Title: The reliability of a modified 505 test and change-of-direction deficit time in elite youth football players.

Article Type: Original Article

Authors: Jonathan M. Taylor[1], Louis Cunningham[2], Peter Hood[1], Ben Thorne[1], Greg Irvin[2] and Matthew Weston[3]

[1] Middlesbrough Football Club, Middlesbrough, UK

[2] Sunderland AFC, Sunderland, UK

[3] Department of Psychology, Sport and Exercise, Teesside University, Middlesbrough, UK

Corresponding author: Jonathan M. Taylor, Middlesbrough Football Club, Middlesbrough, UK

Email: Jonathan.Taylor@mfc.co.uk

Telephone: +44 1325 722002

Abstract word count: 221

Text only word count: 3070

Number of tables: 3

Number of figures: 2

Abstract:

Change-of-direction ability is an important performance skill in football. Consequently, several testing protocols are used to assess this component. This study assessed the test-retest reliability of a modified 505 test (M505) and the change-of-direction deficit (CODD) in elite youth football players. Data were collected from 110 players from the Under [u] 12-18 years age groups (u18 n = 26, u16 n=26, u14 n=39, u12 n=19) within two English Premier League Category 1 Football Academies. Players completed the M505 twice in 7-days, in addition to a 10-m sprint test to allow CODD to be calculated. Reliability was assessed with respect to chronological and biological age (according to PHV status). Typical error (%), minimal detectable change (MDC%) and intraclass correlation coefficients (ICC) were calculated. Typical errors in M505 were moderate to large (2.0 to 3.2%), while intraclass correlation coefficients (ICC) ranged from low to high (r = 0.26 to r = 0.82). Typical errors in CODD were moderate to large (7.1 to 12.0%), with ICC's ranging from low to high (r = 0.19 to 0.79). Minimal detectable changes were 5.5 to 8.9% in M505 and 17.7 to 33.3% in CODD. The typical errors and minimal detectable changes observed here indicate that the M505 and CODD tests have limited practical utility in the evaluation of change of direction ability in elite youth football players.

Keywords: Peak height velocity, soccer, fitness testing, variability

Introduction

The ‘Elite player performance plan’ (EPPP), was introduced by the English Premier League with the aim of increasing the number, and quality of home grown players graduating from English football academies (EPPP guidelines, 2011). In this respect, standardised ‘benchmark performance testing’ is completed nationwide, using a testing battery that includes a ‘modified 505 test’ (M505) to profile change-of-direction ability in elite youth players. Change-of-direction ability is an important component of performance in football, with high-speed changes of direction occurring around key moments in match-play. Specifically, ~10% of goals are preceded by a change-of-direction sprint, while it is possible that change-of-direction ability is of greater importance in defensive situations (Faude et al. 2012). Consequently, several change-of-direction tests have been developed, with versions of the ‘505 test’ popular in football, due to the simplicity of this test and isolated nature of the turn (Svensson and Drust, 2005).

It has been suggested that short-sprint performance and change-of-direction ability are related in several currently used protocols, potentially skewing data interpretation (Sayers, 2015; Gabbett et al. 2008). Specifically, changing direction is reported to account for only 31% of total 505 time (Nimphius et al. 2013). This has led to the emergence of the change-of-direction deficit (CODD), defined as ‘the additional time that one directional change requires when compared to a linear sprint of equivalent distance’. The CODD is suggested to be a superior way of isolating change-of-direction ability (Nimphius et al. 2016).

To date, the reliability of the M505 and CODD tests in elite youth football players has not been determined, which is noteworthy given their potential use for talent-identification and monitoring purposes (Buchheit and Mendez-Villanueva, 2013). This is highly relevant, as understanding the short-term reliability of a test during a period where no true changes in measurement should occur (i.e. test-retest reliability) is critical to effective data interpretation

(Hopkins, 2000; Atkinson and Nevill, 1998). Knowledge of the random and systematic error (i.e. the noise), can allow the smallest meaningful change in performance (i.e. the signal, also known as the minimal detectable change) to be identified appropriately (Hopkins, 2000). While several measures are used to assess reliability, such as the intraclass correlation coefficients (a measure of relative reliability), perhaps the most frequently used reliability measure for assessing the variability in repeated-tests in athletes is the typical error (an absolute measure of reliability); also termed the standard error of measurement, due to the simplicity of this measure (Weir, 2005; Hopkins, 2000). The typical error can also be used to calculate the minimal detectable change (MDC), which provides information on the change in performance required for a practitioner to have confidence (95%) that a real change has occurred (Bernards et al. 2018)

The reliability of a change-of-direction test is particularly relevant in youth athletes around ‘Peak-height Velocity’ (PHV) (Beunen and Malina, 1988). While performance would generally be expected to improve with age, circa PHV, motor control and co-ordination is negatively affected through ‘adolescent awkwardness’ (Lloyd et al. 2015; Philipaerts et al. 2006). While maturation was reported to have no effect on the test-retest reliability of sprint, countermovement jump and aerobic performance (Buchheit and Mendez-Villanueva, 2013), it is more likely to impact the consistency of change-of-direction performance which requires greater co-ordination/motor-control ability.

This study assessed the test-retest reliability of the M505 and CODD in elite youth football players, while exploring the effects of maturation on performance, with the aim of facilitating practitioners to make better informed judgments on the change-of-direction ability of young football players.

Materials and Methods

Participants

A total of 110 players from Under [u] 12-18y age groups were recruited from two English Premier League Category 1 Academies. Table 1 displays player descriptive data along with best 10-m sprint time during the testing window. All players were registered with the academies, and as part of their registration documentation completed informed consent and medical screening forms (Parental consent was obtained for players aged 16 and under). The testing was part of routine practice, therefore ethical approval was not necessary (Winter and Maughan, 2009). The study was, however, conducted in accordance with the declaration of Helsinki.

Experimental approach to the problem

To assess the test-retest reliability of the M505 (Figure 1) and CODD, all players completed the test on two occasions with 7-days between testing bouts. All testing was carried out in accordance with EPPP guidelines. The players completed 4 trials (turning on each leg twice) with full recovery (~3 mins) allowed between each (Bogdanis et al. 1995). With this in mind, M505L and CODDL refers to trials where players turned on their left leg, whilst M505R and CODDR refers to trials where players turned on their right leg. The players started in a two-point athletic stance, with their preferred foot on the start-line. On instruction, the players accelerated as quickly as possible, before decelerating and touching the turning line with the correct foot and then accelerating back through the starting gates. The players also completed speed testing with 10-m sprint times recorded. To calculate CODD, best 10-m sprint time recorded was subtracted from best M505 time. Players had completed the tests as part of in-season fitness testing previously, and were therefore familiar with the procedures. A standardised warm-up protocol consisting of a general aerobic warm-up/ dynamic flexibility work (~8 mins), followed by three 20-m strides (at 80%, 90% and 100% of maximal effort) and two practice trials of the M505 (at 90% and 100% of maximal effort respectively), was completed before all tests. All testing was completed at the same time of day for each age-group, on the same indoor artificial field-turf training facility. Training structure in the days preceding testing was similar in all groups, with a minimum of 72 hours between match-play and testing. All testing was completed prior to training. Performance times were recorded to the nearest 0.01 s using Brower speed trap 2 light sensitive timing gates (Brower timing gate systems, USA), with the player's best time turning on each limb included in analysis. This timing system is suitable for tracking changes in short-sprint performance, with no marked systematic bias reported previously ($p<0.05$) (Shalfawi, et al. 2012). The height of the timing gates was set according to EPPP guidelines, at 75 cm for the u12 and u14 age groups, and 95 cm for the u16 and u18 age-groups i.e. approximately hip height (Haugen and Buchheit, 2016).

Figure 1 here

Table 1 here

Maturation, Performance and Reliability

Biological age was estimated as maturity offset in years from peak-height velocity (PHV) derived from sitting height, stretch stature, body mass and date of birth, recorded within the 7 days prior to the first trial (Mirwald, et al. 2002). Given that the majority of the players had been within the academy system for at least one year prior to testing (95%), this single measurement was assessed against serial measurements when data were available. Peak-height velocity was used to provide an indicator of somatic maturity, and players were subcategorised into Pre- (-0.5 years), At- (-0.51 to 0.5 years) and Post (≥0.51 years) peak-height velocity (PHV) (Wright et al. 2016). This allowed reliability to be assessed with respect to biological age.

Statistical Analysis

Descriptive statistics are displayed as mean ± standard deviation. Data were analysed with respect to chronological and biological age. All data were log-transformed to reduce the effect of non-uniformity of error. A custom-made reliability spreadsheet was used to calculate the typical error (expressed as coefficient of variation % [CV]) and intraclass correlation coefficients (ICC 3,1 with absolute agreement) (Shrout and Fleiss, 1979) in M505 and CODD performance (Hopkins, 2015). Subsequently, the minimal detectable change (MDC) was calculated as a percentage for each variable (in Microsoft excel 2016) as: typical error$\times 1.96 \times \sqrt{2}$ (Bernards et al. 2018). Qualitative inferences in intraclass correlation coefficients were based on the following thresholds: >0.99, extremely high; 0.99–0.90, very high; 0.75–0.90, high; 0.50–0.75, moderate; 0.20–0.50, low; <0.20, very low (Malcata et al. 2014). Precision in estimates are shown as 95% confidence intervals (CI).

Between-group pairwise comparisons (i.e. consecutive age-groups) in M505 and CODD performance were carried out using a customised spreadsheet for comparing group means (Hopkins, 2007) with effect sizes and a 95% confidence interval calculated. Effects were quantified using standardized thresholds (i.e. <0.2, 0.2, 0.6 and 1.2 standard deviations) derived from the harmonic mean of the group standard deviations (Hopkins, 2007). Magnitude based inferences were subsequently applied (Hopkins, 2007). Differences in performance were evaluated mechanistically, with clear inferences qualified using the following scale: 25% to 75%, possibly; 75% to 95%, likely; 95% to 99.5%, very likely; and >99.5%, most likely (Batterham and Hopkins, 2006).

Results

Test-retest reliability

Performance times for the M505 and CODD are displayed in Figure 2. Reliability data are presented in Tables 2 and 3. Typical errors were moderate to large in M505 (2.0 to 3.2%) and moderate to large in CODD (7.1 to 12.0%). Minimal detectable changes were 5.5 to 8.9% in M505 and 17.7 to 33.3% in CODD. Intraclass correlation coefficients (ICC) ranged from low to high in M505 (r = 0.26 to r = 0.82) and CODD (r = 0.19 to 0.79).

Between-group comparisons

For M505L possibly large (-4.6%;± 95% Confidence Interval 2.1%) effects between the u12-u14 groups, and possibly large (-5.1;±1.6%) effects between the u14-u16 groups were observed. Likely large (-5.1;±1.7%), likely moderate (-4.3;±2.4%) and likely small (-1.3;±1.8%) effects were observed between the u14-u16, u12-u14 and u16-u18 groups in M505R. Likely moderate (-10.2;±5.9%) and possibly moderate (-5.6;±5.9%) effects were observed between the u14-u16 and u12-u14 groups respectively in CODDL. A likely moderate effect in CODDR (-10.1;±4.9%) was observed between the u14-u16 groups.

Very likely large (-6.4;±2.0%) and possibly moderate (-2.2;±2.2%) effects in M505L were observed At-Post PHV and Pre-At PHV respectively. Likely large (-5.9;±2.0%) and possibly moderate (-2.5;±2.4%) effects were observed in M505R At-Post PHV and Pre-At PHV respectively. Possibly large effects in CODDL (-15.1;±6.5%) and CODDR (-12.6;±6.5%) were observed At-Post PHV. Where effects were observed between chronological and biological age-groups, older players were quicker. All reported between groups effects were smaller than the group specific MDC.

Table 2. here

Table 3. here

Figure 2. here

Discussion

Establishing the reliability of a physical test is critical to ensure that changes in performance are interpreted appropriately (Hopkins et al. 2001). Here, we assessed the test-retest reliability of the modified 505-test (M505) that is used as part of the ‘English premier league elite player performance plan (EPPP)’ benchmark performance testing in football youth academies. This study also assessed the reliability of the change-of-direction deficit (CODD) in elite youth players. In general, the M505 and the CODD elicit moderate to large typical errors and low to moderate relative reliability (ICC’s) in elite youth players. Maturity does not affect the reliability of the M505 and CODD in youth football players, as indicated by the magnitude of the typical errors and mean changes in performance. Importantly, our results suggest that a ~6-9% change in M505, and an ~18-33% change in CODD performance is required for a practitioner confident that a true change has occurred. Our findings are highly relevant given the widespread use of the M505 in talent identification and player monitoring in football, and the recent suggestion that CODD represents a better way of assessing change-of-direction ability (Nimphius et al. 2016).

Change-of-direction ability is highly relevant to football performance (Faude et al. 2012; Bloomfield et al. 2007), and this supposition underpins the use of the M505. While the M505 is used routinely in youth football, our findings are novel, and contrast with some of the existing work exploring the reliability of other modified 505 protocols. Previously, the reliability of a modified version of the 505 test was explored in multi-directional sport athletes with small typical errors (CV ~2.8%) and very high relative reliability (ICC’s r>0.90) reported (Dos Santos et al. 2017). Similarly, excellent between session relative reliability for a modified 505 test was reported in elite female team-sport players (r >0.96) (Barber et al. 2016). The typical errors we observed were generally greater in magnitude (moderate-large), while the relative (ICC) reliability was lower than that previously reported. The MDC in M505 that we reported was also greater than that reported previously (~3%) (Barber et al. 2016).

This disparity between our findings and previous work, and the trend towards improvement in performance on test 2 across the groups might be explained through our participant demographics. The onset of adolescent awkwardness around PHV here might have increased biological variance globally across our sample; specifically, the ability to accelerate, decelerate and change direction in a consistent manner (Lloyd et al. 2015; Philipaerts et al. 2006). Additionally, increased systematic bias as a consequence of a learning effect or athlete motivation, could have had a greater impact on our findings (Hopkins, 2000). With this in mind, it is possible that the scope for familiarization/learning was increased, explaining the potential improvements with subsequent tests. It is also possible that differences in testing procedures and equipment/running surface might explain some of the disparity between our findings and that of previous work.

The CODD has emerged as a potentially useful method of assessing change-of-direction performance (Dos Santos et al. 2018; Nimphius et al. 2016). This study is the first to explore the reliability of this measure (with respect to the M505) in elite youth football players. Our findings suggest that the CODD elicits moderate-large typical errors (7-12%), and has less than satisfactory relative reliability (ICC's) (ranging between 0.19 and 0.66) across all age groups. Furthermore, the MDC's of 18-33% reported suggest that a considerable change in performance would be necessary to be termed a real change. Given that the CODD time was short in this study (Nimphius et al. 2013), it is likely that systematic bias through learning might have been magnified due to the highly technical component of turning, perhaps explaining the greater typical errors and MDC reported in comparison to the M505.

With respect to performance on the M505 test, there was trend towards older and more mature players recording quicker times, while CODD performance was generally better in older players. Several between groups effects were unclear in CODD, this was particularly evident between the u16 and u18 age-groups. The observation that the MDC and typical errors did not differ across chronological and biological age groups with respect to the magnitude is somewhat surprising. Relative reliability (ICC's) was also similar across the groups, with the exception of M505R and CODDR in the 'At-PHV' group. It would be expected that players who are 'At-PHV' would produce less stable performance, due to the associated negative effects on co-ordination and motor-control (Philipaerts et al. 2006). It has been suggested that being highly trained can offset the impact of this 'adolescent awkwardness' on performance in young players, which might provide some explanation for our findings (Buchheit and Mendez-Villanueva, 2013) however, we acknowledge that this is supposition is speculative.

Given the trend towards improved performance on the M505 and CODD in test 2, extensive familiarization appears necessary to reduce systematic bias through learning (Hurst et al. 2018; Hopkins, 2000). Several tests over the course of subsequent days/weeks would likely be needed to gain a true understanding of the players ability to change-direction effectively. For example, if each player had completed four tests, the 'noise' of the test would be halved (Hopkins, 2000). This would be unfeasible in many situations however, given the time-constraints placed on physical training/ performance testing with technical and tactical training often taking priority (Turner et al. 2011). Furthermore, there is a trade-off between practitioner time availability and number of required tests needed to minimise noise in the test (Ehrenbrusthoff et al. 2016).

The information provided here on the minimal detectable change and typical errors in M505 and CODD performance indicate limited practical utility and suggest that these tests might not be suitable for use in this population (Bernards, 2018; Hopkins, 2000). Specifically, the MDC reported for the M505 would suggest that a change in performance of >0.16 s would be required for a change to be accepted with 95% confidence. The MDC reported in CODD indicates that a change of up to 0.3 s would be required for a change to be accepted with 95% confidence. In both instances a change of this magnitude would be unlikely in elite youth soccer players with test-retest intervals of ~12 weeks commonly used. Our findings provide further evidence of the difficulties in assessing worthwhile changes in change-of-direction ability, due to the lack, and or questionable reliability of change-of-direction measures. Consequently, alternative testing protocols might be considered to assess change-of-direction ability in young football players. If practitioners insist on using these tests, the group specific typical errors and minimal detectable change presented should be used to identify meaningful changes in performance. (Buchheit, 2016; Hopkins, 2000). Changes in performance that are smaller than the minimal detectable change should considered with caution, as it cannot be stated with 95% confidence that these changes are substantial.

Our findings are highly relevant and carry practical application within physical profiling of youth football players, yet, this study is not without limitations. A key limitation is the fact that players completed the testing on two occasions only. Undoubtedly, implementing further tests would have presented more powerful data. Given that this testing was completed in season, within two elite academies it was unfeasible to test on more occasions due to team-training schedules. Despite this, our work maintains strong practical application due to the population used, and the implications that this has with respect to the EPPP guidelines in English youth academies. Another limitation pertains to the assessment of biological status. It has been suggested that PHV status using the equation used here may be overestimated (Mills et al. 2017). Furthermore, it has been suggested that the data used in the original study validating this equation was outdated and therefore has questionable applicability. Despite this limitation, this method is commonplace within elite youth football where technology to perform more advanced methods is unavailable. The overestimation of PHV status may be offset to some extent by taking serial measurements (i.e. 2-3 per annum), which was considered here when data were available.

In conclusion, while a gold-standard change-of-direction test in football has not been identified to date, the M505 and CODD should be used with caution for assessment of change-of-direction ability in elite youth football players. The test-retest reliability of the M505 and CODD tests does not appear to be affected by maturation status in this population.

Disclosure statement

The authors report no conflict of interest.

References:

1. Atkinson G, Nevill AM. 1998. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 26(4): 217-238.
2. Barber OR, Thomas C, Jones PA, McMahon JJ, Comfort P. 2016. Reliability of the 505 change-of-direction test in netball players. Int J Sports Physiol Perf. 11(3):377-380.
3. Bernards JR, Sato K, Haff GG, Bazyler CD. 2018. Current research and statistical practices in sport science and a need for change. Sports. 5(87): 1-10.
4. Beunen GP, Malina, RM. 1988. Growth and physical performance relative to the timing of the adolescent spurt. Exerc Sport Sci Rev. 16: 503-540.
5. Bloomfield J, Polman R, O'Donoghue P. 2007. Physical demands of different positions in FA Premier League soccer. J Sports Sci Med. 6(1): 63.
6. Bogdanis GC, Nevill ME, Boobis LH, Lakomy HK, Nevill AM. 1995. Recovery of power output and muscle metabolites following 30 s of maximal sprint cycling in man. J Physiol. 482(2): 467-480.
7. Buchheit M., 2016. Chasing the 0.2. Int J Sport Physiol Perf. 11: 417-418.
8. Buchheit M, Mendez-Villanueva A. 2013. Reliability and stability of anthropometric and performance measures in highly-trained young soccer players: effect of age and maturation. J Sports Sci. 31(12): 1332-1343.
9. Dos Santos T, Thomas C, Jones PA, Comfort P. 2018. Asymmetries in isometric force-time characteristics are not detrimental to change of direction speed. J Strength Cond Res. 32(2): 520-527.
10. Dos Santos T, Thomas C, Jones PA, Comfort P. 2017. Mechanical determinants of faster change of direction speed performance in male athletes. J Strength Cond Res. 31(3): 696-705.

11. Ehrenbrusthoff K, Ryan CG, Gruneberg C, Wolf U, Krenz D, Atkinson G, Martin D. 2016. The intra- and inter-observer reliability of a novel protocol for two-point discrimination in individuals with chronic low back pain. Physiol Meas. 37(7): 1074-1088.
12. Faude O, Koch T, Meyer T. 2012. Straight sprinting is the most frequent action in goal situations in professional football. J Sports Sci. 30(7): 625-631.
13. Gabbett TJ, Kelly JN, Sheppard JM. 2008. Speed, change of direction speed, and reactive agility of rugby league players. J Strength Cond Res, 22(1):174-181.
14. Haugen T, Buchheit M. 2016. Sprint running performance monitoring: methodological and practical considerations. Sports Med, 46(5): 641-656.
15. Hopkins WG. 2015. Spreadsheets for analysis of validity and reliability. Sportscience. 19: 36–42. sportsci.org/2015/ValidRely.htm
16. Hopkins WG. 2007. A spreadsheet to compare the means of two groups. Sportscience. 11: 22-23. www.Sportsci.org.
17. Hopkins WG, Schabort EJ, Hawley JA. 2001. Reliability of power in physical performance tests. Sports Med. 31(3): 211-234.
18. Hopkins WG. 2000. Measures of reliability in sports medicine and science. Sports Med. 30(1): 1-15.
19. Hurst C, Batterham AM, Weston KL, Weston M. 2018. Short- and long-term reliability of leg extensor power measurement in middle-aged and older adults. J Sport Sci. 36(9): 970-977.
20. Lloyd RS, Oliver JL, Radnor JM, Rhodes BC, Faigenbaum AD, Myer GD. 2015. Relationships between functional movement screen scores, maturation and physical performance in young soccer players. J Sport Sci. 33(1): 11-19.
21. Malcata RM, Vandenbogaerde TJ, Hopkins WG. 2014. Using Athletes' World Rankings to Assess Countries' Performance. Int J Sports Physiol Perf. 9(1): 133-138.

22. Mills K, Baker D, Pacey V, Wollin M, Drew MK. 2017. What is the most accurate and reliable methodological approach to predicting peak height velocity in adolescents? A systematic review. J Sci Med Sport. 20: 572-577.

23. Mirwald RL, Baxter-Jones AD, Bailey DA, Beunen GP. 2002. An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc. 34(4): 689-694.

24. Nimphius S, Callaghan SJ, Spiteri T, Lockie RG. 2016. Change of direction deficit: A more isolated measure of change of direction performance than total 505 time. J Strength Cond Res. 30(11): 3024-3032.

25. Nimphius S, Geib G, Spiteri T, Carlisle D. 2013. Change of direction" deficit measurement in division I american football players. J Aust Strength Cond. 21(S2): 115-117.

26. Philippaerts RM, Vaeyens R, Janssens M, Van Renterghem B, Matthys D, Craen R, Bourgois J, Vrijens J, Beunen G, Malina RM. 2006. The relationship between peak height velocity and physical performance in youth soccer players. J Sports Sci. 24(3): 221-230.

27. Sayers MG. 2015. Influence of test distance on change of direction speed test results. J Strength Cond Res. 29(9): 2412-2416.

28. Shalfawi SA, Enoksen E, Tønnessen E, Ingebrigtsen J. 2012. Assessing test-retest reliability of the portable Brower speed trap II testing system. Kinesiology. 1:24-30.

29. Shrout PE, Fleiss JL. 1979. Intraclass correlations: uses in assessing rater reliability. Psychol Bulletin. 86(2): 420-424.

30. Svensson, M., Drust, B. 2005. Testing soccer players. J Sports Sci. 23(6):601-618.

31. 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 Cond J. 33(5): 29-39.

32. Weir JP. 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 19(1): 231-240.

33. Winter EM, Maughan RJ. 2009. Requirements for ethics approvals. J Sports Sci. 27(10): 985-986.
34. Wright MD, Hurst C, Taylor JM. 2016. Contrasting effects of a mixed-methods high-intensity interval training intervention in girl football players. J Sports Sci, 34(19), pp 1808-1815.
35. The English Premier League. Elite Player Performance Plan. In; 2011

Table 1. Descriptive statistics of elite youth football players (Mean ± SD)

	Age (y)	Height (cm)	Body Mass (kg)	10-m Sprint (s)
	Chronological age-groups			
U12 (n=19)	12.0 ± 0.3	153.8 ± 8.3	41.5 ± 5.7	1.87 ± 0.05
U14 (n=39)	13.6 ± 0.5	165.9 ± 10.6	52.1 ± 8.5	1.79 ± 0.08
U16 (n=26)	15.5 ± 0.5	175.8 ± 5.2	63.7 ± 6.9	1.74 ± 0.07
U18 (n=26)	17.4 ± 0.6	180.2 ± 6.0	72.4 ± 6.3	1.72 ± 0.05
	Biological age-groups			
Pre-PHV (n=33, 30%)	12.5 ± 0.7	154.6 ± 8.4	42.8 ± 0.6	1.86 ± 0.06
At-PHV (n=16, 15%)	13.8 ± 0.7	167.5 ± 4.6	51.7 ± 5.2	1.79 ± 0.07
Post-PHV (n=61, 55%)	16.2 ± 1.4	177.6 ± 5.0	67.6 ± 7.3	1.73 ± 0.03

Table 2. Chronological age-group ICC's, Typical error % and minimal detectable change % (MDC) for the M505 and CODD tests.

		U12	U14	U16	U18
Performance measures					
M505L	ICC	.48 (0.05-0.76)	.58 (0.33-0.76)	.31 (-0.08-0.62)	.37 (-0.01-0.66)
	Inference	Low	Moderate	Low	Low
	Typical error	2.8 (2.1-4.2)	2.8 (2.3-3.6)	3.2 (2.5-4.5)	2.8 (2.2-3.8)
	Inference	Large	Large	Large	Large
	MDC	7.8	7.8	8.9	7.8
M505R	ICC	.82 (0.59-0.93)	.51 (0.24-0.71)	.57 (0.24-0.78)	.68 (0.40-0.84)
	Inference	High	Moderate	Moderate	Moderate
	Typical error	2.2 (1.6-3.2)	2.8 (2.3-3.7)	2.5 (2.0-3.5)	2.0 (1.6-2.8)
	Inference	Moderate	Large	Large	Moderate
	MDC	6.1	7.8	6.9	5.5
CODD-L	ICC	.22 (-0.26-0.61)	.43 (0.14-0.66)	.19 (-0.21-0.53)	.44 (0.07-0.70)
	Inference	Low	Low	Very Low	Low
	Typical error	9.7 (7.3-14.7)	10.0 (8.1-13.0)	12.0 (9.3-17.0)	9.8 (7.6-13.7)
	Inference	Large	Large	Large	Large
	MDC	26.9	27.7	33.3	27.2
CODD-R	ICC	.71 (0.39-0.88)	.44 (0.14-0.66)	.40 (0.02-0.68)	.66 (0.37-0.83)
	Inference	Moderate	Low	Low	Moderate
	Typical error	7.5 (5.6-11.4)	9.7 (7.8-12.6)	9.0 (7.0-12.6)	7.1(5.5-9.9)
	Inference	Moderate	Large	Large	Moderate
	MDC	20.8	26.9	24.9	19.7

CI – Confidence interval

Table 3. Biological age-group ICC's, Typical error % and minimal detectable change % (MDC) for the M505 and CODD tests.

		Pre-PHV	At-PHV	Post-PHV
Performance measures				
M505L	ICC	.54 (0.25-0.74)	.68 (0.30-0.88)	.26 (0.01-0.48)
	Inference	Moderate	Moderate	Low
	Typical error	2.6 (2.1-3.5)	2.4 (1.7-3.7)	3.2 (2.7-3.9)
	Inference	Large	Moderate	Large
	MDC	7.2	6.7	8.9
M505R	ICC	.65 (0.40-0.81)	.78 (0.47-0.92)	.54 (0.33-0.70)
	Inference	Moderate	High	Moderate
	Typical error	2.7 (2.2-3.6)	2.0 (1.5-3.2)	2.6 (2.2-3.2)
	Inference	Moderate	Moderate	Large
	MDC	7.5	5.5	7.2
CODD-L	ICC	.33 (-0.01-0.60)	.60 (0.17-0.84)	.26 (0.01-0.48)
	Inference	Low	Moderate	Low
	Typical error	9.1 (7.3-12.2)	8.4 (6.1-13.3)	11.8 (10.0-14.6)
	Inference	Large	Large	Large
	MDC	25.2	23.3	32.7
CODD-R	ICC	.50 (0.20-0.72)	.79 (0.51-0.92)	.47 (0.24-0.64)
	Inference	Moderate	High	Low
	Typical error	9.2 (7.3-12.3)	6.4 (4.7-10.0)	9.3 (7.8-11.4)
	Inference	Large	Moderate	Large
	MDC	25.5	17.7	25.8

CI – Confidence interval.