Fitzpatrick, J., Hicks, K., Russell, M. and Hayes, P.R. (2019). The reliability of potential fatigue monitoring measures in elite youth soccer players. *The Journal of Strength and Conditioning Research*.

THE RELIABILITY OF POTENTIAL FATIGUE MONITORING MEASURES IN ELITE YOUTH SOCCER PLAYERS

Running Head: Reliability of Potential Fatigue Monitoring Measures

John F. Fitzpatrick[1,2], Kirsty M. Hicks[1], Mark Russell[3], Philip R. Hayes[1]

Author Affiliations:

[1]Department Sport, Exercise and Rehabilitation, Northumbria University, Newcastle-upon-Tyne, UK.

[2]Sports Science and Medical Department, Newcastle United Football Club, Newcastle-upon-Tyne, UK.

[3]School of Social and Health Sciences, Leeds Trinity University, UK.

Corresponding Author:

Mr John F. Fitzpatrick, Sports Science and Medical Department, Newcastle United Football Club, Greenlee Drive, Little Benton, Newcastle upon Tyne, NE7 7SF; Tel: 0191 2153200; Email: john.fitzpatrick@nufc.co.uk

THE RELIABILITY OF POTENTIAL FATIGUE MONITORING MEASURES IN ELITE YOUTH SOCCER PLAYERS

ABSTRACT

Monitoring fatigue is of vital importance to practitioners, however, logistics and concerns about reliability may impede the use of certain measures. This study aimed to quantify the reliability of potential measures of fatigue; a subjective wellness questionnaire, jump performance tests and tri-axial accelerometer variables derived during sub-maximal shuttle running in elite youth soccer players. A secondary aim was to establish the minimum test duration that could be used for the sub-maximal shuttle run while maintaining good reliability. Seventeen male youth team players (age: 17.4 ± 0.5 years) were assessed on two occasions, spaced seven days apart. Typical error (TE), coefficient of variation (CV%), interclass correlation (ICC) and minimum detectable change (MDC) were calculated for a subjective wellness questionnaire, countermovement jump (CMJ), squat jump (SJ) and drop jump contact time (DJ-CT), jump height (DJ-JH), and reactive strength (DJ-RSI). A novel sub-maximal shuttle running test was also used to assess tri-axial accelerometer data reliability. Results suggest that CMJ, SJ, DJ-CT and DJ-RSI have good test re-test reliability (CV% = 4.5 – 7.7; ICC = 0.80 – 0.88), however DJ-JH did not show acceptable reliability (CV% = 6.0; ICC = 0.76). Good reliability was found for all tri-axial accelerometer variables during a 3 min (2 min analysis) sub-maximal shuttle run (CV% = 2.4 – 8.0; ICC = 0.81 – 0.95), except for % PlayerLoad™ anterior–posterior (%PL_{AP}) (CV% = 7.2; ICC = 0.63). The subjective wellness questionnaire demonstrated poor reliability for all items (CV% = 11.2 – 30.0; ICC = 0.00 – 0.78). The findings from this study provide practitioners with valuable information about the reliability of a range of potential fatigue monitoring measures. This can be used to help make accurate decisions about the magnitude of change in these assessments when used in practice.

Key Words: reliability; subjective wellness; jump performance; tri-axial accelerometer; sub-maximal testing.

INTRODUCTION

Team sport activity has been shown to elicit fatigue commensurate with performance decrements and increased injury risk in youth and senior players (23, 25). Therefore, the ability to monitor and manage training and fatigue is of vital importance to coaches and practitioners (14). In an attempt to make informed decisions about readiness to train and training prescription, practitioners seek methods that attempt to quantify the magnitude of fatigue throughout the competitive week (1, 20). The broad use of the term fatigue within the literature presents a challenge as this can encompass several different phenomena that are the consequence of different physiological and perceptual processes (12). Practitioners have therefore used a number of methods in an attempt to monitor fatigue in an applied setting; self-report measures, autonomic nervous system function, physical performance tests and biochemical markers to name a few (14, 30). However, logistical feasibility and concerns about the reliability may impede the use of such methods on a regular basis (1).

An important factor to consider when selecting a potential monitoring tool is measurement reliability. The reliability of a test refers to an acceptable level of consistency between repeated tests within a practically relevant timeframe (2). A test with poor reliability will be unsuitable for tracking changes in fatigue due to an inability to detect a true change in the measure (16). Factors that influence reliability include the protocol, measurement device used to collect the data and any systematic or random changes in the mental or physical state of the individual between trials (2).

Self-report measures are widely used in team sports (29), however there has yet to be a consensus on the most appropriate questionnaire to be used. Profile of Mood States (22), Recovery-Stress Questionnaire (18) and Daily Analyses of Life Demands (27) are just some of the assessment tools which have been used within the literature. However, their length, narrow focus or lack of specificity to the sporting context has led many sports programs to develop their own questionnaires (28). Subsequently, practitioners and researchers have incorporated customised, shortened questionnaires into their monitoring practices and research (15, 21), although, reliability and sensitivity of these shortened wellness questionnaires has yet to be established. By contrast, tests of jump performance are well established and demonstrate good reliability, with reported coefficients of variations (CV%) of 5% for the countermovement jump (CMJ) (10), 3% for the squat jump (SJ) (13) and 5-8% for variables derived from a drop jump (DJ) (3).

A survey showed that 61% of elite European soccer teams regularly use a sub-maximal, non-exhaustive performance test to assess autonomic function (1). However, research suggests that due to the variability of heart rate measures this approach offers limited meaningful information (19). Notwithstanding issues with reliability, validity and sensitivity, a possible solution maybe to utilise the sub-maximal performance tests that are already widely used in practice, by analysing other data streams that can be collected during this assessment. The use of tri-axial accelerometers, such as those integrated into micro-electro-mechanical systems (MEMS), have demonstrated an ability to detect fatigue post exercise (24), with vertical acceleration showing changes under fatigue during both match-play (9) and training (26). More recently, Buchheit et al. (7) assessed the reliability of stride variables derived from MEMS devices during treadmill running. They found that measures of contact time, fight time and vertical stiffness have good to moderate reliability (4-16% CV). These data give preliminary insights into the ability of accelerometer data to monitor fatigue, however the reliability of measures derived from MEMS devices during sub-maximal field tests has yet to be established.

Establishing the reliability of measures used to monitor athletes is an imperative aspect of applied research and practice. Further, field-based, in situ reliability assessments are required in order to quantify “real” changes in potential monitoring tools in athletes within their normal training environment and across time periods that are typically used to quantify the effects of any intervention (2). Therefore, the aim of this study was to quantify the test re-test reliability of a subjective wellness questionnaire, assessments of jump performance and tri-axial accelerometer variables derived during sub-maximal shuttle running in elite youth soccer players. A secondary aim is to establish the minimum test duration that can be used for the sub-maximal shuttle run in order to maintain good reliability.

METHODS

Experimental Approach to the Problem

This study was completed at the beginning of the 2015-16 season (October) and consisted of two testing sessions spaced seven days apart. All players were in full training during the study, completing around 10.5 h per week of pitch (8.5 h) and gym (2 h) based activity. All players were familiarized with the experimental procedures prior to commencing the study. Each testing session consisted of morning ratings of subjective wellness (n = 17), three different assessments of jump performance; CMJ, SJ and DJ (n = 17) and a sub-maximal shuttle run test used to assess accelerometer variables (n = 15), in respective order. Training load was monitored carefully throughout the study period ensure limited differences between training weeks (sRPE; Week 1 = 2247, Week 2 = 2280, CV% = 9.9%). Both testing sessions were preceded by 48 h of rest and were conducted at the same time of day to limit the influence of possible circadian variation.

Subjects

Seventeen youth soccer players (Age: 17.4 ± 0.5 years [range: 16-18 years], Height: 176.7 ± 5.2 cm, Body Mass: 72.1 ± 9.2 kg), competing in the English Under-18 Premier League agreed to participate in the present study. Participants were given full details of the study procedures and informed of the risks and benefits of the study prior to any data collection. Participants provided personal, and when under 18, parental or guardian, written informed consent before participation. Institutional ethical approval was gained prior to any study involvement. Prior to inclusion in the study, all participants were deemed fit and free of illness or injury by the soccer club's medical staff.

Procedures

Subjective wellness

A psychometric questionnaire based on previous recommendations was collected each day to assess general indicators of player wellness (15). Participants recorded their scores each morning, in private, on an electronic device using a custom made application, as soon as they entered the training ground. The questionnaire was composed of 5 questions relating to fatigue, sleep quality, muscle soreness, stress and mood. Each question was scored on a 5-point Likert scale with 1-point increments (scores of 1–5, with 1 and 5 representing very poor and very good, respectively) (21). Additionally, the summation of all 5 scores provided a total wellness score between 5 - 25.

Jump performance

A standardised warm up consisting of three minutes light aerobic activity on a cycle ergometer at a self-selected pace (Keiser, Fresno, CA, USA), followed by dynamic mobility exercises and three submaximal practice jumps was conducted prior to each testing session. Players then performed three different tests to assess jump performance; CMJ, SJ and DJ. The CMJ was executed to a self-selected depth with the hands placed on the hips. Players were instructed to jump as high as possible with no knee or hip flexion during the flight phase. The same instructions were given for the SJ however, players were instructed to hold their self-selected depth for a four second count. The DJ was performed from a 30 cm box with hands on their hips. Players were instructed to step off the box, rebound off the floor as quickly as possible and jump as high as possible. All participants were well drilled and familiarised with each test. Each assessment consisted of four attempts, separated by one minute of rest. Jumps were performed in a randomized counterbalanced manner to reduce order effects. All jumps were completed using an optical timing system (Optojump, Microgate, Italy), the validity of which has been previously established (13). Jump height was recorded for CMJ and SJ, while for DJ, contact time (DJ-CT), jump height (DJ-JH) and reactive strength index (DJ-RSI) were recorded.

Sub-maximal shuttle running

A sub-maximal shuttle running test was used to assess players' mechanical loading. All players were fitted with a MEMS device (MinimaxX S4, Catapult Sports, Melbourne, Australia) worn between the scapular in a tight-fitting vest to reduce movement artefact. Devices contained a tri-axial piezoelectric linear accelerometer (Kionix: KXP94) sampling at a frequency of 100 Hz. Following this, a continuous 20 m shuttle run was performed for a 5 min period, at an average speed of 12 $km \cdot h^{-1}$, on an artificial 3G surface. Pacing was controlled using a custom audio track played over a loudspeaker. Data were downloaded using the manufacturer's software (Catapult Sprint, Version 5.1.7) and raw data were exported to Microsoft Excel. The first minute of data was discarded as a stabilization period, the subsequent 2, 3 and 4 minutes of the collection period were used for statistical analysis. Combined tri-axial accelerometer data were presented as PlayerLoad (PL), which is a modified vector magnitude expressed as the square root of the sum of the squared instantaneous rates of change in acceleration in each of the three planes divided by 100 (5). Individual component planes of PL, anterior-posterior [PL_{AP}], mediolateral [PL_{ML}], and vertical [PL_{V}] were recorded and expressed in arbitrary units (au). The percentage contribution of each component plane to overall PL was also calculated.

Statistical Analysis

Test re-test reliability was calculated for each variable and expressed as a typical error (TE), coefficient of variation (CV%) and interclass correlation (ICC), and calculated using a custom spreadsheet (17). To assess reliability, thresholds of $\leq$ 10% for CV% and $\geq$ 0.80 for ICC were set (16). To assess the ability of each variable to assess "real" change the minimum detectable change (MDC; 75% confidence level) was also calculated (31). To evaluate the internal consistency of the subjective wellness questionnaire Cronbach's Alpha (α) was calculated (11), with a threshold of >0.7 being set for an acceptable α (4) and inter-item correlations also considered.

RESULTS

Subjective Wellness

Reliability statistics for each subjective wellness measure are shown in Table 1. The TE for individual subjective wellness questions ranged from 0.30 to 0.60 and was 1.59 for total wellness. When expressed as a CV% values ranged from 11.2% to 30.0%. Therefore no subjective wellness measures met the threshold for acceptable reliability of ≤ 10%. Similarly, ICCs for subjective wellness measures ranged from -0.01 to 0.78 meaning no measures met the threshold for an acceptable ICC of ≥ 0.80. The MDC for each subjective wellness measure is displayed in Table 1. As an additional measure of reliability for the subjective wellness questionnaire, Cronbach's Alpha was assessed, this analysis resulted in $\alpha = 0.45$ meaning the internal consistency of the 5 items within this subjective wellness questionnaire is poor. Further analysis into the inter-item correlations is shown in Table 2.

INSERT TABLES 1 AND 2 AROUND HERE

Jump Performance

Reliability statistics for each jump test are shown in Table 3. The TE for CMJ, SJ and DJ-JH was 1.6 cm, 1.5 cm and 1.8 cm respectively. When expressed as a CV% these values were 4.8%, 4.5% and 6.0% respectively. The TE and CV% for DJ-CT was 0.01 s and 4.9% and for DJ-RSI was 0.11 and 7.7%. Therefore all measures derived from jump assessments met the threshold for acceptable reliability of ≤ 10%. Similarly, ICCs for the jump assessments ranged from 0.76 to 0.88 meaning all measures expect for DJ-H (0.76) met the threshold for an acceptable ICC of ≥ 0.80. The MDC for each jump test is displayed in Table 3.

INSERT TABLE 3 AROUND HERE

Sub-Maximal Shuttle Running

A summary of reliability for the accelerometer data from a sub-maximal shuttle run are shown in Table 4. Reliability was consistent across all analysis time frames with all measures meeting the threshold of ≤ 10% for an acceptable CV%. Similarly, ICCs for each accelerometer variable across all time frames ranged from 0.63 to 0.96 meaning that all variables expect for % PL_{AP} (0.63) met the threshold for an acceptable ICC of ≥ 0.80. The MDC for each accelerometer variable is displayed in Table 4.

INSERT TABLE 4 AROUND HERE

DISCUSSION

The aim of the present study was to quantify the test re-test reliability of a range of potential fatigue monitoring tools in a group of elite youth soccer players. The key findings from this study were: (1) a short 3 minute sub-maximal shuttle run from which tri-axial accelerometer data was collected showed good reliability, (2) good reliability was shown for a number of jump tests, (3) a psychometric questionnaire used to assess subjective wellness showed poor reliability.

A novel aspect of the present study was the use of a sub-maximal test to assess accelerometer data gained during shuttle running. All variables displayed good reliability across all time frames, expect for % PL_{AP} (ICC = 0.63-0.75). These result are in accordance with recent research which assessed the validity and reliability of measures of vertical stiffness and peak loading forces collected via GPS-embedded accelerometers during treadmill running, and found good to moderate reliability (4-16% CV) (8). However, the measures used in the aforementioned study require specialist software for analysis that utilizes proprietary detection algorithms to recognize foot strikes. These variables are not widely available to practitioners using manufacturer software, therefore the data presented in the present study may provide a more practical alternative. This is the first study to assess these variables during a sub-maximal field test, something that is essential when performing this test in an applied environment. A secondary aim of this study was to assess the minimum test duration that could maintain good reliability. A three minute assessment, during which the first minute is discarded as a stabilization period and the final two minutes used for analysis, is shown to be an acceptable time frame. This provides a key finding for practitioners who may want to assess a large squad of players simultaneously in a small time period, for example as part of a pre training warm-up.

The assessment of fatigue via jump testing represents a popular method by which to monitor neuromuscular function in the field and has been shown to be associated with objective measures of peripheral fatigue (6). Indeed, a survey indicated that 39% of a sample of 41 elite European soccer teams utilize a form of jump testing on a weekly basis (1). Despite this widespread use, there are only a limited number of studies that have evaluated the test re-test reliability of various jump tests. Findings from the present study indicate good reliability for all jump measures apart from DJ-JH (ICC = 0.76), which did not meet the criteria from an acceptable ICC (≥ 0.80). These results are in agreement with previous observations that have reported CV% of 5% for the countermovement jump (CMJ) (10), 3% for the squat jump (SJ) (13) and 5-8% for variables derived from a drop jump (13). An important consideration when reviewing the literature on reliability of jump assessments is the range of testing modalities used (contact mats, force platforms and photoelectric technology). The present study used the OptoJump system which has previously shown good reliability and validity when assessing CMJ and SJ (13).

In the present study all individual subjective wellness items and overall total wellness did not meet the criteria for acceptable reliability. These findings suggest that the current subjective wellness questionnaire is not reliable enough to track changes in fatigue from a week to week basis. To our knowledge, this is the first study to examine the reliability of this type of short psychometric questionnaire that is regularly used in the applied environment. Other, larger questionnaires such as the Recovery-Stress questionnaire (76-questions) have shown large test re-test correlations (r = 0.79) (18). The factor/s mediating the poor reliability found in the present study may be due to the simplicity of the questionnaire used. The categorical nature of a 5-point Likert style question means a 1 point change, e.g. from 5 to 4, is the equivalent to a 20% decrease. This makes the suggested criteria of a CV% ≤ 10% difficult to meet.

Another aspect of the reliability of a psychometric questionnaire is the internal consistency of each question, this was assessed via Cronbach’s α, with results indicating that the internal consistency of this questionnaire is poor (α = 0.45). This has implications for the composite total wellness score which is the summation of all five questions. As each item has relatively low inter-item correlations (Table 2) it could be suggested that a composite score for total wellness should be used with caution. Future research should look to amend which items are included in the composite score in order to improve the internal consistency. In conclusion, given the high CV% and low ICC of each variable, and the poor internal consistency of the composite total wellness score, in order to make this subjective wellness questionnaire more reliable and robust in an applied environment the low categorical nature of the Likert scale should be addressed, perhaps by increasing the number of point within the scale to > 5. Additionally the internal consistency of the 5 items needs to be improved in order to make total wellness a viable measure of fatigue.

In conclusion, this study has established good reliability for CMJ, SJ and DJ variables. Tri-axial accelerometer data; PL, PL_{AP}, PL_{ML}, PL_{V}, $\%PL_{ML}$ and $\%PL_{V}$, gained during sub-maximal shuttle running also displayed good reliability. However, results suggest that subjective wellness assessed via a short 5 item psychometric questionnaire has poor test re-test reliability and internal consistency, therefore, caution must be taken when assessing changes in subjective wellness as an indicator of a players fatigue status. These findings suggest that measures with good week-to-week test re-test reliability may provide the greatest potential as markers of fatigue in an applied environment, with future research looking to establish the sensitivity of these measure to fatigue.

PRACTICAL APPLICATIONS

If looking to assess levels of fatigue in youth soccer players, all accelerometer variables, expect for % PL_{AP}, display good test re-test reliability from 3 minutes (2-minutes analysis) of sub-maximal shuttle running, suggesting this may be a novel way of assessing the fatigue levels of a large group of players in a short amount of time. Assessments of CMJ, SJ, DJ-CT and DJ-RSI have displayed good reliability on a week-to-week basis and therefore may provide sound estimates of a player's physical performance level. Finally, the MDC calculated for each variable provides researchers and practitioners with thresholds for what may be considered a "real" change, allowing practitioners to make accurate decisions about the magnitude of fatigue.

AKNOWLEDGEMENTS

The authors would like to thank Craig Musham, Adam Rowan and James Newton for their help and expertise during the testing procedures.

Table 1.
Reliability statistics for subjective wellness measures.
Data are presented as group means (± SD) for each trial, typical error (TE), interclass correlation (ICC), coefficient of variation (CV%), and minimum detectable change (MDC) (75% confidence level).

	Trial 1 (SD)	**Trial 2 (SD)**	**TE (90% CI)**	**ICC (90% CI)**	**CV% (90% CI)**	**MDC**
Fatigue (au)	3.2 (0.5)	3.1 (0.7)	0.3 *(0.2, 0.4)*	0.78 *(0.56, 0.90)*	14.9 *(11.4, 21.8)*	0.5
Sleep Quality (au)	3.8 (0.5)	3.6 (0.6)	0.6 *(0.5, 0.8)*	-0.01 *(-0.41, 0.39)*	21.0 *(16.1, 31.1)*	0.9
Muscle Soreness (au)	3.0 (0.9)	2.7 (0.8)	0.6 *(0.5, 0.9)*	0.54 *(0.17, 0.77)*	30.0 *(22.7, 45.1)*	0.9
Stress (au)	3.4 (0.7)	3.1 (0.9)	0.6 *(0.5, 0.8)*	0.46 *(0.08, 0.73)*	22.7 *(17.3, 33.6)*	0.9
Mood (au)	3.9 (0.6)	3.9 (0.7)	0.6 *(0.5, 0.8)*	0.15 *(-0.27, 0.52)*	19.2 *(14.7, 28.2)*	0.9
Total Wellness (au)	17.3 (1.9)	16.5 (2.0)	1.6 *(1.2, 2.3)*	0.35 *(-0.06, 0.66)*	11.2 *(8.6, 16.2)*	2.5

An acceptable threshold for reliability was set at 0.80 for ICC and 10% for CV%

* ICC ≥ 0.80
** CV% ≤ 10%

Table 2.
Subjective wellness inter-item correlation matrix

	Fatigue	Sleep	Soreness	Stress	Mood
Fatigue	-	0.12	0.25	0.16	0.04
Sleep	-	-	0.13	0.35	0.18
Soreness	-	-	-	-0.10	-0.12
Stress	-	-	-	-	0.70
Mood	-	-	-	-	-

Table 3.
Reliability statistics for jump tests; countermovement jump height (CMJ), squat jump height (SJ), drop jump contact time (DJ-CT) drop jump height (DJ-JH) and drop jump reactive strength index (DJ-RSI).
Data are presented as group means (± SD) for each trial, a typical error (TE), interclass correlation (ICC), coefficient of variation (CV%), and minimum detectable change (MDC) (75% confidence level).

	Trial 1 (SD)	**Trial 2 (SD)**	**TE (90% CI)**	**ICC (90% CI)**	**CV% (90% CI)**	**MDC**
CMJ (cm)	35.3 (4.5)	35.0 (4.0)	1.6 *(1.3, 2.3)*	0.88 * *(0.73, 0.94)*	4.8 ** *(3.7, 6.9)*	2.5
SJ (cm)	34.7 (4.3)	34.4 (3.8)	1.5 *(1.2, 2.1)*	0.88 * *(0.75, 0.95)*	4.5 ** *(3.5, 6.4)*	2.3
DJ-CT (s)	0.20 (0.02)	0.21 (0.02)	0.01 *(0.01, 0.01)*	0.85 * *(0.69, 0.93)*	4.9 ** *(3.8, 7.0)*	0.01
DJ-JH (cm)	29.8 (3.5)	29.9 (3.4)	1.8 *(1.4, 2.5)*	0.76 *(0.51, 0.89)*	6.0 ** *(4.6, 8.6)*	2.8
DJ-RSI ($m.s^{-1}$)	1.49 (0.23)	1.43 (0.23)	0.11 *(0.08, 0.15)*	0.80 * *(0.59, 0.91)*	7.7 ** *(6.0, 11.1)*	0.16

An acceptable threshold for reliability was set at 0.80 for ICC and 10% for CV%

* ICC ≥ 0.80
** CV% ≤ 10%

Table 4.
Summary of reliability statistics for 2, 3 and 4 minutes for PlayerLoad (PL), individual component planes; anterior-posterior (PL_{AP}), mediolateral (PL_{ML}), and vertical (PL_V), and the % contribution of each plane. Data are presented as group means (± SD) for each trial, typical error (TE), interclass correlation (ICC), coefficient of variation (CV%), and minimum detectable change (MDC) (75% confidence level).

		Trial 1 (SD)	**Trial 2 (SD)**	**TE (90% CI)**	**ICC (90% CI)**	**CV% (90% CI)**	**MDC**
2 min	PL	39.6 (3.8)	39.5 (3.6)	0.9 *(0.7, 1.4)*	0.95 * *(0.87, 0.98)*	2.4 ** *(1.8, 3.5)*	1.4
	PL_{AP}	14.8 (2.3)	14.5 (2.5)	1.1 *(0.9, 1.6)*	0.81 * *(0.59, 0.92)*	8.0 ** *(6.1, 11.8)*	1.7
	PL_{ML}	14.4 (2.0)	13.9 (1.9)	0.7 *(0.5, 1.0)*	0.89 * *(0.74, 0.95)*	5.7 ** *(4.4, 8.5)*	1.1
	PL_V	28.1 (3.1)	28.1 (3.0)	0.9 *(0.7, 1.3)*	0.93 * *(0.84, 0.97)*	3.2 ** *(2.5, 4.8)*	1.3
	% PL_{AP}	25.8% (3.0%)	25.6% (2.7%)	1.8% *(1.4%, 2.6%)*	0.63 *(0.28, 0.83)*	7.2 ** *(5.5, 10.6)*	2.8%
	% PL_{ML}	25.2% (2.8%)	24.6% (3.1%)	1.3% *(1.0%, 1.9%)*	0.83 * *(0.63, 0.93)*	5.9 ** *(4.5, 8.7)*	2.0%
	% PL_V	49.0% (2.8%)	49.8% (2.9%)	1.3% *(1.0%, 1.9%)*	0.83 * *(0.62, 0.93)*	2.7 ** *(2.0, 3.9)*	2.0%
3 min	PL	59.3 (5.8)	59.1 (5.4)	1.5 *(1.1, 2.1)*	0.94 * *(0.86, 0.98)*	2.5 ** *(1.9, 3.7)*	2.2
	PL_{AP}	22.3 (3.2)	21.9 (3.4)	1.4 *(1.1, 2.0)*	0.85 * *(0.66, 0.94)*	6.3 ** *(4.8, 9.3)*	2.1
	PL_{ML}	21.7 (3.1)	21.0 (2.8)	0.9 *(0.7, 1.3)*	0.93 * *(0.82, 0.97)*	4.3 ** *(3.3, 6.4)*	1.3
	PL_V	42.3 (4.8)	42.5 (4.6)	1.5 *(1.2, 2.3)*	0.91 * *(0.79, 0.96)*	3.9 ** *(2.9, 5.7)*	2.3
	% PL_{AP}	25.8% (2.8%)	25.5% (2.3%)	1.4% *(1.1%, 2.0%)*	0.73 *(0.45, 0.88)*	5.4 ** *(4.2, 8.0)*	2.1%
	% PL_{ML}	25.2% (2.8%)	24.7% (3.1%)	1.0% *(0.8%, 1.4%)*	0.91 * *(0.78, 0.96)*	4.3 ** *(3.3, 6.3)*	1.5%
	% PL_V	49.0% (2.7%)	49.8% (2.9%)	1.1% *(0.9%, 1.7%)*	0.85 * *(0.67, 0.94)*	2.4 ** *(1.8, 3.5)*	1.7%
4 min	PL	79.4 (7.9)	78.9 (7.2)	1.7 *(1.3, 2.4)*	0.96 * *(0.90, 0.98)*	2.1 ** *(1.6, 3.1)*	2.5
	PL_{AP}	30.1 (4.3)	29.4 (4.6)	1.7 *(1.3, 2.5)*	0.87 * *(0.71, 0.95)*	5.9 ** *(4.5, 8.7)*	2.6
	PL_{ML}	29.0 (4.0)	28.1 (3.8)	1.2 *(0.9, 1.8)*	0.92 * *(0.80, 0.96)*	4.8 ** *(3.6, 7.0)*	1.8
	PL_V	56.5 (6.4)	57.0 (6.2)	1.9 *(1.4, 2.7)*	0.93 * *(0.83, 0.97)*	3.4 ** *(2.6, 5.0)*	2.8
	% PL_{AP}	26.0% (2.7%)	25.6% (2.4%)	1.3% *(1.0%, 2.0%)*	0.75 *(0.48, 0.89)*	5.3 ** *(4.1, 7.9)*	2.1%
	% PL_{ML}	25.2% (2.8%)	24.6% (3.1%)	1.1% *(0.8%, 1.6%)*	0.88 * *(0.73, 0.95)*	4.9 ** *(3.7, 7.2)*	1.7%
	% PL_V	48.8% (2.8%)	49.8% (3.0%)	1.2% *(0.9%, 1.8%)*	0.85 * *(0.67, 0.94)*	2.5 ** *(1.9, 3.7)*	1.8%

An acceptable threshold for reliability was set at 0.80 for ICC and 10% for CV%

* ICC ≥ 0.80
** CV% ≤ 10%

1

References

1. Akenhead R and Nassis GP. Training load and player monitoring in high-level football: current practice and perceptions. *Int J Sports Physiol Perform* 11: 587-593, 2016.
2. Atkinson G and Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 26: 217-238, 1998.
3. Beattie K and Flanagan EP. Establishing the reliability & meaningful change of the drop-jump reactive strength index. *J Aust Strength Cond* 23: 12-18, 2015.
4. Bland JM and Altman DG. Statistics notes: Cronbach's alpha. *Br Med J* 314: 572, 1997.
5. Boyd LJ, Ball K, and Aughey RJ. The reliability of MinimaxX accelerometers for measuring physical activity in Australian football. *Int J Sports Physiol Perform* 6: 311-321, 2011.
6. Brownstein CG, Dent JP, Parker P, Hicks KM, Howatson G, Goodall S, and Thomas K. Etiology and recovery of neuromuscular fatigue following competitive soccer match-play. *Front Physiol* 8: 831, 2017.
7. Buchheit M, Gray A, and Morin J-B. Assessing stride variables and vertical stiffness with GPS-embedded accelerometers: preliminary insights for the monitoring of neuromuscular fatigue on the field. *Journal of Sports Science and Medicine* 14: 698, 2015.
8. Buchheit M, Gray A, and Morin J. Assessing stride variables and vertical stiffness with GPS-embedded accelerometers: preliminary insights for the monitoring of neuromuscular fatigue on the field. *J Sports Sci Med* 14: 698, 2015.
9. Cormack SJ, Mooney MG, Morgan W, and McGuigan MR. Influence of neuromuscular fatigue on accelerometer load in elite Australian football players. *Int J Sports Physiol Perform* 8: 373-378, 2013.
10. Cormack SJ, Newton RU, McGuigan MR, and Doyle TL. Reliability of measures obtained during single and repeated countermovement jumps. *Int J Sports Physiol Perform* 3: 131-144, 2008.
11. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 16: 297-334, 1951.
12. Enoka RM and Duchateau J. Muscle fatigue: what, why and how it influences muscle function. *The Journal of physiology* 586: 11-23, 2008.
13. Glatthorn JF, Gouge S, Nussbaumer S, Stauffacher S, Impellizzeri FM, and Maffiuletti NA. Validity and reliability of Optojump photoelectric cells for estimating vertical jump height. *J Strength Cond Res* 25: 556-560, 2011.
14. Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med* 44: 139-147, 2014.
15. Hooper SL and Mackinnon LT. Monitoring overtraining in athletes. *Sports Med* 20: 321-327, 1995.
16. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 30: 1-15, 2000.
17. (sportsci.org/2015/ValidRely.htm).
18. Kellmann M and Kallus KW. *Recovery-stress questionnaire for athletes: User manual.* Human Kinetics, 2001.
19. Lamberts R, Swart J, Capostagno B, Noakes T, and Lambert M. Heart rate recovery as a guide to monitor fatigue and predict changes in performance parameters. *Scand J Med Sci Sports* 20: 449-457, 2010.
20. McCall A, Carling C, Nedelec M, Davison M, Le Gall F, Berthoin S, and Dupont G. Risk factors, testing and preventative strategies for non-contact injuries in professional football: current perceptions and practices of 44 teams from various premier leagues. *Br J Sports Med* 48: 1352-1357, 2014.

21. McLean BD, Coutts AJ, Kelly V, McGuigan MR, and Cormack SJ. Neuromuscular, endocrine, and perceptual fatigue responses during different length between-match microcycles in professional rugby league players. *Int J Sports Physiol Perform* 5: 367-383, 2010.
22. McNair D, Lorr M, and Droppleman L. Profile of mood states: EdITS manual. *San Diego, CA: Educational and Industrial Testing Service*, 1981.
23. Mohr M, Krustrup P, and Bangsbo J. Fatigue in soccer: a brief review. *J Sports Sci* 23: 593-599, 2005.
24. Patterson M, McGrath D, and Caulfield B. Using a tri-axial accelerometer to detect technique breakdown due to fatigue in distance runners: a preliminary perspective. Presented at Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, 2011.
25. Rampinini E, Bosio A, Ferraresi I, Petruolo A, Morelli A, and Sassi A. Match-related fatigue in soccer players. *Med Sci Sports Exerc* 43: 2161-2170, 2011.
26. Rowell AE, Aughey RJ, Clubb J, and Cormack SJ. A Standardised Small Sided Game Can Be Used to Monitor Neuromuscular Fatigue in Professional A-League Football Players. *Front Physiol* 9: 1011, 2018.
27. Rushall BS. A tool for measuring stress tolerance in elite athletes. *J Appl Sport Psychol* 2: 51-66, 1990.
28. Saw AE, Main LC, and Gastin PB. Monitoring athletes through self-report: factors influencing implementation. *J Sports Sci Med* 14: 137, 2015.
29. Taylor K, Chapman D, Cronin J, Newton MJ, and Gill N. Fatigue monitoring in high performance sport: a survey of current trends. *J Aust Strength Cond* 20: 12-23, 2012.
30. Twist C and Highton J. Monitoring fatigue and recovery in rugby league players. *Int J Sports Physiol Perform* 8: 467-474, 2013.
31. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19: 231-240, 2005.