Enhancing the Evaluation and Interpretation of Fitness Testing data within Youth Athletes

Till, K. A.; Morris, R.; Emmonds, S.; Jones, B.; Cobley, S.

Abstract:

Fitness testing is common practice within youth athletes. However, the interpretation of fitness data often occurs within chronological annual-age categories, resulting in athletes being (dis)advantaged due to age or maturity discrepancies. Instead, evaluating fitness performance against rolling averages may be more appropriate. This article presents a novel method for analyzing fitness testing data in youth athletes using Z-scores according to rolling averages for both chronological age and maturity status. This analysis technique allows the dual ability to interpret youth fitness performance according to age and maturation, enhancing accuracy of data interpretation for talent identification, development and strength and conditioning programming.

Key Words: Anthropometric, Fitness, Youth, Adolescent, Maturity, Strength & Conditioning

Introduction

Fitness testing of youth athletes is common practice whether within schools (18) or in sport academy programmes (6, 23). Sport scientists and strength and conditioning professionals implement a range of fitness tests to assess the anthropometric (e.g., height, mass) and fitness (e.g., speed, strength, power) qualities of youth athletes. The purpose and use of such measurement and assessment is to determine the fitness characteristics of an athlete (5, 8), evaluate an athlete's strengths and weaknesses (15, 21), provide objective data for talent identification and development (20, 27) and evaluate the effectiveness of training interventions (3).

There are a range of studies across multiple sports that have assessed the anthropometric and fitness qualities of youth athletes (e.g., Basketball, 28; Gymnastics, 19; Tennis, 7). However, a challenge for practitioners is to be able to accurately interpret such data to provide meaningful information to inform their practices (e.g., evaluating an athlete's needs). Recently, several articles (2, 9, 14, 21, 25, 26) have recommended analysis techniques to more accurately interpret such data. Collectively, these studies suggest the implementation of Z-scores to interpret an individual's testing and performance assessment. The advantage of Z-scores are that they provide an estimate on a given measurement relative to others who performed the same test. In other words, a score relative to the mean and standard deviation (SD) of a dataset. Z-scores on multiple parameters permit the ability to identify general or specific strengths and weaknesses related to an individual profile.

A Z-score is calculated as:

$$\text{Z-score} = \frac{(\text{athletes score} - \text{mean score})}{\text{SD}}$$

When interpreting a Z-score, a score of 0 on a given measure (e.g., agility, speed) represents a score reflecting the mean of the broader group. Therefore, the athlete would be performing on average with the comparison group. An athlete with a Z-score of plus or minus 1 reflects a score 1 SD above or below the mean respectively (2, 21), containing 68% of the group comparative scores. Meanwhile, a Z-score of ±2 represents a score of 2 SD's above or below the group mean, reflecting 95% of the comparative data (25). Therefore, positive or negative Z-score results show whether the athlete is performing above or below the average compared to the comparative group. The calculation and reporting of Z-scores are a useful technique for evaluating individual data relative to a specific broader population, whether that be a training group, team, representative academy or whole populations of participants.

While the implementation of Z-scores has increased in recent years based on recent recommendations (2, 9, 14, 21, 25, 26), there are however several limitations to consider with this analysis method. Firstly, Z-scores require a mean and SD to be derived from a group population and usually within youth athletes this might be presented within age categories. For example, research within youth rugby league players (21) used Z-scores by comparing individual athletes against a broad sample of data spanning three age categories (Under 13s to 15s). This was then extended by Cobley and colleagues (2) who – to improve Z score accuracy – made individual estimations relative to each annual-age category consecutively (i.e., Under 13s, 14s and 15s). In other words, a 14 year old player was compared with an age-matched Under 14 sample of scores. However, such methods are limited as they result in athletes being assessed within annual-age groups which may be confounded by other factors such as relative age (1, 24). Relative age refers to an individual's chronological age relative to a cut-off date applied to create chronological annual-age categories (e.g., 1st September in the United Kingdom). For example, players competing within the same age category are compared against the same comparative data even though their chronological age may differ by 364 days (e.g., 1st September vs. 31st August births in the United Kingdom). This annual-age grouping process has resulted in relative age effects (RAEs) whereby relatively older individuals are advantaged in participation and selection opportunities within sport due to their advanced age (1).

A second limitation with the current use of Z-scores relates to accurately tracking Z-scores over longer periods of time as mean and SD values change on an annual basis due to improvements in fitness performance alongside growth (6). Further, in youth sport, accuracy could be further confounded by maturity status, referred to as the timing and tempo of progress towards the mature adult state (11). Like RAEs, maturity biases exist within sport favouring earlier maturing individuals (12, 20). Therefore, due to the potential for inter-athlete variability within chronological age groups associated with maturation, alongside the known relationships between maturation and physical performance (12, 22), recent recommendations suggest comparing individuals relative to maturity status (bio-banded) groups (4). These methods begin to address limitations associated with purely age-based comparisons, by grouping according to maturity offset (16) or percentage of predicted adult height (10; reflecting maturity status). Although research presenting fitness performance according to maturity status is recently emerging (4, 13, 22), at this stage it may be most beneficial to assess individuals according to both age and maturity status matched cohorts to more accurately assess an athlete's current performance and development on youth sport parameters, a method which is uncommon to date.

To account for the changing trajectories that occur with age and partially account for potential maturity dynamics on performance, Romann & Cobley (17) recently applied the technique of corrective adjustments to better estimate 60 m sprinting performance in a large broader sample of Swiss male sprinters. The corrective numerical adjustments applied to individuals considerate of their decimal age on the day of the respective sprint event, meant that a 'net' corrected estimate could potentially correct for age disparities. Romann & Cobley (17) propositioned that corrective adjustments (i.e., rolling averages) could also be applied to team sport athlete's fitness testing profiles to better inform athlete evaluation. To date, no study has applied such a method to analyse and evaluate youth athlete's fitness testing profiles, which would support the interpretation of fitness testing data for talent identification, development and strength and conditioning purposes and potentially address some of the previously identified limitations of z-score analysis.

Therefore, the purpose of this article was to present a novel method of analyzing fitness testing data in youth athletes by establishing rolling averages according to both chronological age and maturity status. To achieve this, a hypothetical fitness testing dataset on youth athletes aged between 12 and 16 years old were used with case studies for ten hypothetical players presented to compare their fitness testing profiles relative to the broader dataset and display the potential of this method for evaluating player fitness performance.

Fitness Testing Battery

A hypothetical fitness testing dataset for two hundred and fifty male youth athletes aged between 12 and 16 years (13.55 ± 1.40 years) was developed to establish a comparative dataset (for establishing means and SDs). The fitness testing battery included anthropometric (i.e., standing height, sitting height and body mass) and fitness (i.e., speed, agility, lower body power and strength) characteristics. Speed was reported at 10 and 30 m distances; agility was reported based on the arrowhead agility run; lower body power was reported using countermovement jump (CMJ) height (m) and impulse ($N \cdot s^{-1}$); and strength was reported using the isometric mid-thigh pull (IMTP) peak force (N) and relative peak force ($N \cdot kg^{-1}$) measures.

Alongside the fitness testing scores, maturity offset (Years from peak height velocity; YPHV) was calculated for each participant using the Mirwald method (16). The Mirwald method (16) uses a gender specific prediction equation including height, sitting height, leg length (height – sitting height), body mass and chronological age and their interactions to calculate maturity offset (i.e., YPHV). The Mirwald regression equations are shown below or alternatively coaches can use the following website to input height, sitting height, body mass and date of birth calculate maturity offset https://kinesiology.usask.ca/growthutility/phv_ui.php.

In boys, the predictive equation was: Maturity Offset = -9.236 + 0.0002708·Leg Length and Sitting Height interaction - 0.001663·Age and Leg Length interaction + 0.007216·Age and Sitting Height interaction + 0.02292·Weight by Height ratio.

In girls, the predictive equation was: Maturity Offset = -9.376 + 0.0001882·Leg Length and Sitting Height interaction + 0.0022·Age and Leg Length interaction + 0.005841·Age and Sitting Height interaction - 0.002658·Age and Weight interaction + 0.07693·Weight by Height ratio

All data were stored in an Excel spreadsheet and used for further analysis. To show the analysis undertaken, ten hypothetical players were chosen as case studies with their fitness testing data presented in Table 1.

Insert Table 1 near here

Establishing Rolling Averages

To establish a rolling average for all anthropometric and fitness measures according to chronological age and maturity offset, scatter plots with regression equations were produced. Table 2 presents the regression equations for all anthropometric and fitness measures according to both chronological age and maturity offset. Figure 1 shows the scatterplot for chronological age vs. height. Figure 2 shows the scatterplot for maturity offset vs. height. These regression equations estimate a rolling average based on the broader sample and permit evaluation of a given fitness testing score specific to an exact chronological age or maturity offset time-point. For example, Player E's chronological age = 13.62 years and their estimated maturity offset = -1.61 years. The average height at this chronological age is 163.3 cm ((6.71 x 13.62) + 71.9) and maturity offset is 155.2 cm ((7.52 x -1.61) + 167.3). These regression equations allow individuals to be compared to the rolling average at their specific chronological age and maturity offset when calculating Z-scores rather than a group mean (see calculating Z-scores section).

Insert Table 2 near here

Insert Figure 1 near here

Insert Figure 2 near here

Establishing Standard Deviations

To calculate a Z-score for an athlete, a SD is required alongside the mean (or regression equation). Standard deviations were calculated for all fitness measures for the whole sample of youth soccer players (see Table 3). However, due to the large ranging dataset, with players aged 12-16 years, this accentuated the SD compared to the SD at each age group. This affects the calculation of a Z-Score resulting in a greater likelihood of an individual representing a Z-score between ±1. Therefore, the SD of each age group was calculated with an average of the SDs used as the SD in establishing Z-scores for each fitness test opposed to the SD of the whole dataset (see Table 3). Although there may be limitations of this method, the authors feel it is the best approach to establishing SD when using rolling averages for a wide range of ages as used in the current example.

Insert Table 3 near here

Calculating Z-Scores

To calculate Z-Scores the following formula was used:

$$\text{Z-score} = \frac{\text{(athletes score – mean score)}}{\text{SD}}$$

However, when applying rolling averages, the mean score was replaced by the regression equation shown in Table 2 for each fitness measure according to chronological age or maturity offset. For example, the Z-score equations for height according to chronological age and maturity offset are:

$$\text{Height Z-Score for Chronological Age} = \frac{\text{(athlete score – (6.71 x Age) + 71.8))}}{7.8}$$

$$\text{Height Z-Score for Maturity Offset} = \frac{\text{(athlete score – (7.52 x YPHV) + 167.3))}}{7.8}$$

Z-scores were calculated for all fitness measures for each case study player (Player A to J). Table 4 summarises the Z-scores for all players applying the individual scores (Table 1), regression equations (Table 2) and the standard deviations (Table 3) according to chronological age and maturity offset.

Insert Table 4 near here

Presenting and Interpreting Z-scores

Table 4 demonstrates the varying Z-score profiles for each case study player according to chronological age and maturity offset. These findings demonstrate the importance of considering both chronological age and maturity offset in the interpretation and evaluation of youth fitness testing data. To provide further analysis and examples of the presentation and interpretation of Z-scores, Player E (Figure 3) and F (Figure 4) are presented graphically with their results discussed below. This modified and novel way of analysing and summarising individual data allows the sport scientist or strength and conditioning professional to visualise an individual's anthropometric and fitness profile against chronologically aged and maturity matched values based on the regression equations from a similar broader population.

For Player E (Figure 3), the data shows that when compared for chronological age, scores are average (Z-score ~ 0) for 10 and 30 m sprint, CMJ jump height and IMTP relative peak force. However, below average performance is shown for the other measures with Z-scores of between -1 and -1.5. However, when compared to maturity offset, the Z-scores for all measures were above those of chronological age with 10 and 30 m sprint and CMJ jump height above average (Z-score ~ 0.5). This data suggests that Player E is a later maturing player who performs worse than his age-group peers but when maturity status is considered, is performing on par relative to players of an equivalent maturity status.

For Player F (Figure 4), the data shows that when compared to chronological age, the athlete is tall and heavy for his age (Z-score > 2) and performs above average on all measures except IMTP relative peak force. However, when compared against maturity offset, Z-scores in all measures are lower than those of chronological age. Height and mass are still above average given his maturity status (Z-score > 1); while sprint, agility and CMJ impulse are around the average score (Z-score ~ 0). However, for CMJ jump height and IMTP strength, athlete F is below average with relative peak force very low. The data for Player F indicates the player is tall and heavy for his age and maturity status with most performance indices above average for his age. However, his power production and strength are comparatively poor for age and maturity status and should therefore be a priority for onward strength and conditioning programming, particularly given his large body size.

Overall, the procedural methods highlighted allow youth athletes at their various development stages (chronological ages and maturity status) to be more accurately analysed and interpreted given broader data that might be available for comparison. The above example also illustrates how comparison of athletes by chronological age may be erroneous when working with youth athletes. Therefore, this data analysis has additional value in interpreting and evaluating fitness data in youth athletes and could be used as a talent identification tool or by strength and conditioning coaches in their needs analysis procedures to ensure that optimal programming can be applied to individual athletes.

Practical Applications

The above explanation has detailed how rolling averages for anthropometric and fitness qualities can be used to establish Z-score estimates to enable a direct individual to individual, individual to data-set (i.e., mean and SD), or individual to longitudinal trajectory comparisons. Below is a summary of the procedures for practically applying this method whether working as a coach, sport scientist or strength and conditioning professional in practice.

1. Develop and conduct a fitness testing battery with your athletes. A larger dataset allows for the better establishment of rolling averages.
2. Includes a measure of maturity status within the fitness testing battery when working with youth athletes. This maturity measure could include the Mirwald method (16) presented in this article whereby height, sitting height, body mass and date of birth can be entered into a regression equation to estimate maturity offset.
3. Instead of calculating age grouped means, establish a regression equation (i.e., rolling average) for chronological age and maturity status by applying the trendline function to a scatterplot in Microsoft Excel. This can be achieved by selecting the data within your Excel spreadsheet (e.g., chronological age and 20 m speed), selecting 'Insert' then 'Scatterplot'. When the scatterplot is shown, right click a point within the graph and select 'Add Trendline' and then select 'Display equation on chart'. This will then show your regression equation for your dataset.
4. Calculate the SD of your dataset. Consider the range of data within your sample before deciding which SD is appropriate to use. Standard deviations can be calculated in Excel using the '=STDEV' function.
5. Calculate Z-scores for your data using the formula:

   $$\text{Z-Score} = \frac{(\text{athletes score} - \text{regression equation})}{\text{SD}}$$

6. Use graphs (bar or radar) to visually summarise your data according to chronological age and maturity status for individual athletes. These can be done by selecting the z-score data for your athletes and then selecting 'Insert' and then the 'Graph Type'. This helps present a visual representation of your Z-Score analysis as in Figure 3 and 4 within this manuscript.

Limitations

Although the method proposed in this article progresses on current practices for evaluating and interpreting fitness testing data within youth athletes, practitioners should still consider the limitations. Firstly, a reasonably large dataset is required to calculate rolling averages over multiple age categories. However, such data collection may heighten the SD of a dataset and therefore the consideration of the calculation of the SD to calculate Z-scores should be considered. Secondly, the current method calculated maturity status using maturity offset, which may have limitations and an alternative could be to use percentage of predicted adult height based on the recommendations by Cumming and colleagues (4) utilising the Khamis-Roche method (10). The Khamis-Roche method includes collection of parental heights and may be considered as a measure of maturation status in addition or alternatively to the Mirwald method (16).

Summary

With the increase in popularity of fitness testing in youth athletes along with the increased understanding about maturational variability, it is more important than ever that practitioners appropriately analyze and accurately evaluate and interpret data related to youth or developmental athletes. This article proposes a novel methodology to more accurately assess anthropometric and fitness variables, considerate of continuous chronological age and maturity offset, using a rolling average method rather than traditional age grouping processes. This analysis technique allows the dual ability of both chronological age and maturity offset to be considered to give detailed insight of a youth athletes performance relative to a large and broader cohort, especially during this key developmental stage. The authors therefore recommend this method of evaluating and interpreting fitness testing data within a range of sports and settings (i.e., schools and academy programmes) to provide greater accuracy within talent identification, player development as well as individualised strength and conditioning programming.

Conflicts of Interest

There are no conflicts of interest

References

1. Cobley, S, Baker, J, Wattie, N, and McKenna, J. Annual age-grouping and athlete development. *Sports Med* 39:235-256, 2009.
2. Cobley, S, Till, K, O'Hara, J, and Cooke, C. Variable and changing trajectories in youth athlete development: Further verification in advocating a long-term inclusive tracking approach. *J Strength Cond Res* 28:1959-1970, 2014.
3. Contreras, B, Vigotsky, AD, Schoenfeld, BJ, Beardsley, C, McMaster, DT, Reyneke, JH, and Cronin, J. Effects of a Six-Week Hip Thrust vs. Front Squat Resistance Training Program on Performance in Adolescent Males: A Randomized Controlled Trial. *J Strength Cond Res* 31:999-1008, 2017.
4. Cumming, S, Lloyd, R, Oliver, JL, Eisenmann, J, and Malina, R. Bio-banding in Sport: Applications to Competition, Talent Identification, and Strength and Conditioning of Youth Athletes. *Strength Cond J* 39,34-47, 2017.
5. Darrall-Jones, J, Jones, B, and Till, K. Anthropometric and Physical Profiles of English Academy Rugby Union Players. *J Strength Cond Res* 29:2086-2096, 2015.
6. Emmonds, S, Till, K, Jones, B, Mellis, M, and Pears, M. Anthropometric, speed and endurance characteristics of English academy soccer players: Do they influence obtaining a professional contract at 18 years of age? *Int J Sports Sci Coach* 11:212-218, 2016.
7. Fernandez-Fernandez, J, Ulbricht, A, and Ferrauti, A. Fitness Testing of tennis players: how valuable is it? *Br J Sports Med 48:2*2-31, 2014.
8. Jones, B, Till, K, Barlow, M, Lees, M, O'Hara, J, and Hind, K. Anthropometric and three-compartment body composition differences between Super League and Championship rugby league players: Considerations for the 2015 season and beyond. *PLoS One*. 10, https://doi.org/10.1371/journal.pone.0133188, 2015.
9. Jones, B, Till, K, Manley, A, and McGuigan, MR. A Multidisciplinary Approach to the Profiling and Interpretation of Fitness Testing Data: A Case Study Example. *J Aust Strength Cond* 25: 2017.
10. Khamis, HJ, and Roche, AF. Predicting adult stature without using skeletal age—the Khamis-Roche Method. *Pediatrics* 95:504–507, 1995.
11. Malina, RM, Bouchard, C, and Bar-Or, O. *Growth, Maturation, and Physical Activity* (2nd ed.). Champaign, IL: Human Kinetics, 2004.
12. Malina, RM, Rogol, AD, Cumming, SP, Coelho e Silva, MJ, and Figueiredo, AJ. Biological maturation of youth athletes: assessment and implications. *Br J Sports Med* 49:852-859, 2015.
13. McCunn, R, Weston, M, Hill, JKA, Johnston, RD, and Gibson, NV. Influence of Physical Maturity Status on Sprinting Speed Among Youth Soccer Players. *J Strength Cond Res* 31:1795-1801. 2017.

14. McGuigan, MR, Cormack, SJ, and Gill, ND. Strength and power profiling of athletes: selecting tests and how to use the information for program design. *Strength Cond J* 35:7-14, 2013.
15. McGuigan, M. Evaluating Athletic Capacities. In: *High Performance Training for Sports*. Joyce D, and Lewindon D, eds. Champaign, Illinois: Human Kinetics, 3-14, 2014.
16. Mirwald, RL, Baxter-Jones, ADG, Bailey, DA, and Beunen, GP. An assessment of maturity from anthropometric measurements. *Med Sci Sport Exerc* 34:689-694, 2002.
17. Romann, M, and Cobley, S. Relative age effects in athletic sprinting and corrective adjustments as a solution for their removal. *PLoS One* 10: e0122988. doi: 10.1371/journal.pone.0122988, 2015.
18. Sawczuk, T, Jones, B, Scantlebury, S, Weakley, J, Read, D, Costello, N, Darrall-Jones, J, Stokes, K, and Till, K. Between-Day Reliability and Usefulness of a Fitness Testing Battery in Youth Sport Athletes: Reference Data for Practitioners. *Measure Phys Ed Exerc Sci* 22:11-18, 2017.
19. Sleeper, MD, Kenyon, LK, and Casey, E. Measuring fitness in female gymnasts: The gymnastics functional measurement tool. *Int J Sports Phys Ther* 7:124-138, 2012.
20. Till, K, Cobley, S, O'Hara, J, Brightmore, A, Chapman, C, and Cooke, C. Using Anthropometric and Performance Characteristics to Predict Selection in Junior UK Rugby League Players. *J Sci Med Sport* 14:264-269, 2011.
21. Till, K, Cobley, S, O'Hara, J, Chapman, C, and Cooke, C. An individualized longitudinal approach to monitoring the dynamics of growth and fitness development in adolescent athletes. *J Strength Cond Res* 27:1313-1321, 2013.
22. Till, K, and Jones, B. Monitoring anthropometry and fitness using maturity groups within youth rugby league. *J Strength Cond Res* 29:730-736, 2015.
23. Till, K, and Jones, B. Fitness Testing for Rugby. Chapter 2. In Till, K. & Jones, B. eds. *The Science of Sport: Rugby*. Crowood Press. 32-40, 2015.
24. Towlson, C, Cobley, S, Midgely, AW, Garrett, A, Parkin, G, and Lovell, R. Relative Age, Maturation and Physical Biases on Position Allocation in Elite-Youth Soccer. *Int J Sports Med* 38:201-209, 2017.
25. Turner, A. Total Score of Athleticism: a strategy for assessing an athlete’s athleticism. *Prof Strength Cond J* 23:14-17, 2014.
26. Turner, A, Brazeir, J, Bishop, C, Chavda, S, Cree, J, and Read, P. Data Analysis for Strength and Conditioning Coaches: Using Excel to Analyze Reliability, Differences, and Relationships. *Strength Cond J* 37:76-83, 2016.

27. Woods, CT, Cripps, A, Hopper, L, and Joyce, C. A comparison of the physical and anthropometric qualities explanatory of talent in the elite junior Australian football development pathway. *J Sci Med Sport* 20:684-688, 2017.
28. Ziv, G, and Lidor, R. Physical attributes, physiological characteristics, on-court performances and nutritional strategies of female and male basketball players. *Sports Med* 39: 547-568, 2009.

Table Legend

Table 1. Fitness Testing Data for Case Study Players

Table 2. Regression equations for chronological age and maturity offset within 12-16 year old youth soccer players

Table 3. Standard deviations of the fitness tests for all players and at each annual-age group

Table 4. Z-Scores for Case Study Players using the rolling average approach

Figure Legend

Figure 1. Scatterplot for chronological age vs height for youth soccer players aged 12-16 years

Figure 2. Scatterplot for maturity offset (YPHV) vs height for youth soccer players aged 12-16 years

Figure 3. Z-scores for Player E according to chronological age and maturity offset

Figure 4. Z-scores for Player F according to chronological age and maturity offset