Short Communication

Reliability and validity of the running anaerobic sprint test (RAST) in soccer players

Katherine Burgess, Thomas Holt, Steven Munro, Paul Swinton

Objectives: To investigate the validity and relative and absolute reliability of the Running Anaerobic Sprint Test (RAST) in amateur soccer players.

Design: Cross-sectional experimental design with an element of repeated measures.

Methods: Twenty three males completed the RAST on two occasions and a Wingate test (WAnT) as criterion measure of anaerobic power.

Results: Criterion validity for the RAST was strong for peak power ($r = 0.70$, $p < 0.001$) and average power ($r = 0.60$, $p = 0.002$); however, the RAST significantly underestimated peak power compared to WAnT. The RAST showed very good relative reliability for average power, ICC = 0.88 and good relative reliability for peak power, ICC = 0.72. Assessment of absolute reliability highlighted that although when averaged across a group, test and re-test scores will be similar, when monitoring individuals an individual's retest score may range between 0.81 and 1.2 times the original value for peak power and between 0.9 and 1.16 for average power.

Conclusion: The RAST is a practicable field test to estimate levels of average anaerobic power. However, the results show that the RAST is not sensitive enough to detect strongly individual changes below 20 % and is therefore not recommended to continually monitor an individual's anaerobic power. Also, if true measures of peak power are required the RAST test is limited.

(*Journal of Trainology* 2016;5:24-29)

Key words: field test ■ anaerobic power ■ specificity ■ football

INTRODUCTION

During most field based team sports, athletes are required to perform repeated sprint efforts. In soccer, sprinting bouts generally occur every 90 s and each of these sprints last approximately 2-4 s.[1,2] During a game of soccer 1 % to 11 % of the distance covered is done so whilst sprinting and each outfield player performs 1000 to 1400 high intensity short duration activities.[3] Although aerobic metabolism dominates the energy delivery during a soccer game, the most decisive actions are engendered by means of anaerobic metabolism [3]. In addition, it has been shown that elite professional players cover greater cumulative distances by high-intensity running and sprinting than professional players of a lower standard during the course of a game.[4] This critical reliance on a soccer player's ability to generate anaerobic power creates interest in methods that can be used to objectively assess this ability.

The Running Anaerobic Sprint Test (RAST) was developed in 1997 by Draper and Whyte to provide a means of determining anaerobic power, which was both inexpensive and simple to implement and thus accessible to coaches for players of all levels.[5] In addition, the test was founded on the basis of providing a more specific assessment of power for running based sports as it utilises flat sprinting rather than other modes of exercise such as cycling or staircase running.[6,7] The RAST comprises performance of six 35 m sprints with 10 s rest intervals. The power produced during each sprint is calculated based on the mechanical principle that power is the product of force and velocity. Average velocity for each 35 m sprint is calculated using the sprint time and known distance (velocity = displacement/time). Acceleration is then calculated through change in velocity/time, whereby average velocity is used to represent the change in the quantity due to the observation that the initial velocity is equal to zero. Consequently, force is calculated based on Newton's second law (force = mass × acceleration). As a result, power calculated from performance in the RAST is based on assumed uniform motion of the individual.

Since its conception in 1997, the RAST has become widely used by sports teams despite limited research evaluating the effectiveness of the test. Hodson and Jones were the first to report data concerning the reliability of the RAST when they included the test to investigate the effects of caffeine ingestion on repeated sprint ability.[8] The authors reported reliability coefficients ranging between $r = 0.92$ and $r = 0.97$ for all test performance measures, although they did not specify how the reliability data were determined or the specific r value for each of the measures. Subsequently, reliability of the test has been reported by others[9-11] through the use of test-retest correlations or Intraclass Correlation Coefficients (ICCs). Both statistics provide an indication of the test's relative reliability but not its absolute reliability. In order to present a more robust account of the measurement error in a performance test it is recommended that statistics measuring both relative and absolute

Received June 22, 2016; accepted August 5, 2016
From the School of Health Sciences, Robert Gordon University, Aberdeen, AB107QG, UK (K.B., T.H., S.M., P.S.)
Communicated by Takashi Abe, PhD
Correspondence to Dr. Katherine Burgess, School of Health Sciences, Robert Gordon University, Aberdeen, AB107QG UK. Email: k.burgess@rgu.ac.uk

reliability be provided[12]. In a recent study conducted by Zagatto et al.[11] relative and absolute reliability of outcome measures taken from the RAST were assessed using ICCs and Bland-Altman plots, respectively. However, the authors failed to meaningfully interpret data from the Bland-Altman plots, in particular with regards to the sensitivity of the test to detect real changes that may occur due to training. Therefore, as one of the main uses of the RAST is to monitor training progress[5], it is important that reliability statistics assist coaches to determine the likely measurement error in a test score and the magnitude of change that is required in order to be confident that different test scores reflect a change in physical status rather than simply random variability.

In addition to assessing reliability of the RAST researches have also investigated the validity of the test through comparison with a criterion measure. Although there is no universally agreed ‘gold standard’ for measuring anaerobic power, the Wingate test[6] is widely accepted as a criterion and has been used to investigate the validity of a range of different anaerobic tests[13-15]. Significant correlations between the WAnT and RAST have been reported for peak power (ranging from $r = 0.46$ to $r = 0.90$) and mean power (ranging from $r = 0.53$ to $r = 0.975$)[16,11]. Conversely, non-significant correlations between these measures have also been reported.[17] However, the study by Keir et al.[17] was conducted on small number of participants ($n = 8$) and did not adhere to standard protocols incorporating an unloaded acceleration phase immediately prior to the initiation of the 30 s Wingate test. Of the four previous studies which reported significant correlations, three of these did not provide full details of their methods[16,10,18] leading to caution over their interpretation. The fourth study[11] utilised members of the armed forces as participants. Army personnel typically undergo physical training with a focus on aerobic and muscular endurance and strength and rarely include intervals of less than 100 m in their training.[19,20] Conversely, soccer players’ training (and match play) includes the performance of repeated short sprints.[1] This element of training and testing specificity could lead to differences in the criterion validity of these tests between these differing populations. In addition, all four of the aforementioned studies either completed the RAST in a controlled indoor facility or did not state the test location. Although using controlled indoor conditions enhances the internal validity of the studies, it could potentially limit the applicability to outdoor sports who would conduct this test at their outdoor training facilities.[21,22] Therefore, the aim of this study is to investigate the criterion validity, relative reliability and absolute reliability of the RAST in soccer players when conducted in an outdoor environment.

METHODS

Participants and Study Design

Twenty three male amateur soccer players (age 24 ± 3 years, mass 75.4 ± 5.9 kg, and height 180 ± 5 cm) participated in the study. All participants played/trained at least three times a week and the group had an average of 7 ± 4 years’ experience of playing at club level. The investigation was approved by the University Institutional Ethics Committee, and all participants gave their written informed consent to participate in the study. The study is in agreement with the declaration of Helsinki of the World Medical Association. All participants performed three testing sessions, all a minimum of two days and a maximum of seven days apart. The participants were asked to refrain from partaking in strenuous exercise for a minimum of 24 hours prior to testing. There were instructed to maintain their normal diet, ensure they had eaten on the day of testing, but were asked to refrain from eating a full meal in the 2 hours prior to the testing sessions.

During two of the testing sessions participants completed the RAST and during the other participants completed the WAnT. The two RAST conditions were always carried out in consecutive sessions but the first test (WAnT or RAST) was randomised. For each participant all testing sessions were carried out at the same time of day following a standardised warm up. The warm up included five minutes of pulse raising activities (jogging, high knees, heel flicks and lunges) and two practice sprints at 75 % perceived maximal efforts (35 m running sprints prior to the RAST test and 5 s cycle sprints prior to the WAnT). At the beginning of all testing sessions clothed body mass was measured using digital scales and standing height was recorded prior to the first test only. Prior to the testing sessions participants completed a familiarisation session in which they performed both RAST and WAnT tests once.

Running Anaerobic Sprint Test

To complete the RAST participants were required to perform six maximal 35 m sprints on an AstroTurf pitch with 10 s rest periods between each sprint. Players were instructed to wear their normal training footwear (this was moulded football boots in all cases). The time for each sprint was recorded using a Brower timing gate system (Brower Timing Systems, USA) with photocells positioned 35 m apart at approximately waist height. The participant started each sprint 0.3 m behind the timing gate[23] (see Figure 1) and performed repeated sprints in alternate directions. The 10 s rest periods were timed using a stop watch and a tester gave the participant a 3 s count down prior to each sprint. Weather conditions during testing were dry and cold (3-6°C) with little wind.

The power produced during each sprint was determined by the following formula: Power = (Body Mass × Distance^2)/Time^3. Peak power was defined as the power obtained during the fastest sprint and average power (for all six sprints) was calculated by taking the mean.

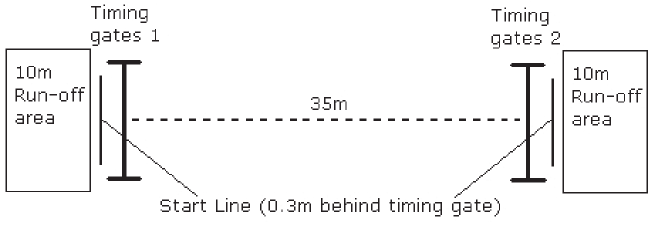

Figure 1 Diagramatical representation of RAST test set up.

The Wingate Test

The WAnT test required participants to cycle at maximum cadence on a Monark cycle ergometer (Monark 894E ergomedic peak bike) for 30 s against a resistance equivalent to 7.5 % of their body mass. Pedal revolution rate and consequently power output was measured using Monark Anaerobic test software. Participants accelerated up to maximum cadence against zero resistance before the load was applied and the test began. Power was determined over 1 s time intervals with peak power measured as the maximum value obtained and mean power calculated over the full test (30 s).

Statistical Analysis

Criterion validity was assessed with Pearson correlation coefficients to quantify the relationship between power values measured during the RAST (mean of tests 1 & 2) and WAnT, whilst paired t-tests were used to compare differences in the magnitude of power values calculated. Correlation coefficients ranging from 0.4 to 0.59 were categorised as indicating a moderate linear relationship, 0.6 to 0.79 were categorised as strong, and 0.8+ were categorised as very strong.[24] Relative reliability of test and retest scores of outcome variables measured from the RAST were assessed by intra-class correlation coefficients ($ICC_{2,1}$) using a 2-way random model with absolute agreement and 95 % CIs. ICC values were interpreted using the following guidance: 0.41 to 0.60 as moderate reliability, 0.61 to 0.80 as good reliability and 0.81 + as very good reliability.[25] Absolute reliability of the same data was quantified using the 95 % limits of agreement (LOA) method originally described by Bland and Altman[26]. Firstly, tests of systematic bias between test and re-test scores were assessed using paired t-tests. No evidence of systematic bias was found for any of the comparisons made. However, as recommended by Atkinson and Nevil[21] the 95 % LOA were still to be expressed as $\overline{X}_{diff} \pm (1.96 \times S_{diff})$ where $\overline{X}_{diff}$ is the difference between the average of the test and re-test scores and S_{diff} is the standard deviation of the difference scores. Expressed in this way, the 95 % LOA provide a measure of total error (bias ± random error) where the bias is only slight. Prior to presentation of these results, occurrence of heteroscedasticity in the data was investigated for each dependent measure by calculating the Pearson correlation coefficient between the mean of participants tests scores and the absolute value of the differences. Relatively large positive correlation coefficients were obtained for all variables ($r = 0.31$ to 0.55) indicating that the amount of random error increased as the measured values increased (i.e. data were heteroscedastic). As a result of these findings, the original test data were log-transformed using the natural logarithm and the LOA procedure was performed using the transformed data.[26] Dimensionless ratios were calculated by taking the antilog of the bias $\exp\{\overline{X}_{diff(ln)}\}$; where $\overline{X}_{diff(ln)}$ is the difference between the average of the log-transformed test and retest scores and the antilog of the random error component $\exp\{1.96 \times S_{diff(ln)}\}$; where $S_{diff(ln)}$ is the standard deviation of the difference of the log-transformed scores.[26] As a result, 95 % of the ratios of test scores (i.e. test /retest) should lie between the antilog of the bias multiplied and divided by the antilog of the random error component. Finally, the minimum difference (MD) statistic which can be used as a guide for the required change in the RAST test after a period of training to detect a 'real' change was computed. MD was calculated by multiplying the standard error of the mean (SEM) by 1.96 and $\sqrt{2}$. The SEM was calculated using the following equation: $SEM = \sqrt{1 - ICC_{2,1}}$, where *SD* is the standard deviation of all scores from the test.[27]

RESULTS

Validity

Criterion validity was strong for peak power ($r = 0.70$, $p < 0.001$) and average power ($r = 0.60$, $p = 0.002$). Paired t-tests revealed the average value for peak power was significantly greater in the WANT compared with the RAST ($t(22) = 11.570$, $p < 0.001$). Conversely, the average value obtained when measuring average power was not significantly different ($t(22) = 0.565$, $p = 0.578$) between tests (Figure 2). Bland and Altman plots illustrating the distribution of the difference scores between tests for peak and average power are illustrated in Figure 3.

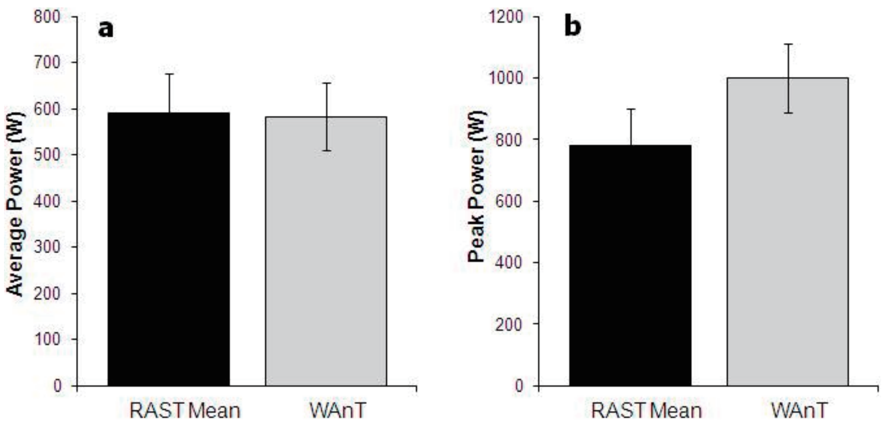

Figure 2 a) Average power and b) Peak power recorded during the RAST (mean of RAST 1&2) and WANT. Data are mean ± SD.

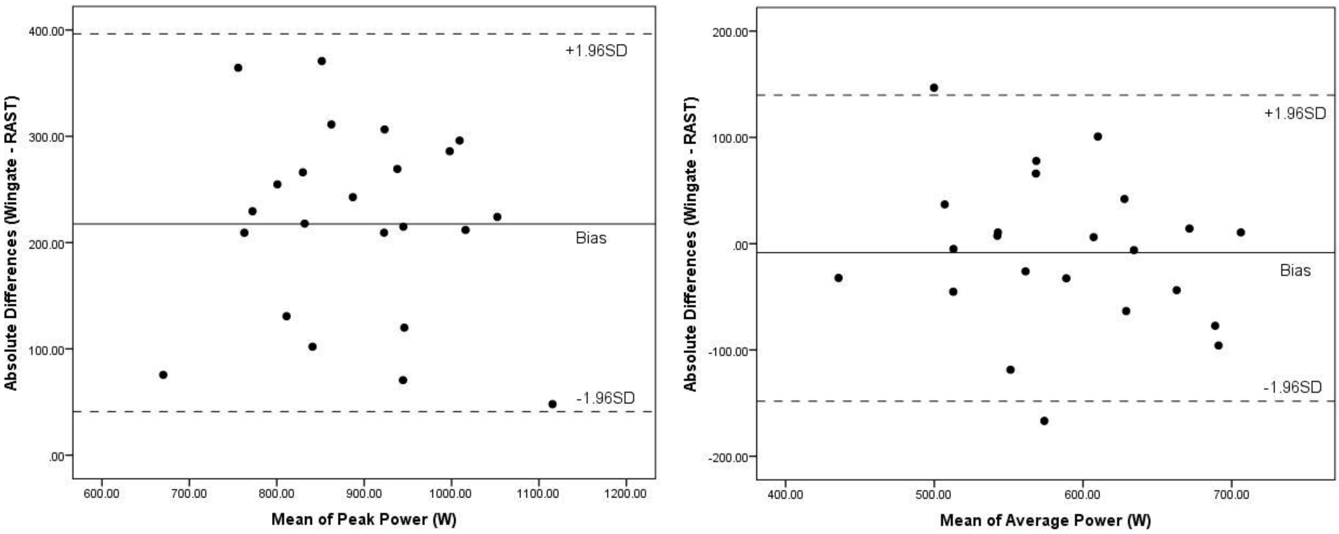

Figure 3 Bland and Altman plots illustrating the distribution of the difference scores between tests for peak and average power

Reliability

Relative reliability

The RAST showed very good relative reliability for average power, ICC = 0.88 (0.74 - 0.95: 95% CI) and good relative reliability for peak power, ICC = 0.72 (0.44 - 0.87: 95% CI) (see Table 1 for power values and Figure 4 for sprint times).

Absolute reliability

For 95% LOA analyses evidence of heteroscedasticity was obtained for peak power and average power ($r = 0.55$, $p = 0.007$; $r = 0.32$, $p = 0.137$, respectively). Log-transformation of test and retest data substantially reduced estimates of heteroscedasticity ($r = 0.12$, $p = 0.650$; $r = 0.08$, $p = 0.717$ for peak power and average power, respectively). The 95% LOA for the log-transformed data are displayed in Table 2. Minimum difference for identification of a 'real' change for average power was calculated as 83 W for average power and 160 W for peak power.

Table 1 Peak power and average power obtained during test 1 and test 2.

	RAST 1	RAST 2
Peak Power (W)	766 ± 106	776 ± 114
Average Power (W)	596 ± 86	584 ± 87

Data are mean ± SD

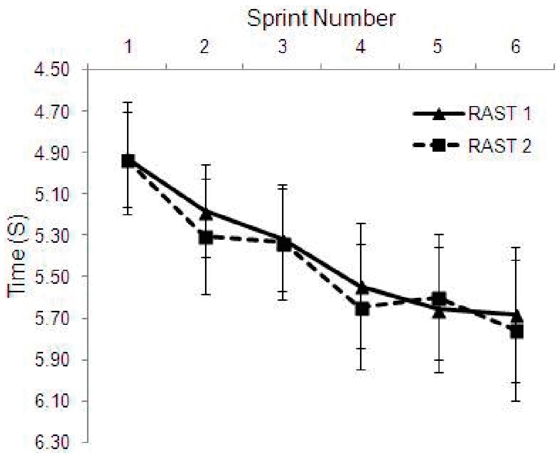

Figure 4 Time to complete each of the six RAST sprints during test 1 and test 2. Data are mean ± SD.

Table 2 Outcome of limits of agreement analyses for heteroscedastic data.

Variable	Bias	Lower, upper 95% LOA	Antilog of bias	Antilog of lower, upper 95% LOA	Interpretation of antilog values	
					Bias	Lower, upper 95% LOA
Peak Power	-0.01	-0.21, 0.18	0.99	0.81, 1.20	Averaged across a group, test and re-test scores will be very similar	For an individual a retest score may range between 0.81 and 1.2 times the original value
Average Power	0.02	-0.11, 0.15	1.02	0.90, 1.16	Averaged across a group, test and re-test scores will be very similar	For an individual a retest score may range between 0.9 and 1.16 times the original value

Bias = mean of differences in test and re-test scores. The antilog values are dimensionless ratios, where 1 represents equality.

DISCUSSION

The aim of this study was to investigate the criterion validity, relative reliability and absolute reliability of the RAST in amateur soccer players when conducted in an outdoor environment. The results have shown that criterion validity of the RAST was strong for peak power ($r = 0.70$, $p < 0.001$) and average power ($r = 0.60$, $p = 0.002$); however, the RAST significantly underestimated peak power in comparison to the WAnT. The RAST showed very good relative reliability for average power, ICC = 0.88 (0.74 - 0.95: 95% CI) and good relative reliability for peak power, ICC = 0.72 (0.44 - 0.87: 95% CI). Assessment of absolute reliability highlighted that although when averaged across a whole group test and re-test scores will be very similar (for both average and peak power), when used to monitor individuals an individual's retest score for peak power may range between 0.81 and 1.2 times the original value and an individual's retest score for average power may range between 0.9 and 1.16 times the original value.

Previously reported individual scores from the RAST have ranged from 367-1092 W for peak power and 319-927 W for average power.[9-11,17,27] In addition, previous research has reported large ranges in group peak power (599-810 W) and average power (451-665 W) values.[9-11,17,28] This current study's group means for peak and mean power lie within these ranges (781 ± 121 W and 591 ± 85 W). Similarly, the values obtained here for soccer player's peak and average power during the WAnT are comparable to values previously reported. Typical values of group means for soccer players measured during a WAnT are in the order of 740-860 W for peak power and 350-700 W for average power[10,17,29,30], with individual values for both variables varying from 400-1434 W[10] and 218-900 W[10,29]. The values obtained for the RAST test are in line with those previously reported from indoor environments. Although 'outdoor' testing is often considered to be less reliable due to potentially fluctuating conditions the conditions remained relatively stable throughout this studies testing period and hence resulted in relative reliability scores which fell within the range of those previously reported. This would support the idea that as long as the environment remains relatively stable outdoor testing can be just as reliable as indoor testing.

The significant correlations for both peak and mean power ($r = 0.70$, $p < 0.001$; $r = 0.60$, $p = 0.002$) reported in the present study between the RAST and WAnT correspond with correlation values reported elsewhere. Previous studies have reported significant r values of 0.46-0.90 for peak power and 0.53-0.98 for average. To the authors' knowledge the only previous study conducted with soccer players performing the RAST in an outdoor environment is that by Keir et al.[17] who found no significant correlations between peak and mean power in the RAST and WAnT tests which contradicts the findings reported here. However, the study by Keir et al.[17] was conducted on only eight participants and they did not include an unloaded acceleration phase immediately prior to the initiation of the 30 s Wingate tests which is contrary to the standard protocol[6]. The results presented here also show that whilst power produced during the RAST is related to power production during the WAnT, comparatively the RAST significantly underestimates peak power values. This finding is similar to that reported by Zagatto et al.[11] who also found that the RAST produced significantly lower peak power scores than the WAnT. A potential explanation of this finding is that peak power in the WAnT is a more instantaneous measure of power (1 s average) which usually occurs in the first three seconds of the test following the unloaded acceleration phase. Conversely, for the RAST peak power is determined from the fastest sprint which equates to an average power for an approximately 5-6s time period. The RAST also includes an element of acceleration with the standing start and therefore an average value will be reduced by this initial low velocity period. However, when this acceleration phase was also removed by Keir et al.[17] during the WAnT they still found the WAnT test to produce significantly higher peak power than the RAST.

The relative reliability of the RAST reported here lie within the range previously reported for both average power (ICCs between 0.72 and 0.97) and peak power (ICCs between 0.58 and 0.92).[8-11] The major novel finding of this study is the absolute reliability statistics associated with this test demonstrating the high random variability that exists and that the variation is heteroscedastic. Heteroscedasticity is a relatively common feature of test-retest scores in sport and exercise science but has implications in detecting real changes in an athlete's fitness, particularly those with already well developed attributes and high test scores. In the current study heteroscedasticity was reduced by applying a log-transformation and once the antilogs were taken, to return values to the original scale, the 95% LOA represented ratio limits of agreement. That is, instead of the potential change in a retest score being calculated by adding and subtracting a total error value, the potential change is found by multiplying by an error value. For example, If a new athlete representative of the population studied here were to perform the RAST and obtain an average power score of 440 W, we expect with approximate 95% probability that the second score will be between the range of 396-510 W (e.g. 440 × 0.9 = 396 W and 440 × 1.16 = 510 W, with error values taken from the antilog of lower, upper 95% LOA in Table 2). The range in these possible retest values may be considered by some already to be too large and therefore unacceptable. However, for an athlete with higher power production the 95% LOA for a retest value will be even wider due to the heteroscedastic nature of the data. If, for example, the second athlete obtained an average power score of 680 W, there is an approximate 95% probability that the second score will be between the range of 612-788 W. For peak power the potential range of scores that could be expected on a retest are shown here to be even greater than average power, with changes of up to approximately 20% for peak power within the 95% probability level. It is clear that with the above criteria the RAST would be of limited use in detecting changes in the majority of pre-post training designs used by coaches. One potential strategy to mitigate this problem is to incorporate narrower LOA, and therefore, the magnitude required of any change in test score to be considered indicative of altered ability will be lower. However, with this approach the probability that a different test score will incorrectly be considered a true change is increased. A less abstract and more readily understood practice is to calculate the minimum difference statistic which provides a single minimum value for all players which a test score must change by after a period of training to reflect a ‘real’ change. The results from the study suggest that individuals would have to increase their average power by at least 83 W and their peak power by at least 160 W in the RAST test to be confident that the difference reflected a training related increase in muscular power. These values equate to approximately 14% and 27% of the average values produced, further illustrating the low sensitivity of the RAST test to detect training related changes.

CONCLUSIONS

In conclusion the findings of this study have shown that the RAST is a relatively reliable, practicable field based test which can be used by coaches to estimate their soccer player’s level of average anaerobic power. However the test is not sensitive enough to detect individual changes below approximately 15 to 20% and is therefore not recommended to be used to continually monitor individual performance. The test can, however, be used to monitor the anaerobic power of a team as a whole. In addition if true peak power measures are required this test also has limitations.

ACKNOWLEDGEMENTS

To the authors knowledge there are no conflicts of interest concerning this manuscript.

REFERENCES

1. Bangsbo J, Mohr M, Krustrup P. Physical and metabolic demands of training and match-play in the elite football player. *J Sports Sci* 2006;24:665-674.
2. Reilly T, Thomas V. A motion analysis of work-rate in different positional roles in professional football match-play. *J Hum Mov Stud* 1976;2:87-97.
3. Stølen T, Chamari K, Castagnaand C et al. Physiology of Soccer An Update. *Sports Med* 2005;35:501-536.
4. Mohr M, Krustrup P, Bangsbo J. Match performance of high-standard soccer players with special reference to development of fatigue. *J Sports Sci* 2003;21:519–528.
5. Draper N, Whyte G. Anaerobic performance testing. *Peak Perform* 1997;97:3-5.
6. Bar-Or O. The Wingate anaerobic test: An update on methodology, reliability and validity. *Sports Med*, 1987;4:381-394.
7. Margaria R, Aghemo P, Rovelli E. Measurement of muscular power (anaerobic) in man. *J Appl Physiol* 1966;221:1662-1664.
8. Hodson B, Jones MA. The effect of caffeine ingestion on repeated sprint performance. *J Sports Sci* 2004;22:281.
9. Balčiūnas M, Stonkus S, Abrantes C et al. Long Term Effects of Different Training Modalities on Power, Speed, Skill and Anaerobic Capacity in Young Male Basketball Players. *J Sports Sci Med* 2006;5:163-170.
10. Reza AB, Rastegar M. Correlation between Running-based Anaerobic Sprint Test (RAST) field tests, Sargent jump and 300 yard shuttle run tests with laboratory anaerobic Wingate test in evaluation of indoor soccer player’s anaerobic readiness. *Ann Bio Res* 2012;3:377-384.
11. Zagatto AM, Beck WR, Gobatto CA. Validity of the running anaerobic sprint test for assessing anaerobic power and predicting short-distance performances. *J Strength Cond Res* 2009;23:1820–1827.
12. Atkinson G, Nevill A. Selected issues in the design and analysis of sport performance research. *J Sports Sci*, 2001;19:811-827.
13. Cooper SM, Baker JS, Eaton ZE et al. A simple multistage field test for the prediction of anaerobic capacity in female games players. *Br J Sports Med* 2004;38:784-789.
14. Haj-Sassi R, Dardouri W, Gharbi Z et al. Reliability and validity of a new repeated agility test as a measure of anaerobic and explosive power. *J Strength Cond Res* 2011;25:472-480.
15. Nummela A, Alberts M, Rijntjes RP et al. Reliability and validity of the maximal anaerobic running test. *Int J Sports Med* 1996;17:S97-S102.
16. Abbasian S, Golzar S, Onvani V et al. The predict of RAST test from WANT test in elite athletes. *Res J Recent Sci* 2012;1:72-75.
17. Keir DA, Theriault F, Serresse O. Evaluation of the Running-Based Anaerobic Sprint Test as a Measure of Repeated Sprint Ability in Collegiate-Level Soccer Players. *J Strength Cond Res* 2013;27:1671-1677.
18. Zacharogiannis E, Paradisis G, and Tziortzis S. An Evaluation of Tests of Anaerobic Power and Capacity. *Med Sci Sports Exerc* 2004;36:S116.
19. British Army (2013). Physical training and tests. Retrieved from http://www.army.mod.uk
20. US army (2013). Army Physical Readiness Training (PRT) (FM 3-22.20) Retrieved from http://www.army.mil/
21. Atkinson G, Nevill A. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;26:217–238.
22. Svensson M, Drust B. Testing soccer players *J Sports Sci* 2005;23:601-618.
23. Glaister M, Howatson G, Pattison JR et al. The reliability and validity of fatigue measures during multiple-sprint work: an issue revisited. *J Strength Cond* Res 2008;22:1597–1601.
24. Evans JD. Straightforward statistics for the behavioural sciences. California, Brooks/Cole Publishing, 1996.
25. Altman DG. Inter-rater agreement, Practical statistics for medical research, London, Chapman and Hall, 1991.
26. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* 1986;327:307–310.
27. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficinet and the SEM. *J Strength Cond Res* 2005;19:232-240.
28. Kaminagakura EI, Zagatto AM, Redkva PE et al. Can the Running-Based Anaerobic Sprint Test be used to Predict Anaerobic Capacity? *J Exerc Physiol online*, 2012;15:90-99.
29. Miller DK, Kieffer HS, Kemp HE et al. Off-season physiological profiles of elite National Collegiate Athletic Association Division III male soccer players. *J Strength Cond Res* 2011;25:1508-1513.
30. Popadic Gacesa JZ, Barak OF, Grujic NG. Maximal anaerobic power test in athletes of different sport disciplines. *J Strength Cond Res* 2009;23:751-755.