International Journal of Sports Physiology and Performance, 2009, 4, 41-53

Reproducibility of Physiological and Performance Measures from a Squash-Specific Fitness Test

Michael Wilkinson, Damon Leedale-Brown, and Edward M. Winter

Purpose: We examined the reproducibility of performance and physiological responses on a squash-specific incremental test. ***Methods:*** Eight trained squash players habituated to procedures with two prior visits performed an incremental squash test to volitional exhaustion on two occasions 7 days apart. Breath-by-breath oxygen uptake (Vo_2) and heart rate were determined continuously using a portable telemetric system. Blood lactate concentration at the end of 4-min stages was assessed to determine lactate threshold. Once threshold was determined, test speed was increased every minute until volitional exhaustion for assessment of maximal oxygen uptake (Vo_{2max}), maximum heart rate (HR_{max}), and performance time. Economy was taken as the 60-s mean of Vo_2 in the final minute of the fourth stage (below lactate threshold for all participants). Typical error of measurement (TEM) with associated 90% confidence intervals, limits of agreement, paired sample *t* tests, and least products regression were used to assess the reproducibility of scores. ***Results***: Performance time (TEM 27 s, 4%, 90% CI 19 to 49 s) Vo_{2max} (TEM 2.4 $mL \cdot kg^{-1} \cdot min^{-1}$, 4.7%, 90% CI 1.7 to 4.3 $mL \cdot kg^{-1} \cdot min^{-1}$), maximum heart rate (TEM 2 $beats \cdot min^{-1}$, 1.3%, 90% CI 2 to 4 $beats \cdot min^{-1}$), and economy (TEM 1.6 $mL \cdot kg^{-1} \cdot min^{-1}$, 4.1%, 90% CI 1.1 to 2.8 $mL \cdot kg^{-1} \cdot min^{-1}$) were reproducible. ***Conclusions:*** The results suggest that endurance performance and physiological responses to a squash-specific fitness test are reproducible.

Keywords: squash, fitness, reproducibility, endurance

Squash movements are characterized by rapid accelerations and decelerations over short distances and involve turning, lunging, and side-stepping.[1] These specific movement patterns provide a unique challenge to physiologists attempting to assess elements of fitness relevant to squash performance. The challenge is to combine laboratory control with the ecological validity of tests involving sport-specific movement patterns. Improvements in elite sport performance arise mainly

Wilkinson is with the Division of Sport Science, Sports Performance Research Group, Northumbria University, Newcastle-upon-Tyne, England; Leedale-Brown is with Reflex Squash and Fitness, Wilmington, DE; and Winter is with the Centre for Sport and Exercise Science, Sheffield Hallam University, Sheffield, England.

from increased training quality that could be assessed with sport-specific tests.[2] This necessitates the development of valid and reproducible sport-specific procedures that can assess players' strengths and weaknesses and track training adaptations that might be missed by less sensitive nonspecific procedures.

Elite squash is a high-intensity intermittent activity with mean rally durations of 16 to 21 seconds and recovery between rallies lasting 10 to 16 seconds.[3–5] Despite the short duration of these rallies,[6] heart rate reaches a steady state ranging between 80% and 95% of age-predicted maximum.[5,7] Oxygen uptake reaches mean values of approximately 54 $mL \cdot kg^{-1} \cdot min^{-1}$ (≈86% Vo_{2max}) and mean lactate concentrations of 8 $mmol \cdot L^{-1}$ have been reported recently.[5] These responses suggest that energy is provided largely via intramuscular phosphates, glycolysis, and myoglobin O_2 stores that are replenished by oxidative metabolism during the short recovery periods.[8] The Vo_{2max} values of 62 to 66 $mL \cdot kg^{-1} \cdot min^{-1}$ in elite male players confirm the importance of high aerobic power at the highest standards of play.[9,10]

Recent attempts to produce controlled tests to replicate squash-specific physiological demands have focused on simulation of match play rather than assessments of squash-specific fitness.[11,12] Only two previous papers describing on-court squash protocols developed for assessment purposes have been published.[13,14] A third study by Kingsley et al[15] described a squash-specific incremental test procedure that *could* be used to assess squash-specific fitness. However, the study focused on the use of the protocol in the development of a squash simulation protocol and did not address the validity of the incremental test for assessment of squash-specific fitness.

The procedure of Steininger and Wodick[13] was devised to mimic physiological demands and techniques of squash movement but in defined increments to allow the assessment of squash-specific endurance fitness. Ranked performance data from the test correlated with ranked playing fitness measures estimated from competitive results and coaches' subjective estimates of match fitness ($r = .9$, $P < .05$) highlighting the ability of the test to assess physiological capacities in squash-specific movements. However, the test is challenged by the need to replicate the stochastic nature of squash movement. The neuromuscular ability to accommodate rapid accelerations and decelerations is a crucial performance characteristic in squash but is likely to go undetected by a test that uses predictable movement sequences.[16]

The squash-specific test described by Girard et al[14] overcomes some limitations of Steininger and Wodick's[13] test by including uncertainty of movement direction. A strong correlation ($r = -0.96$, $P < .001$) between time to exhaustion and player world ranking in the squash test, and higher Vo_{2max} scores on the squash test than on a treadmill test suggest that the Girard et al[14] protocol is a valid and specific test for squash players. However, the evaluation of reproducibility for squash-specific protocols is lacking and requires further investigation.

Previous attempts to develop valid and controlled tests of squash-specific fitness are challenged by the stochastic nature of match play and the need for reproducibility of scores. Any valid sport-specific test devised for assessment purposes must also demonstrate good reproducibility if it is to be of value in tracking improvements in fitness and performance with training.[17–19] Accordingly, the purpose of this study was to assess reproducibility of measures from a squash-specific incremental test that comprised randomized movements.

Methods

Participants

With institutional ethics approval, 8 trained squash players (age mean ± SD), 29.6 ± 9.4 years, stature 1.77 ± 0.05 m, and body mass 69.4 ± 6.7 kg, who were fully habituated to the procedures, participated. Habituation involved two visits to the laboratory on separate days where participants performed five submaximal stages of the incremental squash-specific test wearing test equipment, but with no data being collected. The players were regular and current competitors in the premier or first division of their regional leagues, with at least 5 years of playing experience at this standard. All participants were instructed to report for testing well rested, well hydrated, and well nourished and to have refrained from eating at least 2 hours before testing. Participants were also instructed to abstain from drinking alcohol and avoid stimulants such as caffeine for at least 8 hours before testing.

Experimental Design

Participants performed an incremental squash test (ST) to volitional exhaustion on two occasions, 7 days apart. Tests were conducted under similar environmental conditions (temperature 18.9 ± 3.4°C, relative humidity 49 ± 8%, barometric pressure 1016 ± 11 mb) at the same time of day and in the same footwear and clothing.

Overview of the Incremental Squash-Specific Test

The ST involved squash-specific movement patterns to and from four marked positions (two front corners and two back corners) on a squash court floor beginning from a central T position (Figure 1). Movements were performed randomly with the order and frequency controlled by an audio signal of a number corresponding to one of the four marked and numbered targets. Whereas individual movements were administered randomly, the proportions minute-by-minute reflected those seen in match play as identified from match analysis (74% back corner movements, 26% front corner movements).[3] The movement distances and mean movement speeds involved were encompassed in the ranges reported in previously published match analysis studies.[1,3] Validity of the incremental squash-specific test was assessed in a separate study.[20] Participants were required to move to specified court positions, place one foot on the marked target, mimic a forceful shot down the nearest side wall of the court, and return to the T position in time for the next audio signal.

Assessment Protocol

Phase 1: Determination of Lactate Threshold and Movement Economy. Participants completed between six and ten 4-minute stages with 1-minute rest intervals between stages for collection of capillary blood from a finger tip. Breath-by-breath oxygen uptake (V_{O_2}) and heart rate (HR) were continuously determined and recorded using a portable telemetric system (Metamax 3B, Cortex Biophysik,

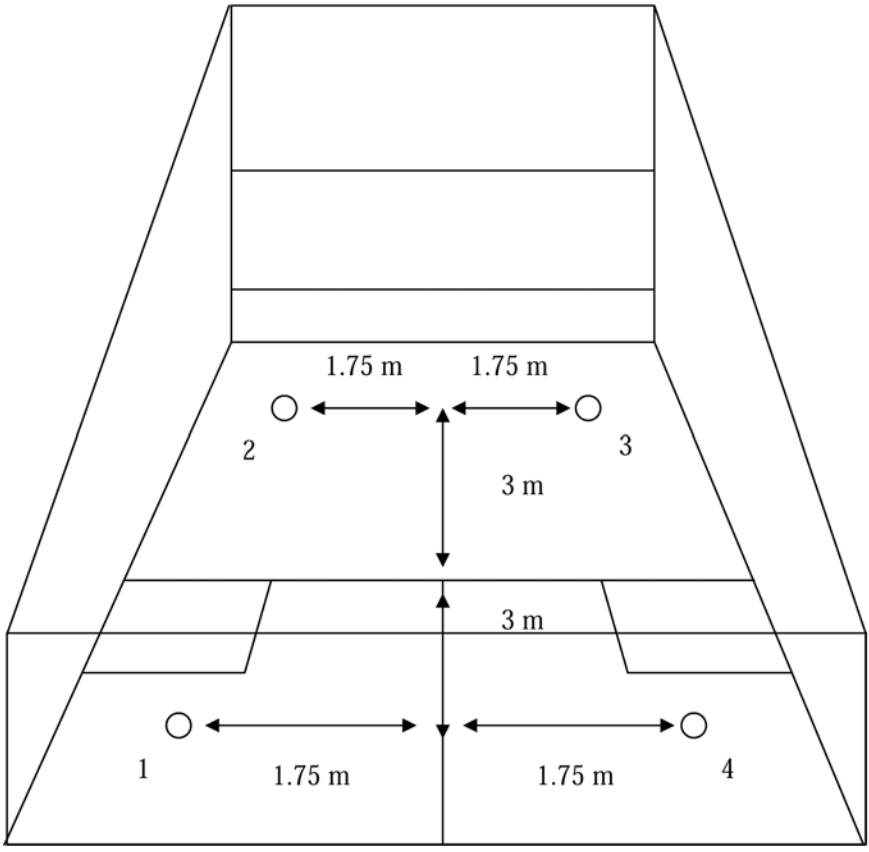

Figure 1 — Set-up and dimensions of the squash-specific incremental test. Numbered court positions are indicated (1–4). Arrows indicate distances for the location of the numbered court positions relative to court markings not the route to the court positions.

Leipzig, Germany) that was calibrated according to manufacturer guidelines before each test.

Lactate threshold was identified from visual inspection of lactate values plotted against test stage and was taken as the test stage before the first sudden rise in blood lactate concentration. Blood lactate concentration was assessed by an electrochemical method in triplicate using 25-μL samples (YSI 1500, Yellow Springs Instruments, Yellow Springs, Ohio, USA). Before testing, the analyzer was calibrated with a lactate standard of known concentration (5 mmol·L^{-1}) and linearity was checked with standards of 15 and 30 mmol·L^{-1}. Once participants reached a blood lactate concentration of ≥4 mmol·L^{-1}, phase 1 ceased and participants were allowed a 10- to 15-minute rest period before beginning phase 2 of testing.

Following the determination of lactate threshold, phase-1 $\dot{V}O_2$ data were used to determine movement economy, which was taken as the 60-s mean of $\dot{V}O_2$ in the final minute of the fourth stage (below lactate threshold for all participants).

Phase 2: Determination of Maximal Oxygen Uptake. Following recovery from phase 1, participants completed incremental 1-minute stages in a continuous manner commencing at stage 1 (10 moves per minute), with speed increased by one movement per minute every minute until volitional exhaustion. Breath-by-breath oxygen uptake ($\dot{V}O_2$) and heart rate (HR) were continuously determined as previously described. This phase of testing ended when the participant voluntarily stopped exercising or was stopped by the experimenter if after two warnings they were unable to place a foot on the correct court mark in time with the audio signals.

The $\dot{V}O_{2max}$ was calculated using 30-s retrograde, stationary time mean with $\dot{V}O_{2max}$ taken as the highest 30-s mean during the final stages of each test. The HR_{max} was taken as the highest 30-s mean during the final stages of each test. Attainment of a plateau in $\dot{V}O_2$ ($\leq$2.1 mL·kg^{-1}·min^{-1} rise with an increase in exercise intensity), respiratory exchange ratio (RER) >1.1, posttest blood lactate concentration >8 mmol·L^{-1}, HR within 10 beats·min^{-1} of age-predicted maximum, and participant subjective reporting of maximal effort were used as criteria to judge whether test performances were truly maximal.[21] If a participant failed to satisfy three or more of these criteria, the test result was deemed to be a peak rather that a maximum value. Performance time to exhaustion was recorded to the nearest second using an electronic stop watch (FastTime 1, Click Sports, Cambridge, UK).

Statistical Analysis

Precisely which metric of reproducibility to use is the subject of enthusiastic debate because each has its detractors and supporters.[18,19] We used the following methods: typical error of measurement (TEM),[19] limits of agreement (LOA),[22] least products regression (LPR),[23] and paired sample *t* tests. Version 12 of SPSS (SPSS Inc., Chicago, IL) software was used to generate descriptive statistics and undertake the analysis for LPR. The TEM (and the 90% confidence intervals thereof) and LOA were calculated using the Microsoft Excel spreadsheet of Hopkins[24] and, together with LPR, were used to assess the reproducibility of scores. Paired sample *t* tests were used to assess systematic bias between test and retest scores. Before LOA analysis, the assumption of homoscedasticity was confirmed using Pearson's correlation coefficient to examine relationships between the individual mean of scores on each trial and the absolute individual difference between scores on consecutive trials for each variable.

Results

Movement speed at lactate threshold varied between players in a range from stage 4 (13 moves per minute) to stage 8 (17 moves per minute). However, each player achieved identical movement speeds at lactate threshold across both test sessions, so movement speed at the lactate threshold was not subjected to reproducibility analysis.

The descriptive statistics and reproducibility of other physiological and performance measures are shown below in Tables 1 and 2.

The data show a small increase in mean Vo_{2max} (1.2 mL·kg^{-1}·min^{-1}) and mean performance time (23 seconds) between the first and second test in conjunction with slight reduction in mean HR_{max} (2 beats·min^{-1}). There was also an improvement in economy with a reduced mean oxygen cost of movement at stage four of the ST (3.9 mL·kg^{-1}·min^{-1}).

The LOA values for performance time and Vo_{2max} support the trends in the descriptive data, showing a small positive systematic bias between test 1 and test 2, although paired *t* tests showed that these were not significant ($t_7 = -1.69$, $P = .1$; $t_7 = -0.27$, $P = .8$ for performance time and Vo_{2max} respectively). Similarly, the trend was supported for HR_{max} and economy with small negative systematic biases between test 1 and test 2. Paired *t* tests showed that the systematic bias was not significant for HR_{max} ($t_7 = 2$, $P = .1$), but was significant for economy ($t_7 = 4.5$, $P = .003$). The random error component on all variables was low. Relative TEM (%) showed similar test–retest variation for performance time (4%), Vo_{2max} (4.7%), HR_{max} (1.3%), and economy (4.1%). The 90% confidence intervals of the TEM scores were narrow for all variables (performance time 19 to 49 s; Vo_{2max} 1.7 to 4.3 mL·kg^{-1}·min^{-1}; HR_{max} 2 to 4 beats·min^{-1}; economy 1.1 to 2.8 mL·kg^{-1}·min^{-1}). The use of LPR showed some variation in the quantification of reproducibility in comparison with other measures. For example using TEM, performance time had low test–retest variation (4%). However, the LPR values for slope (1.14) and intercept (−71) were some way from the values of 1 and 0 that reflect perfect reproducibility. Bland–Altman and LPR plots for all measures are shown in Figures 2 and 3 respectively.

Discussion

This study examined the reproducibility of physiological and performance measures in an incremental squash test devised to mimic squash movement patterns while replicating the stochastic nature of movement in match play.

The results show good although varying degrees of reproducibility in performance time, HR_{max}, Vo_{2max}, and economy depending on which metric of reproducibility is favored. The relative TEM (%) for performance time (4%) was higher than that reported for the Girard et al[14] protocol (0.9%) and HR_{max} TEM (1.3%) was lower in comparison (1.8%). The TEM of Vo_{2max} in this study (4.7%) was within the range reported in other studies using treadmill running as the exercise

Table 1 Performance time, Vo_{2max}, HR_{max}, and Economy from Two Trials of the ST Performed 7 Days Apart (Values Are Mean ± SD)

	Time (s)	Vo_{2max} (mL·kg^{-1}·min^{-1})	HR_{max} (beats·min^{-1})	Economy at test stage 4 (mL·kg^{-1}·min^{-1})
Test 1 (n = 8)	692 ±148	50.8 ±6.5	189 ±10	32.6 ±5.2
Test 2 (n = 8)	715 ±168	51.2 ±6.9	187 ±10	28.7 ±3.5

Table 2 Reproducibility of Performance Time, Vo_{2max}, HR_{max}, and Economy on the ST (n = 8)

	Time (s)	Vo_{2max} ($mL \cdot kg^{-1} \cdot min^{-1}$)	HR_{max} ($b \cdot min^{-1}$)	Economy ($mL \cdot kg^{-1} \cdot min^{-1}$)
Limits of Agreement	14 ± 62	0.2 ± 5.1	−2 ± 6	−3.9 ± 4.3
Typical Error	27 (4%)	2.4 (4.7%)	2 (1.3%)	1.55 (4.1%)
Least Products Regression:				
Slope	1.1	1.1	0.99	0.7
Intercept	−71	−2.9	−0.5	6.8

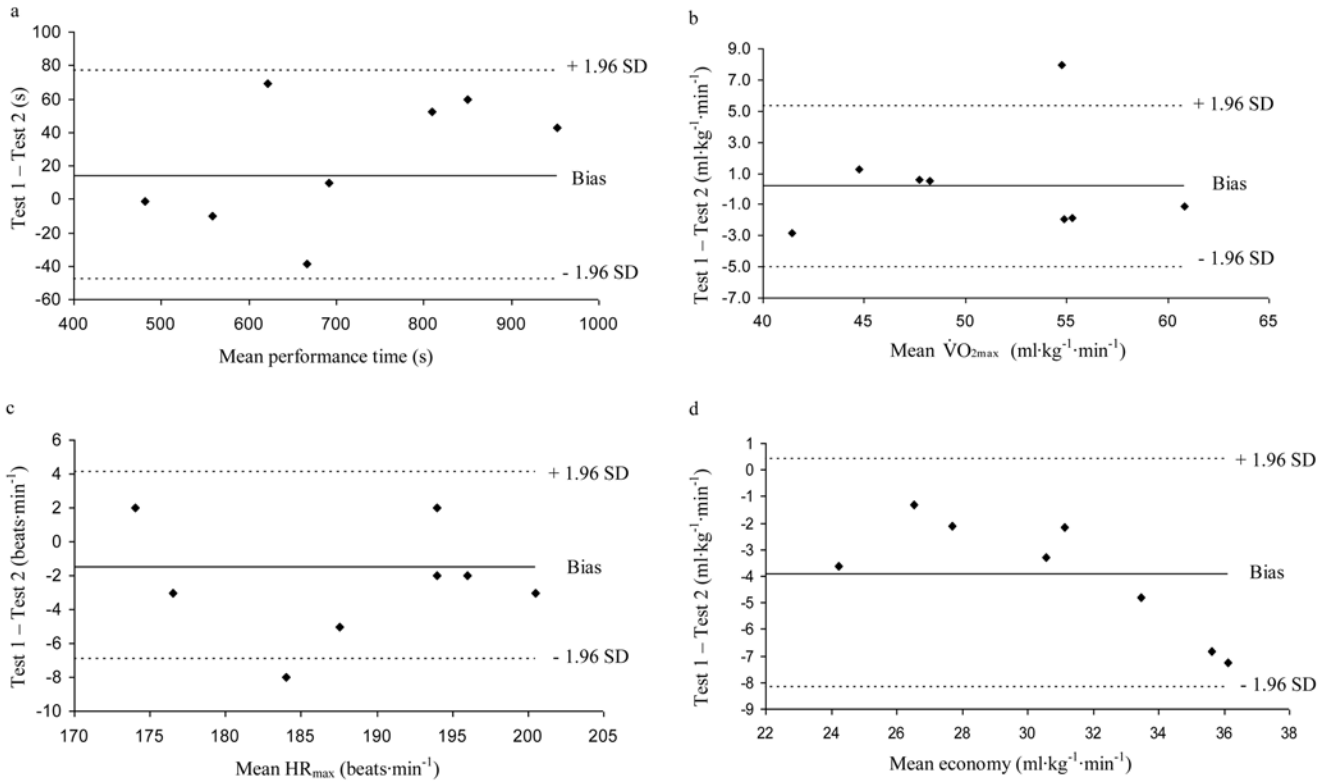

Figure 2 — Bland–Altman plots for (a) performance time; (b) Vo_{2max}; (c) HR_{max}; and (d) movement economy, measured in two trials of the ST performed 7 days apart.

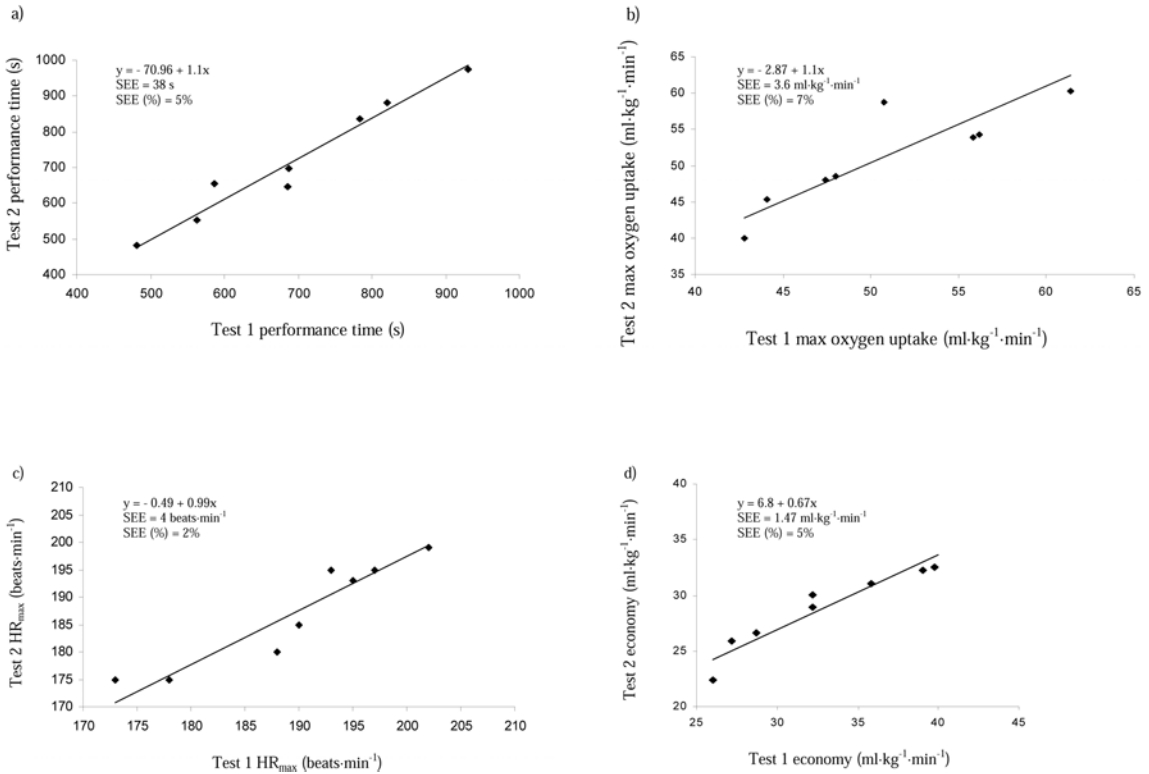

Figure 3 — LPR plots for (a) performance time; (b) Vo_{2max} ; (c) HR_{max}; (d) movement economy measured in two trials of the ST performed 7 days apart.

mode (3% 25; 5.6% 26). The results suggest both endurance performance and physiological measures from the ST are reproducible, although there are no comparative values published for squash-specific economy.

Systematic Bias Between Tests

The positive (but nonsignificant) bias evident in the LOA values for performance time and Vo_{2max} suggest improved test 2 performance especially when viewed in conjunction with the negative (nonsignificant) bias for HR_{max} and the negative (significant) bias for economy. This suggests that a learning effect occurred despite two habituation visits. However, the magnitude of these differences should be considered in the light of normal biological variation and the size of the absolute TEM scores for these variables. An examination of these values indicates that the small positive and negative biases in the LOA analyses are well within previously discussed test–retest variability for these measures.[25,26]

Which Reproducibility Measure Should Be Favored?

Methods for assessing reproducibility are debated, with some authors favoring LOA analysis and others recommending test–retest coefficient of variation (TEM), and still others preferring LPR.[18,19,23] Typical error of measurement represents approximately 68% of the error actually present in the repeated measurement of an individual in the sample, whereas LOA represents 95% of the likely variation in scores between repeated tests of a population.[18] Ludbrook[23] argues for the use of LPR analysis as it minimizes the sum of the products of horizontal and vertical distances of *x* and *y* values from the regression line. However, Atkinson and Nevill[18] point out that reproducibility analysis does not generally possess a predictor and response variable (an assumption of regression analysis) and that the assumption of a homogenous sample is not always met. The arguments that each of the authors presents for the use of their preferred analysis method all have merits, but it is beyond the scope of this paper to discuss the statistical benefits of one method over another, or their application to particular study designs. Nevertheless, a common factor in all the methods discussed and used in the current study is that the interpretation of reproducibility requires the researcher to judge (based on proposed use of the test) whether the test–retest error is small enough for the test to be of practical use.[18] To make this judgment, the researcher must possess knowledge of the smallest worthwhile change in a performance or physiological variable, and then assess whether the test is sensitive enough to detect such a change.[19] We suggest that TEM analysis best suits this purpose. This is due to the simplicity of interpretation (absolute and percentage error) and the accompanying confidence intervals, the upper value of which can be used (if the typical error and size of the CI is small) as an estimate of the lower limit for a *meaningful change* in a variable with repeat testing.[19] Moreover, the anticipated value for TEM is independent of sample size and does not suffer from the bias that can occur when LOA are calculated with small degrees of freedom (ie, small sample sizes and few repeat tests).[19]

Physiological Profiling Using the Incremental Squash Test

Laboratory-based exercise tests are challenged by the need to reflect the specific muscular, metabolic, and technical demands of a particular sport. Success in squash depends on technical, tactical, and motor skills.[27] However, owing to the nature of the game at the highest standard, aerobic fitness is an essential attribute.[5,8,9] Previous studies have demonstrated the specificity of aerobic fitness in squash players and the efficacy of specific training as preparation for match play.[11,14] However, appropriate training intensities based on prior physiological assessment are key to the success of training.[2] It is common practice in other endurance sports to train in heart rate zones defined by proximity to the lactate threshold and movement speeds that correspond to $\dot{V}O_{2max}$. The ST described in this study allows collection of all the data necessary to provide a full aerobic physiological profile of a player ($\dot{V}O_{2max}$, lactate threshold, economy, movement speed at $\dot{V}O_{2max}$, etc.) from which training intensities in squash-specific movement patterns can be derived. Furthermore, the reproducibility reported provides further support for the use of the ST as an assessment tool. The confidence intervals reported could also be used to assess whether a training intervention has resulted in a meaningful change in endurance performance or physiological responses on the ST. However, future studies should examine reproducibility over longer test–retest durations to confirm the usefulness of the test for the tracking of training adaptations in fitness and performance. Test–retest variability should also be established for other samples of squash players such as females, juniors, and subelite groups.

The importance of aerobic fitness[5,8,9] and the value of sport-specific assessment of this attribute for squash are well documented.[14] As such, squash-specific aerobic profiling using a test sensitive enough to track training-induced changes is likely to be a useful addition to the fitness assessment of squash players. We suggest that the ST described in this study could provide these benefits. However, it should be noted that aerobic profiling requires test sessions of approximately 1 hour (including participant preparation). This needs to be considered when planning the schedule of test batteries.

Conclusions

The results suggest that the squash-specific incremental test described produces reproducible measures for the assessment of squash-specific fitness and performance capabilities. Further testing is required to establish measurement error over longer test–retest durations and thus confirm the value of the test for tracking adaptations over extended training periods.

References

1. Vučković G, Dežman B, Erčulj F, Perš J. Differences between the winning and the losing players in a squash game in terms of distance covered. In: Lees A, Kahn J, eds. *Science and Racket Sports III*. Oxon: Routledge; 2004:202–207.

2. Müller E, Benko U, Raschner C, Schwameder H. Specific fitness and testing in competitive sports. *Med Sci Sports Exerc*. 2000;32:216–220.
3. Hughes M, Robertson C. Using computerised notational analysis to create a template for elite squash and its subsequent use in designing hand notation systems for player development. In: Lees A, Maynard I, Hughes M, Reilly T, eds. *Science and Racket Sports II*. London: E & FN Spon; 1998:227–234.
4. Montpetit RR. Applied Physiology of Squash. *Sports Med*. 1990;10:31–41.
5. Girard O, Chevalier R, Habrad M, Sciberras P, Hot P, Millet GP. Game analysis and energy requirements of elite squash. *J Strength Cond Res*. 2007;21:909–914.
6. Sharp NCC. Physiological demands and fitness for squash. In: Lees A, Maynard I, Hughes M, Reilly T, eds. *Science and Racket Sports II*. London: E & FN Spon; 1998:1–13.
7. Brown D, Winter EM. Heart rate responses in squash during competitive matchplay. *J Sports Sci*. 1995;14:68–69.
8. Glaister M. Multiple Sprint Work: Physiological responses, mechanisms of fatigue and the influence of aerobic fitness. *Sports Med*. 2005;35:757–777.
9. Brown D, Weigland DA, Winter EM. Maximum oxygen uptake in junior and senior elite squash players. In: Lees A, Maynard I, Hughes M, Reilly T, eds. *Science and Racket Sports II*. London: E & FN Spon; 1998:15–19.
10. Chin M, Steininger K, So RCH, Clark CR, Wong ASK. Physiological profiles and sport specific fitness of Asian élite squash players. *Br J Sports Med*. 1995;29:158–164.
11. Todd MK, Mahoney CA, Wallace WFM. The efficacy of training routines as a preparation for competitive squash. In: Lees A, Maynard I, Hughes M, Reilly T, eds. *Science and Racket Sports II*. London: E & FN Spon; 1998:91–96.
12. Sherman RA, Creasey TJ, Batterham AM. An on-court, ghosting protocol to replicate physiological demands of a competitive squash match. In: Lees A, Kahn J, eds. *Science and Racket Sports III*. Oxon: Routledge; 2004:3–8.
13. Steininger K, Wodick RE. Sports-specific fitness testing in squash. *Br J Sports Med*. 1987;21(2):23–26.
14. Girard O, Sciberras P, Habrard M, Hot P, Chevalier R, Millet GP. Specific incremental test in elite squash players. *Br J Sports Med*. 2005;39:921–926.
15. Kingsley M, James N, Kilduff LP, Dietzig RE, Dietzig B. An exercise protocol that simulates the activity patterns of elite junior squash. *J Sports Sci*. 2006;24:1291–1296.
16. Behm DG. Plyometric training for squash. *J Strength Cond Res*. 1992;14:26–28.
17. Alricsson M, Harms-Ringdahl K, Werner S. Reliability of sports related functional tests with an emphasis on speed and agility in young athletes. *Scand J Med Sci Sports*. 2001;14:229–232.
18. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998;26:217–238.
19. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30:1–15.
20. Wilkinson M, Leedale-Brown D, Winter EM. Validity of a squash-specific fitness test. *Int J Sports Physiol Perform.* 2009;4:29–40.
21. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. *Sports Med*. 2007;37:1019–1028.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between tow methods of clinical measurement. *Lancet*. 1986;I:307–310.
23. Ludbrook J. Comparing methods of measurement. *Clin Exp Pharmacol Physiol*. 1997;24:193–203.

24. Hopkins WG. Reliability from consecutive pairs of trials (Excel spreadsheet). In *A new view of statistics.* Sportsci.org: Internet Society for Sport Science, sportsci.org/resource/stats/xrely.xls, 2000.
25. Bagger M, Petersen PH, Pederson PK. Biological variation in variables associated with exercise training. *Int J Sports Med*. 2003;24:433–440.
26. Katch VL, Sady SS, Freedson P. Biological variability in maximum aerobic power. *Med Sci Sports Exerc*. 1982;14:21–25.
27. Lees A. Science and the major racket sports: a review. *J Sports Sci*. 2003;21:707–732.