Reliability of 'in-season' fitness assessments in youth elite soccer players: a working model for practitioners and coaches

Journal:	*Science and Medicine in Football*
Manuscript ID	RSMF-2017-0027.R1
Manuscript Type:	Short Communications/Technical Reports
Keywords:	youth-soccer, reliability, in-season, yo-yo test, sprint, strength

Abstract: Purpose: To assess the reliability of a battery of field tests when performed around habitual training during one micro-cycle of the in-season period in youth elite soccer players. Methods: n=19 English Premier League academy players (mean ± SD: age, 18.3±0.2 years; stature, 1.80±0.05 m; body mass, 76.5±7.5 kg; V̇O2 max; 62.3±4.38 ml·kg-1·min-1 Sum of 8 skinfolds; 64.8±17.4 mm) performed; '1RM half-back squat (HBS)', 'vertical jump' (VJ), 'Yo-Yo IR2', '5, 10 & 20 m, sprint', 'the agility T-test' and a 'repeated sprint ability' assessments around their habitual 'in-season' training and match-play on two occasions. Typical error (TE), coefficient of variation (CV), the smallest worthwhile change (SWC) was then calculated between trials. Test usefulness was then calculated by comparing the 'SWC' to the TE and rated as; good, OK or marginal. The smallest difference needed to be considered real (MD), was also calculated using the equation (TE×1.96×√2). Results: Most assessments demonstrated relatively high levels of reproducibility (CV; 0.3 - 4.3 %) to witness an "almost certain beneficial change" (i.e.,>MD) changes of approximately 5% (RSAbest, RSAmean, 10m & 20m sprint, 'agility T test' and 1RM HBS), and 11.5% (VJ, 5m sprint, and Yo-Yo IR2) are needed. Conclusions: The present training and testing 'model' is reliable and could be used when evaluating the fitness of highly trained youth elite soccer players during the 'in-season' period.

Title: Reliability of 'in-season' fitness assessments in youth elite soccer players: a working model for practitioners and coaches.

Abstract

Purpose: To assess the reliability of a battery of field tests when performed around habitual training during one micro-cycle of the 'in-season' period in youth elite soccer players. **Methods:** n=19 English Premier League academy players (mean ± SD: age, 18.3±0.2 years; stature, 1.80±0.05 m; body mass, 76.5±7.5 kg; $\dot{V}O_{2\ max}$; 62.3±4.38 ml·kg^{-1}·min^{-1} Sum of 8 skinfolds; 64.8±17.4 mm) performed; '1RM half-back squat (HBS)', 'vertical jump' (VJ), 'Yo-Yo IR2', '5, 10 & 20 m, sprint', 'the agility T-test' and a 'repeated sprint ability' assessments around their habitual 'in-season' training and match-play on two occasions. Typical error (TE), coefficient of variation (CV), the smallest worthwhile change (SWC) was then calculated between trials. Test usefulness was then calculated by comparing the 'SWC' to the TE and rated as; good, OK or marginal. The smallest difference needed to be considered real (MD), was also calculated using the equation (TE×1.96×√2). **Results:** Most assessments demonstrated relatively high levels of reproducibility (CV; 0.3 - 4.3 %) to witness an "almost certain beneficial change" (i.e.,>MD) changes of approximately 5% (RSA_{best}, RSA_{mean}, 10m & 20m sprint, 'agility T test' and 1RM HBS), and 11.5% (VJ, 5m sprint, and Yo-Yo IR2) are needed. **Conclusions:** The present training and testing 'model' is reliable and could be used when evaluating the fitness of highly trained youth elite soccer players during the 'in-season' period.

INTRODUCTION

Elite football players require a range of dynamic and discrete physical qualities and capacities (e.g., strength, speed, aerobic capacity and agility). As such, the implementation of systematic fitness assessments across the season has been recommended (Reilly, 2007). If implemented in the correct manner, this process can allow the coaching staff to establish 'baseline data' for their players, highlight individuals' and teams strengths and/or weaknesses, assess the effectiveness of training interventions, track players across time and inform future training strategies (Maud & Foster 2006). Alternatively, if assessments are not reliable or conducted frequently enough, they are less likely to inform the coaches decision making process (Cross & Lyle 1999). Therefore, it is vital that assessments are reliable and implemented in a way that has synergy with the coaches' philosophy and can be performed at various time-points throughout the season with minimal impact upon other important aspects, such as training and match-play related activities.

Performing a 'battery' of fitness tests (using field based tools) in one day has been recommended (Turner et al. 2011). In this regard, assessments of strength (e.g., repetition maximum [RM] tests), speed (e.g., 5 to 30 meter sprints), power (e.g., counter-movement jump tests), sports specific aerobic/anaerobic capacity (e.g., Yo-Yo tests, repeated sprint assessments), and agility tests (e.g., T-Test, 5-0-5 agility) are recommended for youth professional soccer players (Hulse et al. 2013). A 'testing battery' can take one training day to complete (n=15-20 players), (usually across a morning and afternoon session), where, power and speed based movements are conducted in the morning and anaerobic and strength related tests are administered in the afternoon (Walker and Turner 2009).

Although, removing even one day from the weekly 'in-season' training schedule can be a challenge in professional football. As a compromise, it is possible that tests could be separated across the training week and be performed around the soccer-specific training that needs to be completed. If found to be reliable this approach would be advantageous as it could allow data to be collected without the need to remove a training day from the weekly schedule, thus, having minimal impact upon 'match-play', 'technical and/or tactical' training.

The aim of this study was to assess the reliability of a battery of field tests ('1RM half-back squat test', 'vertical jumping assessments', 'Yo-Yo IR2 test', '5, 10 & 20 m, sprint', 'the agility T-test' and a 'repeated sprint ability test') when performed around one micro-cycle of the 'in-season' training in youth elite soccer players. We aimed to highlight the 'usefulness of the test' battery by investigating the 'typical error' (TE), 'smallest worthwhile change' (SWC), and minimal change required to be considered 'real' (MD) using methods previously described by (Buchheit et al. 2010).

METHODS

Participants. Nineteen players from a Premier League academy in England participated in this study (mean ± SD: age, 18.3±0.2 years; stature, 1.80±0.05 m; body mass, 76.5±7.5 kg; $\dot{V}O_{2\ max}$; 62.3±4.38 $ml \cdot kg^{-1} \cdot min^{-1}$ Sum of 8 skinfolds; 64.8±17.4 mm; training history; 4-6 $d \cdot wk^{-1} \cdot 2$ years). Each player had previously completed all tests on 5 or more occasions. All participants were given oral and written information concerning any possible risks associated with protocols prior to the study. All participants gave their written informed consent to participate in the study. The study was approved by the Institutional Review Board (IRB) at a University within the United Kingdom.

Study design

The present experimental study followed a test-re-test design where a battery of fitness assessments where performed around one week of 'in-season' training on two occasions separated by 7 days. The scheduling of testing and training is presented in Table 1. Where possible, tests followed an order that allowed the completion of both testing and training with as much recovery time as achievable in this environment.

Experimental procedures.

The Yo-Yo Intermittent Recovery test 2 (Yo-Yo IR2). Participants ran a series of 40-m shuttle runs at progressive speeds until voluntarily exhaustion or when the participant failed to reach the end line in time with the audio cue on two occasions (Krustrup and Bangsbo, 2001).

Agility 'T-Test'. Participants sprinted to the centre cone (10 m), turn to the right to the right cone (5 m), then turned and run back to the left cone (10 m), and then turn and run back to centre cone (5 m) and then ran straight back through the start/finish photocell (Brower Timing Systems, Utah, USA).

Vertical Jump Height. Following a standardised dynamic warm-up each participant performed three squat jumps (SJ) and three countermovement jumps (CMJ) using standardised methods previously described by Acero et al., (2011). Each jump height was measured using 'jump mat' (FLS electronics, Cookstown, Northern Ireland). The highest SJ and CMJ height (cm) was recorded for further analysis.

Acceleration and sprint speed. Participants completed three maximal 20 meter sprints. Photocells were placed at the '0 m', '5 m' and '10 m' and '20 m' (Brower Timing Systems, Utah, USA). The lowest sprint time for each distance was recorded for analysis.

Repeated Sprint Ability. The repeated sprinting protocol consisted of 7, 30 m sprints with 25 seconds of active recovery between each sprint (Reilly et al. 2000). The start position was located 1 meter from the first photocell. Between each sprint the participants were required to jog back to the start position within the allocated time. Using Photocells (Brower Timing Systems, Utah, USA) the fastest sprint time (RSA_{best}), average sprint time (RSA_{mean})

and the sprint decrement (expressed as a percentage) (RSA_{Dec}[%]) ((mean sprint time / best sprint time × 100) – 100) were used for analysis (Buchheit et al. 2010).

Training procedures. Training was designed and by the head football coach. To ensure reproducibility between training sessions the areas and/or dimensions of each 'small-sided game' or 'training drill' was measured with a 50 meter measuring tape and marked with cones prior all pitch based training sessions. The durations of each session, the rules of each small-sided game (e.g. number of players per team and scoring format), the time intervals between efforts within each small-sided game, breaks in play for water and level encouragement was matched in each week.

Training load. The players perceived training load for each exercise bout was measured using the rating of perceived exertion method (sRPE-TL) (Foster et al., 1995). Global positioning devises were also placed on the players during each 'pitch-based' training session (STATSports Group Limited). Resistance-training volume was characterised using the volume load (VL) method (reps x sets x weight [kg]) (Schoenfeld et al., 216).

Statistical Analyses. To examine the reliability of the test over the two consecutive trials the magnitude of differences between consecutive trials was expressed as a standardised mean difference. Here, the effect size (ES) was calculated from the ratio of the mean difference to the pooled standard deviation. The magnitude of the ES was classified using guidelines outlined by Hopkins (2001) (less than 0.2; trivial, 0.2-0.5; small, 0.5-0.8; moderate, more than 0.8; large). To allow comparison the typical error of measurement (TE) was also calculated and expressed as a coefficient of variation (CV) (Atkinson & Nevill 1998). The smallest worthwhile change (SWC) was calculated by multiplying the smallest worthwhile effect (i.e., 0.2) by the pooled standard-deviation (Hopkins 2011). The usefulness of the test was assessed by comparing the 'SWC' with the TE and was rated using indicators outlined by Buchheit and colleagues (2010) (TE < SWC; good; TE ~ SWC; OK; TE > SWC; marginal). The smallest difference needed to be considered real (MD), was calculated using the equation (TE × 1.96 × √2) (Weir 2005). Statistical significance (P) was set at ≤ 0.05. All data in text, and tables are presented as means ± SD and where possible are reported with 90% confidence intervals (90% CI). Microsoft Excel (Microsoft, Redmond, WA, USA) was used for all statistical procedures.

INSERT TABLE 1 HERE

RESULTS

Training load agreement between experimental weeks. Internal and external training load data for each category of training is presented in Table 2. Each training day was matched and compared between training weeks. Whilst some differences in 'perceived training load' (sRPE-TL) existed, there were no statistical differences between weeks and differences between weeks and were classified as 'trivial'. External training load metrics were also similar between experimental weeks and considered as 'trivial' or 'small'.

INSERT TABLE 2 HERE

All reliability data is presented in table 3. Differences in all measurements between repeated trials displayed "trivial" ES. The TE and CV was considered low for all assessment (<5%) whereas the CV for RSA_{Dec} was considered very high (28%). The 'usefulness' of each test was calculated by comparing the SWC with the TE and was rated using indicators outlined by Buchheit and colleagues (2010). This then allowed the researchers to establish the minimum change needed to be observed for it to be considered 'real'. Changes in performance likely to be "real" ranged from 1.9% to 11.5%.

INSERT TABLE 3 HERE

DISCUSSION

The aim of this study was to assess the reliability of a battery of field tests performed around habitual 'in-season' training and competition in youth elite soccer players. The findings suggest that the present training and testing 'model' is reliable and could therefore be used in the evaluation of highly trained youth elite soccer players who have been suitably familiarised and have sufficient training experience. Most assessments demonstrated relatively high levels of reproducibility (CV; 0.3 - 4.3 %). For practitioners who decide to use the present training and testing model, we have calculated the minimum values needed to be able to distinguish between a negligible change and an "almost certain beneficial change". Here, practitioners should seek to observe changes of approximately 5% (RSA $_{best}$, RSA $_{mean}$, 10m & 20m sprint, 'agility T test' and 1RM half-back squat), and 11.5% (CMJ, SQ, 5m sprint, and Yo-Yo IR2) to attain an "almost certain beneficial" increase in performance. This might be useful information for practitioners who would like to be able to distinguish between changes in physical performance across the season. Although, the 'repeated sprint decrement' (CV; 28%, MD; 60%) demonstrated a high degree of variability, was considered unreliable and therefore should be treated with caution by practitioners.

In the present study, all participants had been training within the academy on a full time basis for approximately 2 years and had completed the tests on 5 previous occasions as part of their habitual training regime. As a result, athletes' training background and degree of familiarisation might explain the high level of reproducibility observed. In this regard, a similar models could be implemented in other professional training environments, where the players have a similar training history and exposure to testing protocols (e.g., 'English premier league academies'). Whilst, the sample size used in this study could be considered as small (Hopkins et al. 2001) a football team typically comprises of 15-20 athletes. Therefore, is would be difficult to replicate this study with a larger group of athletes (>40) (i.e., and match all aspects of training and competition). Although, as a good level of reproducibility was evident, it suggests that increasing the sample size would not have influenced the results in a meaningful manner.

Conclusion

This study offers a unique insight into how fitness testing procedures could be organised across an 'in-season' micro-cycle in youth professional football. Given that good levels of reliability were evident, the arrangement of each test and the day-by-day training-load planned before and after each training session could be suggested as a guideline that practitioners could use when designing 'in-season' training and testing. Practitioners who choose to use the present training and testing paradigm should aim to implement them at pre-determined times-points of the season in line with the annual plan. Indeed, the addition of fitness testing to a typical training week would inevitably increase the total 'training-load' for that particular week. As such, if planned appropriately this could be used to help the coach manipulate the intensity and volume of the training week and thus could be implemented as a 'heavy' or 'high' training week which might precede a 'low' or 'moderate' training week. Therefore the timing of each testing week should be taken in the context of other training and competition goals and implemented with the specific aims of each individual.

REFERENCE LIST

Acero, R. M., Olmo, M. F., Sánchez, J. A., Otero, X. L., Aguado, X., & Rodríguez, F. A. (2011). Reliability of Squat and Countermovement Jump Tests in Children 6 to 8 Years of Age. *Pediatric Exercise Science*, 2011, 23, 151-160

Alricsson, M., Harms-Ringdahl, K. & Werner, S. (2001). Reliability of sports related functional tests with emphasis on speed and agility in young athletes. *Scandinavian journal of medicine & science in sports*, 11 (4), pp.229–232.

Atkinson, G. & Nevill, A.M., 1998. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26(4), pp.217–238.

Buchheit, M., Spencer, M. & Ahmaidi, S. (2010). Reliability, usefulness, and validity of a repeated sprint and jump ability test. *International journal of sports physiology and performance*, 5 (1), pp.3–17.

Cross, N. & Lyle, J. eds., (1999). The coaching process : principles and practice for sport, Oxford: Butterworth-Heinemann.

Duthie, G.M. Pyne D.B, Ross A.A, Livingstone S.G, Hooper S.L (2006). The reliability of ten-meter sprint time using different starting techniques. *Journal of Strength and Conditioning Research*, 20 (2), pp.246–251.

Ferrari Bravo D, Rampinini E, Sassi R, Bishop D, Sassi A, Tibaudi A, Impellizzeri FM. Ecological validity of a repeated sprint ability test and its reproducibility in soccer (2005). *European College of Sport Science – 10th Annual Congress. Belgrade, Serbia:* 5: 267

Hood, P.E., Barnes, C. & Gregson, W. (2002). Reliability of a battery of soccer-specific field tests. *Journal of Sports Sciences*, 20 (1), pp.20–21.

Hopker, J.G. Coleman, D.A, Wiles, J., Galbrait, A. (2009). Familiarisation and Reliability of Sprint Test Indices During Laboratory and Field Assessment. *Journal of Sports Science & Medicine*, 8 (4), pp.528–532.

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

Hopkins, W.G., Schabort, E.J. & Hawley, J.A. (2001). Reliability of power in physical performance tests. *Sports medicine* (Auckland, N.Z.), 31 (3), pp.211–234.

Hulse, M.A. Morris JG, Hawkins RD, Hodson A, Nevill AM, Nevill ME. (2013). A field-test battery for elite, young soccer players. *International Journal of Sports Medicine*, 34 (4), pp.302–311.

Impellizzeri, F.M., Rampinini E, Castagna C, Bishop D, Ferrari Bravo D, Tibaudi A, Wisloff U. (2008). Validity of a repeated-sprint test for football. *International Journal of Sports Medicine*, 29 (11), pp.899–905.

Krustrup, P. Mohr M, Nybo L, Jensen JM, Nielsen JJ, Bangsbo J. (2006). The Yo-Yo IR2 test: Physiological response, reliability, and application to elite soccer. *Medicine and Science in Sports and Exercise*, 38 (9), pp.1666–1673.

Little, T. & Williams, A.G., (2006). Effects of differential stretching protocols during warm-ups on high-speed motor capacities in professional soccer players. *Journal of strength and conditioning research*, 20 (1), pp.203–207.

Markovic, G. Dizdar D, Jukic I, Cardinale M. (2004). Reliability and factorial validity of squat and countermovement jump tests. *Journal of strength and conditioning research*, 18 (3), pp.551–555.

Maud, P. & Foster, C. eds., (2006). Physiological assessment of human fitness., IL: Human Kinetics.

McHugh, M. P. (2003). Recent advances in the understanding of the repeated bout effect: the protective effect against muscle damage from a single bout of eccentric exercise. Scandinavian Journal of Medicine & Science in Sports, 13(2), 88–97.

Mirkov, D. Nedeljkovic A, Kukolj M, Ugarkovic D, Jaric S. (2008). Evaluation of the reliability of soccer-specific field tests. *Journal of strength and conditioning research*, 22 (4), pp.1046–1050.

Oberacker, L.M. Davis SE, Haff GG, Witmer CA, Moir GL. (2012). The yo-yo IR2 test: physiological response, reliability, and application to elite soccer. *Journal of Strength and Conditioning Research*, 26 (10), pp.2734–2740.

Peterson, M.D. Pistilli E, Haff GG, Hoffman EP, Gordon PM. (2011). Progression of volume load and muscular adaptation during resistance exercise. *European Journal of Applied Physiology*, 111 (6), pp.1063–1071.

Rampinini, E. Alberti G, Fiorenza M, Riggio M, Sassi R, Borges TO, Coutts AJ. (2015). Accuracy of GPS Devices for Measuring High-intensity Running in Field-based Team Sports. *International Journal of Sports Medicine*, 36 (1), pp.49–53.

Rampinini, E. Bishop D, Marcora SM, Ferrari Bravo D, Sassi R, Impellizzeri FM (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 (3), pp.228–235.

Reilly, T. (2005). An ergonomics model of the soccer training process. *Journal of Sports Sciences*, 23 (6), pp.561–572.

Reilly, T. (2007). The science of training - soccer : a scientific approach to developing strength, speed and endurance, London: Routledge.

Reilly, T. Bangsbo, J. & Franks, A. (2000). Anthropometric and physiological predispositions for elite soccer. J Sports Sci, 18 (9), pp.669–683.

Schoenfeld, B. J., Ogborn, D., Contreras, B., Cappaert, T., Silva Ribeiro, A., Alvar, B. A., & Vigotsky, A. D. (2016). A Comparison of Increases in Volume Load Over 8 Weeks of Low-Versus High-Load Resistance Training. *Asian Journal of Sports Medicine*, *7*(2),

Semenick D. (1994) Testing protocols and procedures. In: Baechle T, editor. Essentials of strength training and conditioning. 1st ed. Champaign (IL): Human Kinetics. p. 258–73.

Seo, D.-I. Kim E, Fahs CA, Rossow L, Young K, Ferguson SL, Thiebaud R, Sherk VD, Loenneke JP, Kim D, Lee MK, Choi KH, Bemben DA, Bemben MG, So WY (2012). Reliability of the one-repetition maximum test based on muscle group and gender. *Journal of sports science & medicine*, 11 (2), pp.221–225.

Spencer, M., Fitzsimons 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(1–2), pp.181–184.

Tanner, R & Gore, C. (2013). Australian Institute of Sport. *Physiological tests for elite athletes*. Champaign, IL : Human Kinetics.

Turner, A., Walker, S., Stembridge, M., Coneyworth, P., Reed, G., Birdsey, L., Barter, P., Moody, J. (2011). A Testing Battery for the Assessment of Fitness in Soccer Players. *Strength and conditioning journal*, 33(5), pp.29–39.

Walker, S. & Turner, A., (2009). A One-Day Field Test Battery for the Assessment of Aerobic Capacity, Anaerobic Capacity, Speed, and Agility of Soccer Players. *Strength and Conditioning Journal*, 31 (6), pp.52–60.

Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of strength and conditioning research*, 19(1), pp.231–240.

Table 1: The organisation of physiological testing and habitual training/match-play performed during each experimental week.

Day	Morning schedule (~08:00 - 12:30)	Afternoon schedule (~14:00 - 16:00)
Saturday	U18 Premier League Match (K.O 11 am)	*Rest/travel*
Sunday	*Rest/day off*	*Rest/day off*
Monday	Rest / video analysis session	Speed & Agility Testing *followed by* Soccer-Specific Training
Tuesday	Yo-yo IR2 *followed by* Soccer-Specific Training	Resistance Training
Wednesday	Educational day	Educational day
Thursday	Vertical Jumps / Repeated Sprint Test *followed by* Soccer-Specific Training	1RM data collection followed by Upper Body Resistance Training
Friday	Soccer-Specific Training	*Rest*

Note: a 90-minute recovery phase was allocated between the morning and afternoon schedule from 12:30 to 14:00h

Table 2: Internal and external training load for each training session during experimental week 1 and 2

sRPE-TL (AU)	Week 1	Week 2	ES	*t*-test (*P*)
Match-play				
Saturday	405 ± 238	660 ± 191	0.50 (small)	0.08
Soccer-specific training				
Monday (MD+2)	183 ± 41	177 ± 41	0.15 (trivial)	0.61
Tuesday (MD-4)	634 ± 61	630 ± 65	0.07 (trivial)	0.83
Thursday (MD-2)	630 ± 65	630 ± 65	0.00 (trivial)	1
Friday (MD-1)	177 ± 41	171 ± 44	0.14 (trivial)	0.67
Resistance-training				
Tuesday	295 ± 35	295 ± 27	0.00 (trivial)	1
Thursday	435 ± 38	438 ± 39	-0.08 (trivial)	0.83
TD (m)				
Monday (MD+2)	2970 ± 448	2977 ± 336	-0.02 (trivial)	0.94
Tuesday (MD-4)	3946 ± 429	4119 ± 455	-0.39 (small)	0.27
Thursday (MD-2)	3334 ± 317	3354 ± 714	-0.04 (trivial)	0.91
Friday (MD-1)	2225 ± 352	2378 ± 313	-0.46 (small)	0.17
HS (m)				
Monday (MD+2)	49 ± 30	50 ± 26	-0.04 (trivial)	0.89
Tuesday (MD-4)	99 ± 63	109 ± 45	-0.18 (trivial)	0.61
Thursday (MD-2)	119 ± 54	111 ± 45	0.17 (trivial)	0.64
Friday (MD-1)	33 ± 28	35 ± 26	-0.07 (trivial)	0.83
SD (m)				
Monday (MD+2)	3 ± 5	5 ± 7	-0.19 (trivial)	0.56
Tuesday (MD-4)	20 ± 20	14 ± 11	0.35(small)	0.42
Thursday (MD-2)	11 ± 20	11 ± 13	0.03 (trivial)	0.93
Friday (MD-1)	5 ± 10	4 ± 6	0.14 (trivial)	0.68
Resistance-training VL [AU]				
Tuesday	3190 ± 571	3190 ± 571	0	0
Thursday	6475 ± 984	6750 ± 727	-0.31 (small)	0.67

Match-day (MD), Sessional rating of perceived exertion-training load (sRPE-TL), arbitrary units (AU), Total distance (TD), sprint distance (SD) (> 6.9m/s), high speed distance (HS) (>5.5 m/s), volume load (VL), Effect Size (ES).

Table 3: Reliability statistics for a battery of field tests performed around in-season habitual training in youth elite soccer players.

	Test	Re-test	Systematic bias (90% CI)	TE (90% CL)	CV (90%CL)	ES (rating)	SWC (%) (rating)	MD
Aerobic/Anaerobic related capacity								
Yo-yo IR2 (m)	920 ± 156	910 ± 158	-9.41 (-29.75-10-92)	34.0 (26.5-48.1)	4.2% (3.2-5.9)	0.06 (trivial)	31.0 (3.4) (marginal)	94.1 (10.3%)
RSA (best) (s)	4.33 ± 0.19	4.31 ± 0.18	-0.02 (-0.05-0.01)	0.05 (0.04-0.07)	1.2% (0.9-1.6)	0.11 (trivial)	0.04 (1%) (marginal)	0.14 (3.2%)
RSA (mean) (s)	4.58 ± 0.15	4.59 ± 0.17	0.00 (-0.02-0.02)	0.03 (0.02-0.05)	0.7% (0.6-1.0)	0.02 (trivial)	0.03 (1%) (ok)	0.09 (1.9%)
RSA (Dec) (%)	10.6 ± 4.3	8.6 ± 2.9	-2.0 (-3.2-0.7)	2.1 (1.6-2.9)	28% (21-42)	0.55 (moderate)	0.75 (7.8) (marginal)	5.76 (60%)
Speed, acceleration and agility								
5m (s)	1.01 ± 0.06	1.02 ± 0.05	0.01 (-0.01-0.03)	0.03 (0.03-0.05)	3.3% (2.5-4.7)	-0.12 (trivial)	0.01 (1%) (marginal)	0.09 (9%)
10m (s)	1.77 ± 0.07	1.76 ± 0.07	-0.01 (-0.02-0.01)	0.03 (0.02-0.04)	1.7% (1.3-2.5)	0.09 (trivial)	0.01 (1%) (marginal)	0.08 (4.8%)
20m (s)	3.09 ± 0.12	3.10 ± 0.11	0.01 (-0.02-0.03)	0.04 (0.03-0.06)	1.4% (1.1-2.0)	-0.06 (trivial)	0.02 (1%) (marginal)	0.12 (4%)
T-test (s)	9.17 ± 0.21	9.18 ± 0.14	0.01 (-0.03-0.05)	0.06 (0.05-0.09)	0.7% (0.05-1.0)	-0.07 (trivial)	0.04 (1%) (marginal)	0.18 (2%)
Vertical jump								
CMJ (cm)	40.8 ± 4.2	41.4 ± 3.5	0.61 (-0.41-1.62)	1.69 (1.32-2.40)	4.3% (3.3-6.1)	-0.16 (trivial)	0.76 (1.85%) (marginal)	4.7 (11.5%)
SJ (cm)	38.7 ± 4.3	39.0 ± 3.5	0.23 (-0.61-1.07)	1.41 (1.10-1.99)	3.7% (2.9-5.3)	-0.08 (trivial)	0.77 (2%) (marginal)	3.9 (10%)
Muscle strength								
1-RM half back squat (kg)	106.5 ± 14.7	107.7 ± 15.4	1.20 (-0.24-2.50)	2.0 (1.5-3.0)	1.8% (1.4-2.8)	-0.08 (trivial)	2.95 (2.8%) (ok)	5.52 (5.1%)

Key; Typical error of measurement (TE), Effect size (ES) smallest worthwhile change (SWC) minimal difference needed to be considered as "real" (MD) calculated for each assessment. TE < SWC; good; TE ~ SWC; OK; TE > SWC; marginal