©Journal of Sports Science and Medicine (2009) **8,** 77-82
http://www.jssm.org

Research article

Familiarization, reliability, and comparability of a 40-m maximal shuttle run test

Mark Glaister ✉, Hanna Hauck, Corinne S. Abraham, Kevin L. Merry, Dean Beaver, Bernadette Woods and Gillian McInnes
St Mary's University College, Strawberry Hill, Twickenham, UK.

Abstract
The aims of this study were to examine familiarization and reliability associated with a 40-m maximal shuttle run test (40-m MST), and to compare performance measures from the test with those of a typical unidirectional multiple sprint running test (UMSRT). 12 men and 4 women completed four trials of the 40-m MST (8 × 40-m; 20 s rest periods) followed by one trial of a UMSRT (12 × 30-m; repeated every 35 s); with seven days between trials. All trials were conducted indoors and performance times were recorded via twin-beam photocells. Significant between-trial differences in mean 40-m MST times were indicative of learning effects between trials 1 and 2. Test-retest reliability across the remaining trials as determined by coefficient of variation (CV) and intraclass correlation coefficient (ICC) revealed: a) very good reliability for measures of fastest and mean shuttle time (CV = 1.1 – 1.3%; ICC = 0.91 – 0.92); b) good reliability for measures of blood lactate (CV = 10.1 – 23.9%; ICC = 0.74 – 0.82) and ratings of perceived exertion (CV = 5.3 – 7.6%; ICC = 0.79 – 0.84); and c) poor reliability for measures of fatigue (CV = 38.7%; ICC = 0.59). Comparisons between performance indices of the 40-m MST and the UMSRT revealed significant correlations between all measures, except pre-test blood lactate concentration (r = 0.47). Whilst the 40-m MST does not appear to provide more information than can be gleaned from a typical UMSRT, following the completion of a familiarization trial, the 40-m MST provides an alternative and, except for fatigue measures, reliable means of evaluating repeated sprint ability.

Key words: repeated sprint ability, intermittent, agility, multiple sprint work.

Introduction

Tests of multiple sprint performance are becoming increasingly popular as a way of evaluating the performance capabilities of athletes involved in field and court sports. Based on the results of several time-motion analyses, these tests typically comprise repeated bouts (≤ 20) of short (≤ 6 s) unidirectional sprints interspersed with relatively short (≤ 60 s) recovery periods (Spencer *et al.*, 2005). However, in many sporting events, athletes are required to beat or challenge opponents by rapidly and repeatedly changing direction (Sheppard and Young, 2006). As such, multiple sprint tests incorporating changes of direction would appear to provide a more ecologically valid way of assessing the ability to excel in various multiple sprint sports; particularly since unidirectional sprints and sprints incorporating changes of direction are reported to be poorly correlated (Sheppard and Young, 2006). One test which incorporates these characteristics is the 40-m maximal shuttle run test (40-m MST) (Baker et al., 1993).

For the 40-m MST to be used effectively as a way of quantifying athletic ability and evaluating the effects of various experimental interventions, it is important to establish the reliability of the key performance outcomes of the test. Although Baker et al. (1993) reported a test-retest Pearson correlation coefficient of 0.86 for fastest shuttle run time, no corresponding value was reported for the other main performance outcome, fatigue. Moreover, since the sample size of the study was relatively small (n = 10), the Pearson correlation coefficient is likely to have overestimated the true strength of the relationship (Hopkins, 2000). Whilst the use of an intraclass correlation coefficient (ICC) would provide a better measure of reliability, in comparison with the coefficient of variation (CV), this approach still fails to provide an indication of the magnitude of the variability between trials (Hopkins, 2000).

In addition to the lack of information regarding the reliability of the 40-m MST, there are a number of factors that are likely to impact on the reliability of the test outcomes. First, previous research used hand timing to measure the duration of each shuttle run (Baker et al., 1993); thereby introducing another source of biological variability and error to the data (Moore et al., 2007). Secondly, the procedure for determining measures of fatigue conflicts with the one developed by Fitzsimmons et al. (1993) which recent research has suggested to be the most valid and reliable in unidirectional multiple sprint tests (Glaister et al., 2008). Thirdly, there has been no discussion of possible learning effects associated with the test; an important consideration given the skill and co-ordination required to change direction at high speed.

The aims of the present study therefore were to address the above issues by introducing electronic timing to the test and subsequently to: a) determine the number of familiarization trials required to minimise the impact of learning effects on test-retest reliability of the 40-m MST; b) establish test-retest reliability of the key performance outcomes and associated measures of the 40-m MST once familiarization has occurred; and c) compare performance outcomes of the 40-m MST with a typical unidirectional test of multiple sprint running.

Methods

Subject

Sixteen physically active sport science students (men: n = 12; age 21 ± 2 years, height 1.81 ± 0.05 m, mass 76.5 ±

Received: 03 September 2008 / Accepted: 05 January 2009 / Published (online): 01 March 2009

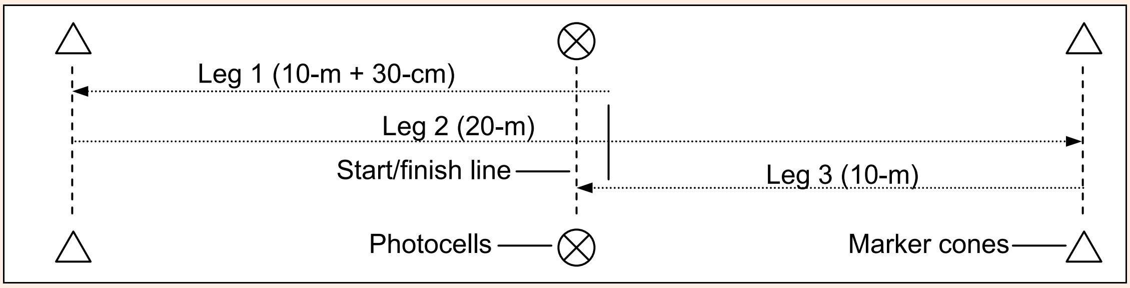

Figure 1. Schematic illustration of the set-up for the 40-m maximal shuttle run test.

9.3 kg; women: n = 4; age 20 ± 1 years, height 1.64 ± 0.10 m, mass 61.0 ± 8.5 kg) (mean ± *s*) from a variety of sporting backgrounds volunteered for the study, which was approved by St Mary's University College Ethics Committee. Prior to testing, subjects received written and verbal instructions regarding the nature of the investigation and completed a training history questionnaire, which indicated that all had been actively involved in sport for approximately 13 years and that most (*n* = 13) regularly participated in some form of multiple sprint sport. Mean times spent training and competing each week were reported as 9.4 ± 5.2 and 5.1 ± 5.6 hours, respectively. Prior to commencement, all subjects completed a health-screening questionnaire and provided written informed consent.

Procedures

To provide sufficient data for familiarization analysis and to improve the precision of reliability estimates, all subjects first completed four trials of the 40-m MST, which consisted of 8 × 40-m shuttles, interspersed with 20 s recovery intervals. To enable performance in the 40-m MST to be compared with that of a standard multiple sprint test, subjects subsequently completed a 12 × 30-m unidirectional multiple sprint test with sprints repeated at 35 s intervals (Glaister et al., 2007). All trials were completed at approximately the same time of day with seven days between each trial. Subjects were instructed to avoid food and drink in the hour before testing and to avoid strenuous exercise and caffeine consumption 24 hours prior to each trial.

All testing was conducted indoors on a synthetic running surface. Prior to each multiple shuttle/sprint test, subjects performed a standardised warm-up (approximately five-minutes) comprising 400-m of jogging (self-selected pace), a series of sprint drills (high-knees, heel-flicks, and walking lunges), and three practice sprints. Following the warm-up, subjects were given five minutes to stretch and prepare themselves for each test. For all tests, sprints were initiated from a line 30-cm behind the start line (to prevent false triggering of the first timing gate) and times were recorded electronically via twin-beam photocells (Swift Performance Equipment, Lismore, Australia). In the 40-m MST, subjects were required to run between two lines placed 20-m apart, with the start/finish line (and the photocells) placed at the midpoint of the course (Figure 1). On instruction, each subject sprinted 10 m from the start/finish line to the end of the course, turned 180°, sprinted 20-m to the other end of the course, turned 180°, and sprinted 10-m back through the start/finish line. Subjects were instructed to place at least one foot over the line at the end of each shuttle, the adherence of which was monitored to ensure full compliance. In the unidirectional multiple sprint test, subjects were required to sprint between two lines 30-m apart, with alternate sprints performed in the opposite direction to maximise the available recovery time. Recovery times for both protocols were recorded using a digital stopwatch and a 5 s countdown was provided prior to the start of each shuttle/sprint. During each trial, subjects were verbally encouraged to give a maximal effort. To provide an indication of the anaerobic contribution to the exercise, one minute before the start of each trial and immediately after, capillary blood samples were drawn from a hyperaemised earlobe for the subsequent analysis (via an automated analyser) of blood lactate (Biosen C-Line, EKF Diagnostic, Ebendorfer Chaussee 3, Germany). The analyser was calibrated prior to each trial in accordance with the manufacturer's instructions. Ratings of perceived exertion (RPE) were recorded midway and at the end of each trial using a 15-point scale (Borg, 1970). Fatigue during each trial was calculated from 40-m shuttle and 30-m sprint times, respectively, using the percentage decrement calculation (Fitzsimmons et al., 1993):

Percentage decrement calculation

Fatigue = (100 x (total sprint time ÷ ideal sprint time)) – 100

Where total sprint time = sum of sprint times from all sprints. Ideal time = number of sprints × fastest sprint time.

Statistical analyses

All statistical analyses were conducted using the Statistical Package for the Social Sciences (SPSS for Windows, SPSS Inc., Chicago, IL). Measures of centrality and spread are presented as means ± standard deviation (s). To examine the process of familiarization, differences in the performance measures of both fastest and mean 40-m shuttle times were evaluated with one-way repeated measures analysis of variance (ANOVA). If significant ($p < 0.05$) between-trial differences were observed, a Bonferonni-adjusted post-hoc analysis was used to determine where those differences occurred. After determining the number of trials required to limit the effects of familiarization, measures of within-subject variation (CV) were

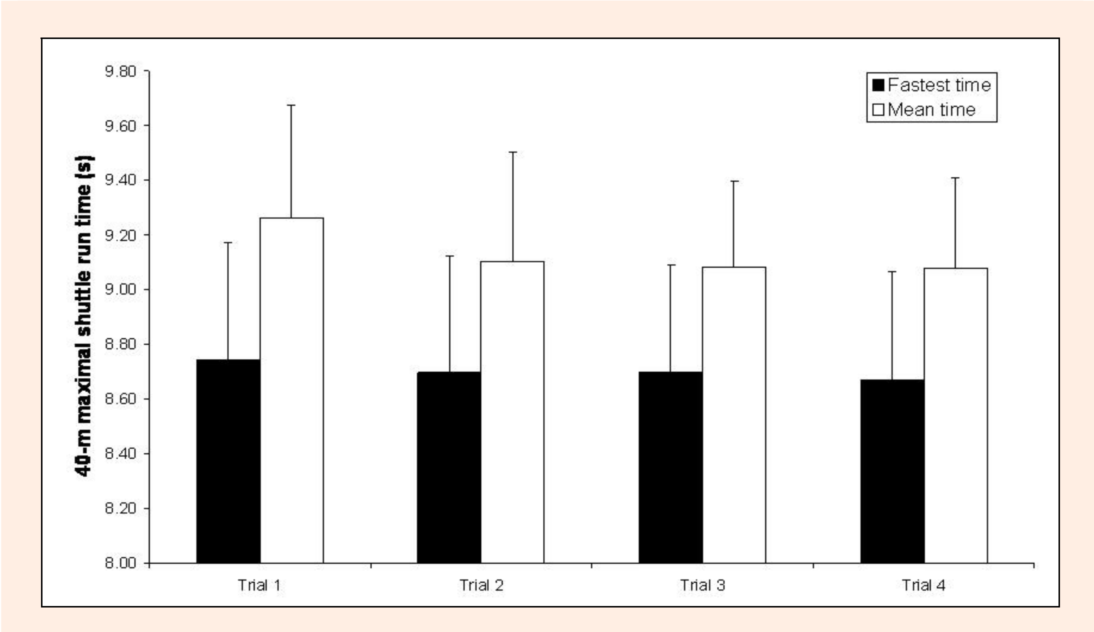

Figure 2. Fastest and mean sprint times from four trials of a maximal shuttle run test (8 × 40-m; 20 s rest periods). Values are presented as mean ± SD.

derived from a two-way ANOVA on the remaining trials using the procedure described by Schabort et al. (1999). Fastest time, mean time, fatigue, blood lactate concentration, and RPE were the dependant variables in each model, with subject identity included as a random effect and trial number as a fixed effect. Retest correlations were derived from the ANOVA as ICC using the method described by Bartko (1966). Confidence limits (95%) for CV and ICC were calculated using Chi-square, and McGraw and Wong (1996) estimates, respectively. Comparisons between the dependent variables of the 40-m MST and the 30-m multiple sprint test were performed using Pearson Product Moment Correlations, with trial 4 of the 40-m MST used as the representative trial. Correlation coefficients were interpreted in accordance with the following scale of magnitudes as devised by Cohen (1988): $r < 0.1$, trivial; $0.1 \le r < 0.3$, small; $0.3 \le r < 0.5$, moderate; $r \ge 0.5$, large.

Results

Familiarization

The pattern of responses for measures of fastest and mean sprint time during the 40-m MST are presented in Figure 2. Whilst measures of fastest time showed no significant difference between trials ($F_{(3,45)} = 0.904$; $p = 0.447$), the same was not true for mean time ($F_{(2.0,29.9)} = 9.805$; $p = 0.001$). Further analysis of mean time revealed significant differences only in those contrasts that involved Trial 1. Since learning effects influence the reliability of the performance-related measures, reliability was subsequently evaluated across trials 2 – 4.

Reliability

Reliability data for all measures are presented in Table 1. Measures of fastest and mean sprint time showed very good test-retest reliability as evidenced by the low coefficients of variation and high intraclass correlation coefficients. In contrast, measures of fatigue showed poor reliability. Test-retest reliability for blood lactate concentration was moderate and was better for post-test measures. Test-retest reliability for RPE was high and was greater at Sprint 8 than at Sprint 4.

Correlations

Results from the 30-m multiple sprint test are presented in Table 2, with results of the correlation analysis comparing corresponding data from the 40-m MST presented in Table 3. With the exception of pre-test measures of blood lactate concentration, results revealed significant correlations between all contrasts.

Discussion

The aims of the present study were to examine familiarization and reliability associated with the 40-m

Table 1. Reliability data from three trials of a repeated maximal shuttle run test (8 × 40-m; 20 s rest periods).

	Fastest time (s)	Mean time (s)	Fatigue (%)	Blood lactate (mmol·L^{-1})		RPE	
				Pre-test	*Post-test*	*Sprint 4*	*Sprint 8*
Mean (SD)	8.69 (0.40)	9.09 (0.34)	4.66 (2.09)	3.96 (1.52)	11.61 (2.51)	14 (2)	17 (2)
CV (%)	1.3	1.1	38.7	23.9	10.1	7.6	5.3
95% CL	1.1 – 1.8	0.9 – 1.5	30.9 – 51.7	19.1 – 32.0	8.1 – 13.5	6.0 – 10.1	4.3 – 7.1
ICC	0.92	0.91	0.59	0.74	0.82	0.79	0.84
95% CL	0.83 – 0.97	0.82 – 0.97	0.31 – 0.81	0.51 – 0.89	0.64 – 0.93	0.58 – 0.91	0.67 – 0.93

CV = coefficient of variation; CL = confidence limits; ICC = intraclass correlation coefficient.

Table 2. Performance measures and related data from a unidirectional multiple sprint running test (12 × 30-m; repeated at 35 s intervals).

	Fastest time (s)	Mean time (s)	Fatigue (%)	Blood lactate (mmol·L^{-1})		RPE	
				Pre-test	*Post-test*	*Sprint 6*	*Sprint 12*
Mean (±SD)	4.57 (0.25)	4.78 (0.23)	4.57 (2.32)	3.59 (1.21)	9.53 (2.92)	13 (1)	16 (2)

RPE = rating of perceived exertion.

MST, and to compare performance and associated parameters in the test with those of a typical unidirectional multiple sprint test. With regards to familiarization, the results revealed evidence of learning effects between the first two trials. In other words, performance improved from Trial 1 to Trial 2, and remained relatively consistent thereafter. In effect, despite the subjects being well-trained, it appears that the demands of the test were sufficiently novel to inhibit optimal performance in the first trial. Whilst the same phenomenon could also potentially be explained by a training effect, it is unlikely that such a brief stimulus would be sufficient to produce the observed response, particularly given the training background of the subjects. Moreover, the idea of a training effect would fail to account for the absence of any further significant improvements in the remaining trials. Learning effects in tests of multiple sprint work have previously been observed in tests involving cycle ergometry (Capriotti et al., 1999; Glaister et al., 2003), though not in unidirectional sprint running (Glaister et al., 2007); the latter being explained by the competency of the subjects to perform this type of exercise as a result of frequent exposure.

After minimising the influence of learning effects, analysis of the remaining trials showed good test-retest reliability for two of the key performance outcomes of the test, namely fastest and mean time. These values are comparable with those reported in tests of unidirectional multiple sprint running (Fitzsimmons et al., 1993; Glaister et al., 2007). In contrast, and in-line with previous research in multiple sprint work (Fitzsimmons et al., 1993; Glaister et al., 2007), measures of fatigue showed poor reliability. Reasons for this anomaly have been addressed elsewhere (Glaister et al., 2008), but in short, are probably related to the fact that fatigue is derived, rather than measured, from data which have their own inherent variability.

Reliability statistics for the remaining performance-related parameters of blood lactate concentration and RPE were high when considered in context with mean values. For example, test-retest variability of less than 1.0 mmol·L^{-1} would seem reasonable to expect from a pre-test measure of blood lactate. Furthermore, the reliability of blood lactate data compares well with that in unidirectional multiple sprint work (Glaister et al., 2007) and, as with measures of fastest and mean sprint time, reflects the degree of precision with which the effects of various experimental interventions on these measures can be evaluated.

The final aim of the present study was to compare the key performance indicators of the 40-m MST with those of a typical unidirectional multiple sprint running test. Previous research, comparing performance time in a change of direction test with that of a unidirectional sprint test has reported moderate to low correlations (Sheppard and Young, 2006). In fact, in their review of the agility literature, Sheppard and Young (2006) reported that, in contrast to anecdotal beliefs, no evidence could be found of a strong relationship between unidirectional sprinting speed and change of direction sprinting speed. Whilst the present study involved repeated bouts of maximal exercise, the fact that measures of fastest shuttle/sprint time were strongly correlated suggests that this is an issue requiring further investigation, ideally with a larger sample size given the spread of the associated confidence limits. Nevertheless, the fact that all measures, with the exception of pre-test blood lactate concentration, were strongly correlated reflects the large degree of similarity in the overall physical and metabolic demands of the two protocols. Indeed, even measures of fatigue were strongly correlated, despite their poor level of reliability.

Conclusion

Tests of multiple sprint work are designed to evaluate the attributes of speed and endurance required to excel in multiple sprint sports. The results of the present study show that the 40-m MST provides a relatively easy and, except for fatigue measures, reliable way of evaluating these attributes. Whilst the 40-m MST does not appear to provide any more information than can otherwise be gleaned from a typical unidirectional multiple sprint running test, it does provide practitioners and coaches with a viable alternative, particularly when constrained by limited testing space. However, before the test can be used to accurately evaluate the effects of various interventions, it is important that athletes perform at least one practice trial to minimise the influence of learning effects.

Acknowledgments

The authors would like to express their gratitude to Jon Goodwin, Glyn Howatson, Anthony Lockey, and John Pattison for their help in carrying out the tests.

References

Baker, J., Ramsbottom, R. and Hazeldine, R. (1993) Maximal shuttle running over 40 m as a measure of anaerobic performance.

Table 3. Pearson correlation coefficients between several key performance determinants in a 40-m maximal shuttle run test (8 × 40-m; 20 s rest periods) and a 30-m multiple sprint test (12 × 30-m; repeated at 35 s intervals).

	Fastest time	Mean time	Fatigue	Blood lactate		RPE	
				Pre-test	*Post-test*	*Mid-test*	*End-test*
Pearson *r*	0.83	0.83	0.71	0.47	0.77	0.51	0.85
95% CL	0.57 – 0.94	0.57 – 0.94	0.33 – 0.89	-0.03 – 0.79	0.45 – 0.92	0.01 – 0.80	0.61 – 0.95

CL = confidence limits; RPE = rating of perceived exertion.

British Journal of Sports Medicine **27**, 228-232.

Bartko, J.J. (1966) The intraclass correlation coefficient as a measure of reliability. *Psychological Reports* **19**, 3-11.

Borg, G. (1970) Perceived exertion as an indicator of somatic stress. *Scandinavian Journal of Rehabilitation Medicine* **2**, 92-98.

Capriotti, P.V., Sherman, W.M. and Lamb, D.R. (1999). Reliability of power output during intermittent high-intensity cycling. *Medicine and Science in Sports and Exercise* **31**, 913-915.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences.* 2nd edition. Lawrence Erlbaum, New Jersey.

Fitzsimmons, M., Dawson, B. Ward, D. and Wilkinson, A. (1993) Cycling and running tests of repeated sprint ability. *Australian Journal of Science and Medicine in Sport* **25**, 82-87.

Glaister, M., Howatson, G., Lockey, R.A., Abraham, C.S. Goodwin, J.E. and McInnes, G. (2007) Familiarisation and reliability of multiple sprint running performance indices. *Journal of Strength and Conditioning Research* **21**, 857-859.

Glaister, M., Howatson, G., Pattison, J.R. and 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.

Glaister, M., Stone, M.H., Stewart, A.M., Hughes, M. and Moir, G.L. (2003) Reliability of power output during short-duration maximal-intensity intermittent cycling. *Journal of Strength and Conditioning Research* **17**, 781-784.

Hopkins, W.G. (2000) Measures of reliability in sports medicine and science. *Sports Medicine* **30**, 1-15.

McGraw, K.O. and Wong, S.P. (1996) Forming inferences about some intraclass correlation coefficients. *Psychological Methods* **1**, 30-46.

Moore, A.N., Decker, A.J., Baarts, J.N., Dupont, A.M., Epema, J.S., Reuther, M.C., Houser, J.J. and Mayhew, J.L. (2007) Effect of competitiveness on forty-yard dash performance in college men and women. *Journal of Strength and Conditioning Research* **21**, 385-388.

Schabort, E.J., Hawley, J.A., Hopkins, W.G. and Blum, H. (1999) High reliability of performance of well-trained rowers on a rowing ergometer. *Journal of Sports Sciences* **17**, 627-632.

Sheppard, J.M. and Young, W.B. (2006) Agility literature review: classifications, training and testing. *Journal of Sports Sciences* **24**, 919-932.

Spencer, M., Bishop, D., Dawson, B. and Goodman, C. (2005) Physiological and metabolic responses of repeated-sprint activities: specific to field-based team sports. *Sports Medicine* **35**, 1025-1044.

Key points

- Tests of multiple sprint performance are a popular means of evaluating repeated sprint ability.
- Multiple sprint tests incorporating changes of direction may be more ecologically valid than unidirectional protocols.
- The 40-m maximal shuttle run test is a reliable way of evaluating repeated sprint ability following the completion of one familiarization trial.
- The 40-m maximal shuttle run test shows no clear advantage over a standard unidirectional multiple sprint test.

AUTHORS BIOGRAPHY

Mark GLAISTER
Employment
Senior Lecturer, School of Human Sciences, St Mary's University College
Degree
PhD
Research interests
Physiological responses to multiple sprint work and ergogenic aids to multiple sprint performance
E-mail: glaistem@smuc.ac.uk

Hanna HAUCK
Employment
Undergraduate student, School of Human Sciences, St Mary's University College
Degree
Studying for BSc in History and Human Biology
E-mail: 060342@students.smuc.ac.uk

Corinne S. ABRAHAM
Employment
Lecturer, School of Human Sciences, St Mary's University College
Degree
MSc
Research interests
Multiple sprint work and mechanics of martial arts kicks
E-mail: abrahamc@smuc.ac.uk

Kevin L. MERRY
Employment
Lecturer, School of Human Sciences, St Mary's University College
Degree
BSc
Research interests
Multiple sprint work and methods of vVO_{2max} determination
E-mail: merryk@smuc.ac.uk

Dean BEAVER
Employment
Lecturer, School of Education, St Mary's University College
Degree
BSc
Research interests
3D analysis of weightlifting techniques
E-mail: beaverd@smuc.ac.uk

Bernadette Woods
Employment
Lecturer, School of Human Sciences, St Mary's University College
Degree
PhD
Research interests
Multiple sprint work and anxiety/performance concerns in team sports
E-mail: woodsb@smuc.ac.uk

Gillian McInnes
Employment
Academic Director, School of Human Sciences, St Mary's University College
Degree
MSc
Research interests
Physiological responses to multiple sprint work
E-mail: mcinnesg@smuc.ac.uk

✉ **Dr Mark Glaister**
School of Human Sciences, St Mary's University College, Waldegrave Road, Strawberry Hill, Twickenham, UK.