Age-related reference intervals for physical performance test outcomes relevant to male youth Middle Eastern football players

Heading title: Age-specific physical performance ranges

Authors: Lorenzo Lolli[1,2], Warren Gregson[2], Daniele Bonanno[1], Sami Kuitunen[1], Valter Di Salvo[1,3]

[1] Aspire Academy, Football Performance & Science Department, Doha, Qatar
[2] Department of Sport and Exercise Sciences, Institute of Sport, Manchester Metropolitan University, Manchester, UK
[3] Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome, Italy

Corresponding Author:
Lorenzo Lolli, Football Performance and Science Department, Aspire Academy, PO Box 22287, Doha, State of Qatar.
e-mail: Lorenzo.Lolli@aspire.qa

Abstract word count: 197
Text only word count: 3777
Tables: 4
Figures: 6

Abstract

Purpose: To develop age-specific reference intervals for physical performance test outcomes relevant to male youth Middle Eastern football players. **Methods:** We analysed mixed-longitudinal data (observations range: 1751 to 1943 assessments) from a sample of 441 male, youth outfield football players (chronological age range: 11.7 to 18.4 y) as part of the Qatar Football Association and Aspire Academy development programme over fourteen competitive seasons. Semi-parametric generalized additive models for location, scale and shape (GAMLSS) estimated age-specific reference centiles for 10-m sprinting, 40-m sprinting, countermovement jump (CMJ) height, and maximal aerobic speed variables. **Results:** The estimated reference intervals indicated that the distribution of the physical performance test scores increased monotonically and non-linearly with advancing chronological age for sprinting and CMJ outcome measures, reaching a plateau after 16 years common to each of these performance variables. The maximal aerobic speed median score increased substantially until ~14.5 y, with the non-linear trend flattening off towards relatively older chronological ages. **Conclusions:** We developed age-related reference intervals for physical performance test outcomes relevant to youth Qatari football players. Country-wide age-specific reference intervals can assist in the longitudinal tracking of the individual player's progresses over time against benchmark values derived from the reference population.

Keywords: Football; player tracking; Middle East; CMJ, sprint; GAMLSS

Introduction

Insights on organisational processes and working practices in youth academies of professional associational football (soccer) clubs from around the world substantiated how youth football academies strive for developing players for the first team.[1] Long-term athlete development is a multifaceted process characterised by different phases in the pursuit of high performance.[2] Applied sports science service provision, with a particular reference to physical conditioning, performance assessment, player monitoring and computer-based match-analysis become pivotal to support talent detection, identification, selection, and development processes.[2,3] Talent *detection* refers to the general process of discovery of potential prospects who are currently not involved in any sports programme, whereas talent *identification* denotes recognizing youth subjects already competing in sports with the potential to become professional athletes.[2] More generally, talent *development* relates to the provision of optimal conditions for youth athletes to realize their sporting potential.[2] In this context, ongoing assessment of anthropometric, physiological, and physical performance attributes is central to support the long-term athlete development process.[4]

From a real-world perspective, information from physical performance testing batteries can serve as guidance for coaches and key stakeholders in professional team-sport environments to guide decisions relevant to optimal youth player developmental strategies.[2] The general attention devoted to the tracking of anthropometric, physical, and physiological determinants of football performance stems from the need to obtain data that can be utilized throughout the course of any developmental phase to continually gauge performance levels towards the youth-to-senior career transition phase.[2,5] Context-specific physical performance assessments, together with ongoing screening of growth and development,[6,7] can offer an advantage for early recruitment strategies while fine-tuning talent identification and development processes to guarantee principled investments and remain competitive on a sport level.[2,3,8] Importantly, the longitudinal tracking of anthropometric, physical, and physiological determinants of football performance can support the development of age-specific reference intervals to benchmark individual player test scores against the reference population.[9] Nevertheless, most youth football research has examined developmental changes in physical performance outcomes from small-scale samples of players over relatively curtailed age ranges.[10-12] Likewise, larger-scale investigations that examined participants over the entire typical age range of football development programmes were limited to estimating only single mean trajectories for different proxy measures of physical performance.[13,14] In particular, studies in this realm did not explore the development of age-related reference intervals for benchmarking physical test performance.

In a sporting academy setting, the construction of reference intervals is relevant to facilitate the interpretation of real-world physical performance data for tracking the individual youth player at a given chronological age throughout different developmental stages.[15] With examples from the biomedical literature in mind, reference intervals can also assist coaches and practitioners in rationalizing whether a new player meets minimum criteria for entry in the academy and conducting principled interpretations of progresses during the developmental programme.[16] Nevertheless, the definition of age-related reference intervals in the sports and exercise sciences remains limited to physical fitness outcomes in general and clinical populations from Western countries.[17-20] A recent study illustrated physical fitness standards in a sample of 765 campus football children aged 9 to 11 from China, although of limited generalisation and relevance to the context of professional football academies environments.[21] No study, however, illustrated reference intervals for a population of male, youth academy football players. Therefore, we aimed to develop the first age-related reference intervals for physical performance test outcome measures relevant to male youth football players from the Middle East.

Methods

Participants

The study sample included physical performance assessments data available for a sample of n=441 male, full-time, youth outfield football players enrolled as academy student-athletes (chronological age range: 11.7 to 18.4 y; standing height range: 134.3 to 190.3 cm, body mass range: 28.9 to 78.7 kg) over fourteen competitive seasons. The ethnicity of student-athletes in the present investigation was predominantly Middle Eastern Arab (i.e., ~94% of study participants).[22] The general schedule for this sample of full-time, student-athletes consisted of six school classes from 07:30 until 15:30 and double training sessions from 10:30 to 12.00 and 16:00 to 18:00 on Sunday, Monday, and Wednesday. School classes from 08:00 until 15:20 and one training session in the afternoon from 15:30 to 17:30 were scheduled on Tuesday, and school classes from 07:25 until 13:30 only on Thursday. Study participants competed in official matches with their respective clubs during weekend days, with a duration of 90 min for U16, 80 min for U15, and 60 min for U14 and lower player age categories. Medical, anthropometric, and performance test outcome measurements collected in student-athletes as part of the regular annual screenings were retrieved from the Academy records, anonymized, and analyzed to address the purpose of this investigation.[6,7] Parents and guardians signed an informed consent form at the beginning of each academy season prior to any routine medical and performance screening collection process to permit the use of data for both service provision and research purposes. This retrospective study was approved by the Aspire Zone Foundation Institutional Review Board, Doha, State of Qatar (protocol number: E202008009).

Design and procedures

The present retrospective investigation examined mixed-longitudinal field-based physical performance testing data collected from youth outfield football players measured on a least one occasion (annual screening range: 1 to 12 assessment visits). A mixed-longitudinal design combines features of both cross-sectional and longitudinal designs and represents a valuable option to maximising practical benefit for the estimation of measurement distance standards in terms of time and cost resources.[23]

Youth players in the present study sample were assessed on distinct occasions every three months over the competitive season. Tapering of training programmes was scheduled 3–5 days preceding each testing session.[22] All assessments took place under standardized clothing, running or sport-specific shoes depending on the assessment task and venue, and environmental conditions as described for previous measurement reliability evaluations involving participants from this population.[24] Following a standardized warm-up, linear speed was evaluated by recording 10-m split times measured to the nearest 0.01s to determine the best time from 2 maximal 40-m trials using electronic timing gates (Swift Performance Equipment, Lismore, Australia).[24] Players were instructed to start with their front foot half one metre behind the first timing gate and to sprint as fast as possible over the full 40-m distance.[24] Lower limb explosive strength was assessed using a force plate (Kistler 9286AA, Kistler Instrument Corp., Winterthur, Switzerland), with countermovement jump (CMJ) height selected as a proxy measurement of interest. CMJ height was determined from the best of three trials separated by 25-s of passive recovery.[24] Each player was instructed to keep their hands on their hips with the depth of the counter movement self-selected. Each trial was validated by visual inspection to ensure each landing was without significant leg flexion prior to final test score determination. Maximal aerobic speed (MAS) was determined using a continuous incremental field running test assessment.[24] The assessment begins at a starting running speed of 8.5 km·h^{-1} and increasing by 0.5 km·h^{-1} each minute until volitional exhaustion. The average velocity of the last stage a player achieved was recorded as the performance score, with the MAS (km·h^{-1}) calculated as follow: MAS = S + (t/60 × 0.5), where S is the last completed speed in km·h^{-1}and t is the time expressed in units of seconds, if the stage was not completed. The estimated standard error of the measurement for 10-m sprinting, 40-m sprinting, CMJ height, and MAS was ± 0.042 s (95% confidence interval [CI], 0.036 to 0.051 s), ± 0.102 s (95% CI, 0.086 to 0.124 s), ± 1.6 cm (95% CI, 1.4 to 1.9 cm), and ± 0.51 km·h^{-1} (95% CI, 0.43 to 0.61 km·h^{-1}), respectively.[7] The assessment venues were an indoor synthetic track (i.e., concrete overlaid with rubber) and soft artificial turf of synthetic fibres in the last season only of the fourteen competitive seasons examined in this study. With a subsample of n=25 student-athletes from this population assessed twice in a random order, one week apart, either first on an indoor synthetic track or soft artificial turf and *vice versa*, two one-sided tests (TOST) sensitivity analyses anchored against minimal detectable change values for 10-m sprinting, 40-m sprinting, and CMJ height performance[7] suggested measurement equivalence regardless of assessment venue. The mean difference in soft artificial turf *versus* indoor synthetic track 10-m sprinting, 40-m sprinting, and CMJ height performance was 0.021 s (95% CI, - 0.004 to 0.046 s), 0.002 s (95% CI, - 0.044 to 0.048 s), -0.4 cm, (95% CI, -1.7 to 0.9 cm), respectively.

Statistical analysis

Informed by the criteria and recommendations relevant to method selection for growth standards development,[25] semi-parametric generalized additive models for location, scale and shape (GAMLSS) estimated age-related reference interval for physical performance test outcomes.[26] The *lms* function determined the smoothing degrees of freedom and the distribution of physical performance data based on the model minimising the global deviance score.[26] Models estimated nine reference centiles at 0.38th, 2.27th, 9.12th, 25.25th, 50th, 74.75th, 90.88th, 97.72th, and 99.62th values spaced ⅔ of a standard deviation score apart.[26] Postestimation diagnostics were conducted according to the visual inspection of the worm plot prior to final reference intervals estimation.[26] Analyses were conducted using R (version 3.6.3, R Foundation for Statistical Computing) and reference intervals were estimated using the *gamlss* package.[26]

Table 1 about here

Figure 1 about here

Table 2 about here

Figure 2 about here

Results

Age-related reference intervals for 10-m sprinting, 40-m sprinting, CMJ height, and maximal aerobic speed mixed-longitudinal data are presented in Figures 1-4, respectively. The functions for the GAMLSS models estimated references intervals following Box-Cox Cole-Green distributions for 10-m sprinting and 40-m sprinting variables, and Box-Cox power exponential distributions for CMJ and MAS variables (Tables 1-4). Test performance scores for sprinting and CMJ outcome measures increased monotonically and non-linearly with advancing chronological age, reaching a plateau after 16 y common to each variable. The MAS median score increased substantially until ~14.5 y, with the non-linear trend flattening off towards relatively older chronological ages. Model residuals for each model were well-behaved, and visual inspection of the worm plots indicated adequate model fit (Figure 5).

Table 3 about here
Figure 3 about here
Table 4 about here
Figure 4 about here

Discussion

We provided the first age-related reference intervals for physical performance test outcomes relevant to male youth football players from Qatar. Within the context of national football federation and sporting academy settings, the construction of reference standards can serve as a tool for coaches and support staff involved in long-term development processes to define performance benchmarks for talent identification and facilitate the longitudinal tracking of the youth football player.

Figure 5 about here

Outcomes from recent investigations exploring practices on youth player identification and development from football academies worldwide revealed professional organizations strive to develop players for the first team[1] driven by a club-based strategy.[5] As part of their strategy, football academies embraced the integration of applied sports science methods to advance talent identification and developmental processes.[1,3] The definition of reference benchmarks following methodical, state-of-the-art statistical methodologies[25] relevant to address these processes remained unexplored in this field of research. In the medical realm, the construction of age-related reference intervals generally aids medical decisions in the diagnosis and treatment pathways for subjects whose measurement relative standing exceeds nor falls below a cut-point of clinical relevance.[16] In sports, age-related reference intervals for physical performance test outcomes can be useful for coaches, managers, and executives to inform more objective value judgments on an individual player developmental pathway.[15] In the context of our study, recommendations from a group of biostatisticians and growth experts convened by the World Health Organization informed considerations on method selection for developing age-related reference intervals for physical performance test outcomes.[25] Following a comprehensive review of 30 existing statistical procedures, GAMLSS, fractional polynomials and exponential transformations, and the system of frequency curves by methods of translation procedures met methodological criteria for adequate growth standards construction.[25] However, no previous investigation formally developed reference intervals for physical performance test outcomes relevant to male youth football players from a national sporting academy following GAMLSS procedures.[25] Accordingly, we could not compare our results with findings from other investigations in other populations of youth football players due to the lack of similar information based on similar physical assessment methods, chronological age ranges, and examination of different types of distributions and link functions within GAMLSS. Researchers in sports and exercise sciences constructed reference intervals mainly for cardiorespiratory and physical fitness outcome measures in adolescents, with more recent illustrations of GAMLSS procedures applied to summarise measurements from general and clinical populations.[17-20] The only investigation in male youth football is limited to the definition of physical fitness standards in a sample of 765 campus football children aged 9 to 11 from China whose generalisability is of limited relevance to the context of professional football academies environments.[21] The recent study of Datson et al.[15] illustrated another application of GAMLSS methods to estimate reference standards for key performance test outcomes in a female population as part of the English Football Association's national development programme. Collectively, our first description of age-related reference intervals for physical performance test outcomes in a population of male youth football players provided a meaningful contribution to the existing knowledge base that highlighted the value of defining reference standards for aiding talent identification and development processes.

Our study is novel for providing insights valuable to guide rationalised interpretations of longitudinal development patterns for proxy measures of physical performance such as 10-m sprinting, 40-m sprinting, CMJ height, and MAS. Test performance scores for sprinting and CMJ performance reached a plateau after 16 y common to each variable, whereas the median of the distribution of maximal aerobic speed scores flattened off at relatively earlier developmental stages (Figures 1-4). Considering outcomes for 10-m sprinting and CMJ performance, the plateau we observed in our study appeared to coincide with breakpoints described in European samples of youth football players.[13,14] Nevertheless, the fact these investigations used alternative statistical procedures involving linear segmented models precludes any direct comparison with our study outcomes. Other investigations in paediatric exercise sciences explored longitudinal growth curves for physiological and physical performance measures in relation to changes in body size over time. Reports for children and adolescents from general populations revealed that the largest increases in physiological attributes, such as maximal ($\dot{V}O_{2max}$) or peak ($\dot{V}O_{2peak}$) oxygen uptake, occurred approximately at the time of the adolescent growth spurt or age at peak height velocity.[27] Youth football players from Belgium (n=36) reached peak development in explosive strength, cardiorespiratory endurance and anaerobic capacity at peak height velocity followed by a plateau in the rate of growth for approximately 12 to 18 months subsequent to the adolescent growth spurt event.[28] Knowledge of the height growth spurt and radius-ulna-short bones ossification timings generally occurring at 13.6 y (95% CI, 13.5 to 13.7 y)[6] and 15.1 y (95% CI, 14.9 to 15.3 y)[29] in this population may, in part, provide a logical explanation to the nature of the trends we observed in our study. In applied settings similar to our study context, formal evaluation and understanding of the trends for each proxy measure of physical performance, as we described in our study, can, therefore, aid coaches and support staff in defining differential player developmental plans that could promote adaptations beyond concurrent growth-mediated effects.[29]

Figure 6 about here

When generalising these insights from an applied perspective, formal benchmarking of a player physical performance is fundamental to the definition of reference standards for establishing minimum criteria for the individual player to pursue a professional career.[15] To illustrate this from a practical standpoint, we shall consider physical performance data for a subject from our study sample assessed on multiple visits between 12.4 to 17.1 y (Figure 6). The skeletal age determined as per the Tanner-Whitehouse II (TW-II) protocol[6] at the time of the first assessment was 11.9 y. This student-athlete followed a normal pattern of development, with TW-II skeletal ages of 12.4 y, 14.7 y, 15.5 y, and 16.5 y at chronological ages of 13.4 y, 14.5 y, 15.4 y, and 16.5 y, respectively. In keeping with generalisations in the clinical literature, the 91st centile for sprinting variables and the 9th centile for jumping and aerobic endurance variables shall define *minimum acceptable* test scores for this population. In simple words, these centiles define normal ranges encompassing ~90% of future similar test scores in this population whereby an individual test score beyond these cut-points would occur in fewer than 10 student-athletes in 100 *above* nor *below* these limits, respectively. Visual inspection of this student-athlete test scores for 10-m sprinting and 40-m sprinting revealed performances within normal ranges for the reference population (Figure 6 A, B). When entering the academy as student-athlete at 12.4 y, the 10-m sprinting and 40-m sprinting performance score where below nor around the 25th centile for both measurements. Translated into lay terms for coaches and key stakeholders in professional team-sport environments, this player recorded an individual 10-m sprinting or 40-m sprinting performance that would occur in fewer than 25 student-athletes in 100 *below* this limit. Likewise, the CMJ height score of 31 cm at 12.4 y lied approximately above the 75th centile suggesting that this student-athlete recorded an individual performance that would occur in fewer than 25 players in 100 *above* this limit (Figure 6C). Sprinting for this player generally showed a pattern of improvement until ~15.5 y, with CMJ height score increasing steadily until reaching an *atypically high* performance of 51.9 cm that exceeded the 91st centile at 17.1 y. Conversely, the trend for MAS showed a relatively flat pattern (Figure 6D) consistent with considerations at the population level. From a coaching perspective, assessment of test scores for each performance variable at a given visit suggested the performance for this student-athlete was well-within the normal ranges and reached atypically high CMJ scores towards the youth-to-senior transition phase. Taken together, visual inspection of raw data scores interpreted against population-specific reference intervals can be of practical relevance for coaches and practitioners to inform optimal youth athlete physical conditioning strategies from entry in the academy until the transition from youth-to-senior competition.

Notwithstanding the fact that we followed methodological recommendations for reference charts development,[25] our study is not without limitations. First, we constructed reference intervals using performance test data gathered from players selected for a country-wide development programme based on criteria beyond sole physical performance aspects. This may have contributed to render our estimations prone to any effect of talent *selection* and *identification* practices per se, although the nature of our dataset itself is fundamental to the definition of reference standards consistent with the notion of sporting excellence relevant to our context. Nevertheless, we highlighted our study intended to provide tracking solutions limited to support developmental processes for youth players from our reference population whose extent is not generalizable to match performance considerations nor applications beyond our study context. The development of reference centiles for performance test outcomes from other populations of youth football players thus warrants considerations to allow contextual comparisons between different countries.[15] Second, and despite performance assessments data gathered over fourteen competitive seasons, we derived age-related reference intervals over a relatively curtailed chronological age range. With the example of Datson et al.[15] in female football, availability of consistent youth-to-senior physical performance data could have maximised the practical value of our estimations and broadened their spectrum of application. Nonetheless, we performed formal estimation of physical performance reference intervals using measurements gather over a chronological age range consistent with the reality of a professional football academy environment. While a particular strength of our study, the development of age-related reference intervals based on performance assessments collected over more than a decade deserves consideration of practical factors that pose a number of challenges for future similar studies in this field. Specifically, coping with the recent and ongoing evolution in testing equipment solutions and the inevitable staff turnover requires pragmatic adherence to standardized testing protocols with no changes in processing procedures to ensure integrity and methodological continuity of performance data collection for pursuing the development of reference charts that could be translated into advanced business solutions at country-level as in the present study. Practical aspects regarding performance assessment facility and terrain characteristics also deserve attention as concrete examples relevant to the last season only of the fourteen competitive seasons examined in this study. Nonetheless, in our context, the equivalence testing outcomes from our sensitivity analyses revealed no evidence of a tangible influence of terrain characteristics on sprinting and CMJ performance.

Likewise, the limited amount of data for proxy measures of biological maturation that could be matched with serial physical performance assessments[7] was not adequate for a precise estimation of reference intervals expressed by a different predictor variable other than chronological age.[9,30] However, evidence from a recent investigation exploring the integration of skeletal age and performance outcome measures in this population of youth Middle Eastern football players suggested differences in relative skeletal maturity, determined as TW-II skeletal age *minus* chronological age, accounted for ~ 1% to 9% only of the between-subject variability in 10-m sprinting, 40-m sprinting, CMJ, and MAS performance.[7] Third, our sampling composition and characteristics deserve consideration. To address the practical demands of our context, we used a mixed-longitudinal dataset to estimate age-related references intervals for the physical performance test outcomes selected according to academy-based criteria.[3] With the availability of information for subjects measured once and others more than once, using mixed-longitudinal data is a particular strength of our study to address ethical, cost-related, and practical limitations of typical cross-sectional and longitudinal study designs[23] and meet fundamental sample size requirements for reference charts development.[9] The adoption of a cross-sectional research design requires a relatively larger number of study participants and provides only information about distance that may be comparable to estimations conducted in smaller-scale studies.[23] Despite these advantages, the nature of our sample composition precluded from conducting a formal estimation of unbiased pointwise confidence bands for each centile curve.[30] Methodological studies in applied biostatistics illustrated bootstrapping procedures using GAMLSS yet applicable to cross-sectional research designs only.[30] Incorrect treatment of mixed-longitudinal data can result in estimating spuriously inflated standard errors yielding potentially over-precise confidence bands for a given centile curve.[30] Therefore, the lack of established procedures limited a robust description of the uncertainty surrounding the point centile estimates at a given chronological age.

Practical applications

- The construction of age-related reference intervals can leverage interpretations of physical performance assessments useful for coaches, managers, and executives to inform more objective value judgments on an individual player development.
- Age-specific intervals can assist coaches and performance staff in the longitudinal tracking of the individual player's progresses towards the transition to senior competition against benchmark values derived from the reference population.
- Reference charts can be translated into dynamic business solutions to facilitate ongoing tracking of the individual academy player.

Conclusions

We illustrated the first age-related reference intervals for physical performance test relevant to male youth football players from a Middle Eastern country. From a real-world perspective, country-wide age-related intervals for physical performance test outcomes can serve as a tool to address the practical demands of national federations and, accordingly, professional football academies in the pursuit of developing players for the first team.

Acknowledgments
The authors would like to express their gratitude to the staff of the Football Performance & Science Department at Aspire Academy for the invaluable support in the collection of the routine assessments data along the years. The authors wish to thank the two anonymous referees whose comments and suggestions were invaluable to improve the manuscript.

Disclosure No potential conflict of interest was reported by the author(s).

Funding No funding is associated with this study.

References

1. Ford PR, Bordonau JLD, Bonanno D, et al. A survey of talent identification and development processes in the youth academies of professional soccer clubs from around the world. *J Sports Sci.* 2020;38(11-12):1269-1278.
2. Lidor R, Côté J, Hackfort D. ISSP position stand: To test or not to test? The use of physical skill tests in talent detection and in early phases of sport development. *Int J Sport Exerc Psychol.* 2009;7(2):131-146.
3. Vaeyens R, Lenoir M, Williams AM, Philippaerts RM. Talent identification and development programmes in sport : current models and future directions. *Sports Med.* 2008;38(9):703-714.
4. Hulse MA, Morris JG, Hawkins RD, Hodson A, Nevill AM, Nevill ME. A field-test battery for elite, young soccer players. *Int J Sports Med.* 2013;34(4):302-311.
5. Lundqvist C, Gregson W, Bonanno D, Lolli L, Di Salvo V. A worldwide survey of perspectives on demands, resources, and barriers influencing the youth-to-senior transition in academy football players. *Int J Sports Sci Coach.* 2022;0(0):17479541221135626.
6. Lolli L, Johnson A, Monaco M, Cardinale M, Di Salvo V, Gregson W. Tanner–Whitehouse and Modified Bayley–Pinneau adult height predictions in elite youth soccer players from the Middle East. *Med Sci Sports Exerc.* 2021;53(12):2683-2690.
7. Lolli L, Johnson A, Monaco M, Di Salvo V, Gregson W. Relative skeletal maturity and performance test outcomes in elite youth Middle Eastern soccer players. *Med Sci Sports Exerc.* 2022;54(8):1326-1334.
8. Gregson W, Carling C, Gualtieri A, et al. A survey of organizational structure and operational practices of elite youth football academies and national federations from around the world: A performance and medical perspective. *Front Sports Act Living.* 2022;4:1031721.
9. Altman DG, Ohuma EO. Statistical considerations for the development of prescriptive fetal and newborn growth standards in the INTERGROWTH-21st Project. *BJOG.* 2013;120 Suppl 2:71-76, v.
10. Figueiredo AJ, Gonçalves CE, Coelho ESMJ, Malina RM. Youth soccer players, 11-14 years: maturity, size, function, skill and goal orientation. *Ann Hum Biol.* 2009;36(1):60-73.
11. Gouvea M, Cyrino ES, Ribeiro AS, et al. Influence of skeletal maturity on size, function and sport-specific technical skills in youth soccer players. *Int J Sports Med.* 2016;37(6):464-469.
12. Teixeira AS, Valente-dos-Santos J, Coelho ESMJ, et al. Skeletal maturation and aerobic performance in young soccer players from professional academies. *Int J Sports Med.* 2015;36(13):1069-1075.

13. Fransen J, Bennett KJM, Woods CT, et al. Modelling age-related changes in motor competence and physical fitness in high-level youth soccer players: implications for talent identification and development. *Sci Med Footb.* 2017;1(3):203-208.
14. Towlson C, Cobley S, Parkin G, Lovell R. When does the influence of maturation on anthropometric and physical fitness characteristics increase and subside? *Scand J Med Sci Sports.* 2018;28(8):1946-1955.
15. Datson N, Weston M, Drust B, Atkinson G, Lolli L, Gregson W. Reference values for performance test outcomes relevant to English female soccer players. *Sci Med Footb.* 2022;6(5):589-596.
16. Cole TJ, Stanojevic S, Stocks J, Coates AL, Hankinson JL, Wade AM. Age- and size-related reference ranges: a case study of spirometry through childhood and adulthood. *Stat Med.* 2009;28(5):880-898.
17. Blagus R, Jurak G, Starc G, Leskošek B. Centile reference curves of the SLOfit physical fitness tests for school-aged children and adolescents. *J Strength Cond Res.* 2023;37(2):328-336.
18. Cardoso FM, Almodhy M, Pepera G, Stasinopoulos DM, Sandercock GR. Reference values for the incremental shuttle walk test in patients with cardiovascular disease entering exercise-based cardiac rehabilitation. *J Sports Sci.* 2017;35(1):1-6.
19. Eisenmann JC, Laurson KR, Welk GJ. Aerobic fitness percentiles for U.S. adolescents. *Am J Prev Med.* 2011;41(4 Suppl 2):S106-110.
20. Nevill AM, Myers J, Kaminsky LA, Arena R, Myers TD. Comparing individual and population differences in minute ventilation/carbon dioxide production slopes using centile growth curves and log-linear allometry. *ERJ Open Res.* 2021;7(3).
21. Jia H, Wan B, Bu T, et al. Chinese physical fitness standard for campus football players: A pilot study of 765 children aged 9 to 11. *Front Physiol.* 2022;13:1023910.
22. Buchheit M, Mendez-Villanueva A, Mayer N, et al. Locomotor performance in highly-trained young soccer players: does body size always matter? *Int J Sports Med.* 2014;35(6):494-504.
23. Cole TJ. The international growth standard for preadolescent and adolescent children: statistical considerations. *Food Nutr Bull.* 2006;27(4 Suppl Growth Standard):S237-243.
24. Buchheit M, Mendez-Villanueva A. Reliability and stability of anthropometric and performance measures in highly-trained young soccer players: effect of age and maturation. *J Sports Sci.* 2013;31(12):1332-1343.
25. Borghi E, de Onis M, Garza C, et al. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Stat Med.* 2006;25(2):247-265.
26. Stasinopoulos DM, Rigby RA, Heller GZ, Voudouris V, De Bastiani F. *Flexible regression and smoothing: using GAMLSS in R.* Chapman and Hall/CRC; 2017.
27. Geithner CA, Thomis MA, Vanden Eynde B, et al. Growth in peak aerobic power during adolescence. *Med Sci Sports Exerc.* 2004;36(9):1616-1624.
28. Philippaerts RM, Vaeyens R, Janssens M, et al. The relationship between peak height velocity and physical performance in youth soccer players. *J Sports Sci.* 2006;24(3):221-230.
29. Lolli L, Johnson A, Monaco M, Di Salvo V, Gregson W. Skeletal maturation in male elite youth athletes from the Middle East. *Am J Hum Biol.* 2023:e23906.
30. Cole TJ. Sample size and sample composition for constructing growth reference centiles. *Stat Methods Med Res.* 2021;30(2):488-507.

Tables Legends

Table 1. Predicted reference intervals for 10-m sprinting time (s) by chronological age (N=441, n=1943)
Table 2. Predicted reference intervals for 40-m sprinting time (s) by chronological age (N=435, n=1820)
Table 3. Predicted reference intervals for CMJ height (cm) by chronological age (N=441, n=1964)
Table 4. Predicted reference intervals for MAS score ($km \cdot h^{-1}$) by chronological age (N=430, n=1751)

Figures Legends

Figure 1. Predicted reference intervals for 10-m sprinting time (s) by chronological age (N=441, n=1943).
Figure 2. Predicted reference intervals for 40-m sprinting time (s) by chronological age (N=435, n=1820).
Figure 3. Predicted reference intervals for CMJ height (cm) by chronological age (N=441, n=1964).
Figure 4. Predicted reference intervals for MAS score ($km \cdot h^{-1}$) by chronological age (N=430, n=1751).
Figure 5. Worm plots diagnostics for the 10-m sprinting (A), 40-m sprinting (B), CMJ (C), and MAS (D) models.
Figure 6. Individual-athlete physical performance data superimposed on age-related reference intervals relevant to 10-m sprinting (A), 40-m sprinting (B), CMJ (C), and MAS (D) assessments.