

Relationship among field test measures and physical match performance in elite-standard soccer referees

Running title: Field testing in soccer referees

Words Counts:

Abstract = 197 words

Manuscript = 3838 words

Figures: 2

Tables: 2

Key words: soccer, match-analysis, repeated-sprint ability, fitness, intermittent exercise

Abstract

The aim of this study was to assess the extent to which measures derived from the new FIFA referees fitness tests can be used to monitor a referee's match-related physical capability. Match-analysis data were collected (Prozone®, Leeds, UK) from 17 soccer referees for 5, $s=1.7$ FA Premier League matches per referee during the first four months of the 2007/08 season. Physical match performance categories were: total distance covered; high-intensity running distance (speed $>5.5 \text{ m}\cdot\text{s}^{-1}$); and sprinting distance ($>7.0 \text{ m}\cdot\text{s}^{-1}$). The tests involved a repeated 40-m sprint test followed by a 150-m interval test. Interval test heart rate load was correlated to match total distance covered ($r=-0.70$, $P=0.002$) and high-intensity running ($r=-0.57$, $P=0.018$). Fastest 40-m sprint was related to total distance covered ($r=-0.69$, $P=0.002$), high-intensity running ($r=-0.76$, $P=0.000$) and sprinting distance ($r=-0.75$, $P=0.001$) and mean time for the 40-m sprints was related to total distance covered ($r=-0.70$, $P=0.002$), high-intensity running ($r=-0.77$, $P=0.000$) and sprinting distance ($r=-0.77$, $P=0.000$). The referees with the 'best' interval test heart rate load and $40\text{m}_{\text{fastest}}$ produced the best physical match performances. However, only the 40-m test and in particular the $40\text{m}_{\text{fastest}}$ sprint time showed appropriate construct validity for the physical assessment of soccer referees.

Key words: soccer, match-analysis, repeated-sprint ability, fitness, intermittent exercise

Introduction

National and international soccer referees' associations routinely assess the fitness of elite-standard officials as high fitness is considered necessary to cope with the physical stress imposed on referees during matches (Castagna, Abt & D'Ottavio, 2002). Prior to 2005 referees had to pass a battery of field-based tests that included 2x 50-m and 2x 200-m sprints, followed by the 12-minute Cooper test (Eissmann & D'Hooghe, 1996). However, research has illustrated that these tests are poor measures of referee's match-related physical capability. Specifically, Castagna *et al.*, (2002) reported that the 12-minute run had only a low correlation with high-intensity and maximal speed running ($r=0.51$ and $r=0.32$, respectively). The 50-m and 200-m sprint times were not correlated with distances covered at high-intensity and maximal running speeds. Similar findings were reported by Mallo, Navarro, Garcia-Aranda, Gilis & Helsen (2007) who concluded that measures from the Fédération Internationale de Football Association (FIFA) fitness tests were not correlated with match activities and the evaluation of elite referees should be more specific and related to activities performed during matches.

In an attempt to overcome this problem, FIFA recently introduced two new fitness tests: a repeated 6x 40-m sprint test and a high-intensity 150-m interval test (FIFA, 2008). Minimum requirements have been set by FIFA so that should the referees achieve this standard, their fitness is deemed sufficient to cope with match demands. However, the use of these tests as a measure of match-related fitness is unclear as Mallo, Navarro, Aranda & Helsen (2009) have recently reported the 40-m mean and best sprint times to be poorly related to distance covered at high-speed running during matches ($r=-0.02$, $P=0.96$ and $r=-0.13$, $P=0.74$, respectively). Also,

mean interval test heart rate was not correlated to total distance covered ($r=-0.07$, $P=0.85$) or time spent running at high-speed during matches ($r=-0.49$, $P=0.18$).

These discrepancies could be attributable to technical issues associated with the new tests. For example, the determination of minimum standards requires empirical validation and in the current test procedures, there is no performance discrimination on the interval test and no precise resting time for the repeated-sprint test. If some of these methodological issues were addressed, i.e., the use of heart rate analysis to give a performance discriminator on the interval test and the standardisation of the repeated-sprint recovery time, then it would be possible to examine the extent to which measures derived from the tests could be used to assess referees' match-related physical capability. Also, to increase the test demands FIFA decreased the recovery time on the interval test from 40 seconds to 35 seconds (Mallo *et al.*, 2009) and the validity of this altered protocol has yet to be examined.

With these issues taken into consideration, we hypothesised that 1) physiological performance (i.e. heart rate response) on the interval test would be related to total distance covered and the distance covered at high-intensity running during matches, 2a) fastest 40-m sprint test scores would be related to the distance covered at maximal speeds during matches; and, 2b) mean 40-m sprint would be related to the distance covered at high-intensity running during matches.

Methods

Participants

Subjects were 17 English Football Association (FA) Premier League referees. The data were collected for FA Premier League matches during the first four months (August-November) of the 2007/08 English soccer season. The mean age and body mass of the referees was 40.0, $s=5.1$ years and 82.8, $s=10.0$ kg, respectively. Written informed consent was received from all referees after verbal and written explanations of the experimental design. The local Institutional Review Board approved this study design.

Experimental design

Physical Match Performances

The referees' physical match performances were examined for 85 matches (mean: 5.0, $s=1.7$ matches per referee). Each match was examined using a computerised, semi-automatic video match-analysis image recognition system (ProZone[®], Leeds, England). This method uses eight cameras positioned at roof height around football pitches to record positional data of every player and official, every 0.1-s during matches. The data captured are then systematically analysed using proprietary software to provide comprehensive data on each individual, including distances covered and intensity of movement for each player and official. Recent findings have demonstrated that the ProZone[®] match-analysis system provides valid and reliable analyses of movement patterns of footballers during matchplay (Di Salvo, Collins, McNeill & Cardinale, 2006; DiSalvo, Gregson, Atkinson, Tordoff & Drust, 2009).

The objective measures of physical match performance selected for analysis were: 1) total distance covered (m); 2) high-intensity running distance (running speed >5.5

$\text{m}\cdot\text{s}^{-1}$); 3) sprinting distance (running speed $>7.0 \text{ m}\cdot\text{s}^{-1}$); 4) peak sprinting speed, calculated as the peak speed value recorded during 0.5-s samples ($\text{m}\cdot\text{s}^{-1}$) and, 5) mean distance from the fouls over the entire match (m), which has been suggested to be of technical importance in soccer refereeing as considerable distances from fouls can lead to incorrect decisions (Krustrup & Bangsbo, 2001).

Fitness Tests

The following fitness tests are the official FIFA referee fitness tests and were performed in the week prior to the commencement of the 2007/08 English Premier League season. Each referee arrived at the testing session having rested for a minimum of 48 hours and having followed nutritional guidelines to ensure they were fully hydrated and energised. All referees had performed the tests on at least one occasion prior to the testing session and had not reported any injuries throughout the period of pre-season training that preceded the fitness tests nor throughout the period of match data collection.

6x 40-m Sprint Test

The FIFA test procedures state a maximum of 90 seconds recovery in between each of the 6x 40-m linear sprints. However, the recovery time for the sprints in the present study was standardised at exactly 90 seconds. Therefore, after a 25 minute warm up consisting of low-intensity running, stretching and striding the referees performed the 6x 40-m sprint test, with 90 seconds recovery between each sprint. The referees' starting position was 1.5-m behind the starting line, as per FIFA's procedures, and the sprint times were recorded using photoelectric beams at 0-m and 40-m (Microgate, Italy). The referees were instructed to perform each sprint maximally and each sprint had to be completed in a maximum time of 6.2 seconds.

The sprints were performed on a regular athletics track and the use of spikes was not allowed. If a referee failed one of the sprints they were allowed an extra sprint. However, if the referee failed on two occasions to run below 6.2 seconds, they failed the test (FIFA 2008).

The referees fastest 40-m sprint ($40m_{\text{fastest}}$), average of 6x 40-m sprints ($40m_{\text{mean}}$), and % decrement in performance ($40m_{\text{decrement}}$) were used for analysis. The % decrement in 40-m sprint performance was calculated as per Spencer, Fitzsimmons, Dawson, Bishop & Goodman (2006).

Following a recovery period of 8 minutes the referees performed the interval test.

Interval Test

Again using an athletics track the referees interval test alternated 150-m running, which had to be completed in a maximum of 30 seconds, with 50-m walking, which had to be completed in 35 seconds, around the 400-m athletic track (see fig. 1). During the interval runs and also the recovery periods an audio signal was played over a loud speaker system giving the referees warning when there were 10 seconds remaining (single bleep), 5 seconds (double bleep) and 0 seconds (whistle) (FIFA MP3 package).

The referees were required to complete a minimum of 10 laps to pass the test, which represented 20 high-intensity runs (4000-m in total of which 3000-m consisted of high-intensity running). If a referee failed to complete a 150-m interval in 30 seconds on one occasion, they received a verbal caution. If they then failed to complete another 150-m interval in 30 seconds, they failed the test.

Heart Rate Demand

Presently FIFA's only selection criterion on the interval test is that the referees are required to complete 10 laps in order to successfully complete the test. However, this does not permit any performance individualisation. Therefore, although it is not a requirement of FIFA, individual heart rates during the test were recorded as a measure of physiological performance. Interval test heart rates were recorded via short-range telemetry (Polar S610, Kempele, Finland), with the data being recorded every 5 seconds. Data recording commenced from the start of the first lap and finished upon completion of the 10th lap. Following the test, the data were downloaded onto a computer for analysis (Polar Precision software, version 3.0, Kempele, Finland).

Interval test heart rate load was computed by multiplying the accumulated duration in each of five different heart rate zones by a multiplier for each zone (<60% of maximal heart rate=1; 60-75%=2; 76-85%=3; 86-93%=4; >93%=5) and summing the results (Edwards, 1993). The heart rate zones in this study were based on the individual physiological response to incremental exercise as described by Bourdon (2000). The referees' individual maximal heart rate values were determined from the peak value recorded in any of the 5-second periods observed during matches, training sessions or fitness tests (Helsen & Bultynck, 2004). This methodology was employed in order to obtain a true maximal value for the referees as recent work has reported higher maximal heart rate values during training and competition when compared to laboratory fitness tests (Semin *et al.*, 2008).

Statistical analyses

Data are presented as the mean and standard deviation of the mean (s). Normality assumption was checked with the Shapiro-Wilk W test. Relationships between the referees' fitness test performances and their physical match performances were examined using a Pearson's product moment correlation, with 95% confidence intervals also presented. The following scale of magnitudes proposed by Hopkins (www.sportsci.org) was used to interpret the correlation coefficients: <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. Levene's test for Equity of Variances was computed with no differences being found. A one-way analysis of variance with repeated measures (ANOVA) was performed on the groups' 6x 40-m sprints. The median split technique was used to determine between-group differences. This procedure divides into two groups based on performance on each of the fitness tests. The referees were assigned to either the 'best' or 'worst' group if their score was above or below the median value of the group with the median value being assigned to the 'best' group. 'Best' and 'worst' groups were assigned using the median value for interval test heart rate load, $40m_{\text{fastest}}$, $40m_{\text{mean}}$ and $40m_{\text{decrement}}$. Following this procedure, the physical match performance variables (total distance, high-intensity running, sprinting distance, peak sprinting speed and average distance from fouls) of the best group were compared to those of the worst group using unpaired *t*-tests. The effect sizes (*d*) for the between-group differences were also determined, with values of 0.2, 0.5 and >0.8 representing a small, moderate and large difference, respectively (Vincent, 1995). Statistical significance was set at $P < 0.05$. All calculations were performed using the Statistica statistical analysis software package (Version 6.0, Statsoft Inc., Tulsa, OK, USA).

Results

Fitness Tests

All referees completed exactly 20 interval runs (10 laps) on the interval test (total test time 21 minutes 40 seconds). The mean heart rate load for the 10 laps was 82.3, $s=9.1$ au (range: 67-98 au). The mean 40m_{fastest} sprint was 5.59, $s=0.21$ seconds (range: 5.25-5.87 seconds) and the 40m_{mean} was 5.71, $s=0.19$ seconds (range: 5.37-5.95 seconds). In terms of the 40m_{decrement} the mean decrease was 2.0, $s=0.94\%$ (range: of 0.7-4.3%). Each referee successfully achieved the FIFA standard on all 6x 40-m sprints.

The group demonstrated no change in sprint time across the six repeated-sprints ($F_{5, 101}=0.32$, $P=0.90$). There was a nearly perfect correlation between 40m_{fastest} and 40m_{mean} ($r=0.98$, $P=0.000$) and correlations were also observed for interval test heart rate load with 40m_{fastest} ($r=0.67$, $P=0.003$), 40m_{mean} ($r=0.58$, $P=0.014$) and 40m_{decrement} ($r=-0.70$, $P=0.002$).

Physical Match Performances

The referees' mean total distance was 11478, $s=580$ m (range: 10556-12640 m). The distance covered at different running speeds was 752.8, $s=231.1$ m (range: 497.3-1176.9 m) for high-intensity running and 160.1, $s=77.8$ m (range 66.0-352.7 m) for sprinting. The mean peak sprinting speed and average distance from fouls was 8.83, $s=0.22$ m·s⁻¹ (range: 8.30-9.18 m·s⁻¹) and 14.3, $s=1.1$ m (range: 12.6-16.6 m).

Relationship between fitness tests and the referees' physical match performances

The correlation coefficients between fitness test results and physical match performance variables are presented in Fig. 2 and Table 1. Very large correlations were observed between interval test heart rate load and total match distance, $40m_{\text{fastest}}$ and match high-intensity running and sprinting distance and also $40m_{\text{mean}}$ with total distance, high-intensity running and sprinting distance. A large correlation was observed for interval test heart rate load and high-intensity running and also for $40m_{\text{fastest}}$ and total distance covered. Large correlations were also observed between $40m_{\text{fastest}}$ and $40m_{\text{mean}}$ with peak sprinting speed. The referees distance from fouls was not related to any of the fitness test performance variables.

Table 2 illustrates that the referees who had the 'best' interval test heart rate load and $40m_{\text{fastest}}$ covered more total distance, high-intensity running, and sprinting and also had a greater sprinting speed when compared to those referees with the 'worst' interval test heart rate load and $40m_{\text{fastest}}$. Effect size for these comparisons was large. However, for $40m_{\text{mean}}$ the faster referees covered more total distance but there were no between-group differences for all other variables, even though the effect sizes were large for the variables and moderate for peak sprinting speed. No between-group differences were found for $40m_{\text{decrement}}$.

Discussion

The present study reported that the new FIFA referees fitness tests did demonstrate some measure of a referee's match-related fitness. Specifically, the interval test

heart rates demonstrated a large correlation with total distance covered and match high-intensity running. These results are in contrast to Mallo *et al.* (2009), which could reflect the differences in recovery interval duration as from 2006 onwards FIFA reduced the time from 40 seconds to 35 seconds and the data reported by Mallo *et al.* were collected in 2005. Very large correlations between performance on the 40-m sprint test with match high-intensity running and sprinting were also reported despite Mallo *et al.* (2009) reporting that the best and mean 40-m times were poorly related to the distance covered at high-speed running during matches. It could be that the faster high-intensity speed threshold used in the present study enabled a more intricate assessment of the relationship between test performances with match performance. Also, Mallo *et al.* (2009) did not have a separate classification for sprinting. Furthermore, differences could also be attributed to the statistical power of the studies as Mallo *et al.* reported simultaneous recordings of match-analysis and physical capacity on only 9 matches.

The referees' interval test heart rates correlated to the amount of total distance covered by the referees ($r=-0.70$, $P=0.002$). This finding is in-line with the relationship between match distance and maximal performance on the 12-minute run (Castagna *et al.*, 2002), the Yo-Yo intermittent recovery test (Krustrup & Bangsbo, 2001) and also VO_2 max (Castagna & D'Ottavio, 2001). However, since standing, walking and jogging account for more than 75% of a referee's match activity (Krustrup & Bangsbo, 2001) it is the amount of high-intensity exercise that best indicates the demands of the game and the development of fatigue (Mohr, Krustrup & Bangsbo, 2005). In this regards a large correlation between interval test heart rate load and match high-intensity running was observed ($r=-0.57$, $P=0.018$).

Along with the large correlation between interval test heart rate load and match high-intensity running, table 2 illustrates that those referees who had the 'best' interval test heart rate load had better physical match performances when compared to the 'worst' group. However, a field test cannot be considered sport-specific until a direct association between the most relevant aspects of physical match performance and the field test investigated has been determined (Castagna, Impellizzeri, Rampinini, D'Ottavio & Manzi, 2008). This is evidenced by a common variance of only 33% between the two variables and also the fact that the referees produced wide ranging physical match performances (high-intensity range: 563.3–1529.6m) despite all achieving FIFA's minimum standard of 10 laps. Although as referees' physical match performances are partly related to the intensity of the game itself (Weston, Castagna, Impellizzeri, Rampinini & Abt, 2007) it is acknowledged that inter-game differences in match intensity could account for part of this variation. Also, the interval test could have demonstrated a stronger relationship with physical match performance had the referees run to exhaustion. However, given that the test is not incremental this would involve a high volume of running on behalf of the referees in the best physical condition. At present, the use of individual interval test heart rate load provided an individual performance discriminator and given our results it should be that the referees' heart rate response is included as part of FIFA's selection criteria provided an accurate maximal value can be obtained for all referees prior to the test. Ultimately, further validation and possible modifications on the selection criteria for the interval test are required if it is to be considered a truly valid measure of physical match performance in elite soccer referees.

The Yo-Yo intermittent recovery test (YYIR1) is considered a specific soccer refereeing test as it has been reported to possess construct validity (Castagna, Abt & D'Ottavio, 2005) and training sensitivity (Krustrup & Bangsbo, 2001; Weston, Helsen, MacMahon & Kirkendall, 2004) in elite soccer referees. In the interest of worldwide uniformity FIFA require all fitness tests to be performed on a standard athletics track with a limited amount of specialised testing equipment. However, the YYIR1 is also performed on a consistent surface and requires the same specialist equipment; a pacing signal and audio speakers. Therefore, the use of the YYIR1 should be taken into consideration, although the setting of a FIFA minimum standard would still require empirical validation.

It was hypothesised that the sprint test scores would be related to the distance covered at high-intensity running and sprinting during matches. Very large correlations were found between the referees' performance on the sprint test and the amount of distance covered at high-intensity and sprinting. Performance on the 40-m sprint test correlated with total distance covered and large correlations were also observed for interval test heart rate load with $40m_{fastest}$ and $40m_{mean}$. These findings demonstrate a clear relationship between performance on both tests as those referees whose performance was best on the interval test tended to produce the fastest 40-m sprint times; possibly reflecting a predominance of repeated-sprint training within these particular referees' training schedules as Bravo *et al.* (2008) reported that this type of training is an effective strategy for improving both soccer-specific aerobic and anaerobic fitness.

These findings are also apparent in Table 2 which illustrates that those referees with the 'best' $40m_{fastest}$ had better physical match performances when compared to

the 'worst' group. Interestingly, despite a common variance of 95% between these two measures those referees with the best $40m_{mean}$ only performed better with regards to total distance covered. There were no between-group differences for any other match variables despite large effect sizes and P values approaching conventional levels of statistical significance. This could suggest the occurrence of type II errors and a problem of statistical power.

Through the use of the median split technique Rampinini *et al.* (2007) reported no differences in repeated sprint ability (RSA_{best}) between 'best' and 'worst' groups for high-intensity running and sprinting. However, they did report between-group differences for the same match variables for RSA_{mean} . Whilst a very large correlation between $40m_{mean}$ with the same match variables was observed in the present study, no between-group differences for $40m_{mean}$ were found for high-intensity running and sprinting. Rampinini *et al.* (2007) suggested that their findings justified the use of a repeated-sprint test rather than a single sprint test to assess soccer players. This is because the physiological attributes required to perform the repeated-sprint test may be similar to those required to maintain high-speed and sprinting performance throughout a soccer match (Rampinini *et al.*, 2007). The difference between these findings with those reported in the present study could result from the different methodologies as for the Rampinini *et al.* (2007) study a preliminary single sprint was performed prior to the repeated-sprint test and this trial was used as the criterion score during the subsequent test. This procedure is not included in the FIFA referees sprint test and despite being instructed to perform each sprint maximally it is difficult to assess whether the referees did so, especially as some referees can achieve the selection criterion with sub-maximal efforts. This questions

the rationale behind the minimum standards and therefore the test protocol may not provide an accurate reflection of referees' repeated-sprint ability.

The inclusion of the 40m_{decrement} as a performance test variable is questionable as in the present study this measure showed no relationship with any of the physical match performance variables and studies have shown that the reliability of this measure is very low (Impellizzeri *et al.*, 2008; Spencer *et al.*, 2006).

The recovery time in between each sprint in the present study was 90 seconds, which is longer than that reported in previous studies that have addressed repeated-sprint ability (Rampinini *et al.*, 2007; Spencer *et al.*, 2006). Consequently, the group demonstrated no change in sprint time across the six sprints and subsequently the group 40m_{decrement} was less when compared to previous studies of repeated-sprint performance (Rampinini *et al.*, 2007; Spencer *et al.*, 2006). The absence of any performance decrement with 90 seconds recovery is consistent with Balsom, Seger, Sjodin & Ekblom (1992) who reported that 40-m sprint time did not decrease until the 12th sprint with 60 seconds recovery. Furthermore, a recovery time of 90 seconds equates to a work:rest ratio of 1:14, which is higher than the critical 1:10 previously reported for soccer referees (Abt, Castagna, Belardinelli & McCarthy, 2004). Therefore, in order to accurately measure repeated-sprint ability in soccer referees the recovery time on the FIFA sprint test should be decreased and the test protocol brought more in-line with previously validated tests (Balsom *et al.*, 1992; Rampinini *et al.*, 2007; Spencer *et al.*, 2006). The assessment of fatigue is one of the key outcomes from a multiple-sprint test (Glaister, Howatson, Pattison & McInnes, 2008) and if the current repeated-sprint test does not provide this measure then it's inclusion in the testing procedure is questionable given the

strength of the relationship between 40m_{fastest} with match high-intensity running and sprinting.

In summary, FIFA test sprint measures and the interval test heart rate load were related to physical match performance. However, given the relevance of high-intensity running for a referees' physical performance, only the sprint test and in particular the 40m_{fastest} sprint time showed appropriate construct validity for the physical assessment of soccer referees. As national and international soccer referees associations use the FIFA tests as selection criteria for competitions, modifications of the test protocols should be considered in order to increase the construct and logical validity.

Acknowledgements

We would like to acknowledge Martin Bland (Prozone[®]) for his help with the referees' match-analysis data reduction. We sincerely thank Keith Hackett (FA Premier League Referees Manager) for his help and support in the project. The cooperation of the referees was of invaluable importance.

References

Abt, G., Castagna, C., Belardinelli, R. & McCarthy, J. (2004). The effect of sprint duration on repeated sprint exercise with a constant work to rest ratio. *Journal of Sports Sciences*, 22, 521.

Balsom, P., Seger, J.Y., Sjodin, B. & Ekblom, B. (1992). Maximal-intensity intermittent exercise: effect of recovery duration. *International Journal of Sports Medicine*, 13, 528-533.

Bourdon, P. (2000). Blood lactate transition thresholds: concepts and controversies. In *Physiological tests for elite athletes* (edited by C.J. Gore), pp. 50-66. USA: Human Kinetics.

Bravo, D.F., Impellizzeri, F.M., Rampinini, E., Castagna, C., Bishop, D. & Wisløff U. (2008) Sprint vs. interval training in football. *International Journal of Sports Medicine*, 29, 668-674.

Castagna, C. & D'Ottavio, S. (2001) Effect of maximal aerobic power on match performance in elite soccer referees. *Journal of Strength and Conditioning Research*, 15, 420-425.

Castagna, C., Abt, G. & D'Ottavio, S. (2002). Relationship between fitness tests and match performance in elite soccer referees. *Journal of Strength and Conditioning Research*, 16, 231-235.

Castagna, C., Abt, G. & D'Ottavio, S. (2005). Competitive-level differences in Yo-Yo intermittent recovery and 12 min run test performance in soccer referees. *Journal of Strength and Conditioning Research*, 19, 805-809.

Castagna, C., Impellizzeri, F.M., Rampinini, E., D'Ottavio, S. & Manzi, V. (2008). The Yo-Yo intermittent recovery test in basketball players. *Journal of Science and Medicine in Sport*, 11, 202-208.

Di Salvo, V., Collins, A., McNeill, B. & Cardinale, M. (2006). Validation of Prozone: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport*, 6, 108-119.

- Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P. & Drust, B. (2009) Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine*, 30, 205-212.
- Edwards, S. (1993). High performance training and racing. In *The Heart Rate Monitor Book* (edited by S. Edwards), pp. 113-23. Sacramento, CA: Feet Fleet Press.
- Eissmann, H.J. & D'Hooghe, M.D. (1996). Sports medical examinations. In *The 23rd man: Sport medical advice for football referees*. (edited by H.J. Eissmann), pp. 7-19. Leipzig: Gersöne-Druck.
- FIFA (Fédération Internationale de Football Association). (2008). *Refereeing 2008: International lists*, pp. 246-249. Zurich: FIFA.
- Glaister, M., Howatson, G., Pattison, J.R. & McInnes, G. (2008) The reliability and validity of fatigue measures during multiple-sprint work: an issue revisited. *Journal of Strength and Conditioning Research*, 22, 1597-1601.
- Helsen, W. & Bultynck, J.B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, 22, 179–189.
- Hopkins, W. A new view of statistics. Available at:
<http://www.sportsci.org/resource/stats/index.html> Accessed 22 April 2008.
- Impellizzeri, F.M., Rampinini, E., Bishop, D., Castagna, C., Ferrari Bravo, D., Tibaudi, A. *et al.* (2008) Validity of a Repeated Sprint Test for Football. *International Journal of Sports Medicine*, 29, 899-905.
- Krustrup, P. & Bangsbo, J. (2001). Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *Journal of Sports Sciences*, 19, 881-891.

Mallo, J., Navarro, E., Garcia-Aranda, J.M. & Helsen, W. (2009). Activity profile of top-class association football referees in relation to fitness-test performance and match standard. *Journal of Sports Science*, 27, 9-17.

Mallo, J., Navarro, E., Garcia-Aranda, J.M., Gilis, B. & Helsen, W. (2007). Activity profile of top-class association football referees in relation to performance in selected physical tests. *Journal of Sports Science*, 25, 805-813.

Mohr, M., Krustup, P. & Bangsbo, J. (2005) Fatigue in soccer: A brief review. *Journal of Sports Science*, 23, 593-599.

Rampinini, E., Bishop, D., Marcora, S.M., Ferrari Bravo, D., Sassi, R. & Impellizzeri, F.M. (2007). Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *International Journal of Sports Medicine*, 28, 228-235.

Semin, K., Stahlnecker, A.C., Heelan, K., Brown, G.A., Shaw, B.S. & Shaw, I. (2008) Discrepancy between training, competition and laboratory measures of maximum heart rate in NCAA division 2 distance runners. *Journal of Sports Science and Medicine*, 7, 455-460.

Spencer, M., Fitzsimmons, M., Dawson, B., Bishop, D. & Goodman, C. (2006). Reliability of a repeated-sprint test for field hockey. *Journal of Science and Medicine in Sport*, 9, 181-184.

Vincent, W.J. (1995). *Statistics in Kinesiology*. Champaign, IL: Human Kinetics

Weston, M., Helsen, W., MacMahon, C. & Kirkendall, D. (2004). The impact of specific high-intensity training sessions on football referees' fitness levels. *American Journal of Sport Medicine*, 32, S54-S61.

Weston, M., Castagna, C., Impellizzeri, F. M., Rampinini, E. & Abt, G. (2007). Analysis of physical match performance in English Premier League soccer referees

with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport*, 10, 390-397.

Table 1 Correlation co-efficients between measures derived from the FIFA referees interval test and 6x 40-m sprint test with match physical performance variables ($n = 17$)

	Total Distance (m)	High-Intensity Running (m)	Sprinting Distance (m)	Peak Speed ($\text{m}\cdot\text{s}^{-1}$)
Interval test heart rate	-0.70**	-0.57*	-0.47	-0.27
40m _{fastest}	-0.69**	-0.76***	-0.75**	-0.52*
40m _{mean}	-0.70**	-0.77***	-0.77***	-0.50*
40m _{decrement}	-0.41	-0.41	-0.36	-0.34

*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

Table 2 Distance covered (mean \pm SD) for different match activity classifications in two groups of elite soccer referees (best [$n = 9$] and worst [$n = 8$]) according to the median value during the interval test and 6x 40-m sprint test ($40m_{\text{fastest}}$, $40m_{\text{mean}}$ and $40m_{\text{decrement}}$). Effect sizes (d) are also present.

	Total Distance (m)	High-Intensity Running (m)	Sprinting Distance (m)	Peak Speed ($m \cdot s^{-1}$)	Distance from Fouls (m)
Interval test heart rate					
Best	11832.8 \pm 545.8	882.6 \pm 209.1	205.1 \pm 78.3	8.93 \pm 0.12	14.0 \pm 0.8
Worst	11079.3 \pm 289.7	606.8 \pm 88.2	109.3 \pm 35.7	8.70 \pm 0.24	14.5 \pm 1.3
Difference	753.5	275.8	95.8	0.23	-0.44
<i>P</i> value	0.003	0.004	0.007	0.031	0.422
<i>d</i>	1.30	1.29	1.23	1.09	0.42
$40m_{\text{fastest}}$					
Best	11832.8 \pm 545.8	882.6 \pm 209.1	205.1 \pm 78.3	8.93 \pm 0.12	14.0 \pm 0.8
Worst	11079.3 \pm 289.7	606.8 \pm 88.2	109.3 \pm 35.7	8.70 \pm 0.24	14.5 \pm 1.3
Difference	735.5	275.8	95.8	0.23	-0.44
<i>P</i> value	0.003	0.004	0.007	0.031	0.422
<i>d</i>	1.30	1.29	1.23	1.09	-0.42
$40m_{\text{mean}}$					
Best	11741.3 \pm 580.4	838.2 \pm 228.7	191.3 \pm 90.6	8.87 \pm 0.21	14.3 \pm 1.1
Worst	11182.1 \pm 440.1	656.8 \pm 155.2	124.9 \pm 42.4	8.77 \pm 0.22	14.1 \pm 1.0
Difference	559.2	181.4	66.4	0.11	0.19
<i>P</i> value	0.040	0.074	0.073	0.332	0.716
<i>d</i>	0.96	0.85	0.85	0.49	0.18
$40m_{\text{decrement}}$					
Best	11308.4 \pm 568.4	684.4 \pm 202.4	144.1 \pm 86.9	8.73 \pm 0.23	14.3 \pm 1.3
Worst	11669.2 \pm 565.6	829.7 \pm 210.2	178.0 \pm 67.2	8.93 \pm 0.15	14.2 \pm 0.8
Difference	-360.7	-145.3	-33.9	-0.20	0.06
<i>P</i> value	0.210	0.169	0.380	0.055	0.913
<i>d</i>	0.62	0.68	0.44	0.91	0.05

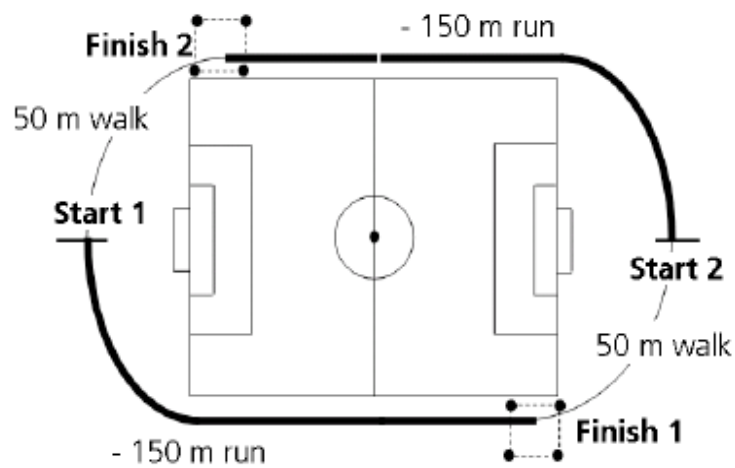


Figure 1 Schematic illustration of the FIFA referees' interval-test

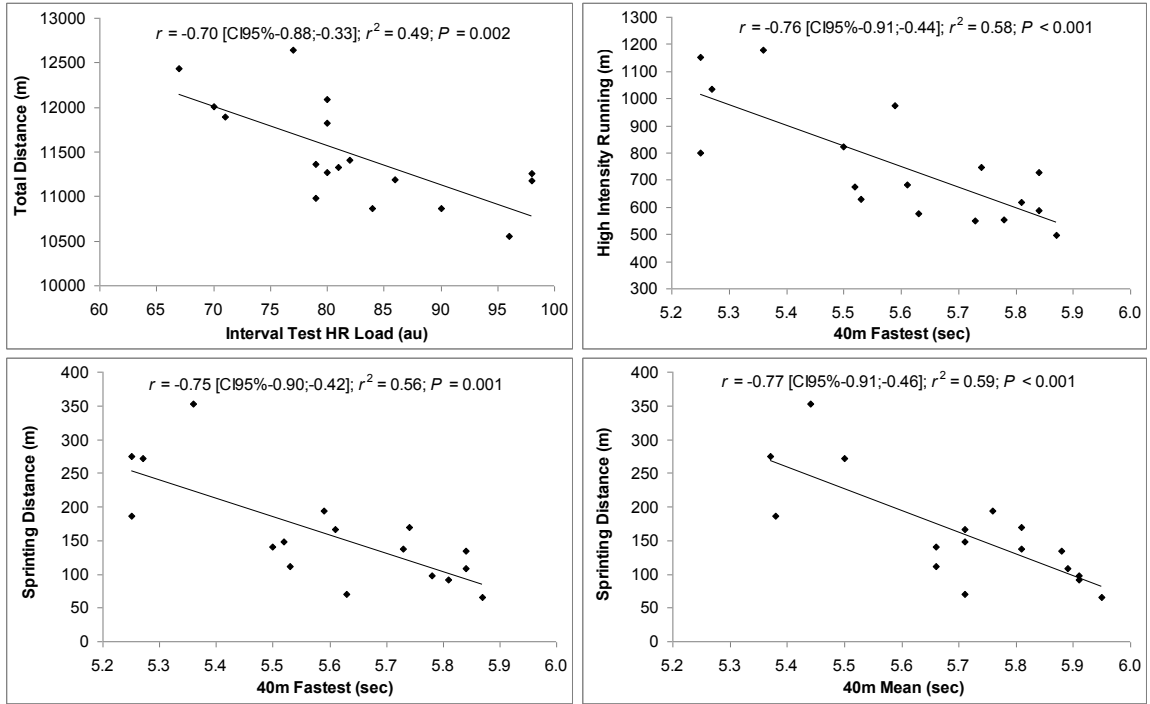


Figure 2 Scatterplots of the very large correlations between measures derived from the FIFA referees' interval-test and 6x 40-m sprint-test with match physical performance variables ($n = 17$)