LETTER TO THE EDITOR

Methodological rigor in reference chart development: A comment on "*Normative Reference Centiles for Sprint Performance in High-Level Youth Soccer Players: The Need to Consider Biological Maturity*"

Running head: Age-specific physical performance ranges

Lorenzo Lolli[1]

[1] Department of Sport and Exercise Sciences, Institute of Sport, Manchester Metropolitan University, Manchester, UK

Address for Correspondence:

Lorenzo Lolli

Department of Sport and Exercise Sciences,

Institute of Sport,

Manchester Metropolitan University,

99 Oxford Rd, Manchester M1 7EL,

United Kingdom

email: lollilorenzo0@gmail.com

In a recent study in *Pediatric Exercise Science*, Ruf et al. (13) illustrated novel reference charts for sprint performance outcome measures by chronological age and ultrasound-based Tanner-Whitehouse (TW) II skeletal age. Notwithstanding the study conclusions (13) advocating the construction of alternative reference charts against proxy measures of skeletal maturation in mind, the design of these reference charts (13) requires consideration of methodological requirements for valid development and evaluation before any implementation (5, 12, 15).

Measurement accuracy is essential for constructing appropriate reference charts (6, 12). Data for dependent variables collected in less-standardized settings are prone to larger measurement error (14) that, if appreciably implausible, might lead to misclassifying an individual's *relative standing* on a reference chart (6). The integrity of sprint performance-on-skeletal age centiles (13) *also* depended on the independent variable measurement accuracy determined using ultrasound procedures (13). Monitoring measurement processes is important for reference chart development, and it involves, for example, application of the Bland-Altman method (12).

In this context, the width of the limits of agreement from a study in male and female adolescent athletes (7) revealed true differences for 95% of pairs of future manually-rated FELS (*criterion*) *versus* ultrasound-based (*alternative*) skeletal ages could be as high as ~2.55 y. Despite the study design (7), the lack of information relevant to the exact TW-II protocol (7), and considerations that methods provided skeletal ages that were comparable (13), the width of the limits of agreement (7), and *not* the mean bias, informs between-method comparison interpretations contingent on well-defined analytical goals (2). Accordingly, the width of these limits of agreement (7) was broad enough to suggest between-method *disagreement* and the potential for ultrasound assessments to produce also inaccurate TW-II skeletal ages for any sprint performance-on-skeletal age centile chart to be of practical value (2, 6, 12).

More evidence for instrumentation-related measurement unreliability comes from descriptive statistics for TW-II skeletal age in Table 1 (13). Specifically, the 95% prediction intervals for the TW-II skeletal age in the U16 and U17 age groups, calculated using the reported mean and standard deviation estimates (1, 4, 13), ranged from ~14.7 to 18.7 y and ~15.6 to 18.8 y, respectively. Note that the upper limits of these intervals exceeded the TW-II skeletal age at full maturity for boys of 18.2 y and indicated the measurement range included individual data points that do *not* exist on the original TW-II scale (3) despite the fitted range (13).

The construction of reference charts also deserves conceptual and contextual considerations (5, 6, 12). First, it was unclear whether this study (13) aimed to pursue a *prescriptive* or *descriptive* approach (12). These centiles seem reconcilable with a *descriptive* approach given the study *Methods* (13), yet probably inconsistent even to define *local* references as unlikely representative of the broader German population of youth football players (6). Second, any advantage for selecting skeletal age, in place of chronological age, as a measure of time requires concrete evidence for justification (16) given the shape of the reference curves differed *between*-performance outcomes only (13). *Prescriptive* skeletal age centiles development may be difficult and can also limit a consistent tracking of the individual player over different career stages (8) considering biological maturation assessment seems unnecessary towards and beyond youth-to-senior transition phases. Despite also alternative illustrations of 95% prediction intervals for performance test outcomes by manually-rated TW-II skeletal age (10), chronological age centiles remain informative *if* contextualized against population-based growth velocity events (9, 11).

Reference charts necessitate formal scrutiny before implementation and can be misleading if inconsistent with fundamental measurement requirements for valid development (5, 6, 12, 15).

References

1. Altman DG, Bland JM. Standard deviations and standard errors. BMJ. 2005;331(7521):903.
2. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-38.
3. Cameron N. Measuring maturity. In: Molinari L, Cameron N, Hauspie RC, editors. Methods in Human Growth Research. Cambridge Studies in Biological and Evolutionary Anthropology. Cambridge: Cambridge University Press; 2004. p. 108-40.
4. Chinn S. Statistics in respiratory medicine. 1. Ranges, confidence intervals, and related quantities: what they are and when to use them. Thorax. 1991;46(5):391-3.
5. Cole TJ. Commentary: Methods for calculating growth trajectories and constructing growth centiles. Stat Med. 2019;38(19):3571-9.
6. Cole TJ. The use and construction of anthropometric growth reference standards. Nutr Res Rev. 1993;6(1):19-50.
7. Cumming SP, Pi-Rusiñol R, Rodas G, Drobnic F, Rogol AD. The validity of automatic methods for estimating skeletal age in young athletes: a comparison of the BAUSport ultrasound system and BoneXpert with the radiographic method of Fels. Biol Sport. 2024;41(1):61-7.
8. 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):589-96.
9. Lolli L, Gregson W, Bonanno D, Kuitunen S, Di Salvo V. Age-related reference intervals for physical performance test outcomes relevant to male youth Middle Eastern football players. Int J Sports Physiol Perform. 2023;18(11):1283-95.
10. 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):1326-34.
11. 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;35(8):e23906.
12. Ohuma EO, Altman DG. Design and other methodological considerations for the construction of human fetal and neonatal size and growth charts. Stat Med. 2019;38(19):3527-39.
13. Ruf L, Altmann S, Kloss C, Härtel S. Normative reference centiles for sprint performance in high-level youth soccer players: the need to consider biological maturity. Pediatr Exerc Sci. 2024:1-9.
14. Wright CM, Haig C, Harjunmaa U, Sivakanthan H, Cole TJ. Assessing the optimal time interval between growth measurements using a combined data set of weights and heights from 5948 infants. Arch Dis Child. 2022;107(4):341-5.
15. Wright CM, Williams AF, Cole TJ. Advances in growth chart design and use: the UK experience. World Rev Nutr Diet. 2013;106:66-74.
16. Zemel BS. From growth charts to growth status: how concepts of optimal growth and tempo influence the interpretation of growth measurements. Ann Hum Biol. 2023;50(1):236-46.