Monitoring Elite Youth Football Players’ Physiological State Using a Small-Sided Game: Associations With a Submaximal Running Test

Kobe C. Houtmeyers, 1 Werner F. Helsen, 1 Arne Jaspers, 1 Sjaantje Nanne, 2,3 Shaun McLaren, 4,5 Jos Vanrenterghem, 1 and Michel S. Brink 2

1Faculty of Movement and Rehabilitation Sciences, KU Leuven, Leuven, Belgium; 2Center for Human Movement Sciences, University of Groningen, Groningen, the Netherlands; 3Football Club Groningen, Groningen, the Netherlands; 4Department of Sport and Exercise Sciences, Durham University, Durham, United Kingdom; 5Newcastle Falcons Rugby Club, Newcastle upon Tyne, United Kingdom

Purpose: To examine the utility of a standardized small-sided game (SSG) for monitoring within-player changes in mean exercise heart rate (HRex) when compared with a submaximal interval shuttle-run test (ISRT). Methods: Thirty-six elite youth football players (17 (±1) y) took part in 6 test sessions across an in-season period (every 4 wk). Sessions consisted of the ISRT (20-m shuttles, 30%:15% work:rest ratio, 70% maximal ISRT) followed by an SSG (7v7, 80 × 56 m, 6 min). HRex was collected during both protocols, with SSG external load measured as high-speed running distance (>9.19 km·h⁻¹) and acceleration distance (>2 m·s⁻²). Data were analyzed using linear mixed-effect models. Results: Controlling for SSG external load improved the model fit describing the SSG–ISRT HRex relationship (χ² = 12.6, P = .002). When SSG high-speed running distance and SSG acceleration distance were held constant, a 1% point change in SSG HRex was associated with a 0.5% point change in ISRT HRex (90% CI: 0.4 to 0.6). Inversely, when SSG HRex was held constant, the effects of a 100-m change in SSG high-speed running distance and a 21-m change in SSG acceleration distance on ISRT HRex were −1.0% (−1.5 to −0.4) and −0.6% points (−1.1 to 0.0), respectively. Conclusions: An SSG can be used to track within-player changes in HRex for monitoring physiological state. Given the uncertainty in estimates, we advise to only give meaning to changes in SSG HRex >2% points. Additionally, we highlight the importance of considering external load when monitoring SSG HRex.

Keywords: submaximal fitness tests, external load, exercise intensity, team sports
subsequent SSGs was again observed. This could be explained by a lack of standardization because games with different rules, durations, number of players, and surface areas were used, and these were performed at different time points within the training session or week. Therefore, it may be appropriate to standardize the SSG in terms of format and timing.

Accordingly, our study aims to examine the within-player associations in HRex between 2 types of SMFT: (1) a traditional intermittent-incremental protocol (ISRT) and (2) a standardized intermittent-variable protocol (SSG).

Materials and Methods

Participants

Thirty-six elite youth football players participated in this study (age: 17 [1] y, height: 177 [8] cm, body weight: 67 [9] kg). All players were part of the same club that competed in the first league in The Netherlands (Tier 3 Participant Classification Framework). Players typically engaged in 4 training sessions and 1 competitive game per week. The study was conducted according to the requirements of the Declaration of Helsinki and was part of a research project that was approved by the KU Leuven Ethics Committee (s57732).

Design

We organized 6 test sessions during the 2019–2020 season from September until January. All sessions took place during the recovery week of a 4-week periodization cycle. A session consisted of 2 SMFT. The ISRT was performed directly at the start of the session. The standardized SSG was performed after the ISRT. We collected the HRex during both protocols, and we conducted a within-player analysis to determine whether between-session changes in ISRT HRex were concurrent with changes in SSG HRex.

Methodology

Submaximal ISRT. Previous research showed that the relative reliability of this test is high (ie, intraclass correlation coefficient > .90) and showed that the outcome of the maximal version of this test is moderately associated with the maximal oxygen uptake of football players. Other studies also demonstrated good levels of absolute reliability for the HRex response to intermittent-incremental SMFT such as the Yo-Yo Intermittent Recovery Test Level 1 (typical error: 1%–2% maximal HR).

During the test, players repeatedly perform 30 seconds of 20-m running bouts (shuttles with 180° change of direction), interceded by 15-second passive rest periods. Running pace was indicated via beep signals, starting at 10 km·h⁻¹ and increasing every 90 seconds with 1 km·h⁻¹ until a running speed of 13 km·h⁻¹ was reached. Then, the pace increased every 30 seconds with 0.5 km·h⁻¹. The test ended when players reached 70% of the maximal amount of shuttles they performed during a maximal ISRT, which was organized before the start of the data collection (August 2019).

The mean HRex over the last 60 seconds of the test was used to determine the physiological state. This value was expressed as a percentage of the maximal HR reached during the maximal ISRT. Players wore the same HR unit throughout the entire study to avoid between-unit error (Polar Team System [1 Hz], Polar Electro), and all HR files were visually inspected for spikes. The session rating of perceived exertion (RPE) was collected using the Borg RPE 6 to 20 scale. The scale was visually presented to the players using an A4 paper sheet. Players responded verbally by indicating the number that matched their perceived exertion. All RPE data were collected at the end of the ISRT, including 2 minutes of passive recovery. Because only one researcher was involved in the data collection, small differences in timing may have been present. Players were assumed to be appropriately anchored to the RPE scale as they routinely provided ratings after training sessions.

Standardized SSG. We standardized the SSG in terms of format, the number of players, relative pitch area, duration, and the timing within the session and the training week. The SSG consisted of 2 teams with 6 outfield players and 1 goalkeeper. Teams were randomly selected based on player positions by the researchers on each session. All SSGs were performed on an artificial pitch of 80 × 56 m, which corresponds to the same relative pitch area as a full-sized match (ie, individual space per player on the pitch = 320 m²). Rules were similar to official games, with the exception of penalty kicks that were taken after the SSG. Three coaches acted as referees to govern the rules. To stimulate high-intensity gameplay, other coaches, including the head coach, encouraged the players verbally to keep pressure on the opponent during the whole game. Both teams were instructed to win the game with a goal difference as high as possible, and a spare ball was immediately given when the ball went out of play. Multiple games were played with each game lasting for 6 minutes. Only the first game of each player was used for analysis. Goalkeepers were not included in the analysis.

External load was measured via a local position measurement system (version 05.91 T, Inmotiotec GmbH). This system has acceptable accuracy for measuring speed and acceleration during team sport activity. Data were sampled at a variable frame rate (depending on number of transponders used during the session, always higher than 25 Hz). Data were then filtered (weighted Gaussian average filter set at 85%; speed frame interval set at 50 ms) using Inmotiot software (version 6.0.0.375, Inmotiotec GmbH). External load variables were the distance covered at >19.8 km·h⁻¹ (high-speed running distance [HSRD]) and >2 m·s⁻² (acceleration distance [ACCSD]). These variables were selected through an exploratory factor analysis on the entire data set (included as Supplementary Material S1 [available online]), as well as discussions between the authors of the paper. Both HRex (mean of the entire SSG) and session RPE were collected as internal intensity measures, in an identical manner to the ISRT. We did not collect any SSG measures of internal load (eg, HR-derived training impulse) because our aims were to compare physiological responses with the ISRT, where SMFT protocols traditionally use markers of intensity, not load, as outcome measures.

Statistical Analyses

Describing Within-Player Changes in ISRT and SSG. We used a general linear model (1-way analysis of variance) to describe the within-player variability in all outcome measures across the 6 sessions. This was expressed as the typical error of measurement in both raw and relative units (ie, coefficient of variation, %). Test-to-test changes in HRex and session RPE from the SSG and ISRT, as well as SSG external load measures, were then extracted from the model pairwise comparisons. All estimates of variability and change were presented with 90% confidence (compatibility) intervals (CIs), representing ranges of values compatible with our data, assumptions, and statistical models.

The practical relevance of test-to-test changes in each outcome measure was evaluated using minimum-effects tests (MET) and...
equivalence tests (ie, two one-sided tests [TOST]). This was given as a probability value derived from the t distribution of the change in relation to a region of practical equivalence (ROPE). A 1% point change in HRex and a 1-arbitrary unit change on Borg RPE 6 to 20 scale were considered practically meaningful and used as ROPEs. In the absence of a known real-world anchor for a meaningful change in SSG external load, we used 0.2 of the observed SD for HSRD and ACCD as ROPEs. Otherwise, the probability value from the TOST was used to describe “practical equivalence” (P_{TOST}). Decisive interpretation was then based on a conventional alpha level of .05, which corresponds to a 90% CI that falls completely outside or within the ROPE for practical significance and equivalence, respectively.

**Modeling the Within-Player Association Between ISRT HRex and SSG HSRD**

Previous investigations examining the internal response to exercise attempted to control for external load or intensity by dividing the former by the latter, creating ratio. Biological ratios have complex statistical properties that are not always obvious, as well as many assumptions that are rarely met when simply expressed as numerator over denominator. The lack of validity in ratio metrics can lead to inaccurate interpretation of data and erroneous conclusions. A solution is to use linear regression techniques, which do not violate statistical assumptions and achieve the desired outcome (ie, investigating the effects of a numerator while holding the denominator constant). The efficacy of this approach when using SSG outcomes as monitoring tools has yet to be examined, however. Accordingly, we used linear mixed-effects models to determine:

1. The extent that SSG HRex can be used to track concurrent within-player changes in ISRT HRex (model 1).
2. If controlling for SSG external load improves the ISRT–SSG HRex model (model 2).
3. If changes in SSG external loads can be used to track concurrent within-player changes in ISRT HRex when controlling for SSG HRex (model 2).

Models were fit via maximum likelihood and with the Satterthwaite denominator degrees of freedom method, using a random effect (intercept) for Player ID to appropriately account for the hierarchical data structure. In model 1, mean-centered SSG HRex was the sole predictor of ISRT HRex. In model 2, mean-centered SSG HSRD and SSG ACCD were specified as additional predictors. Model fit was expressed using the Akaike information criterion and Pseudo R² (marginal and conditional). Residual sum of squares were then compared with the chi-squared statistic (χ²) from an analysis of variance. To provide a practical interpretation of results, both models were visualized with 90% prediction intervals for the variance. To provide a practical interpretation of results, both models were visualized with 90% prediction intervals for the variance.

**Results**

Figure 1 shows the mean (SD) internal intensity and external load measures for each session. The overall mean values and estimates of the within-player variability across the 6 sessions are shown in Table 1. The typical errors of measurements were 0.70% to 0.85% for SSG HSRD and SSG ACCD, 1.9 to 3.0 arbitrary units for SSG RPE and ISRT RPE, and 11 to 50 m for SSG external load measures. Consecutive pairwise mean group changes (±90% CI) are shown in Figure 2. There was a practically significant increase in HRex at session 2 when compared with session 1 for the ISRT (P_{\text{MET}} = .02) and SSG (P_{\text{MET}} = .04). The change in ISRT HRex between sessions 3 and 4 was equivalent (P_{\text{TOST}} = .02), but this was not conclusive for SSG HRex (P_{\text{TOST}} = .11). Changes in RPE were equivalent between sessions 2 to 6 for the SSG (P_{\text{TOST}} ≤ .006) and sessions 2 to 5 for the ISRT (P_{\text{TOST}} ≤ .004; sessions 5–6 P_{\text{TOST}} ≤ .09). There were no other clear changes between any other sessions or for measures of SSG external load.

Models describing the within-player association between ISRT HRex and SSG HRex are presented in Figure 3 and Table 2. The model adjusting for SSG external load (model 2) had a lower Akaike information criterion, a higher marginal R², and a lower residual sum of squares (χ² = 12.6, P = .002) compared with the unadjusted model (model 1; Table 2).

When SSG HRex, SSG HSRD, and SSG ACCD were 87.6%, 235 m, and 36 m, respectively, the associated ISRT HRex was 92.7% (intercept, Table 2). For every 1% point increase in SSG HRex, the associated increase in ISRT HRex was, on average, 0.5% point (90% CI, 1.0% to 0.0%); days 2 to 6 for the SSG (P_{\text{TOST}} ≤ .004) and sessions 2 to 5 for the ISRT (P_{\text{TOST}} ≤ .006). The change in ISRT HRex was, on average, −0.1% (90% CI, −1.1 to 0.0), respectively (Figure 4).

Within-player associations of SSG HRex with SSG HSRD and SSG ACCD were r = .49 (.34 to .61) and .30 (.13 to .45), respectively (Supplementary Material S1 [available online]).

**Discussion**

Football players’ physiological state is often determined based on an intermittent-incremental SMFT such as the ISRT. In this study, we explored whether more integrated intermittent-variable SMFT (ie, standardized SSG) can be used for this purpose by examining the association of HRex in this test with the HRex during the ISRT.

The first primary finding was the positive within-player association between ISRT and SSG HRex (Figure 3). As anticipated, this demonstrates that a test-to-test change in SSG HRex is likely reflected by a concurrent change in ISRT HRex. This was affirmed by the slope value of −0.5% points (Table 2). We assume that the reported changes in HRex are reflective of changes in players’ physiological state, which are mainly indicative of performance capacity in our studied context (eg, aerobic fitness). This is because all tests were performed during the recovery week of a periodization cycle to limit the influence of short-term fatigue from preceding training or game sessions. However, we cannot fully exclude the influence of fatigue on the HRex response, as well as the influence of contextual variables such as players’ hydration status and the weather conditions.

We expected the strength of the association in HRex between ISRT and SSG to be influenced by SSG external load. Although we standardized the SSG in format, it is unrealistic to completely standardize its external load. Indeed, we reported high coefficients of variation in Table 1 for SSG external load measures (25%–31%). Our results showed an improved model fit when taking into account SSG external loads. This highlights the importance of considering the external load when interpreting changes in SSG HRex.
Table 1  Estimates of Within-Player Variability in Internal Intensity (HRex and RPE) During the SSG and ISRT and in External Load (HSRD and ACCD) During the SSG

<table>
<thead>
<tr>
<th>Test format</th>
<th>Variable</th>
<th>Unit</th>
<th>Mean (SD)*</th>
<th>TE</th>
<th>CV</th>
<th>Estimate</th>
<th>90% CI</th>
<th>Estimate</th>
<th>90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISRT</td>
<td>HRex</td>
<td>%HRmax</td>
<td>87.5 (2.3)</td>
<td>0.70</td>
<td>0.63 to 0.80</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RPE</td>
<td>AU</td>
<td>92.8 (1.6)</td>
<td>3.0</td>
<td>2.7 to 3.5</td>
<td>3.7</td>
<td>3.3 to 4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSG</td>
<td>HRex</td>
<td>%HRmax</td>
<td>15.0 (0.1)</td>
<td>0.85</td>
<td>0.75 to 0.97</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RPE</td>
<td>AU</td>
<td>16.5 (0.7)</td>
<td>1.9</td>
<td>1.7 to 2.1</td>
<td>2.0</td>
<td>1.8 to 2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HSRD</td>
<td>m</td>
<td>235 (29)</td>
<td>50</td>
<td>45 to 58</td>
<td>25</td>
<td>22 to 29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACCD</td>
<td>m</td>
<td>36 (2.4)</td>
<td>11</td>
<td>10 to 12</td>
<td>31</td>
<td>27 to 36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ACCD, acceleration distance; AU, arbitrary units; CV, coefficient of variation; HRex, mean exercise heart rate; HRmax, maximal heart rate; HSRD, high-speed running distance; ISRT, interval shuttle-run test; RPE, rating of perceived exertion; sRPE, session RPE; SSG, small-sided game; TE, typical error of measurement.

*Pure between-player.

Figure 1 — Observed internal intensity and external load measures across the 6 test sessions (mean [SD]). ACCD indicates acceleration distance; HRex, mean exercise heart rate; HSRD, high-speed running distance; ISRT, interval shuttle-run test; RPE, rating of perceived exertion; sRPE, session RPE; SSG, small-sided game.
As shown in Figure 4, a negative association between SSG external load and ISRT HRex was evident. This may indicate that players have the capacity to perform more SSG HSRD and SSG ACCD when they are fitter. The strength of this relationship was stronger for SSG HSRD (−1.0% points of ISRT HRex; 90% CI, −1.5 to −0.4) compared with SSG ACCD (−0.6% points; −1.1 to 0.0). This suggests that SSG HSRD might be a more pertinent variable to monitor and account for in this context; a finding further supported by the moderate to large within-player association between SSG HSRD and SSG HRex ($r = .49$; 90% CI, .34 to .61; Supplementary Material S1 [available online]). It is expected that the modalities of the SSG (eg, player numbers, relative pitch area, rules) may affect the extent to which external load variables associate to HRex, which needs further exploration.

To our knowledge, this is the first study to examine the use of training data to monitor players’ physiological state based on a standardized study design. While previous studies analyzed changes in HRex during preseason, we focused on the in-season period, which is known for its congested match schedule and, accordingly, for the limited time available for monitoring. Incorporating testing during training may therefore be relevant especially during in-season, also because it may increase coaches and players’ commitment to testing during these busy periods. Compared with preseason, smaller HRex changes can be expected during in-season (see limited changes in Figures 1 and 2), allowing to better evaluate the sensitivity of the SSG outcomes.58,39

Our study is not without limitations. We were only able to analyze one SSG of 6 minutes per session. While shorter drills are

**Figure 2** — Consecutive paired changes in internal intensity and external load measures across the 6 test sessions (mean ± 90% CI). Gray-shaded areas are the ROPEs (not shown for HSRD [±12 m] or ACCD [±2.2 m] due to differing values). ACCD indicates acceleration distance; HRex, mean exercise heart rate; HSRD, high-speed running distance; ISRT, interval shuttle-run test; ROPE, region of practical equivalence; RPE, rating of perceived exertion; SSG, small-sided game.
interesting in terms of efficiency, they may lower the standardization of external load because of rarely occurring game events, such as counterattacks. In addition, we were also not able to create player-specific profiles of the relationship between SSG external load and SSG HRex. As shown by Lacome et al, the most relevant external load variables may differ between players, which may influence the strength of the observed relationships.

Future research could replicate our study to evaluate whether our findings can be confirmed in a different setting. Because our study did not demonstrate excellent relationships, we encourage researchers to continue exploring test formats to monitor the physiological state using game-based protocols. This may include the analysis of external–internal load relationships based on passing drills that allow standardizing the external load to a greater extent. Passing drills may be an intermediate solution between standardized running drills and SSGs, but care should be taken that the drill elicits a sufficient intensity to interpret HRex changes. Compared with the SSG, passing drills may also be more practical as they are easier to plan and execute in a standardized way. To gain more insights into these practical aspects, future studies should evaluate the (perceived) cost-effectiveness of different SMFT protocols.

## Practical Applications

Our study shows how practitioners can use external and internal load data without using ratio metrics, which are common but
fraught with statistical pitfalls that lead to erroneous data.\textsuperscript{33,34} Practitioners might replicate our methods by examining the external–internal load relationships on an individual level, using linear regression techniques to estimate an “adjusted” HRex based on external load. However, this requires an adequate sample size (repeated measures) that may not always be pragmatic to draw immediate conclusions from. It is assumed that 10 samples per independent variable are sufficient, which would necessitate around 10 repeated measures for a bivariate association on a given individual (eg, SSG HSRD vs SSG HRex).

An alternative to this is using the observed regression coefficients to set simple rules of thumb. When controlling for external load, we demonstrated that for every 1% point increase in SSG HRex, the associated increase in ISRT HRex was, on average, 0.5% point (90% CI, 0.4 to 0.6). This suggests that, on average, a ∼2% point change in SSG HRex would be required to associate with a 1% point change in ISRT HRex, which has been proposed as the minimum practically meaningful change for SMFT HRex.\textsuperscript{30} We acknowledge the uncertainty around this estimate, as given by the moderate predication intervals (Figure 3). These bands are wider than CIs as they incorporate random (player) effects and give more realistic coverage of the likely range for an individual’s data to fall within. We therefore encourage practitioners to consider this fact when interpreting our findings. However, the mean estimate of a 2% point change is a useful start point for the minimum practically meaningful change in SSG HRex, after accounting for external load changes. We highlight that practitioners should also consider the typical error of measurement when interpreting the 2% point change cutoff.

Alternatively, practitioners could evaluate physiological state based on changes in external load when HRex remains the same between tests. A higher SSG external load for the same SSG HRex may reflect a positive physiological state and increased performance capacity. Our regression coefficients show that, on average, a ∼100-m change in SSG HSRD and a ∼35-m change in SSG ACCD are required to associate with a 1% point change in ISRT HRex. Because these changes correspond to 2 within-player SDs (ie, typical error of measurement, as is shown in Table 2), they represent the difference between a typically low and high SSG external load.

\textbf{Conclusions}

Our study provides evidence for the utility of a standardized SSG to monitor in-season changes in elite youth football players’ physiological state. A positive association was observed between the HRex from an ISRT and an SSG. We showed that the model fit was clearly improved when controlling for external load, emphasizing the importance of accounting for external load when interpreting changes in SSG HRex. Based on the regression coefficients of the models, we provided reference values for practitioners to interpret changes in SSG HRex and external load. However, given the uncertainty around model estimates, we advise practitioners to only give meaning to changes in SSG HRex >2% points.

\textbf{Figure 4} — The within-player associations of SSG HSRD and SSG ACCD with ISRT HRex when controlling for SSG HRex. Dots are observed data values. Solid lines are the modeled per-player associations. The thick line is the overall modeled within-player association, with a 90% prediction interval given as the gray ribbon. ACCD indicates acceleration distance; HRex, mean exercise heart rate; HRmax, maximal heart rate; HSRD, high-speed running distance; ISRT, interval shuttle-run test; SSG, small-sided game.
Acknowledgments
The authors would like to thank all players and coaches for their participation. The authors would also like to thank W. Frencken and T. Bergkamp for their support in the development of the test setup and the organization of the test sessions. This work was part of a research project supported by the Agency for Innovation by Science and Technology—IWT, Belgium, under grant IWT HBC.2017.0569 (Houtmeyers). The research grant contains a collaboration between KU Leuven and the company Topsportslab.

References