was found in TD and heart rate, while a *high* variability (~20.8% to ~60.3%) with *poor* to *moderate* reliability (~0.358 to ~0.605) was observed in the other metrics; in submaximal Yo-Yo, heart rate showed *small* variability (~3.7%) with *good* reliability (~0.933 to ~0.916). The SSGs demands showed poor internal and external responsiveness (p>0.05) to the training-induced aerobic adaptations as assessed by submaximal Yo-Yo. The construct validity of SSGs showed overall *large* to *very large* correlations (r=0.53 to 0.90, p<0.05) between SSGs and official match demands across the season.

CONCLUSIONS: SSGs should not replace standardized field tests to detect the training-induced aerobic adaptations. However, SSGs could be confidently used to recreate specific contextual factors in elite football players.

Key words: team sports; performance; locomotor demands; physiological demands; fitness assessment

Introduction

The quantification of physical demands in team sports has become crucial for leading the performance development towards an evidence-based practice.[1,2] Monitoring the stress applied to the athletes using both external and internal load indicators is strongly recommended to maximize the training responses.[2] In football, the training load recorded during both training sessions and matches is used to manipulate the individual workload [1-3] and to monitor the running-based exercises (e.g., high-intensity interval training) and/or football-specific drills (e.g. small-sided games, SSGs).[1,3] Various tracking technologies [e.g., global positioning system (GPS) or semi-automatic video-analysis] are currently used to quantify the physical demands during training [3] and matches [4,5] using different metrics, such as total distance (TD) covered, the distance covered at different running speed and the total distance at high-intensity accelerations and decelerations.[6] Additionally, internal load variables such as the heart rate (HR) and the rating of perceived exertion (RPE) are used to quantify the physiological responses to the imposed external load.[2]

Beyond the importance of quantifying physical and physiological demands, determining the fitness status of each player allows to evaluate the training-induced aerobic adaptations across the season. Several maximal testing protocols performed in the laboratory [7,8] and/or on the pitch [9-11] are a valid and reliable way to determine athletes' cardiorespiratory, metabolic and biomechanical characteristics and to monitor training responses.[7,9] However, due to an extremely growing number of matches, the use of maximal testing protocols would not be easily utilized and repeated across the season, especially within elite football which is faced with congested periods of national and international competitions.

Therefore, submaximal testing protocols seem mandatorily required to detect the fitness status and training-induced adaptations.[12] For instance, submaximal testing assessment has been previously determined as a valid and practical alternative to incremental tests to exhaustion for endurance capacity evaluation in competitive cyclists.[13] In elite football, submaximal testing procedures such as the Yo-Yo submaximal version (Yo-Yo$_{submax}$) were suggested as valid tools to detect intermittent exercise performance,[9,12,14,15] to assess the training-induced aerobic adaptations,[14-17] and to adjust the training load for high-performance development[9,12,18]. However, the time span to test the players is extremely reduced for assessing both maximal and even submaximal testing protocols.

As a supplementary solution, the interest about the day-by-day use of internal and external load to possibly determine the aerobic fitness status is rapidly growing among practitioners.[19-21] In practice, SSGs are largely used since they were shown to effectively replicate the technical[22], tactical[23] and physical football-specific demands.[3,24,25] Additionally, it was recently suggested that the internal and external load recorded during SSGs may possibly represent the aerobic fitness status in professional football players.[19] The authors reported a correlation between TD, high-intensity activities and HR during 5-a-side (25 x 25 m) SSGs and the distance covered during the Yo-Yo intermittent recovery test.[19] Another study has examined the correlation between TD and high-intensity activities during 6-a-side (40 x 34 m) SSGs and the distance covered during Yo-Yo intermittent recovery test, reporting similar findings.[26] Nonetheless, while the former [19] suggested 5-a-side format as a possible supplement for the aerobic fitness assessment, the latter [26] concluded that the 6-a-side format results is too variable in the high-intensity activities and could not be used for fitness assessment.[26] These controversial outcomes maintain an open debate and sustain a growing interest about the possible use of SSGs demands to assess aerobic training status in practitioners.

However, the responsiveness of SSGs to detect the training-induced aerobic adaptations (i.e., internal responsiveness) [29,30] and their longitudinal validity to detect aerobic changes over the time (i.e., external responsiveness) [29,30] have yet to be examined. Therefore, the present study aimed to understand whether or not SSGs could be used to evaluate the aerobic fitness status and the longitudinal training-induced adaptations in football players. Additionally, the capacity of SSGs to recreate the official match demands was investigated.

Materials and Methods

Subjects

Twenty-five elite football players competing in a top-class European championship and in international UEFA tournaments were involved in the present study (age: 26 ± 6 years; body mass: 78 ± 8 kg; body height: 1.85 ± 0.08 m). The goalkeepers were excluded from data collection. The club's medical staff certified the health status of each player. Injured players were excluded from data collection for at least one month after their return to full training. The present data arose from the daily player monitoring in which players' activities are routinely measured over the course of the season. The Local Ethics Committee (protocol #102/14) approved the study. It was performed in accordance with the principles of the Declaration of Helsinki (1975).

Procedures

The present investigation was carried out during the competition period across five months from July to November. Training and matches were monitored across players' traditional weekly routine. A total of 91 individual samples were collected during Yo-Yo$_{submax}$ across five testing sessions (July, T1; August, T2; September, T3; October, T4; November, T5). A total of 3309 individual samples across 20 different SSGs formats were collected. The number of players, pitch area and number of samples for each SSG format are shown in Table I. Only SSGs format with goalkeepers' presence were considered. To gain and maintain the maximal effort during SSGs, a ball was always available by prompt replacement when it went out-of-play and the corners were replaced by a prompt ball-in-game from the goalkeeper.[3] A total of 188 individual samples across 25 official home-matches were recorded; only players who completed the match were inserted in data analysis. A specialized and high-qualified physician staff recommended and monitored the diet regime of each player before and after every training and match session. Each training session and SSGs were performed under the supervision and motivation of several highly-qualified coaches to keep up a high work-rate.[28] The individual cardiovascular fitness was assessed once per month using the submaximal version of the Yo-Yo endurance level 2 (Yo-Yo$_{submax}$).[12]

Training sessions were performed on two grass pitches (105 x 65 m). The matches were played in the home-team' stadium with an official pitch-size of about 105 x 65 m. The Yo-Yo$_{submax}$ was performed on a synthetic playing surface to increase test-by-test reliability.[12] Both natural and synthetic pitches were preserved by qualified operators.

*** Table I***

Small-sided games and match performance analysis

A 10Hz GPS (K-sport, Montelabbate, Italy) was used to collect data during the training sessions [3]. Each device was turned on at least 15-min before each session to allow for acquisition of the satellite signal.[3] To reduce the inter-unit differences, each player wore the same unit for every training session over the whole investigation.[3] The locomotor activities during the official matches were collected using a computerized semi-automated video-based multi-camera image system (Stats Perform, Chicago, Illinois, USA).[3] The systems have previously been shown to provide valid and reliable measurements of the match activity in football.[6] GPS and the video-based multi-camera image system was previoulsy detemined as interchangeable.[3]

During both training sessions and matches, TD, high-speed running distance (HSRD, 15 to 20 $km \cdot h^{-1}$), very high-speed running distance (VHSRD, 20 to 24 $km \cdot h^{-1}$), sprint distance (SPD, >24 $km \cdot h^{-1}$) and the total of high-intensity accelerations and decelerations distance >3 $m \cdot s^{-2}$ (Acc+Dec) [3,5] were measured. TD, HSRD, VHSRD, SPD and Acc+Dec were normalized as relative distance covered in one minute ($m \cdot min^{-1}$) and inserted into the data analysis. Additionally, HR was continuously recorded by a HR monitor (Firstbeat Technologies, Jyväskylä, Finland) during each SSG in order to determine the average HR during each SSG (HR_{SSGs}) inserted into data analysis. To calculate the construct validity, internal and external responsiveness of SSGs, only the SSGs formats repeated during each period were considered (Table I).

The Yo-Yo endurance level 2 submaximal test

The $Yo\text{-}Yo_{submax}$ consisted of repeated 20-m shuttle runs at progressively increasing speeds dictated by an audio beep. The $Yo\text{-}Yo_{submax}$ lasted 4 min (600 m) and it was performed on an artificial surface on a 2x20 m running lane marked by cones.[12] At the end of the test, players were asked to immediately stop running and to recover standing for at least 120-s. Across the season, $Yo\text{-}Yo_{submax}$ was incorporated once per month and it was performed after one day of complete recovery in accordance with the week-by-week periodized training program. The validity and reliability of the $Yo\text{-}Yo_{submax}$ was previously established.[9,12] HR was continuously recorded to determine the peak HR reached at the end of the test (HR_{peak}) and the HR after 60-s of complete recovery (HRR_{60s}). At the end of each test, the rate of perceived exertion (RPE) was collected using a CR10 Borg scale (i.e., from 0 to 10 point). For the purposes of the present study, the test-by-test HR_{peak}, HRR_{60s} and RPE were compared to determine between-periods differences in training status. It ensures to consistently compare variations in locomotor demands using SSGs with individual cardiovascular variations (i.e., different training status).

Statistical Analysis

Statistical analysis was performed using a statistical software package (SigmaPlot v-12.5, Systat Software Inc., San Jose, CA, USA). To check the normal distribution of the sampling, a Shapiro-Wilk test was used.

The validity and the reliability were determined as the between-SSGs and between-matches typical error for TD, HSRD, VHSRD, SPD, and Acc+Dec calculated as coefficient of variation (CV%). The CV% was calculated also for HR_{SSGs} and for HR_{peak}, HRR_{60s} and RPE measured using Yo-Yo_{submax}. The between-measures reliability was calculated for each parameter using interclass correlation coefficient (ICC) and interpreted as follow: <0.50 *poor* reliability, 0.50 to 0.75 *moderate* reliability >0.75: *good* reliability.[31]

The construct validity of SSGs was assessed in T2, T3, T4, T5, with the correlation between the indicator of the average locomotor demands (i.e., TD, HSRD, VHSRD, SPD and Acc+Dec) in SSGs and official matches. Since T1 was a pre-season period with no official matches played, T1 was excluded from the present analysis. The correlation coefficient was interpreted as follows: r =0.00-0.09 *trivial*, 0.10-0.29 *small*, 0.30-0.49 *moderate*, 0.50-0.69 *large*, 0.70-0.89 *very large*, 0.90-0.99 *nearly perfect*.

The internal responsiveness of the Yo-Yo_{submax} and SSGs was described as the magnitude of training-induced aerobic adaptations across the season (i.e., from T1 to T5). One-way analysis of variance (ANOVA) for repeated measures was used to detect the between-periods differences in locomotor demands during both Yo-Yo_{submax} and SSGs. The magnitude of the changes across different periods was assessed using Cohen's d effect size (ES) with 95% confidence intervals (95% CI) and interpreted as follows: <0.20: *trivial*; 0.20-0.59: *small;* 0.60-1.19: *moderate;* 1.20-1.99: *large*; ≥2.00: *very large.* To make a probabilistic mechanic inference about the true between-periods changes/differences the smallest worthwhile change (SWC) was also calculated. The likelihood of between-periods changes for paired comparisons during both Yo-Yo_{submax} and SSGs demands were calculated as follow: *trivial* [0.2*(pretests between-subjects standard deviations, SD)], *possible* [(pretests between-subjects mean/100)*(mean CV%)] and *certain* [2*((pretests between-subjects mean/100)*(mean CV%))].[32]

The external responsiveness of the demands during SSGs was calculated as the correlation between the training-induced aerobic adaptations during both Yo-Yo_{submax} and SSGs across different testing periods (i.e., from T1 to T5). Statistical significance was set at $\alpha < 0.05$. Unless otherwise stated, all values are presented as mean ± standard deviation (SD).

Results

Small-sided games variability and reliability

Locomotor and physiological demands for each parameter during different SSGs format are reported in Supplementary table I. The variability and the reliability for locomotor and physiological demands during each SSG are displayed in Figure 1. On average, TD showed ~6.0% CV with *moderate* ICC (~0.542); HSRD showed ~20.8% CV with *moderate* ICC (~0.546); VHSRD showed ~40.0% CV with *poor* ICC (~0.422); SPD showed ~60.3% CV with *poor* ICC (~0.358); Acc+Dec showed ~14.6% CV and *moderate* ICC (~0.605). On average, the HR_{SSGs} showed ~5.9% CV with *moderate* ICC (~0.663).

Figure 1

Official matches variability and reliability

The locomotor demands during official matches are reported in Supplementary table I. Official matches TD showed 5.2(0.7)% CV with a *moderate* ICC (0.740; CI: 0.450 to 0.920); HSRD showed 16.9(3.4)% CV with *moderate* ICC (0.624; CI: 0.350 to 0.820); VHSRD showed 19.7(3.3)% CV with *moderate*

ICC (0.561; CI: 0.250 to 0.720); SPD showed 28.8(4.7)% CV and a *poor* ICC (0.438; CI: 0.150 to 0.620); Acc+Dec showed a 5.3(0.8)% CV and a *moderate* ICC (0.684; CI: 0.410 to 0.880).

Comparison between official matches and SSGs variability

Comparing match and SSGs variability, TD showed a *small* [ES: 0.43 (-0.13 to 0.99)] higher ($p < 0.001$) CV in SSGs than match. For HSRD, CV was *moderately* [ES: 0.65 (0.08 to 1.22)] higher ($p = 0.023$) in SSGs (~20.8%) than match (~16.9%) while a *largely* to *very largely* (ES: 1.96/2.89) higher ($p < 0.001$) CV in SSGs than match was retrieved for VHSRD (~40.7% vs ~19.7% for SSGs and match, respectively), SPD (~60.3% vs ~28.8%) and Acc+Dec (~14.6% vs ~5.3%).

Yo-Yo$_{submax}$ variability and reliability

During Yo-Yo$_{submax}$, both the HR_{peak} and the HRR_{60s} showed ~3.7% and ~7.2% CV, respectively and *good* ICC (~0.933 and ~0.916, respectively) across overall periods. Conversely, RPE showed ~34.5% CV and *poor* ICC (~0.496).

Construct validity of SSGs

The construct validity of SSGs was verified during T2, T3, T4 and T5, respectively. Correlations between SSGs and official match demands are reported in Table II.

*** Table II ***

Internal responsiveness of SSGs and Yo-Yo$_{submax}$ to training-induced aerobic adaptations

The average between-period SSGs locomotor demand showed *trivial* to *small* (ES: 0.01/0.43) differences for TD, HSRD, VHSRD, SPD, Acc+Dec and HR_{SSGs} with only some exceptions: *moderate* differences were found in T3 *vs* T2 [ES: 1.42 (0.80 to 2.04); $p < 0.001$] for TD, T2 *vs* T1 [ES: 0.91(0.32 to 1.49), $p = 0.002$] for VHSRD, T3 *vs* T2 [ES: 1.14(0.54 to 1.74), $p < 0.001$] and T5 *vs* T4 [ES: -1.18(-1.78 to -0.58), $p < 0.001$] for Acc+Dec. Conversely, the magnitude of the test-by-test seasonal variations during Yo-Yo$_{submax}$ was *trivial* to *large* (ES: 0.01/1.92) for HR_{peak} and HRR_{60s}, and *trivial* to *very large* (ES: 0.08/2.05) for RPE, with the greatest improvement in T2 than T1 (ES: 0.39 to 2.05).

The likelihood of the between-periods changes in both Yo-Yo$_{submax}$ and SSGs demands is shown in Table III.

***Table III ***

External responsiveness between training-induced* adaptations *in SSGs and Yo-Yo$_{submax}$

Table IV shows the external responsiveness of the SSGs demands calculated as the correlation between training-induced aerobic adaptations during both Yo-Yo$_{submax}$ and SSGs across T2 *vs* T1, T3 *vs* T2, T4 *vs* T3 and T5 *vs* T4.

*** Table IV ***

Discussion

The present study investigated whether or not SSGs could be used to assess the training status and the in-season training-induced aerobic adaptations in elite football players. Additionally, the capacity of SSGs to recreate the official match demands was investigated. With the exception of the low variability with *moderate* reliability recorded in TD and HR_{SSGs}, a higher variability with *poor* to *moderate* reliability was observed during SSGs. The SSGs showed overall *large* to *very large* correlation with the official matches demands across the season. Conversely, the SSGs demands showed overall poor internal and external responsiveness to the training-induced aerobic adaptations as assessed by Yo-Yo_{submax} across the periods. As such, SSGs recreate official match demands and seem to be a useful tool to recreate football specific contextual factor during practice. Conversely, SSGs demands should not replace standardized testing procedures to assess the training status and/or to detect the longitudinal training-induced aerobic adaptations in elite football players.

Variability and reliability

Similar to previous findings,[26-28] the present results suggest a high impact of football-specific contextual factors on variability and reliability of SSGs demands, especially for the high-intensity activities. Although with different procedures and speed-zones, ~25% and ~29% variability was previously reported for distance covered at >16 and >22 $km \cdot h^{-1}$ in professional football players,[33] in agreement with the present results. Similarly, 5v5 SSG (pitch size: 25 x 25 m)[19] and 6v6 SSG (pitch size: 40 x 34 m)[26] showed a high variability (~15%) for high-intensity activities in European elite football players. A low reliability for high-speed running activities both in small and large-sided games (i.e. area per player ranging from 110 to 212 $m^2 \cdot$player) were also found in French Ligue 1 football players.[34]

The variability of the match demands is in line with previous findings in Premier League [27,35] in which the match-to-match variability was ~16% and ~30% for distance >19.8 $km \cdot h^{-1}$ and >25.2 $km \cdot h^{-1}$, respectively. Similarly, ~14% was found in European league players for distance >19.8 $km \cdot h^{-1}$ or >21.0 $km \cdot h^{-1}$.[28,36] Interestingly, the SSGs *vs* match demands showed larger variability. Previously, it was suggested that SSGs would show lower variability than match demands because of the more consistent independent variables, such as the same players, rules, pitch sizes, etc..[19] Nonetheless, the present findings contradict this hypothesis, probably because of factors affecting the physical and physiological demands such as technical/tactical contextual factors within the same SSG format across the season.

As expected, a lower variability in HR_{peak} and HRR_{60s} during Yo-Yo_{submax} than SSGs demands were found. The ~3.7% in variability for HR_{peak} during Yo-Yo_{submax} was similar to previous findings showing ~3-7% variability.[18] However, it should be acknowledged that the Yo-Yo_{submax} high reliability [18] and its sensitivity to training-induced aerobic adaptations [12] were previously demonstrated. The present results confirmed these latter findings and highlighted a low variability in aerobic training status across the periods in this elite football population. Lastly, RPE showed a high variability and *poor* reliability, questioning its reproducibility due to many psycho-biological factors possibly affecting the perceived effort.

Construct validity

Large to *very large* correlations were found between the SSGs and the official match demands across the season. Interestingly, the present results showed that the different formats of SSGs used here reflected the locomotor football-specific activities performed during official matches. These findings highlight SSGs as a useful tool to replicate the football-specific patterns during practice. However, it should be remarked that such a relationship does not imply that the SSGs *vs* match demands are similar, and several independent parameters (e.g., area per player, presence of goalkeeper, etc,) may be manipulated to replicate and/or overload official match demands [3] and the peak official match intensities.[4,5]

Internal and external responsiveness in SSGs and Yo-Yo_{submax} demands

Despite the good construct validity, poor internal and external responsiveness to determine training-induced aerobic adaptations were found for the SSGs demands than Yo-Yo_{submax} across each locomotor and physiological parameter. These results remark how much the high-intensity activities during SSGs are affected by technical and tactical factors independently by the athletes' training status. Moreover, the HR responses during SSGs can be possibly influenced by several aspects besides the physical conditioning of the players, such as external temperature, the relative humidity, the atmospheric pressure, the individual hormonal variations (i.e. adrenaline) and the medications' use.[16] Therefore, physical and technical performance are variable *per se* regardless of context.[35] As previously shown in top-class European football players,[26] the present findings demonstrate that SSGs could not be used as a valid and reliable fitness indicator. However, only the construct validity of 6v6 SSGs (pitch size: 40 x 34 m) with regards to Yo-Yo maximal version was previously investigated. To sustain the present hypothesis, the current study added novel findings showing the internal and external responsiveness across a wide range of SSGs formats from 4v4 to 10v10 with a pitch size ranging from ~75 to ~488 m^2·player. Remarkably, for the first time, poor internal and external responsiveness for both external and internal loads was found in a large-samples and a wide range of SSG formats currently used in elite football routine.

Physical SSGs and match performance are essential to provide a platform upon which objective decisions for fitness training and game preparation can be made.[37] However, the use of SSGs as tool to monitor the longitudinal training-induced aerobic adaptations across the season seems questionable, due to their unpredictable and highly variable contextual factors. To our knowledge, no comparison with previous studies about internal and external responsiveness of a wide range of SSGs in elite football players can be made, so future studies are warranted to confirm or contradict the current results. As such, the current findings overall point out that coaches and sport scientists should not replace standardized physiological assessment with the analysis of SSGs demands to detect seasonal training-induced aerobic adaptations. Conversely, the current results confirmed that practitioners should use confidently the Yo-Yo_{submax} to determine the in-season training-induced aerobic adaptations as previously observed in Premier League,[12] elite Dutch [12] and top-level Spanish La Liga football players.[9,14]

Limitations and future perspectives

The present findings come with some strength and limitations opening to future investigations. Firstly, SSGs with goalkeepers could possibly suffer from the greater influence of the tactical behaviors or contextual factors than SSGs without goalkeepers, in which football players are free to move across the pitch to keep ball-possession. Therefore, future studies investigating the relationship between training-induced aerobic adaptations and physical demands during SSGs without goalkeepers are required. Secondly, the SSGs performance can be affected by several physical, physiological and biomechanical factors. The present study mainly analyzed the internal and external responsiveness of SSGs with regards to the training-induced aerobic adaptations (i.e., using Yo-Yo$_{submax}$), while future studies should investigate also training-induced anaerobic and/or neuromuscular adaptations (e.g., repeated-sprint ability, countermovement jump, etc.). Furthermore, the external responsiveness of the official match demands (i.e the correlation between the training-induced aerobic adaptations during both Yo-Yo$_{submax}$ and match demands across different testing periods) has not been investigated. In the real-life context of the elite football, the demands of each player are affected by several confounding parameters (i.e., being starter or non-starter or playing or not a given game) affecting the procedures. Additional, different maximal aerobic testing procedures[7,38,39] could provide higher sensitivity to detect longitudinal training-induced aerobic adaptations that Yo-Yo$_{submax}$; however, it would to be remarked that these information come from elite real-life football practice in which maximal testing procedures should be avoid across the season due to both congested fixture periods and football-specific contextual factors. Lastly, for replication purposes the interchangeability between GPS and computerized semi-automated video-based multi-camera image system needs to be carefully checked, especially when recording high-speed or non-linear movements. However, the present results are based on trivial differences in the metrics and adjusted using a calibration equation as previously investigated using the same technologies.[3] Otherwise, when different technologies are used, specific calibration equations should be assessed.

Practical Applications

The SSGs demands are related to movement patterns required during official matches. Therefore, coaches and sport scientists should utilize SSGs with the aim to manage the training loads depending on the desired football-specific performance goals [3,5] (e.g., high-speed running, sprinting or accelerations and decelerations activities) with the purpose to recreate effectively the official match-demands.[3] This allows replicating/overloading average match demands [3] and/or the most demanding phases of match play [5,40] with regard to the distribution of the maximal intensities [4,41] across weekly routine [42] using different SSGs formats [3,25] with both performance and injury prevention purposes.[2,43] Conversely, practitioners should avoid using SSGs as a procedure to assess the aerobic training status and the longitudinal training-induced aerobic adaptations. The locomotor demands during SSGs appear not related to the seasonal variations in aerobic training status. To this purpose, the use of standardized testing procedures is required, and submaximal testing procedures could be used to monitor the training-induced aerobic adaptations in elite football.

Conclusions

The demands recorded during SSGs should not be used to assess the training status and/or to detect the longitudinal training-induced aerobic adaptations in elite football players. In this context, standardized maximal or sub-maximal testing procedures seem required. However, SSGs recreate the official match demands and could confidently be used to recreate the football-specific contextual factors during practice.

REFERENCES

1. Lacome M, Simpson BM, Cholley Y, Lambert P, Buchheit M. Small-Sided Games in Elite Soccer: Does One Size Fit All? *Int J Sports Physiol Perform.* 2018;13(5):568-576.
2. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and External Training Load: 15 Years On. *Int J Sports Physiol Perform.* 2019;14(2):270-273.
3. Riboli A, Coratella G, Rampichini S, Ce E, Esposito F. Area per player in small-sided games to replicate the external load and estimated physiological match demands in elite soccer players. *PLoS One.* 2020;15(9):e0229194.
4. Riboli A, Esposito F, Coratella G. The distribution of match activities relative to the maximal intensities in elite soccer players: implications for practice. *Res Sports Med.* 2021:1-12.
5. Riboli A, Semeria A, Coratella G, Esposito F. Effect of formation, ball in play and ball possession on peak demands in elite soccer. *Biol Sport.* 2021;38:195-205.
6. Di Salvo V, Baron R, Tschan H, Calderon Montero FJ, Bachl N, Pigozzi F. Performance characteristics according to playing position in elite soccer. *Int J Sports Med.* 2007;28(3):222-227.
7. Riboli A, Ce E, Rampichini S, et al. Comparison between continuous and discontinuous incremental treadmill test to assess velocity at VO2max. *J Sports Med Phys Fitness.* 2017;57(9):1119-1125.
8. Riboli A, Coratella G, Rampichini S, Limonta E, Esposito F. Testing protocol affects the velocity at VO2max in semi-professional soccer players. *Res Sports Med.* 2021:Epub of Print.
9. Bangsbo J, Iaia FM, Krustrup P. The Yo-Yo intermittent recovery test : a useful tool for evaluation of physical performance in intermittent sports. *Sports Med.* 2008;38(1):37-51.
10. Dellal A, Chamari K, Pintus A, Girard O, Cotte T, Keller D. Heart rate responses during small-sided games and short intermittent running training in elite soccer players: a comparative study. *J Strength Cond Res.* 2008;22(5):1449-1457.
11. Buchheit M. The 30-15 intermittent fitness test: accuracy for individualizing interval training of young intermittent sport players. *J Strength Cond Res.* 2008;22(2):365-374.

12. Bradley PS, Mohr M, Bendiksen M, et al. Sub-maximal and maximal Yo-Yo intermittent endurance test level 2: heart rate response, reproducibility and application to elite soccer. *Eur J Appl Physiol.* 2011;111(6):969-978.
13. Sassi A, Marcora SM, Rampinini E, Mognoni P, Impellizzeri FM. Prediction of time to exhaustion from blood lactate response during submaximal exercise in competitive cyclists. *Eur J Appl Physiol.* 2006;97(2):174-180.
14. Bradley PS, Di Mascio M, Bangsbo J, Krustrup P. The maximal and sub-maximal versions of the Yo-Yo intermittent endurance test level 2 are simply reproducible, sensitive and valid. *Eur J Appl Physiol.* 2012;112(5):1973-1975.
15. Bangsbo J, Mohr M, Poulsen A, Perez-Gomez J, Krustrup BR. Training and testing the elite athletes. *J Exerc Sci Fit.* 2006;4:1-10.
16. Dellal A, Diniz da Silva C, Hill-Haas SV, et al. Heart rate monitoring in soccer: interest and limits during competitive match play and training monitored by the heart rate in soccer players. *J Strength Cond Res* 2012;26:2890-2906.
17. Rago V, Krustrup P, Martin-Acero R, Rebelo A, Mohr M. Training load and submaximal heart rate testing throughout a competitive period in a top-level male football team. *J Sports Sci.* 2020;38(11-12):1408-1415.
18. Krustrup P, Mohr M, Amstrup T, et al. The yo-yo intermittent recovery test: physiological response, reliability, and validity. *Med Sci Sports Exerc.* 2003;35(4):697-705.
19. Owen AL, Newton M, Shovlin A, Malone S. The Use of Small-Sided Games as an Aerobic Fitness Assessment Supplement Within Elite Level Professional Soccer. *J Hum Kinet.* 2020;71:243-253.
20. Lacome M, Simpson BM, Buchheit M. Monitoring training status with player-tracking technology. *Aspetar Sports Medical Journal.* 2018;7:54-63.
21. Buchheit M. Monitoring training status with HR measures: do all roads lead to Rome? *Front Physiol.* 2014;5:73.

22. Clemente FM, Wong del P, Martins FM, Mendes RS. Acute effects of the number of players and scoring method on physiological, physical, and technical performance in small-sided soccer games. *Res Sports Med.* 2014;22(4):380-397.

23. Olthof SBH, Frencken WGP, Lemmink K. A Match-Derived Relative Pitch Area Facilitates the Tactical Representativeness of Small-Sided Games for the Official Soccer Match. *J Strength Cond Res.* 2019;33(2):523-530.

24. Castagna C, D'Ottavio S, Cappelli S, Araujo Povoas SC. The Effects of Long Sprint Ability-Oriented Small-Sided Games Using Different Ratios of Players to Pitch Area on Internal and External Load in Soccer Players. *Int J Sports Physiol Perform.* 2019:1265-1272.

25. Riboli A, Olthof SBH, Esposito F, Coratella G. Training elite youth soccer players: area per playe in small-sided games to replicate the match demands. *Biol Sport.* 2022;39(3):579-598.

26. Stevens TG, De Ruiter CJ, Beek PJ, Savelsbergh GJ. Validity and reliability of 6-a-side small-sided game locomotor performance in assessing physical fitness in football players. *J Sports Sci.* 2016;34(6):527-534.

27. 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):237-242.

28. Rampinini E, Coutts AJ, Castagna C, Sassi R, Impellizzeri FM. Variation in top level soccer match performance. *Int J Sports Med.* 2007;28(12):1018-1024.

29. Fanchini M, Schena F, Castagna C, et al. External Responsiveness of the Yo-Yo IR Test Level 1 in High-level Male Soccer Players. *Int J Sports Med.* 2015;36(9):735-741.

30. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.

31. Portney LG, Watkins MP. Validity of Measurements. *Foundations of Clinical Research: Applications to Practice.* 2015.

32. Hopkins WG. How to interpret changes in an athletic performance test. *Sportscience.* 2004;8:1-7.

33. Rago V, Silva JR, Mohr M, Barreira D, Krustrup P, Rebelo AN. Variability of activity profile during medium-sided games in professional soccer. *J Sports Med Phys Fitness.* 2019;59(4):547-554.

34. Dellal A, Chamari K, Payet F, Djaoul L, Wong DP. Reproducibility of physical performance during small- and large-sided games in elite soccer in short period: practical applications and limits. *J Nov Physiother.* 2016;6(6).

35. Bush MD, Archer DT, Hogg R, Bradley PS. Factors influencing physical and technical variability in the English Premier League. *Int J Sports Physiol Perform.* 2015;10(7):865-872.

36. Oliva-Lozano JM, Muyor JM, Fortes V, McLaren SJ. Decomposing the variability of match physical performance in professional soccer: Implications for monitoring individuals. *Eur J Sport Sci.* 2020:1-9.

37. Carling C. Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? *Sports Med.* 2013;43(8):655-663.

38. Riboli A, Coratella G, Rampichini S, Limonta E, Esposito F. Testing protocol affects the velocity at VO2max in semi-professional soccer players. *Res Sports Med.* 2021:1-11.

39. Riboli A, Rampichini S, Cé E, et al. Training status affects between-protocols differences in the assessment of maximal aerobic velocity *Eur J Appl Physiol.* 2021;Epub ahead of print.

40. Oliva-Lozano JM, Martinez-Puertas H, Fortes V, Muyor JM. When do soccer players experience the most demanding passages of match play? A longitudinal study in a professional team. *Res Sports Med.* 2021:1-11.

41. Oliva-Lozano JM, Fortes V, Muyor JM. When and how do elite soccer players sprint in match play? A longitudinal study in a professional soccer league. *Res Sports Med.* 2021:1-12.

42. Oliva-Lozano JM, Gomez-Carmona CD, Rojas-Valverde D, Fortes V, Pino-Ortega J. Effect of training day, match, and length of the microcycle on the worst-case scenarios in professional soccer players. *Res Sports Med.* 2021:1-14.

43. Buchheit M. Managing high-speed running load in professional soccer players: The benefit of high-intensity interval training supplementation. *Sport Perform Sci Reports.* 2019;53:1-5.

Conflicts of interest: The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Funding: The authors report no involvement in the research by the sponsor that could have influenced the outcome of this work.

Authors' contributions: AR have given substantial contributions to the conception of the manuscript. AR and GC have given substantial contributions to the design of the manuscript. AR have given substantial contributions to acquisition, analysis and interpretation of the data. AR and GC have given substantial contributions to writing original draft. AR, AD, FE and GC revised it critically. All authors read and approved the final version of the manuscript.

Acknowledgements: The authors thank Filippo Nardi M.Sc. for his committed help in the data collection.

TABLES

Table I: Small-sided games.

Player N°	Pitch size Length x width	Total area m^2	ApP m^2·player	Total sample N°	Individual Sample N° (SD)	Repeated across periods
10*v*10	105 x 65	6825	341	474	38.0 (21.4)	*
	72 x 65	4680	234	54	9.8 (5.4)	
	70 x 65	4550	228	237	41.5 (18.4)	*
	52 x 65	3380	169	680	64.5 (33.2)	*
9*v*9	52 x 65	3380	188	32	6.0 (4.3)	
	52 x 40	2080	116	64	26.5 (12.5)	*
8*v*8	52 x 65	3380	211	64	9.5 (9.0)	
	65 x52	3380	211	106	13.6 (5.0)	*
	65 x 45	2925	183	62	18.2 (10.6)	
7*v*7	105 x 65	6825	488	360	20.5 (7.1)	
	65 x 52	3380	241	78	13.8 (8.0)	
	52 x 40	2080	149	128	18.0 (12.9)	*
6*v*6	52 x 40	2080	173	182	19.1 (9.4)	*
	40 x 40	1600	133	68	16.0 (10.2)	
5*v*5	40 x 40	1600	160	140	33.0 (14.2)	*
	40 x 32	1280	128	158	44.0 (14.9)	*
	35 x 25	875	88	120	29.0 (18.5)	
4*v*4	40 x 32	1280	160	92	30.5 (18.6)	
	35 x 30	1050	153	104	35.0 (14.1)	*
	30 x 20	600	75	106	34.8 (17.0)	*

The small-sided games are split for the number of players and pitch size (length x width). The total pitch area and area per player (ApP) have been calculated. The total samples and the average number of observations per player [mean (SD)] for each condition are also reported. The SSGs formats repeated across each period (i.e., T1, T2, T3, T4, T5) are also clearly stated (*).

Table II: Correlations with 95% confidence intervals (95% CI) between SSGs and official matches demands across the season.

		Official matches											
		T2			**T3**			**T4**			**T5**		
		r	95% CI	*p*	*r*	95% CI	*p*	*r*	95% CI	*p*	*r*	95% CI	*p*
	TD	**0.651**	**0.228 to 0.867**	**0.006***	**0.782**	**0.378 to 0.936**	**0.002***	**0.898**	**0.618 to 0.980**	**<0.001***	**0.699**	**0.172 to 0.915**	**0.016***
	HSRD	**0.598**	**0.131 to 0.840**	**0.014***	**0.851**	**0.542 to 0.957**	**<0.001***	**0.801**	**0.347 to 0.951**	**0.005***	**0.636**	**0.059 to 0.895**	**0.035***
SSGs	VHSRD	0.427	-0.090 to 0.733	0.099	**0.749**	**0.307 to 0.925**	**0.005***	**0.840**	**0.448 to 0.961**	**0.002***	0.517	-0.120 to 0.853	0.103
	SPD	**0.528**	**0.045 to 0.812**	**0.035***	**0.688**	**0.188 to 0.905**	**0.013***	**0.741**	**0.209 to 0.935**	**0.014***	**0.601**	**0.012 to 0.879**	**0.049***
	Acc+Dec	0.365	-0.159 to 0.729	0.164	0.512	-0.087 to 0.839	0.088	**0.773**	**0.281 to 0.944**	**0.008***	0.533	-0.098 to 0.859	0.091

Abbreviations: SSGs, small-sided games; TD, total distance, HSRD, high speed running distance; VHSRD, very high-speed running distance; SPD, sprint distance; Acc+Dec, acceleration plus deceleration. T1, July; T2, August; T3, September; T4, October; T5, November. Bold text highlights significant correlations. *$p < 0.05$

Table III: The internal responsiveness of the SSGs and Yo-Yo$_{submax}$ is described as the magnitude of training-induced aerobic adaptations across the periods.

	Small-sided games						Yo-Yo$_{submax}$		
	TD	**HSRD**	**VHSRD**	**SPD**	**Acc+Dec**	**HR$_{SSGs}$**	**HR$_{peak}$**	**HRR$_{60s}$**	**RPE**
Unchange (%)	<1.6	<8.8	<12.0	<4.9	<4.9	<1.1	<1.1	<2.6	<3.6
Trivial change (%)	1.6 to 5.9	8.8 to 20.7	12.0 to 39.9	4.9 to 60.2	4.9 to 14.5	1.1 to 5.9	1.1 to 3.6	2.6 to 7.1	3.6 to 34.4
Possible change (%)	6.0 to 12.0	20.8 to 41.6	40.0 to 80.0	60.3 to 120.6	14.6 to 29.2	5.9 to 11.8	3.7 to 7.4	7.2 to 14.4	34.5 to 69.0
Certain change (%)	>12.0	>41.6	>80.0	>120.6	>29.2	>11.8	>7.4	>14.4	>69.0
T2 *vs* T1	↔	↑	↑	↑	↔	↓	↓↓	↓↓↓	↓
T3 *vs* T2	↑↑	↑	↔	↔	↑↑	↓	↓	↓	↓
T4 *vs* T3	↔	↔	↔	↔	↔	↑↑	↓↓	↓↓	↓
T5 *vs* T4	↓	↓	↔	↔	↓	↔	↔	↑	↑

Abbreviations: TD: total distance; HSRD: high speed running distance; VHSRD: very high-speed running distance; SPD: sprint distance; Acc+Dec: acceleration and deceleration distance; HR$_{SSGs}$: heart rate average during small-side games; HR$_{peak}$: heart rate peak; HRR$_{60s}$: heart rate recovery after 60 sec; RPE: rate of perceived exertion. T1: July; T2: August; T3: September; T4: October; T5: November. Values are presented as mean (SD). Likelihood of positive/negative training-induced aerobic changes were calculated for pair comparisons with the SWC and reported as follow:

↑ trivial increase than previous test

↑↑ possible increase than previous test

↑↑↑ certain increase than previous test

↓ trivial decrease than previous test

↓↓ possible decrease than previous test

↓↓↓ certain decrease than previous test

↔ no change than previous test

Table IV: The external responsiveness of the demands during SSGs is showed as the correlation between the training-induced aerobic adaptations during both Yo-Yo$_{submax}$ and SSGs across different testing periods. Correlations are described as mean (95% confidence intervals), *r* square and *p* value.

			TD	HSRD	VHSRD	SPD	Acc+Dec	HR_{SSGs}
HR_{peak}	T2 *vs* T1	*r*	-0.305 (-0.765 to 0.361)	-0.351 (-0.785 to 0.316)	-0.050 (-0.631 to 0.567)	-0.294 (-0.827 to 0.518)	0.789 (0.360 to 0.943)	-0.316 (-0.810 to 0.439)
		r^2	0.093	0.123	0.002	0.086	0.623	0.100
		p	0.362	0.291	0.884	0.480	**0.004***	0.406
	T3 *vs* T2	*r*	-0.556 (-0.847 to -0.007)	-0.771 (-0.928 to -0.383)	-0.621 (-0.873 to -0.106)	-0.410 (-0.784 to -0.181)	-0.381 (-0.770 to 0.216)	-0.518 (-0.853 to 0.118)
		r^2	0.309	0.595	0.386	0.168	0.145	0.268
		p	**0.049***	**0.002***	**0.024***	0.164	0.200	0.103
	T4 *vs* T3	*r*	-0.086 (-0.590 to 0.466)	-0.004 (-0.534 to 0.528)	-0.242 (-0.685 to 0.331)	-0.181 (-0.649 to 0.387)	-0.205 (-0.664 to 0.365)	-0.125 (-0.652 to 0.483)
		r^2	0.007	0.000	0.059	0.032	0.042	0.015
		p	0.770	0.989	0.404	0.535	0.481	0.698
	T5 *vs* T4	*r*	0.243 (-0.418 to 0.736)	0.308 (-0.359 to 0.766)	0.398 (-0.265 to 0.806)	0.860 (0.538 to 0.963)	0.113 (-0.522 to 0.668)	0.514 (-0.228 to 0.878)
		r^2	0.059	0.095	0.158	0.739	0.012	0.264
		p	0.472	0.358	0.225	**0.001***	0.741	0.157
HRR_{60s}	T2 *vs* T1	*r*	-0.274 (-0.750 to 0.390)	-0.308 (-0.767 to 0.358)	0.108(-0.526 to 0.665)	-0.517 (-0.895 to 0.295)	0.518 (-0.119 to 0.858)	-0.101 (-0.717 to 0.603)
		r^2	0.075	0.095	0.015	0.267	0.268	0.010
		p	0.416	0.356	0.752	0.190	0.103	0.795
	T3 *vs* T2	*r*	-0.333 (-0.747 to 0.267)	-0.533 (-0.838 to 0.025)	-0.572 (-0.854 to -0.030)	-0.579 (-0.854 to -0.032)	-0.468 (-0.810 to 0.112)	-0.477 (-0.837 to 0.172)
		r^2	0.111	0.284	0.327	0.328	0.219	0.228
		p	0.267	0.061	**0.041***	**0.041***	0.107	0.138
	T4 *vs* T3	*r*	-0.407 (-0.771 to 0.158)	-0.276 (-0.704 to 0.298)	-0.434 (-0.784 to 0.126)	-0.387 (-0.761 to 0.181)	-0.267 (-0.761 to 0.181)	-0.186 (-0.686 to 0.434)
		r^2	0.166	0.076	0.188	0.149	0.071	0.034
		p	0.149	0.340	0.121	0.172	0.356	0.562

	r	0.310 (-0.356 to 0.767)	0.291 (-0.374 to 0.759)	0.150 (-0.494 to 0.688)	0.522 (-0.113 to 0.854)	0.195 (-0.458 to 0.712)	0.241 (-0.503 to 0.780)
T5 *vs* T4	r^2	0.096	0.085	0.023	0.272	0.038	0.058
	p	0.353	0.385	0.659	0.099	0.565	0.531

Abbreviations: HR_{peak}: heart rate peak; HRR_{60s}: heart rate recovery after 60 sec; T1: July; T2: August; T3: September; T4: October; T5: November; TD: total distance, HSRD: high speed running; VHSRD: very high-speed running; SPD: sprint; Acc+Dec: acceleration/plus deceleration; HR_{SSGs}: average heart rate during SSGs. Bold text highlights significant correlations. * $p < 0.05$

TITLES OF FIGURES

Figure 1. Validity and reliability of the locomotor and physiological demands during small-sided games. The coefficient of variation (CV%) and the interclass correlation coefficient (ICC) for different small-sided games (SSGs) formats and matches are showed for each metric (TD: total distance; HSRD: high-speed running distance; VHSRD: very high-speed running distance; SPD: sprint distance; Acc+Dec: acceleration and deceleration; HR: heart rate). The SSGs formats are display for number of players and pitch size (length x width). Official match pitch size was 105 x 65. Data are presented as mean (SD) and mean (95% CI) for CV% (Panel **A**) and ICC (Panel **B**), respectively. The grey areas represent the reference values for *low* to *moderate* CV% or *moderate* to *very large* ICC.