This is an Accepted Manuscript of an article published by Taylor & Francis Group in *European Journal of Sport Science* on 27 Dec 2018, available online: http://www.tandfonline.com/10.1080/17461391.2018.1556739.

James H. Dugdale, Calum A. Arthur, Dajo Sanders & Angus M. Hunter (2018) Reliability and validity of field-based fitness tests in youth soccer players, *European Journal of Sport Science*, DOI: 10.1080/17461391.2018.1556739

Abstract

This study aimed to establish between-day reliability and validity of commonly used field-based fitness tests in youth soccer players of varied age and playing standards, and to discriminate between players without ("unidentified") or with ("identified") a direct route to professional football through their existing club pathway. Three-hundred-and-seventy-three Scottish youth soccer players (U11–U17) from 3 different playing standards (amateur, development, performance) completed a battery of commonly used generic field-based fitness tests (grip dynamometry, standing broad jump, countermovement vertical jump, 505 (505COD) and T-Drill (T-Test) change of direction, and 10/20m sprint tests) on two separate occasions within 7–14 days. The majority of field-based fitness tests selected within this study proved to be reliable measures of physical performance (ICC=0.83-0.97; p<0.01). However, COD tests showed weaker reliability in younger participants (ICC=0.57-0.79; p<0.01). The field-based fitness testing battery significantly discriminated between the unidentified and identified players; χ^2 (7)=101.646, p<0.001, with 70.2% of players being correctly classified. We have shown field-based fitness tests to be reliable measures of physical performance in youth soccer players. However, results from the 505COD and T-Test change of direction tests may be more variable in younger players, potentially due to complex demands of these tests and the limited training age established by these players. Whilst the testing battery selected in this study was able to discriminate between unidentified and identified players, findings were inconsistent when attempting to differentiate between individual playing standards within the "identified" player group (development vs. performance).

Key Words: Selection, Performance, Discriminate, Profiling, Adolescent, Development

Introduction

Soccer is an intermittent, high intensity sport requiring a broad range of physical abilities in order to achieve competitive success (Stolen et al., 2005). Given the global popularity of competitive soccer and its vast participation levels at grass-roots, there has been a rapid increase in the interest and importance placed on an ability to examine and differentiate between varied competitive standards of youth soccer players (Unnithan et al., 2012). Whilst governing bodies adopt multiple approaches in categorising performance standards specific to their varied infrastructures, it is commonly accepted that the selection/deselection or talent identification process relies on players being signed to a professional soccer academy (identified) comparative to those are not (unidentified) (Unnithan et al., 2012; Murr, Raabe, & Höner, 2018). Historically, selection/deselection within youth soccer employed scouting systems reliant on individual opinion and philosophy (Reilly, Williams, Nevill, & Franks, 2000; Unnithan et al., 2012). However, in recent years this process has been scrutinised due to its subjective nature and potential bias with calls for a more scientific approach (Larkin & O'Connor, 2017; O'Connor, Larkin, & Williams, 2016). Whilst many scholars discuss the multi-dimensional and complex nature associated with assessing ability within youth soccer (Larkin & O'Connor, 2017; Reeves et al., 2018; Sarmento et al., 2018), including measures of physical fitness as a component of assessment and talent identification processes remains prevalent within current selection/deselection processes (Gil et al., 2014; Gonaus & Müller, 2012; Huijgen et al., 2014; Murr, Raabe, & Höner, 2018).

Field-based fitness tests have been widely utilised by practitioners to assess and monitor performance characteristics in soccer players (Deprez et al., 2015; Gil et al., 2014; Huijgen et al., 2014; le Gall et al., 2010). Field-based fitness tests allow for assessing multiple individuals simultaneously, generally requiring low cost equipment and easy accessibility for practitioners and researchers (Hulse et al., 2013; Paul & Nassis, 2015; Pyne, Spencer, & Mujika, 2014). Additionally, a comprehensive field-based testing battery relevant to the multiple physical demands of soccer can generally be conducted within a single session (Hulse et al., 2013; Pyne et al., 2014; Vescovi et al., 2011), therefore proving extremely time effective and practical amongst practitioners. Typically, assessments of aerobic fitness, repeated sprint ability, change of direction, and linear sprinting have been carried out when assessing youth soccer players (Paul & Nassis, 2015). However, due to physical advancements present within modern-day competitive soccer, attributes of explosive power and muscular strength are receiving substantial interest within recent research (Gouvêa et al., 2017; Murr et al., 2018).

A plethora of research has been conducted on the suitability and application of field-based fitness tests to the sport of soccer, however reliability and validity of these measures are mainly demonstrated in senior athletes (Paul & Nassis, 2015). Studies attempting to examine reliability and validity of field-based fitness tests in youth soccer samples have either used limited testing batteries (Rebelo et al., 2013; Thomas, et al., 2009), demonstrated ability to discriminate between players of the same competitive standard (Hulse et al., 2013), or evaluated test reliability for restricted age ranges (Rebelo et al., 2013; Thomas et al., 2009). One of the attractions of a valid fitness testing battery is the potential ability to discriminate between various performance standards

(Murr et al., 2018), however this is still to be conclusively established within a youth sample. Prior research examining performance characteristics of youth soccer players has focussed on senior academy players (U16-U19) (le Gall et al., 2010; Mujika et al., 2009), yet many professional soccer academies register players as young as U9 (Hulse et al., 2013). Due to the influence of physical competency and relative training age on reliability of performance testing (Vandorpe et al., 2012; Vandendriessche et al., 2012), it is important that performance tests are validated across the entire age spectrum they intend to assess. Additionally, increased exposure to training and greater training volume experienced by players of a higher competitive playing level may also influence consistency of testing performance (Rebelo et al., 2013). As a result, we hypothesise that field-based tests for youth soccer players will be, mostly, valid and reliable, although more variable than previously reported in adults. Therefore, the purpose of this study was to: i) examine the between-day reliability of commonly used field-based fitness tests across an appropriate age range (U11-U17) and across multiple performance standards (amateur, development, performance) included within a national governing body for soccer; ii) to assess the construct validity of the testing battery to discriminate between multiple age groups (U11-U17); and iii) evaluate the ability of a battery of commonly used field-based fitness tests to discriminate between playing standards (amateur, development, performance) included within a national governing body for soccer and to discriminate between “unidentified” (amateur) and “identified” (within a progressive pathway to professional soccer) youth soccer players.

Methods

Participants:

Three-hundred-and-seventy-three Scottish youth soccer players (mean±SD: age 13.5±1.8 years; stature 161.1±13.3 cm; body mass 50.8±12.4 kg) from the 3 different playing standards (amateur, development, performance) and 7 age brackets (U11, U12, U13, U14, U15, U16, & U17) identified by the "Club Academy Scotland" (CAS) structure of the Scottish Football Association (SFA) volunteered to participate in this study. Participants were categorised as either "unidentified" (amateur level players – no direct route to professional football), or "identified" (development/performance level players – direct route to professional football through existing club pathway). Additionally, participants were categorised as "amateur" (recreational players), "development" (lower ranked professional academies), and "performance" ("elite" level academies) based upon the SFA CAS structure. Due to the vast positional demands of goalkeepers within soccer, players of this position were excluded from analysis. Prior to conducting any trials, participant and parental/guardian consent was gained alongside providing comprehensive written and oral explanations about the study. Institutional ethical approval was granted.

Design:

Participants completed 2 testing sessions ~1.5hours in length separated by a washout period of 7-14days. Data were collected from 7 field-based fitness tests commonly used as physical performance measures within youth soccer (Paul & Nassis, 2015). All selected tests were identified to be appropriate for implementation across the entire age range of the selected sample, and relevant to the demands of soccer (Paul & Nassis, 2015). To account for circadian variability (Drust et al., 2005), both testing sessions took place at the same time of day and during players' normal training hours. Testing sessions were completed a minimum of 48hours following a competitive game, and in absence of strenuous exercise within 24hours prior. Testing sessions were conducted indoors (~22°C) on a non-slip playing surface. Upon arrival for the initial testing session, participants provided basic descriptive details via a self-report questionnaire including: date of birth, associated playing club, main playing position, and the number of club training sessions completed/week. Prior to both testing sessions, all participants conducted a standardised warm-up protocol consisting of light aerobic activity, dynamic stretching, and progressive sprinting. Testing order and procedures for each test was the same on each occasion. The research team remained constant throughout the data collection process with the same researchers collecting data from the same fitness tests consistently across sessions.

Procedures:

Following the standardised warm-up, participants received verbal instruction and demonstrations from the research team immediately prior to conducting 2-3 familiarisation attempts for each test. When required, guidance and feedback were provided to participants by the research team following each familiarisation attempt, however no guidance was provided to participants between recorded attempts. For tests where electronic timing gates were used, gates were adjusted to an appropriate height as per the mean stature of the sample group, and start positions were standardised as a fixed-position crouch start from 1m behind the start gate (Haugen & Buchheit, 2016). Data were collected using the Brower TC Timing System (Brower

Timing Systems, Draper, UT). Participants completed 3 attempts of each test (unless otherwise stated) with the best attempt being selected for analysis. Recovery intervals between attempts were standardised at 3minutes for each test.

Anthropometrics

Standing stature was assessed using a free-standing stadiometer (Seca, Birmingham, UK) and body mass was assessed using digital floor scales (Seca, Birmingham, UK).

Grip Dynamometry (Grip Strength)

Grip strength was selected as a suitable strength evaluation for implementation across the entire sample within this study (U11-U17) and was examined using an analogue dynamometer (Takei 5001, Takei Scientific Instruments Co., Niigata-City, Japan). Attempts were collected from participants' dominant hand (specified as the participants' writing hand), and with the appropriate hand spacing adjusted for each individual as per manufacturer guidelines. Participants were instructed to hold the dynamometer above head with a locked arm. Participants applied maximum pressure by squeezing the handle of the dynamometer. Over a 5-second period, the dynamometer was lowered in an arc towards the participant's hip maintaining a locked arm.

Standing Broad Jump (SBJ)

SBJ was examined using an open reel tape measure (PerformBetter, Southam, UK) secured to the ground. Participants were instructed to place the front edge of their footwear as close to, but not touching, a designated start line. Without repositioning their feet and utilising countermovement and arm swing, participants jumped forwards maximally landing on both feet. Attempts were disqualified if participants moved their feet upon take-off or landing, or if additional body parts (other than the feet) came in contact with the ground. Measurements were taken from the furthest edge of the designated start line, to the back of the rear landing foot.

Countermovement Vertical Jump (CMJ)

CMJ data were collected using the Just Jump mat (Probiotics, Huntsville, AL). Attempts were conducted adopting the arms akimbo position and utilising a self-selected countermovement depth. Attempts were disqualified if participants abandoned the arms akimbo position or actively flexed at the knee or hip during flight.

505 Change of Direction Test (505COD)

Change of direction ability through the horizontal plane was assessed via the 505COD test. The methodology for the 505COD was conducted as per established methods (Draper & Lancaster, 1985). This involved a 15m linear sprint from a static start, a 180° turn on the nominated leg ensuring contact with a turn line, and a 5m recovery sprint through an identified finish line. The time expired during the final 5m of the 15m linear sprint, turn, and 5m return sprint was recorded. Participants completed two attempts for each turning leg (R/L) with the mean score of the best attempt from each leg being used for analysis.

T-Drill Test (T-Test)

Multi-directional speed and change of direction ability was assessed via the T-Test. The methodology for the T-Test was used as per established methods (Semenick, 1990). This involved completing a pre-planned course touching a series of cones laid out in a

T shape, requiring a combination of maximal sprinting, side shuffling, and backpedalling.

10/20m Sprint

Linear speed and acceleration was assessed over distances of 10/20m as per previously reported match-based observations of youth soccer players (Buchheit, Mendez-Villanueva, Simpson, & Bourdon, 2010).

<u>Statistical Analysis:</u>

Prior to analysis, the assumption of normality was verified using the Shapiro-Wilk test. A two-way random effects intra-class correlation coefficient (ICC) with absolute agreement and coefficient of variation (CV) was used to evaluate relative test-retest reliability. Standardised effect size (ES), reported as Cohen's *d* using the pooled SD as the denominator, was calculated to evaluate the magnitude of the test-retest differences. As per guidelines provided by Atkinson & Nevill (1998) and Hopkins (2000), the tests were deemed as reliable if they met the following criteria: good-excellent ICC (>0.80), moderate CV (≤10%), and a trivial or small effect size (<0.60). For the separate analyses associated with playing standards and age groups, test scores were standardised using within-group z-scores. This involved allocating standardised scores within-age or within-playing standard groups, and then collapsing across levels prior to analysis. This allowed for comparisons between playing standards (using standardised within-age group z-scores) removing potential age effects, and between age groups (using standardised within-playing standard z-scores) removing potential playing standard effects. The mean score of trial 1 and trial 2 was used for comparisons between levels and age. Identified/unidentified players were compared using an independent samples t-test using z-scores, playing standards and age groups were compared via a one-way analysis of variance (ANOVA) using z-scores, with a Bonferroni post-hoc test being implemented to identify differences between groups. Discriminant function analysis was conducted to derive a predictive model for classifying youth players as unidentified or identified based upon fitness test performance. Percentages of correct classification and canonical correlation coefficients were noted. Statistical significance was set at $p \leq 0.05$.

Results

Reliability of physical performance characteristics

Table 1 shows reliability data for all fitness test components at each age group. The majority of field-based fitness tests selected within this study proved to be reliable measures of physical performance across all age groups (ICC=0.83-0.97; p<0.01). However, the 505COD and T-Test tests showed weaker reliability in younger participants (U11/U12) (ICC=0.57-0.79; p<0.01). In addition, the 10m and 20m sprint showed weaker reliability (ICC<0.80) in the U12 (ICC=0.73) and U17 age groups (ICC=0.78). **Figure 1** shows mean performance differences between trials for the 505COD and T-Test across age groups.

(Insert Table 1)

Validity of physical performance characteristics

Identified players were significantly taller (0.11±0.98 vs. -0.18±0.99; p=0.007) and heavier (0.10±1.00 vs. -0.17±0.99; p=0.009) compared to unidentified players. Playing standard comparisons revealed that performance players were significantly taller (0.30±0.12 vs. -0.24±0.13; p=0.044) and heavier (0.31±0.12 vs. -0.26±0.13; p=0.036) than amateur players, however no significant differences were observed for stature or body mass between amateur-development or development-performance player groups. No significant differences were observed between identified and unidentified groups or playing standards for birth month, or birth quarter across all age groups, or between playing position.

(Insert Figure 1)

Figure 2 shows validity data between unidentified and identified player groups. Grip strength (0.16±1.00); SBJ (0.30±0.93); CMJ (0.16±1.01); 505COD (0.23±1.01); and T-Test (0.16±0.98) performance was significantly higher in the identified player group (p<0.001), and also on the 20m sprint test (0.10±0.99; p=0.012). The 10m sprint was the only test demonstrating non-significant differences between the unidentified (-0.01±0.95) and identified (0.01±1.02) player groups (p=0.829).

(Insert Figure 2)

Figure 3 shows validity data between amateur, development, and performance playing standards. The CMJ was the only test that demonstrated significant increases at each of the 3 playing standards in the hypothesised direction (development>amateur, p=0.022; performance>development; p<0.001). Amateur players had significantly (p<0.001) lower grip strength (-0.26±0.93); SBJ (-0.49±0.88); 505COD (-0.37±0.89); and T-Test (-0.26±0.96) compared to both development and performance players. The 20m sprint was significantly slower for amateur (-0.17±0.98) compared to development (0.20±1.12; p=0.005), however not when compared to performance players (0.02±0.85; p=0.127). No significant differences were observed between club levels for the 10m sprint test. SBJ (0.57±0.92) and 505COD (0.61±0.82) tests were significantly (p<0.001) higher for development compared to performance players.

(*Insert Figure 3*)

No significant differences were observed between U11/U12 age groups for any measure; U12/U13 age groups for CMJ (p=0.954) and T-Test (p=0.108), a tendency for U13 players to be faster over the 10m sprint (p=0.056) and 20m sprint (p=0.065); and U16/U17 where no significant differences were observed for any measure except SBJ

(p=0.019). Significant performance differences were observed between all remaining age groups and tests in the hypothesised direction (U17>U11; p<0.001), except for U12/U13 age groups for grip strength (p=0.025) and SBJ (p=0.012). Whilst still significant, these were not observed at the (p<0.001) level as reported for the majority of measures.

Discriminant function analyses indicated that the field-based fitness tests significantly discriminated between the unidentified and identified players; χ^2 (7)=101.646, p<0.001, with 70.2% of players being correctly classified. Inspection of the canonical correlation coefficients revealed that this discrimination was largely due to performance on the SBJ (r=0.75) and 505COD (r=0.54) tests. The additional tests within the testing battery did not make an important contribution to the discriminant function (r<0.40), with 10m sprint performance contributing to group membership least (r=0.02).

Discussion

This study aimed to evaluate the reliability of commonly-used field based fitness tests, and to evaluate the construct validity of a field-based fitness testing battery to discriminate across the entire spectrum of ages and performance levels of male youth soccer players, within a national governing body structure for soccer. The field-based fitness tests used within this study were mostly reliable (acceptable between-day reliability) and valid (able to discriminate between unidentified and identified player groups) in male youth soccer players. However, the lower reliability of the COD tests within younger participants, and the inability of field-based fitness tests to discriminate between development-performance playing standards prevent congruous implementation of this testing battery across the entire sample of the present study.

In agreement with previous findings across comparable age ranges (Hulse et al., 2013), test-retest reliability was reported as "good-excellent" for the majority of age groups across tests, an exception being "acceptable" ICC values reported for the 10/20m sprint tests for U12/U17 age groups. Nevertheless, despite the lower ICC values for U12/U17 age groups, the CV and ES for these groups suggests acceptable between-day test reliability. The lower reliability values observed for the 505COD and T-Test change of direction tests for younger players however, is a novel finding from our study. Whilst the CV values reported for the 505COD are still within range of what is typically considered as "good reliability" (CV<10%) (Atkinson & Nevill, 1998), the moderate ES (*d*=0.75-0.89) and lower ICC (0.57-0.61) within younger age ranges (U11-U12) show higher variability in this test compared to other measures. Due to the physical complexity and demand on eccentric/concentric strength and power of these tests

(Lloyd et al., 2013), it is likely that limited physical development and early stage of maturation associated with the younger players within our study could result in increased variability in test performance (Gil et al., 2007; Paul & Nassis, 2015; Pearson et al., 2006). Lower levels of test-retest reliability, moderate effect sizes and significant differences in performance for the 505COD and T-Test in the young age groups (U11-U12) suggest that these tests may be less suitable as a measure to evaluate change of direction (COD) performance in young soccer players.

In agreement with previous findings, identified players were significantly larger in stature and body mass comparative to unidentified players (Figueiredo et al., 2009). Additionally, six of the seven fitness tests within the testing battery displayed differences between unidentified and identified player groups in the hypothesised direction (identified players scoring better than unidentified). It is possible that these anthropometric and physical discrepancies are due to differences in maturity status between the unidentified and identified player groups. A wealth of previous evidence suggests youth players playing at higher competitive standards often have greater maturity status than their chronologically age-matched peers (Cumming et al., 2018; Gouvêa et al., 2017; Lovell et al., 2015). This has been reported as a significant influencer within the selection/deselection process within youth soccer (Figueiredo et al., 2009; Malina et al., 2010), however there is limited evidence to suggest these maturity and anthropometric differences observed during adolescence contribute to future professional status (le Gall et al., 2010; Ostojic et al., 2014). In addition, performance improvements relative to chronological age groups increased between ages 13-15 years, with inconsistencies in performance improvements observed at the lower (U11-U13) and upper (U16/U17) ranges of our sample population. This finding is supported by previous research, reporting peak physical development transpiring at peak height velocity, typically reported between 13-15 years in active adolescent boys (Cumming et al., 2018; Gouvêa et al., 2017; Lovell et al., 2015; Philippaerts et al., 2006).

Despite the proclaimed importance of acceleration and short sprint ability associated with competitive youth soccer (Gil et al., 2007; le Gall et al., 2010; Rebelo et al., 2013), the 10m sprint was the only test demonstrating no differences between groups. This finding could be due to the convention of younger athletic populations participating in multiple sports simultaneously, and the translation of acceleration and short duration sprint ability across various athletic activities. However, observed differences between playing standards demonstrated that only the CMJ test displayed significant differences at each playing standard in the anticipated direction (performance>development>amateur). Surprisingly, development players scored better than both amateur and performance player groups on the 505COD and SBJ tests. These findings suggest there may be a physical barrier between unidentified and identified player categories, however a physical evaluation alone is insufficient to discriminate between the development and performance players within this study. Additional factors associated with soccer performance must therefore be considered when looking to discriminate between players of different standards within an identified player population.

Interpretation of findings from the discriminant function analysis supports our notion that the field-based testing battery, used in our study, possessed construct validity, demonstrating a 70.2% success rate of correctly classifying unidentified and identified players. Classifications were mostly influenced based on performance on the SBJ and

505COD tests. These findings align with previous suggestions promoting the suitability of a field-based fitness testing battery as a valid and sensitive tool to discriminate between identification status (Unnithan et al., 2012; Vaeyens et al., 2006). Our findings also highlight the importance of muscular power and COD ability within youth soccer.

A limitation of this study was the absence of a field-based measure of cardio-vascular endurance, for example the Yo-Yo Intermittent Recovery Test Level 1 (YYIRT L1). Whilst the YYIRT L1 was within the initial testing battery protocol, it was removed following an unwillingness of development/performance coaches to allow their players to complete this test due to the perceived fatigue associated with completing it. Levels of cardio-vascular fitness are often reported as a key determinant of elite and identified youth soccer players (Buchheit et al., 2010; Gil, et al., 2007; le Gall et al., 2010; Mujika et al., 2009; Stolen et al., 2005), therefore may have added to the discriminative ability of this field-based fitness testing battery. Finally, a lack of measure of maturity status is a further limitation of this study. The effects of progressed maturity status on physical performance tests are well established within current literature (Figueiredo et al., 2009; le Gall et al., 2010). Whilst it is acknowledged that maturity status may be a contributing factor regarding the differences in physical performance demonstrated by this study, our study design is ecologically valid according to the national governing body for soccer appropriate to our sample. Therefore, resultant of the current tendency to categorise players by chronological age, rather than maturity status, our findings suggest that physical ability continues to influence playing standard selection within youth soccer players. Further research adopting a comparable study design is required to verify the results of this study.

Conclusion

Coaches and practitioners should be aware of the potential lower reliability of COD tests in young soccer players. This may result in potential issues when interpreting performance test results with younger age groups (U11/U12) as the magnitude of change in COD performance may be lower than variability within the test. Our results suggest that whilst a comprehensive field-based fitness testing battery can discriminate between two distinct sample groups (unidentified vs. identified), a physical fitness testing battery alone is insufficient to discriminate between players of varied ability within an identified group of youth soccer players.

Acknowledgements

The results of the current study do not constitute the endorsement of any product by the authors or the journal. This paper presents independent research and the views expressed are those of the authors. The authors wish to thank the players, coaches, and staff from the participating teams for their efforts and co-operation during the data collection of this study

References

Atkinson, G., & Nevill, A. M. (1998). Statistical Methods for Assssing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sports Medicine*, *26*(4), 217–238.

Buchheit, M., Mendez-Villanueva, A., Simpson, B. M., & Bourdon, P. C. (2010). Match Running Performance and Fitness in Youth Soccer. *International Journal of Sports Medicine*, *31*, 818–825.

Cumming, S. P., Brown, D. J., Mitchell, S., Bunce, J., Hunt, D., Hedges, C., & Malina, R. M. (2018). Premier League academy soccer players' experiences of competing in a tournament bio-banded for biological maturation. *Journal of Sports Sciences*, *36*(7), 757–765.

Deprez, D., Fransen, J., Boone, J., Lenoir, M., Philippaerts, R., & Vaeyens, R. (2015). Characteristics of high-level youth soccer players: variation by playing position. *Journal of Sports Sciences*, *33*(3), 243–254.

Draper, J., & Lancaster, M. (1985). The 505 test: A test for agility in the horizontal plane. *Australian Journal of Science and Medicine in Sport*, *17*(1), 15–18.

Drust, B., Waterhouse, J., Atkinson, G., Edwards, B., & Reilly, T. (2005). Circadian Rhythms in Sports Performance—an Update. *Chronobiology International*, *22*(1), 21–44.

Figueiredo, A. J., Gonçalves, C. E., Coelho e Silva, M. J., & Malina, R. M. (2009). Characteristics of youth soccer players who drop out, persist or move up. *Journal of Sports Sciences*, *27*(9), 883–891.

Gil, S. Ruiz, F. Irazusta, A. Gil, J, Irazusta, J. (2007). Selection of young soccer players in terms of anthropometric and physiological factors. *Journal of Sports Medicine and Physical Fitness*, *47*, 25–32.

Gil, S. M., Zabala-Lili, J., Bidaurrazaga-Letona, I., Aduna, B., Lekue, J. A., Santos-Concejero, J., & Granados, C. (2014). Talent identification and selection process of outfield players and goalkeepers in a professional soccer club. *Journal of Sports Sciences*, *32*(20), 1931–1939.

Gonaus, C., & Müller, E. (2012). Using physiological data to predict future career progression in 14- to 17-year-old Austrian soccer academy players. *Journal of Sports Sciences*, *30*(15), 1673–1682.

Gouvêa, M. A. De, Cyrino, E. S., Valente-Dos-Santos, J., Ribeiro, A. S., Silva, D. R. P. Da, Ohara, D., & Ronque, E. R. V. (2017). Comparison of Skillful vs Less Skilled Young Soccer Players on Anthropometric, Maturation, Physical Fitness and Time of Practice. *International Journal of Sports Medicine*, *38*(5), 384–395.

Haugen, T., & Buchheit, M. (2016). Sprint Running Performance Monitoring: Methodological and Practical Considerations. *Sports Medicine*, *46*(5), 641–656.

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*, *30*(5), 375–381.

Huijgen, B. C. H., Elferink-Gemser, M. T., Lemmink, K. A. P. M., & Visscher, C. (2014). Multidimensional performance characteristics in selected and deselected talented soccer players. *European Journal of Sport Science*, *14*(1), 2–10.

Hulse, M. A., Morris, J. G., Hawkins, R. D., Hodson, A., Nevill, A. M., & Nevill, M. E.

(2013). A field-test battery for elite, young soccer players. *International Journal of Sports Medicine*, *34*(4), 302–311.

Larkin, P., & O'Connor, D. (2017). Talent identification and recruitment in youth soccer: Recruiter's perceptions of the key attributes for player recruitment. *PLoS ONE*, *12*(4), 1–15.

le Gall, F., Carling, C., Williams, M., & Reilly, T. (2010). Anthropometric and fitness characteristics of international, professional and amateur male graduate soccer players from an elite youth academy. *Journal of Science and Medicine in Sport*, *13*(1), 90–95.

Lloyd, R. S., Read, P., Oliver, J. L., Meyers, R. W., Nimphius, S., & Jeffreys, I. (2013). Considerations for the Development of Agility During Childhood and Adolescence. *Strength and Conditioning Journal*, *35*(3), 2–11.

Lovell, R., Towlson, C., Parkin, G., Portas, M., Vaeyens, R., & Cobley, S. (2015). Soccer player characteristics in English lower-league development programmes: The relationships between relative age, maturation, anthropometry and physical fitness. *PLoS ONE*, *10*(9), 1–14.

Malina, R. M., Peña Reyes, M. E., Figueiredo, A. J., Coelho E Silva, M. J., Horta, L., Miller, R., & Morate, F. (2010). Skeletal age in youth soccer players: Implication for age verification. *Clinical Journal of Sport Medicine*, *20*(6), 469–474.

Mujika, I., Santisteban, J., Impellizzeri, F. M., & Castagna, C. (2009). Fitness determinants of success in men's and women's football. *Journal of Sports Sciences*, *27*(2), 107–114.

Murr, D., Raabe, J., & Höner, O. (2018). The prognostic value of physiological and physical characteristics in youth soccer: A systematic review. *European Journal of Sport Science*, *18*(1), 62–74.

O'Connor, D., Larkin, P., & Mark Williams, A. (2016). Talent identification and selection in elite youth football: An Australian context. *European Journal of Sport Science*, *16*(7), 837–844.

Ostojic, S. M., Castagna, C., Calleja-González, J., Jukic, I., Idrizovic, K., & Stojanovic, M. (2014). The biological age of 14-year-old boys and success in adult soccer: Do early maturers predominate in the top-level game? *Research in Sports Medicine*, *22*(4), 398–407.

Paul, D. J., & Nassis, G. P. (2015). Physical Fitness Testing in Youth Soccer: Issues and Considerations Regarding Reliability, Validity, and Sensitivity. *Pediatric Exercise Science*, *27*(3), 301–313.

Pearson, D. T., Naughton, G. A., & Torode, M. (2006). Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports. *Journal of Science and Medicine in Sport*, *9*(4), 277–287.

Philippaerts, R. M., Vaeyens, R., Janssens, M., Van Renterghem, B., Matthys, D., Craen, R., & Malina, R. M. (2006). The relationship between peak height velocity and physical performance in youth soccer players. *Journal of Sports Sciences*, *24*(3), 221–230.

Pyne, D. B., Spencer, M., & Mujika, I. (2014). Improving the value of fitness testing for football. *International Journal of Sports Physiology and Performance*, *9*(3), 511–

514.

Rebelo, A., Brito, J., Maia, J., Coelho-e-Silva, M. J., Figueiredo, A. J., Bangsbo, J., & Seabra, A. (2013). Anthropometric Characteristics , Physical Fitness and Technical Performance of Under-19 Soccer Players by Competitive Level and Field Position. *International Journal of Sports Medicine*, *34*(4), 312–317.

Reeves, M. J., McRobert, A. P., Littlewood, M. A., & Roberts, S. J. (2018). A scoping review of the potential sociological predictors of talent in junior-elite football: 2000–2016. *Soccer and Society*, (February), 1–21.

Reilly, T., Williams, A. M., Nevill, A., & Franks, A. (2000). A multidisciplinary approach to talent identification in soccer. *Journal of Sports Sciences*, *18*(9), 695–702.

Sarmento, H., Anguera, M. T., Pereira, A., & Araújo, D. (2018). Talent Identification and Development in Male Football: A Systematic Review. *Sports Medicine*, *48*(4), 907–931.

Semenick, D. (1990). The T-test. *Journal of Strength and Conditioning Research*, *12*(1), 36–37.

Stolen, T., Chamari, K., Castagna, C., & Wisloff, U. (2005). Physiology of Soccer: An Update. *Sports Medicine*, *35*(6), 501–536.

Thomas, K., French, D., & Hayes, P. R. (2009). The Effect of Two Plyometric Training Techniques on Muscular Power and Agility in Youth Soccer Players. *Journal of Strength and Conditioning Research*, *23*(1), 332–335.

Unnithan, V., White, J., Georgiou, A., Iga, J., & Drust, B. (2012). Talent identification in youth soccer. *Journal of Sports Sciences*, *30*(15), 1719–1726.

Vaeyens, R., Malina, R. M., Janssens, M., Van Renterghem, B., Bourgois, J., Vrijens, J., & Philippaerts, R. M. (2006). A multidisciplinary selection model for youth soccer: The Ghent Youth Soccer Project. *British Journal of Sports Medicine*, *40*(11), 928–934.

Vandendriessche, J. B., Vaeyens, R., Vandorpe, B., Lenoir, M., Lefevre, J., & Philippaerts, R. M. (2012). Biological maturation, morphology, fitness, and motor coordination as part of a selection strategy in the search for international youth soccer players (age 15-16 years). *Journal of Sports Sciences*, *30*(15), 1695–1703.

Vandorpe, B., Vandendriessche, J., Vaeyens, R., Pion, J., Matthys, S., Lefevre, J., & Lenoir, M. (2012). Relationship between sports participation and the level of motor coordination in childhood: A longitudinal approach. *Journal of Science and Medicine in Sport*, *15*(3), 220–225.

Vescovi, J. D., Rupf, R., Brown, T. D., & Marques, M. C. (2011). Physical performance characteristics of high-level female soccer players 12-21 years of age. *Scandinavian Journal of Medicine & Science in Sports*, *21*(5), 670–678.

Table Captions

Table 1. Between-day test-retest reliability of field-based fitness tests across age groups U11-U17.

Figure Captions

Figure 1. Between trial mean difference comparisons for COD tests across individual age groups. A: 505COD, B: T-Test. * indicates significant differences between U11/U12 and U13-U17 age groups; α indicates significant differences between U11-U13 and U14-U17 age groups ($p<0.01$).

Figure 2. Z-score mean differences on performance tests between different playing standards. A: Grip Strength, B: SBJ, C: CMJ, D: 505COD, E: T-Test, F: 10m Sprint, G: 20m Sprint. * indicates significantly higher than amateur; α indicates significantly higher than development; β indicates significantly higher than performance ($p<0.01$). δ indicates significantly higher than amateur; ϒ indicates significantly higher than development; Ω indicates significantly higher than performance ($p<0.05$).

Figure 3. Z-score mean differences on performance tests between unidentified and identified player groups. A: Grip Strength, B: SBJ, C: CMJ, D: 505COD, E: T-Test, F: 10m Sprint, G: 20m Sprint. * indicates significantly higher than unidentified ($p<0.01$); α indicates significantly higher than unidentified ($p<0.05$).

Table 1.

		U11 (n=26)	U12 (n=51)	U13 (n=75)	U14 (n=59)	U15 (n=81)	U16 (n=46)	U17 (n=35)
Grip Strength	Trial 1 (x ± SD) [kg]	17.8 ± 2.6	18.1 ± 3.6	21.5 ± 4.5	25.3 ± 5.3	33.2 ± 7.5	37.7 ± 7.2	37.5 ± 7.3
	Trial 2 (x ± SD) [kg]	17.7 ± 2.9	18.7 ± 3.3	22.1 ± 4.8	25.9 ± 5.8	33.4 ± 7.2	38.3 ± 6.3	37.8 ± 7.0
	ICC (CI)	0.93 (0.85-0.97)	0.83 (0.71-0.91)	0.85 (0.76-0.90)	0.89 (0.81-0.93)	0.92 (0.87-0.95)	0.92 (0.84-0.96)	0.88 (0.77-0.94)
	CV (%)	4.4	8.2	8.3	6.7	6.4	5.6	5.8
	ES	0.04	0.17	0.13	0.11	0.03	0.09	0.03
SBJ	Trial 1 (x ± SD) [cm]	154.7 ± 12.6	159.3 ± 13.8	171.4 ± 17.8	181.2 ± 17.8	195.6 ± 16.3	201.4 ± 16.5	214.0 ± 21.6
	Trial 2 (x ± SD) [cm]	153.9 ± 13.8	158.4 ± 15.9	174.3 ± 22.0	183.8 ± 17.8	195.8 ± 17.8	203.6 ± 15.7	214.2 ± 22.1
	ICC (CI)	0.93 (0.84-0.97)	0.90 (0.83-0.95)	0.90 (0.84-0.94)	0.96 (0.92-0.97)	0.95 (0.93-0.97)	0.93 (0.86-0.96)	0.97 (0.93-0.98)
	CV (%)	2.5	2.3	3.0	2.5	2.1	2.2	1.9
	ES	0.06	0.06	0.14	0.15	0.01	0.10	0.00
CMJ	Trial 1 (x ± SD) [cm]	36.0 ± 4.3	36.7 ± 4.6	36.8 ± 4.9	40.4 ± 6.6	43.8 ± 4.7	47.6 ± 5.3	48.3 ± 6.5
	Trial 2 (x ± SD) [cm]	35.4 ± 5.1	36.5 ± 4.8	36.8 ± 5.3	41.0 ± 6.4	43.9 ± 4.6	47.5 ± 6.4	48.1 ± 6.1
	ICC (CI)	0.90 (0.78-0.96)	0.94 (0.90-0.97)	0.87 (0.80-0.92)	0.94 (0.91-0.97)	0.90 (0.85-0.94)	0.95 (0.90-0.97)	0.96 (0.92-0.98)
	CV (%)	4.0	3.2	5.0	3.8	3.6	3.2	2.8
	ES	0.13	0.04	0.00	0.09	0.02	0.02	0.05
COD505	Trial 1 (x ± SD) [s]	2.84 ± 0.13	2.76 ± 0.14	2.66 ± 0.16	2.58 ± 0.12	2.48 ± 0.10	2.45 ± 0.13	2.43 ± 0.13
	Trial 2 (x ± SD) [s]	2.96 ± 0.14	2.89 ± 0.20	2.68 ± 0.18	2.60 ± 0.14	2.50 ± 0.12	2.47 ± 0.12	2.42 ± 0.13
	ICC (CI)	0.61 (0.12-0.85)	0.57 (0.08-0.78)	0.91 (0.86-0.94)	0.89 (0.82-0.94)	0.85 (0.77-0.91)	0.86 (0.74-0.93)	0.97 (0.93-0.98)
	CV (%)	3.3	3.6	1.9	1.7	1.9	2.2	1.0
	ES	0.89	0.75	0.12	0.15	0.18	0.16	0.10
T-Test	Trial 1 (x ± SD) [s]	11.83 ± 0.79	11.62 ± 0.52	11.27 ± 0.74	10.88 ± 0.60	10.24 ± 0.46	9.96 ± 0.50	9.86 ± 0.59
	Trial 2 (x ± SD) [s]	12.17 ± 0.70	11.86 ± 0.79	11.45 ± 0.85	10.97 ± 0.67	10.35 ± 0.46	9.98 ± 0.53	10.03 ± 0.74
	ICC (CI)	0.79 (0.46-0.91)	0.75 (0.54-0.86)	0.89 (0.80-0.93)	0.94 (0.89-0.96)	0.87 (0.78-0.92)	0.95 (0.90-0.97)	0.91 (0.76-0.96)
	CV (%)	3.1	3.0	2.3	1.5	1.6	1.3	1.9
	ES	0.46	0.36	0.26	0.14	0.24	0.04	0.25
10m Sprint	Trial 1 (x ± SD) [s]	2.05 ± 0.08	1.99 ± 0.07	1.95 ± 0.11	1.90 ± 0.11	1.82 ± 0.10	1.75 ± 0.09	1.73 ± 0.09
	Trial 2 (x ± SD) [s]	2.05 ± 0.09	2.01 ± 0.09	1.96 ± 0.11	1.93 ± 0.10	1.83 ± 0.10	1.77 ± 0.09	1.76 ± 0.09
	ICC (CI)	0.85 (0.66-0.94)	0.73 (0.42-0.88)	0.90 (0.83-0.94)	0.84 (0.71-0.90)	0.93 (0.88-0.95)	0.94 (0.88-0.97)	0.78 (0.57-0.89)
	CV (%)	1.7	2.2	2.0	2.2	1.7	1.5	2.4
	ES	0.00	0.25	0.09	0.29	0.10	0.22	0.30
20m Sprint	Trial 1 (x ± SD) [s]	3.67 ± 0.14	3.59 ± 0.16	3.50 ± 0.21	3.39 ± 0.20	3.23 ± 0.18	3.09 ± 0.15	3.08 ± 0.15
	Trial 2 (x ± SD) [s]	3.69 ± 0.20	3.64 ± 0.19	3.54 ± 0.21	3.42 ± 0.19	3.26 ± 0.19	3.16 ± 0.17	3.11 ± 0.15
	ICC (CI)	0.85 (0.58-0.92)	0.73 (0.67-0.91)	0.90 (0.89-0.97)	0.83 (0.86-0.96)	0.93 (0.90-0.97)	0.94 (0.63-0.96)	0.78 (0.85-0.97)
	CV (%)	1.8	2.3	1.6	1.8	1.5	2.0	1.4
	ES	0.12	0.28	0.19	0.15	0.16	0.44	0.20

n = sample size; x ± SD = mean ± standard deviation; ICC = intra-class correlation, all p<0.01; CI = confidence interval; CV = coefficient of variation; ES = effect size