External Responsiveness of the Yo-Yo IR Test Level 1 in High-level Male Soccer Players

Authors M. Fanchini[1,2], F. Schena[1], C. Castagna[3,4], A. Petruolo[5], F. Combi[2], A. McCall[7,8], M. Impellizzeri[6]

Affiliations Affiliation addresses are listed at the end of the article

Key words
- association football
- longitudinal validity
- high intensity
- performance

Abstract

The aim of this study was to assess the external responsiveness, construct validity and internal responsiveness of the Yo-Yo Intermittent Recovery test level 1 and its sub-maximal version in semi-professional players. Tests and friendly matches were performed during the preseason and regular season. The distance covered above $15\,km \cdot h^{-1}$ was considered as an indicator of the physical match performance. Construct validity and external responsiveness were examined by correlations between test and physical match performance (preseason and regular season) and training-induced changes. Internal responsiveness was determined as Cohen's effect size, standardized response mean and signal-to-noise ratio. The physical match performance increased after training (34.8%). The Yo-Yo Intermittent Recovery test level 1 improved after training (40.2%), showed longitudinal ($r = 0.69$) and construct validity ($r = 0.73$ and 0.59, preseason and regular season) and had higher internal responsiveness compared to its sub-maximal version. The heart rate at the 6th minute in the sub-maximal version did not show longitudinal ($r = -0.38$) and construct validity ($r = 0.01$ and -0.06, preseason and regular season) and did not significantly change after training (-0.3%). The rate of perceived exertion decreased in the sub-maximal version (-29.8%). In conclusion, the Yo-Yo Intermittent Recovery test level 1 is valid and responsive, while the validity of its sub-maximal version is questionable.

Introduction

Soccer is a physically demanding sport for which players spend a great deal of time training in order to improve their ability to cope with these demands during match-play [3,44]. In order to control the training process, it is important that both the internal training load and the outcomes of this stress are measured [27]. Field tests are commonly used as a viable quantitative strategy to assess training outcomes responses on performance-relevant fitness variables. However, before implementing a test, no matter the objective, it must be appropriately validated by assessing the measurement properties. It has been suggested by Impellizzeri and Marcora [24] that, in sports science, as well as in other scientific areas, at least 5 attributes (i.e., quality criteria) should be evaluated and satisfied: conceptual model, validity, reliability, responsiveness and interpretability [24]. To the best of our knowledge, tests that are commonly used in soccer have not been sufficiently validated following the guidelines of this framework.

The Yo-Yo Intermittent Recovery test level 1 (YYIR1) is the most widespread field test used in soccer. This test displays acceptable ecological validity and it is practical, as large numbers of players can be tested simultaneously. Several studies, [11,13,31,42] have consistently shown that the YYIR1 is correlated ($r > 0.70$) to the high intensity running activity (HIR) performed during a match and thus has shown good evidence of construct validity of the test (i.e., HIR considered as indicator of the physical performance). The YYIR1 has also been shown to be able to discriminate between competitive playing levels and playing positions [4]. In addition, it is reliable and sensitive to training-induced changes (internal responsiveness) [4,9,12,36]. The only attribute not yet examined in soccer players is the longitudinal validity (also known as external responsiveness). External responsiveness refers to "the ability of a tool to detect changes over time in the construct to be measured" [38]. For the YYIR1 this would mean reflecting changes in HIR during the match; in other words: do changes in the test

accepted after revision
January 19, 2015

Bibliography
DOI http://dx.doi.org/10.1055/s-0035-1547223
Published online: 2015
Int J Sports Med

ISSN 0172-4622

Correspondence
Dr. Maurizio Fanchini, PhD
Department of Neurological and Movement Sciences
University of Verona
c/o via Usellini 12
Arona
Italy 28041
Tel.: +39/0322/44 737
Fax: +39/0322/44 737
maurizio.fanchini@gmail.com

reflect similar (same direction) changes in the ability of the players to perform HIR during the match? Ultimately, this is the main question and what fitness coaches expect from a test. To our knowledge, the external responsiveness of the YYIR1 has been provided only in soccer referees (r = 0.77) [4].

As the YYIR1 is a maximal test, which is fatiguing, a sub-maximal version of the YYIR1 (YYIR1-Sub) has been developed as a practical alternative that allows frequent assessment which could be easily incorporated into training or rehabilitation programs [4]. The outcome measure of the YYIR1-Sub is the heart rate recorded at defined recovery time points (i. e., 3–6 min) [7, 8, 31]. The YYIR1-Sub has shown a moderate correlation with the HIR covered during a match (r = −0.48) [4], as well as the ability to change throughout the season [36].

To the best of our knowledge, no studies have examined the external responsiveness of both maximal and sub-maximal versions of the YYIR1 in soccer players. Therefore, the main aim of this study was to examine the longitudinal validity of these 2 YYIR1 versions. We hypothesized that there would be a substantial correlation between the change scores in HIR during the match and performances in the 2 tests. Secondary aims were to examine the construct validity and the internal responsiveness (i. e., sensitivity to changes) of the 2 YYIR1 versions.

Materials and Methods

▼

Subjects and study design

20 outfield players (age 24 ± 6 years; height 176 ± 6 cm; body mass 74 ± 6 kg) from a semi-professional fifth division (Serie D) Italian soccer team participated in the study. Although the fifth division is considered a semi-professional league, training-match frequency (5–6 training sessions and one official match every week) and volume (average training duration 98 ± 2 min) are similar to professional leagues. The study was approved by the Ethics Committee of the University of Verona and meets the ethical standards suggested in sport and exercise science research [20].

At the start of the preseason, players were familiarized with the YYIR1. The training outcomes were assessed using the YYIR1 and the YYIR1-Sub on 2 separate but not consecutive days (i. e., day-one and day-two) during both preseason and regular season sessions (PRE and POST, respectively). On day-one players performed the YYIR1. On day-two (within 2 weeks of day-one) players completed the YYIR1-Sub and a friendly match that was organized with players of the same team. The YYIR1-Sub was performed as part of the warm-up before the friendly matches. POST sessions were organized 17 weeks after the PRE sessions with the same protocol and during a week of no official competition. The training session organized the day before each test and matches consisted of low intensity running, static and dynamic stretching for the lower limbs, core stability, technical and tactical low intensity exercises. Between PRE and POST sessions, training consisted of interval training [25], sprint training [17], small sided games [25, 41] and tactical and technical exercises.

Yo-Yo intermittent recovery test maximal and sub-maximal version

The YYIR1 and YYIR1-Sub were performed as described in a previous study [31]. To summarize, the players completed 20-m shuttle runs at increasing velocity with 10 s of active recovery between runs, until exhaustion. The YYIR1 started with 4 running bouts at 10–13 $km \cdot h^{-1}$ followed by another 7 bouts at 13.5–14 $km \cdot h^{-1}$ which continued with increments of 0.5 $km \cdot h^{-1}$ every 8 running bouts, until exhaustion. The YYIR1 was terminated when the subjects were not able to arrive within the specified time at the marked finishing line on 2 consecutive occasions. Consistent verbal encouragement was given to participants during the YYIR1 test. The YYIR1-Sub consisted of the same protocol as the maximal version with the difference being that the test was terminated after 6 min and the heart rate at the end was recorded. In addition, the heart rate at the beginning of the YYIR1-Sub was measured to exclude different starting values between PRE and POST that could have potentially biased the result of the test. The heart rate was recorded as the average of 5 s (i. e., the default of the system used) at the end of the test at the 6th minute. In line with a previous study [31], the maximal distances reached in YYIR1 and the heart rate attained at the 6th minute in YYIR1-Sub were used as outcomes. Heart rates were presented as beats $\cdot$ min^{-1} (HR6) and as percentage of the peak heart rate (HR6 %). In addition, as presented in previous studies heart rates at different time points (i. e., 2, 4, 5, average 3.40 to 4.00 and 5.30 to 6.00 min) were calculated. The peak heart rate was assessed during the YYIR1. Reproducibility (expressed as percentage coefficient of variation, CV) of the distance covered in YYIR1 and HR6 % in the YYIR1-Sub have been shown to be about 5 and 3 %, respectively [30, 31].

Physical match performance analysis

The physical activity performed during the matches was measured using an SPI-Pro X global positioning system (GPS) (unit mass 76 g, size 48 × 20 × 87 mm, GPSport, Canberra, Australia) sampling at 15 Hz with a 100 Hz accelerometer inside. The GPS system is considered a valid and reliable method to assess activities during team sports [2, 14, 16, 18, 29]. All GPS units were activated outdoors, 20 min before the start of the activity to ensure the signal from the satellites was detected, after which the devices were placed in a back pocket positioned between the shoulder blades of the player. The average number of satellites found to transmit data of positioning during all sessions was 8 ± 2 units. All data were downloaded and analyzed on a personal computer with SPI Ezy software and Team AMS software, respectively (R2 2102, GPSport, Canberra, Australia). The HIR was calculated as the total distance covered at speeds higher than 15 $km \cdot h^{-1}$ during the matches and has been considered as indicator of match physical performance [24]. To control for the inter-match variability due to the opponents activity [40], the same players and tactical module were used in the 2 matches.

Other measures

Heart rate during the matches and YYIR1-sub was measured with a telemetric system (Polar Electro Oy, Kempele, Finland) attached to the thoracic region and compatible with the GPS device. The rating of perceived exertion (RPE) was collected after the YYIR1-Sub and friendly matches with the Borg CR100® scale [6]. Instructions and familiarization with the scale were carried out during the first days in the preseason training period and the RPE used to monitor the load of the players.

Statistical analysis

Data are presented as mean ± standard deviation (mean ± SD). Effects, differences and relationship were presented with their corresponding 90 % confidence intervals (CI). External responsiveness of the tests was examined by assessing the relationship between change scores in the test (i.e., distance and HR6% for YYIR1 and YYIR1-sub, respectively) and change scores in physical performance of matches (i.e., HIR) between the PRE and POST. External responsiveness of the RPE was examined by assessing the relationship between change scores in the test (i.e., YYIR-Sub) and change scores in the match. Construct validity of the tests was assessed both in PRE and POST sessions, with the correlation between the indicator of the physical performance in the matches (HIR) and scores in YYIR1 and YYIR1-sub (distance covered and HR6%, respectively). The Pearson's product-moment of correlation has been calculated in order to assess longitudinal and construct validity. The magnitude of the correlations was determined using the modified scale by Hopkins (http://www.sportsci.org/resource/stats/2002): r<0.1, trivial; 0.1–0.3, small; 0.3–0.5, moderate; 0.5–0.7, large; 0.7–0.9, very large; >0.9, nearly perfect; and 1 perfect. Percentage training-induced changes were also calculated for match and test performances [22]. Furthermore, we calculated the probability of the substantiality of the changes between tests (i.e., larger than the smallest worthwhile change, SWC) [5,26]. Thresholds for assigning qualitative terms of improved/trivial/worsened to the changes were as follows: <1%, almost certainly not; <5%, very unlikely; <25%, unlikely or probably not; <75%, possibly may not; <95%, likely, probable; <99%, very likely, almost certain [26,33]. The SWC was calculated using a distribution-based method, that is as a proportion of the effect size, which represents the magnitude of improvement in a variable as a function of the between-subjects SD of the investigated population (i.e., 0.2 times the between-subject SD) [15]. Internal responsiveness was measured using PRE and POST test scores according to 3 methods: 1) the Cohen's effect size (ES) calculated as the mean difference between the POST and PRE test scores divided by SD of baseline; 2) the standardized response mean (SRM) calculated as the mean difference between the POST and PRE test scores divided by the SD of changes scores [23]; 3) the signal to noise ratio (ES_{TEM}) calculated as the mean difference between PRE and POST test scores divided by the typical error of measurement [1], provided by previous studies [4,31,43]. The mean percentage changes between the PRE and POST tests were considered as the signal and the absolute reliability (expressed as a percentage value) as the noise [1,39]. The modified scale proposed by Hopkins was used to interpret the degree of effect size: trivial, <0.2; small, 0.2–0.6; moderate, 0.6–1.2; or large, >1.2 (http://www.sportsci.org/resource/stats/index.html). Analyses were completed using SPSS software (version 13.0, SPSS Inc., Chicago, IL, USA) and an online spreadsheet provided by Hopkins (www.sportsci.org, http://www.sportsci.org/resource/stats/relycalc.html).

Results

Between PRE and POST tests, 5 players changed teams, 3 players were injured and one player was excluded from the analysis because of lack of commitment during the YYIR1. However, 2 of the injured players were able to participate in the POST match and were included in the analysis of YYIR1-Sub. In addition, the 5 new players engaged in the team and 2 players from the reserve team participated in the POST match.

External responsiveness

From the 27 outfield players involved in the study, the 11 players who participated in all sessions (PRE, POST tests and matches) were included in the analysis to assess the external responsiveness of YYIR1. A large correlation was found between change scores in YYIR1 and HIR (▶ **Fig. 1**). 13 players participated both in PRE and POST matches and completed the YYIR1-Sub. No significant correlation was found between change scores in YYIR1-Sub and HIR (▶ **Fig. 2**). A large correlation was found between change scores in RPE after YYIR1-sub and RPE after matches (▶ **Fig. 3**).

Construct validity

The construct validity of YYIR1 was verified in 16 and 11 players during PRE and POST sessions, respectively. Large to very large correlations were found between YYIR1 and HIR in PRE (r=0.73, 90% CI 0.44, 0.88) and POST (r=0.59, 90% CI 0.10, 0.85), respectively. The construct validity of YYIR1-Sub was examined in 13 players completing both matches and the YYIR1-Sub. No significant correlations were found between YYIR1-Sub and HIR in PRE (r=0.01, 90% CI −0.47, 0.49) and POST (r=−0.06, 90% CI −0.52, 0.43), respectively.

Training-induced changes and internal responsiveness

All of the outcomes, as well as the probability of the substantial changes between tests (percentage chances of improved/trivial/worsened) and the qualitative descriptors of changes between PRE and POST sessions, are presented in ▶ **Table 1**. The distances

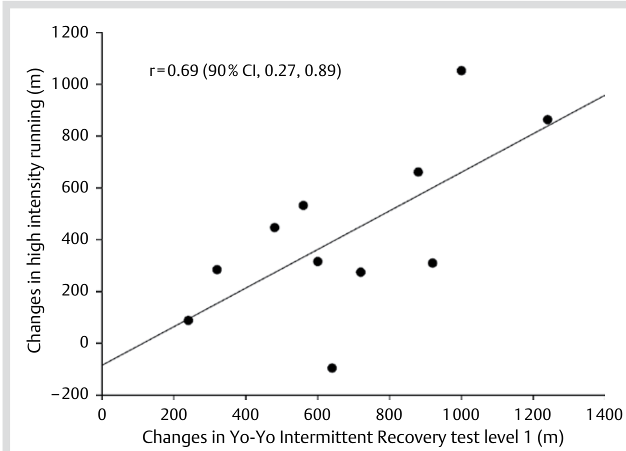

Fig. 1 External responsiveness of the YYIR1.

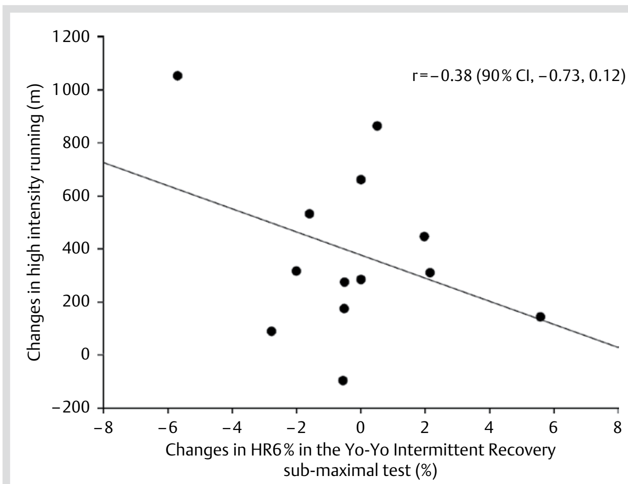

Fig. 2 External responsiveness of the YYIR1-sub.

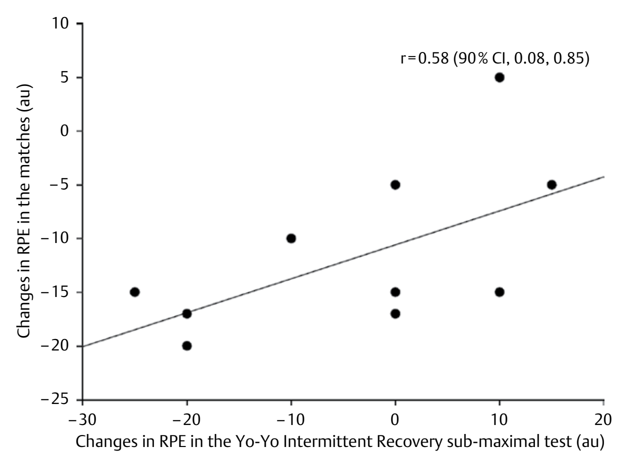

Fig. 3 External responsiveness of the RPE.

covered in YYIR1 and the HIR were improved after training (40.2 %, 90 % CI 30.7, 50.4 and 37.2 %, 90 % CI 19.8, 57.2, respectively). The HR6, HR6 % and HR at the start of the YYIR1-sub showed no significant changes after training (−0.3 % 90 % CI −1.8, 1.2 and −0.3 %, 90 % CI −1.6, 1.1 and −1.9 %, 90 % CI −6.0, 2.2, respectively). The HR at different time points showed no significant changes after training as expected by the result at the 6th minute assessment (data not shown). The peak HR in PRE and POST sessions (190 ± 8 and 189 ± 9 beats · min^{-1}, respectively) showed no significant changes after training (−1 %, 90 % CI −2.5, 0.4). The RPE in the YYIR1-Sub decreased after training (−29.8, 90 % CI −40.6, −17.1). The RPE after the matches showed no significant changes (−4.1 %, 90 % CI −12.2, 4.7). Effect sizes for internal responsiveness are presented in **Table 1**. Large ES and SRM values were found in YYIR1 and HIR between the matches. Trivial ES and SRM values were found for HR6 % and HR6 in YYIR1-sub. Large and small ES and SRM values were found for RPE in YYIR1-Sub and RPE in the matches, respectively. Large, trivial and moderate ES_{TEM} values were found for YYIR1, YYIR1-Sub and RPE in the YYIR1-Sub, respectively.

Table 1 Values, training-induced changes and internal responsiveness of the YYIR1, YYIR1-sub and matches outcomes.

	YYIR1 (n = 11)			YYIR1-Sub (n = 13)		Match HIR (n = 11)	Match RPE (n = 11)
	Distance	**Heart rate start (%)**	**HR6 %**	**HR6**	**RPE (CR100)**	**Distance**	**CR100**
PRE session	1 695 ± 243	50.1 ± 7.8	86.9 ± 4.0	166.2 ± 11.1	36.5 ± 9.9	1201 ± 389	87.7 ± 9.3
POST session	2 385 ± 412	48.1 ± 7.4	86.6 ± 4.5	165.7 ± 12.4	26.4 ± 9.0	1 632 ± 466	84.1 ± 8.9
mean difference POST-PRE (90 % CI)	691 (527, 855)	−1.9 (−6.0, 2.2)	−0.3 (−1.6, 1.1)	−0.5 (−3.0, 2.1)	−10.1 (−14, −6.2)	430 (249, 612)	−3.6 (−11, 3.7)
percentage chances of Improved/trivial/worsened	100/0/0	30/68/2	28/61/11	4/83/13	100/0/0	100/0/0	10/23/66
qualitative probabilistic terms for changes	almost certainly	possibly may not	possibly may not	likely probable	almost certainly	almost certainly	Possibly may not
smallest worthwhile change, SWC (%)	4.7	3.1	0.8	1.4	6.7	6.7	2.1
internal responsiveness (effect size)							
ES (90 % CI)	2.8 (2.1, 3.5)	−0.2 (−0.7, 0.3)	−0.1 (−0.6, 0.4)	−0.04 (−0.3, 0.2)	−1.0 (−1.4, −0.6)	1.1 (0.6, 1.6)	−0.4 (−1.2, 0.4)
SRM (90 % CI)	2.3 (1.7, 2.9)	−0.2 (−0.6, 0.2)	−0.1 (−0.6, 0.4)	−0.1 (−0.6, 0.4)	−1.3 (−1.8, −0.8)	1.3 (0.8, 1.8)	−0.3 (−0.9, 0.3)
ES_{TEM} (90 % CI)	8.2 (6.1, 10.3) *		−0.1 (−0.6, 0.4) **		−0.8 (−1.1, −0.5) #		
	4.6 (3.4, 5.8) §						

YYIR1, Yo-Yo Intermittent Recovery Test level 1; YYIR1-sub, Yo-Yo Intermittent Recovery Test level 1 sub-maximal version; HIR, high intensity running; Total distance; Sprint distance; HR6 %, percentage of peak heart rate at 6th minute of the Yo-Yo Intermittent Recovery Test level 1 sub-maximal version; HR6, heart rate at 6th minute of the Yo-Yo Intermittent Recovery Test level 1 sub-maximal version; RPE, rate of perceived exertion; PRE and POST values are presented as mean ± SD; SWC, smallest worthwhile change; ES, Cohen's effect size, SRM, standardized response mean, ES_{TEM} signal to noise ratio; * from CV 4.9 % (Krustrup, 2003), § from CV 8.7 % (Bangsbo 2008) and # from CV 38.6 % (Scott 2012), ** from CV 3 % (Krustrup 2003)

Discussion

The YYIR1 is a popular test commonly used in soccer and other team sports [4]. Although it is one of the most validated tests available for soccer players, there were a few important measurement properties of the YYIR1 that had not yet been examined. Indeed, the usefulness of a test aimed at assessing physical components relevant for performance is determined by its ability to detect changes in the construct of interest (reference construct), e.g., in the physical match performance (HIR) [24]. The main finding of this study was that changes in the YYIR1 reflected changes in the ability to perform more HIR distance during matches after a period of training, thus confirming the longitudinal validity (external responsiveness) of the test for soccer players. On the other hand, the changes in the YYIR1-Sub were not related to changes in the physical match performance.

External responsiveness

Similarly to a previous study of soccer referees [4, 30], the YYIR1 test provided evidence of longitudinal validity in soccer players (r = 0.77 and 0.69, respectively) when HIR distance covered dur-

ing the matches is used as the reference construct. Although the strength of the correlation is not high enough for predictive purposes, the YYIR1 explained almost half of the variance (48 %) in HIR thus confirming a large association between the physical performance during the test and the physical performance realized during matches. Therefore, when a player changes his performance in the YYIR1, the coaches can reasonably expect a potential variation in the same direction during the game. The results of this study are even more interesting considering that we used a real match instead of, for instance, a simulation. In a real match several confounders may influence the distance covered at various intensities thus obscuring existing relations between tests and match physical performance. However, we wanted to examine the longitudinal validity in an ecologically valid context, to ascertain whether this relation is strong enough to be disclosed in a real game (which was verified). Future studies examining this correlation using match simulation can better describe the quantitative relationship between the changes in the test performance and the ability to run at high intensity during a match.
On the other hand, the YYIR1-Sub did not show evidence of longitudinal validity, suggesting its low ability to detect changes in the HIR during matches. The trivial correlation coefficient may reflect the low data heterogeneity since changes at an individual level were not significant.

Construct validity

Similar to previous studies [11, 31, 32, 37, 42], our results confirmed the construct validity of the YYIR1. The large correlations found between YYIR1 and HIR in PRE test (r = 0.73) were similar to those shown for young (r = 0.77) [11] and adult male (r = 0.71) [31] soccer players, respectively. Recently, the construct validity of the YYIR1 has been assessed also in thermo-neutral (r = 0.76) and hot (r = 0.65) conditions in elite players [37]. The correlation value (r = 0.59) found in our study in the POST-test condition was similar to those shown for young soccer players (r = 0.56) [42]. In the present study, the lower correlation value in POST could be related to the small sample size (n = 11) compared to PRE test (n = 16).
As for longitudinal validity, the YYIR1-Sub did not show adequate construct validity in either PRE or POST tests (r = 0.01 and −0.06, respectively). However, Bangsbo, Iaia and Krustrup [4] showed a significant moderate relationship between HIR during a match and HR6 % in YYIR1-Sub (r = −0.48). Although, a sub-maximal test is certainly interesting and useful from a practical point of view, our data do not support the construct validity of the HR6 % in YYIR1-Sub when match HIR the reference construct.

Training-induced changes and internal responsiveness

Seasonal changes in YYIR1 have been investigated in several studies and reported in a review [4]. The improvements in the YYIR1 found in the current study (40 %) were consistent with the changes (31 %) found in elite male soccer referees [30] and in elite (54 %) and top-class (35 %) soccer referees [46]. However, the changes in the present investigation are greater than those found in players. Previous studies reported improvements in the YYIR1 of 25 % [31] and 13 % [4] in elite and 18 % in top-level players [35]. In sub-elite players, the YYIR1 values increase after 7 weeks of interval (13 %) and repeated sprint (28 %) training [17]. In elite young players an increased performance in YYIR1 was reported after 7 weeks of generic (22 %) and specific (17 %) training [21]. Recently, an improvement of 7–9 % in the YYIR1 was reported after one week of regular season training camp in the heat [10]. The large improvement in the YYIR1 found in our study compared with other reports may be explained by the low fitness of the players at the start of this study.
There was no significant change in YYIR-Sub following the training period in the present study (−0.3 %). This is contrary to the results of Krustrup, Mohr, Amstrup, et al. [31] who found a significant decrease of HR (−5 %) using this test following a period of training. The HR6 % and HR6 values found between PRE and POST in YYIR1-Sub in this study could be interpreted as a lack of changes in aerobic fitness, but this was not the case as there was a significant increase in both HIR during the match and YYIR1 (i. e., maximal version of the test). Nevertheless, the detected changes in HIR and YYIR1 performance may also be due to improvement in anaerobic fitness (i. e., speed endurance), thus not affecting YYIR1-sub [19, 28, 45]. While HR6 % and HR6 did not change, the RPE at the end of the YYIR1-Sub decreased (−29.8 %) after training. Therefore, it could be possible that RPE is more sensitive in detecting improvements in the ability to sustain intermittent sub-maximal running (as in the YYIR1-sub) than heart rate. Indeed, it has been suggested that the RPE may be more accurate compared to heart rate for describing the intensity during intermittent exercises such as soccer drills [34]. However, this was not the main aim of our study and future experimental designs should be used in order to investigate this topic. The RPE collected after the matches did not change. However, this can be the result of the high level of involvement of the players in the match (as demonstrated by the improvement in the physical performance).
The sensitivity of a test to detect changes over a training period (internal responsiveness) is an important characteristic [1, 24] but one that is frequently overlooked. In the present study, there was a large change in YYIR1 performance after training as shown by the large effect sizes. However, the opposite was found for the HR6 % and HR6 that showed trivial effect sizes. Therefore, our results do not support the internal responsiveness of the YYIR1-Sub. Effect size for sensitivity to changes was recently assessed in a few studies. The ESs of the YYIR1 and YYIR1-Sub were similar to previous values reported in the literature. Buchheit and Rabbani [9] and Castagna, Impellizzeri, Chaouachi, et al. [12] found an ES of 1.2 and 2.1, respectively for the YYIR1. Mohr and Krustrup [36] found an ES of 1.2 and 1.4 calculated in a different period of the season in YYIR1-Sub. To obtain a better comparison between the tests, we have reported different statistics (i. e., ES, SRM, ES_{TEM}) that have been recommended for calculating internal responsiveness [23]. The YYIR1 showed greater effect size ES and SRM values compared with its sub-maximal version. The ES_{TEM} reflect the ability of a test to change in relation to its reliability. The YYIR1 showed large ES_{TEM} indicating that the training-induced changes in the YYIR1 are higher than the reliability of the test: acceptable signal-to-noise ratio. On the other hand, the HR6 % in the YYIR1-Sub showed a trivial ES_{TEM} reflecting its low sensitivity to changes. However, the RPE showed a moderate signal-to-noise ratio thus suggesting that the RPE may be an appropriate indicator of changes in YYIR1-Sub. Although the reliability of the CR100® scale has been calculated to be 39 % (typical error is expressed as CV) [43], the acceptable reliability should be interpreted not using benchmarks but in relation to the magnitude of the changes (signal-to-noise ratio). Using this approach, the RPE is sensitive enough when compared to its reliability in detecting changes in the YYIR1-Sub. The results of our

study may be limited by the specificity of the training and internal training loads imposed on the players. However, the training approach adopted (e.g., interval training, sprint training, small-sided games, technical and tactical exercises) is typical of many soccer teams.

Conclusion

The present study provides evidence for the external responsiveness of the YYIR1 and confirms its ability to detect changes induced by training in male soccer players. As coaches are generally more interested in understanding whether their players can perform higher HIR distance in the match than improvement in the tests per se, the changes in the YYIR1 seems to reflect the changes in the ability to run more at high intensity during the match. Finally, although a sub-maximal test would be very useful, especially at the professional level, our results did not provide evidence for the construct validity and longitudinal validity of the YYIR1-Sub. Therefore, before using the sub-maximal version for routine assessment more research addressing its validity is warranted.

Acknowledgements

Authors would like to thank the Clubs involved Gozzano FC and FC Inter Milan, the Coaches of Gozzano FC Mr. Mauro Lesina and Mr. Roberto Bacchin, the Head Fitness Coach of FC Inter Milan Mr. Stefano Rapetti, the Coach of the Youth Team Sergio Zanetti and all the players involved in the study.

Affiliations
[1] Department Neurological, Neuropsychological, Morphological and Movement Sciences, Faculty of Exercise and Sport Science, University of Verona, Verona, Italy
[2] FC Internazionale, Milan, Italy
[3] Football Training and Biomechanics Lab, Italian Football Federation (FIGC), Technical Department, Coverciano (Florence), Italy
[4] School of Sport and Exercise Sciences, University of Rome Tor Vergata, Roma, Italy
[5] Gozzano FC, Gozzano, Italy
[6] Department of Research and Development, Schulthess Clinic, Zurich, Switzerland
[7] Universite de Lille2, Lille, France
[8] LOSC Lille Metropole Football Club, Research & Development Department, France

Conflict of interest: The authors have no conflict of interest to declare.

References

1 *Amann M*, *Hopkins WG*, *Marcora SM*. Similar sensitivity of time to exhaustion and time-trial time to changes in endurance. Med Sci Sports Exerc 2008; 40: 574–578
2 *Aughey RJ*. Applications of GPS technologies to field sports. Int J Sports Physiol Perform 2011; 6: 295–310
3 *Bangsbo J*, *Iaia FM*, *Krustrup P*. Metabolic response and fatigue in soccer. Int J Sports Physiol Perform 2007; 2: 111–127
4 *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: 37–51
5 *Batterham AM*, *Hopkins WG*. Making meaningful inferences about magnitudes. Int J Sports Physiol Perform 2006; 1: 50–57
6 *Borg E*, *Borg G*. A comparison of AME and CR100 for scaling perceived exertion. Acta Psychol (Amst) 2002; 109: 157–175
7 *Bradley PS*, *Bendiksen M*, *Dellal A*, *Mohr M*, *Wilkie A*, *Datson N*, *Orntoft C*, *Zebis M*, *Gomez-Diaz A*, *Bangsbo J*, *Krustrup P*. The application of the Yo-Yo intermittent endurance level 2 test to elite female soccer populations. Scand J Med Sci Sports 2014; 24: 43–54
8 *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: 1973–1975
9 *Buchheit M*, *Rabbani A*. The 30-15 Intermittent Fitness Test versus the Yo-Yo Intermittent Recovery Test Level 1: relationship and sensitivity to training. Int J Sports Physiol Perform 2014; 9: 522–524
10 *Buchheit M*, *Voss SC*, *Nybo L*, *Mohr M*, *Racinais S*. Physiological and performance adaptations to an in-season soccer camp in the heat: associations with heart rate and heart rate variability. Scand J Med Sci Sports 2011; 21: e477–e485
11 *Castagna C*, *Impellizzeri F*, *Cecchini E*, *Rampinini E*, *Alvarez JC*. Effects of intermittent-endurance fitness on match performance in young male soccer players. J Strength Cond Res 2009; 23: 1954–1959
12 *Castagna C*, *Impellizzeri FM*, *Chaouachi A*, *Manzi V*. Preseason variations in aerobic fitness and performance in elite-standard soccer players: a team study. J Strength Cond Res 2013; 27: 2959–2965
13 *Castagna C*, *Manzi V*, *Impellizzeri F*, *Weston M*, *Barbero Alvarez JC*. Relationship between endurance field tests and match performance in young soccer players. J Strength Cond Res 2010; 24: 3227–3233
14 *Coutts AJ*, *Duffield R*. Validity and reliability of GPS devices for measuring movement demands of team sports. J Sci Med Sport 2010; 13: 133–135
15 *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: 54
16 *Duffield R*, *Reid M*, *Baker J*, *Spratford W*. Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. J Sci Med Sport 2010; 13: 523–525
17 *Ferrari Bravo D*, *Impellizzeri FM*, *Rampinini E*, *Castagna C*, *Bishop D*, *Wisloff U*. Sprint vs. interval training in football. Int J Sports Med 2008; 29: 668–674
18 *Gray AJ*, *Jenkins D*, *Andrews MH*, *Taaffe DR*, *Glover ML*. Validity and reliability of GPS for measuring distance travelled in field-based team sports. J Sports Sci 2010; 28: 1319–1325
19 *Gunnarsson TP*, *Christensen PM*, *Holse K*, *Christiansen D*, *Bangsbo J*. Effect of additional speed endurance training on performance and muscle adaptations. Med Sci Sports Exerc 2012; 44: 1942–1948
20 *Harriss DJ*, *Atkinson G*. Ethical standards in sport and exercise science research: 2014 update. Int J Sports Med 2013; 34: 1025–1028
21 *Hill-Haas SV*, *Coutts AJ*, *Rowsell GJ*, *Dawson BT*. Generic versus small-sided game training in soccer. Int J Sports Med 2009; 30: 636–642
22 *Hopkins WG*, *Marshall SW*, *Batterham AM*, *Hanin J*. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc 2009; 41: 3–13
23 *Husted JA*, *Cook RJ*, *Farewell VT*, *Gladman DD*. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000; 53: 459–468
24 *Impellizzeri FM*, *Marcora SM*. Test validation in sport physiology: lessons learned from clinimetrics. Int J Sports Physiol Perform 2009; 4: 269–277
25 *Impellizzeri FM*, *Marcora SM*, *Castagna C*, *Reilly T*, *Sassi A*, *Iaia FM*, *Rampinini E*. Physiological and performance effects of generic versus specific aerobic training in soccer players. Int J Sports Med 2006; 27: 483–492
26 *Impellizzeri FM*, *Rampinini E*, *Castagna C*, *Bishop D*, *Ferrari Bravo D*, *Tibaudi A*, *Wisloff U*. Validity of a repeated-sprint test for football. Int J Sports Med 2008; 29: 899–905
27 *Impellizzeri FM*, *Rampinini E*, *Marcora SM*. Physiological assessment of aerobic training in soccer. J Sports Sci 2005; 23: 583–592
28 *Ingebrigtsen J*, *Shalfawi SA*, *Tonnessen E*, *Krustrup P*, *Holtermann A*. Performance effects of 6 weeks of aerobic production training in junior elite soccer players. J Strength Cond Res 2013; 27: 1861–1867
29 *Jennings D*, *Cormack S*, *Coutts AJ*, *Boyd LJ*, *Aughey RJ*. Variability of GPS units for measuring distance in team sport movements. Int J Sports Physiol Perform 2010; 5: 565–569
30 *Krustrup P*, *Bangsbo J*. Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. J Sports Sci 2001; 19: 881–891
31 *Krustrup P*, *Mohr M*, *Amstrup T*, *Rysgaard T*, *Johansen J*, *Steensberg A*, *Pedersen PK*, *Bangsbo J*. The yo-yo intermittent recovery test: physiological response, reliability, and validity. Med Sci Sports Exerc 2003; 35: 697–705

32 *Krustrup P*, *Mohr M*, *Ellingsgaard H*, *Bangsbo J*. Physical demands during an elite female soccer game: importance of training status. Med Sci Sports Exerc 2005; 37: 1242–1248
33 *Liow DK*, *Hopkins WG*. Velocity specificity of weight training for kayak sprint performance. Med Sci Sports Exerc 2003; 35: 1232–1237
34 *Little T*, *Williams AG*. Measures of exercise intensity during soccer training drills with professional soccer players. J Strength Cond Res 2007; 21: 367–371
35 *Manzi V*, *Bovenzi A*, *Franco Impellizzeri M*, *Carminati I*, *Castagna C*. Individual training-load and aerobic-fitness variables in premiership soccer players during the precompetitive season. J Strength Cond Res 2013; 27: 631–636
36 *Mohr M*, *Krustrup P*. Yo-Yo intermittent recovery test performances within an entire football league during a full season. J Sports Sci 2014; 32: 315–327
37 *Mohr M*, *Nybo L*, *Grantham J*, *Racinais S*. Physiological responses and physical performance during football in the heat. PLoS One 2012; 7: e39202
38 *Mokkink LB*, *Terwee CB*, *Patrick DL*, *Alonso J*, *Stratford PW*, *Knol DL*, *Bouter LM*, *de Vet HC*. 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: 737–745
39 *Norman GR*, *Wyrwich KW*, *Patrick DL*. The mathematical relationship among different forms of responsiveness coefficients. Qual Life Res 2007; 16: 815–822
40 *Rampinini E*, *Coutts AJ*, *Castagna C*, *Sassi R*, *Impellizzeri FM*. Variation in top level soccer match performance. Int J Sports Med 2007; 28: 1018–1024
41 *Rampinini E*, *Impellizzeri FM*, *Castagna C*, *Abt G*, *Chamari K*, *Sassi A*, *Marcora SM*. Factors influencing physiological responses to small-sided soccer games. J Sports Sci 2007; 25: 659–666
42 *Rebelo A*, *Brito J*, *Seabra A*, *Oliveira J*, *Krustrup P*. Physical match performance of youth football players in relation to physical capacity. Eur J Sport Sci 2014; 14 (Suppl 1): S148–S156
43 *Scott TJ*, *Black CR*, *Quinn J*, *Coutts AJ*. Validity and reliability of the session-RPE method for quantifying training in Australian football: a comparison of the CR10 and CR100 scales. J Strength Cond Res 2013; 27: 270–276
44 *Stolen T*, *Chamari K*, *Castagna C*, *Wisloff U*. Physiology of soccer: an update. Sports Med 2005; 35: 501–536
45 *Thomassen M*, *Christensen PM*, *Gunnarsson TP*, *Nybo L*, *Bangsbo J*. Effect of 2-wk intensified training and inactivity on muscle Na+-K+ pump expression, phospholemman (FXYD1) phosphorylation, and performance in soccer players. J Appl Physiol 2010; 108: 898–905
46 *Weston M*, *Helsen W*, *MacMahon C*, *Kirkendall D*. The impact of specific high-intensity training sessions on football referees' fitness levels. Am J Sports Med 2004; 32: 54S–61S