Tanner–Whitehouse and Modified Bayley–Pinneau Adult Height Predictions in Elite Youth Soccer Players from the Middle East

LORENZO LOLLI[1,2], AMANDA JOHNSON[3], MAURICIO MONACO[3], MARCO CARDINALE[4,5], VALTER DI SALVO[1,6], and WARREN GREGSON[1,2]

[1]Aspire Academy, Football Performance & Science Department, Doha, QATAR; [2]Football Exchange, Research Institute of Sport Sciences, Liverpool John Moores University, Liverpool, UNITED KINGDOM; [3]National Sports Medicine Program, Aspetar Orthopaedic and Sports Medicine Hospital, Doha, QATAR; [4]Aspetar Orthopaedic and Sports Medicine Hospital, Doha, QATAR; [5]Institute of Sport, Exercise and Health, University College London, London, UNITED KINGDOM; and [6]Department of Movement, Human and Health Sciences, University of Rome "Foro Italico," Rome, ITALY

ABSTRACT

LOLLI, L., A. JOHNSON, M. MONACO, M. CARDINALE, V. DI SALVO, and W. GREGSON. Tanner–Whitehouse and Modified Bayley–Pinneau Adult Height Predictions in Elite Youth Soccer Players from the Middle East. *Med. Sci. Sports Exerc.*, Vol. 53, No. 12, pp. 2683–2690, 2021. **Purpose:** To provide the first scrutiny of adult height prediction protocols based on automated Greulich–Pyle and Tanner–Whitehouse (TW) skeletal ages applied to elite youth soccer players from the Middle East. **Methods:** We examined the application of modified Bayley–Pinneau (BoneXpert®), TW-II, and TW-III protocols using mixed-longitudinal data available for 103 subjects (chronological age range, 19.4 to 27.9 yr) previously enrolled as academy student-athletes (annual screening range, one to six visits). Random-effects generalized additive models quantified the presence of systematic mean differences between actual versus predicted adult height. Effects were deemed practically equivalent based on the location of the confidence interval (95% CI) against a realistic difference value of $\Delta = \pm 1$ cm. Each model pooled residual standard deviation described the actual precision of height predictions and was used to calculate a 95% prediction interval. **Results:** The BoneXpert® method overpredicted adult height systematically at chronological ages in the range of approximately 13.5 to 14.5 yr (95% CI range, −1.9 to −1 cm) and Greulich–Pyle skeletal ages between 13.5 and 15 yr (95% CI range, −3.1 to −1 cm). Effects based on TW-II were practically equivalent across the chronological and skeletal age measurement ranges, with this protocol yielding adult height predictions with a precision (standard deviation) of approximately ±2.6 cm. The mean TW-III effects indicated systematic adult height overpredictions until the attainment of 14.5 and 15 yr of chronological age (95% CI range, −3.8 to −1.1 cm) and TW-III skeletal age (95% CI range: −5.2 to −2.3 cm), respectively. **Conclusions:** Tanner–Whitehouse-II adult height prediction method provided more consistent estimates and can be considered the method of choice for talent development purposes in youth soccer players from the Middle East. **Key Words:** SKELETAL AGE, HEIGHT, SOCCER, TALENT SELECTION, YOUTH, MATURATION

Ongoing assessment of growth and maturity is an integral part of the elite youth athlete development process. Growth is concerned with any quantitative increase in body size over time (1). Maturation pertains to the process of progressive changes that lead from an undifferentiated or immature state to a highly organized, specialized, mature or adult state (2). In biology, adulthood also refers to functional maturation or the ability for successful procreation (2), with age-related and growth cessation criteria defining the attainment of this status (3). Within the context of growth, the process of maturation is continuous in nature (2). However, a precise definition of maturity status involves examining discrete indicators during the course of development, such as skeletal age, secondary sexual and somatic characteristics, and dental age (4).

A number of studies in soccer assessed maturation status using both invasive (5) and noninvasive (6) methods. In this context, skeletal age received particular attention as an assessment of physiological suitability for a career or a sport (7), yet, in practice, it is not always available due to its invasive nature and cost (8). Skeletal age is a measure of the stage-by-stage metamorphosis of the cartilaginous and membranous skeleton from fetal life to fully ossified bones in the adult individual (2). In pediatrics and other clinical fields, visual assessment of left-hand and wrist roentgenograms is the general method used to determine skeletal age (2). Developmental stages to each epiphyseal center of interest can be assigned according to criteria of atlases and bone-scoring techniques (2).

The Greulich–Pyle (GP) method is an example of an atlas technique (9). In this type of protocol, the skeletal age is the chronological age assigned to the standard consistent with the degree of ossification illustrated in the roentgenogram (2). As an alternative method developed on a different conceptual basis, a bone-specific technique is independent of chronological age (2). Within the full passage from immaturity to maturity, the evaluation of a series of indicators relating to the appearance of each specific bone of the hand and wrist provides a cumulative score determining the skeletal age of the subject (2). Of these techniques, the Tanner–Whitehouse (TW) system (2) evolved from the Oxford method (10), with the original version (TW-I) determining the skeletal age based on maturity indicators for 20 bones. This version was later revised in TW-II (11) and, more recently, TW-III (12) systems, with no alteration in the descriptions and ratings of stages of the bones. The TW-II revision (11) provided skeletal ages based on the assessments of (i) the 20 bones (TW-II 20), (ii) the seven carpals (TW-II Carpal), or (iii) the radius-ulna-short (RUS) bones (TW-II RUS), with the TW-III revision excluding the 20-bone protocol (12). An important distinction between the revised protocols was the definition of maturity status (RUS score = 1000 au) corresponding to a skeletal age of 18.2 yr in TW-II and 16.5 yr in TW-III (5). Manual assessment of skeletal age is, however, prone to intrarater and interrater variabilities (13). Recent advances in digital radiography now permit reliable automated analysis of skeletal age based on the principles of atlas or bone-scoring techniques and free from the influence of intrarater random error (14–16). Limiting any degree of imprecision in skeletal age rating is fundamental to obtain unbiased measurements of indicators relevant to understanding the human growth process, such as predicted adult height (13,16–18).

Assessment of skeletal age combined with anthropometric and demographic data can provide estimates of predicted adult height (19). In athletic populations, obtaining adult height predictions can inform the grouping of athletes for specific competitions and training according to, for example, the biobanding strategy (20). This method uses the percentage predicted adult height (%), calculated as the simple ratio of current height divided by predicted adult height × 100 at the time of observation, as the indicator for grouping youth players into specific bands (20). Furthermore, estimation of predicted adult height can be useful for preliminary evaluations of anthropometric suitability for professional career progression in specific sports and/or positional roles in team sports (7,21). In practice, the notion of adult height prediction relates to deriving an estimation of an unmeasured dimension of interest at some point in time that is expected to reflect the extent of human somatic growth of the adult athlete (22). Bayley and Pinneau illustrated one of the first methods for predicting adult height also with tables providing an estimate of the fraction of adult height achieved as per the GP skeletal age at the time of the roentgenogram (23). Conversely, the general protocol by Tanner and Whitehouse (11,12) provides estimates of predicted adult height according to the present age, the RUS score or skeletal age, and the present height. Importantly, the common Tanner and Whitehouse adult height prediction methods are fundamentally different. The TW-II adult height prediction protocol includes the actual skeletal age in the equation, whereas the TW-III protocol considers the raw maturity score (i.e., RUS) as a predictor variable to address the influence of environmental factors (22).

The most important time window for adult height prediction is deemed between 8 and 13 yr (21,24), with the precision of the predicted values depending on a number of factors. First, the occurrence of certain events during the growth process may introduce prediction errors (19). In the clinical realm, accurate adult height predictions are considered plausible only after a subject passes the pubertal growth spurt (13,19). Second, valid adult height prediction might depend on the applicability of different skeletal age methods to a given population (5,25,26). Researchers in this field suggested to adjust existing skeletal age standards when applied to populations other than those from which they were derived (27). For example, researchers explored the application of the GP atlas to children and adolescents (age range, 2 to 15 yr) from the United Kingdom (25). The results of this investigation suggested that the GP atlas remained valid when applied to that population (25). In contrast, the GP standards might be imprecise if applied to subjects of African, Arab, and Asian ethnicity (28,29). Information regarding the application of atlases or bone-scoring techniques to youth Arab subjects are nevertheless limited, with formal appraisals of different adult height prediction methods conducted in general and clinical populations from Western countries (13,17,19,21,23,26,30–34).

Using automated image analysis methods, we examined the application of adult height prediction protocols based on GP and TW skeletal ages to elite youth soccer players from the Middle East.

METHODS

Participants. We assessed the equivalence between actual versus predicted adult height based on modified Bayley–Pinneau (BoneXpert®), TW-II, and TW-III protocols using data available for a sample of 103 subjects (age range, 19.4 to 27.9 yr; adult standing height range, 157.5 to 189.7 cm) previously enrolled as academy student-athletes (age range, 11.5 to 17.8 yr; standing height range, 137.5 to 187 cm). Adult standing height was defined as the height for a subject older than 18 yr (3,33). The data collection was part of the annual medical screening and a longitudinal growth and maturation project (protocol number: E202008009) involving also regular performance/fitness screenings. Signed parental consent was obtained before each academy season to use data for research purposes. This retrospective study was approved by the Aspire Zone Foundation Institutional Review Board, Doha, State of Qatar.

Design and procedures. With some student-athletes measured once and others more than once (annual screening range, one to six visits), the present investigation adopted a retrospective, mixed-longitudinal study design. Hand x-rays, standing height, and body mass measurements collected in student-athletes over a 14-yr period (N = 876) as part of the annual medical screening were retrieved from the Academy medical records, anonymized, analyzed and used to determine skeletal age at the time of the scan. Standing height was measured using a wall-mounted stadiometer to the nearest 0.1 cm according to the stretch stature protocol (Holtain Limited, Crosswell, Crymych, Pembs., UK), and body weight measurements were obtained using digital scales. Using test–retest data from a subsample of 17 elite youth soccer players (age range: 14 to 14.8 yr) from the available population (N = 876), the standard error of the measurement (SEM) for standing height was 0.34 cm (95% confidence interval [CI], 0.25 to 0.52 cm). The estimated minimal detectable change was approximately 1 cm (95% CI, 0.7 to 1.4 cm), consistent with established growth cessation criteria (3).

With the left hand and wrist placed flat and down on the x-ray plate (Digital Diagnost; Philips, Amsterdam, Netherlands), assessment of skeletal age involved standard radiographs of the radius, ulna, carpals, metacarpals and phalanges (35). Modern technology now allows minimizing the exposure to radiation to as little as 0.001 mSv, which is commensurate to less than natural background radiation walking around a city center, or any radiation associated with a 2-h flight (35).

The automated assessment of roentgenograms involved digital images processing using the computerized BoneXpert® determination method as per the manufacturer recommendations (version 3.1.4, Visiana, Holte, Denmark). This medical device was originally developed, calibrated, and validated on the GP atlas using data from European samples (16). This method reproduces the borders of 15 bones automatically. It then computes intrinsic GP skeletal ages for each of 13 bones (radius, ulna, and 11 short bones) from roentgenograms of the hand (14). A new standard version of the TW skeletal age rating was implemented by the BoneXpert® method and calibrated on manual rating data from the First Zürich Longitudinal Study (15). The BoneXpert® adult height prediction method is founded on a default skeletal age (GP atlas) as a reelaboration of the conventional Bayley–Pinneau protocol (16). With the estimated 50th centile for standing height at 18 yr corresponding to 174.3 cm (unpublished data), modified Bayley–Pinneau adult height predictions considered the Caucasian European South ethnicity option as the most plausible in the context of our study population from Western Asia (16). The TW-II (11) and TW-III (12) adult height predictions were also derived using BoneXpert® readings for skeletal ages in combination with relevant equations. Given the purpose of our investigation, automated ratings for TW-III skeletal ages were back-converted to RUS scores for deriving TW-III predicted adult heights using relevant conversion tables (12).

Test–retest assessment of the manual rating method suggested most of the items in the TW RUS protocol showed moderate to excellent intrarater reliability (1,3), with the estimated SEM of 0.38 yr (95% CI, 0.34 to 0.43 yr) for TW-II and 0.36 yr (95% CI, 0.32 to 0.41 yr) for TW-III (see Figure, Supplemental Digital Content 1, test–retest assessment of the TW RUS protocol, http://links.lww.com/MSS/C365). Comparison of manual and automated TW-II and TW-III skeletal age determination in this population (n = 103) did not reveal the presence of a systematic bias, with estimates not exceeding differences corresponding to approximately ±1.1 yr (see Table, Supplemental Digital Content 2, Estimated marginal means for the difference between manual versus automated skeletal age assessment by chronological age, http://links.lww.com/MSS/C366).

Statistical analysis. Random-effects generalized additive models with restricted maximum likelihood (36) quantified the equivalence between actual *minus* predicted adult height with separate analyses including chronological age or skeletal age at the time of the hand-wrist x-ray scan as the explanatory variable, respectively (23). Models included the raw difference (Δ) as the response variable, with the smooth term for the explanatory variable set at 3, 5, 7, and 9 basis functions plus a subject-specific random effect penalized by a ridge penalty (36). An information-theoretic approach was adopted for optimal smooth model selection (36). Estimated marginal means described the degree of overprediction or underprediction presented with 95% CI and 95% prediction interval (PI) (37). A 95% CI described the likely range of values compatible with the true population parameter (38), whereas a 95% PI indicated the range of values within which 95% of future similar observations may lie (39,40). Estimates of actual *minus* predicted adult height were declared practically equivalent based on the location of the 95% CI for the mean effects (38) interpreted against a realistic difference value of ±1 cm (3). The pooled residual standard deviation (SD) from each model described the actual precision of height predictions and was used to calculate a 95% PI (38). Positive values in sign (+) suggested method underprediction, whereas negative values (−) indicated method overprediction. Given the available sample over the selected observational period (N = 876), empirical guidelines informed the design (age range, 9.9 to 18.3 yr) and sampling (n = 125) to describe the size, timing, and intensity of the pubertal growth spurt in this population of youth athletes (41,42). Using this subsample data, a Superimposition by Translation and Rotation model with 5 degrees of freedom quantified age at peak height velocity and peak height velocity (41). Statistical analyses were conducted using R (version 3.6.3, R Foundation for Statistical Computing).

RESULTS

Descriptive data for this subsample (n = 103) were illustrated in density plots (Fig. 1). Outcomes for comparisons were reported by chronological age (Table 1) and reference skeletal age (Table 2). The estimated age at peak height velocity was 13.62 yr (95% CI, 13.55 to 13.70 yr), and peak height velocity was 9.9 $cm{\cdot}yr^{-1}$ (95% CI, 9.5 to 10.3 $cm{\cdot}yr^{-1}$) according to the growth curve analysis (see Figure, Supplemental Digital Content 3, Population mean height and velocity curves, http://links.lww.com/MSS/C367). The negative relationship between

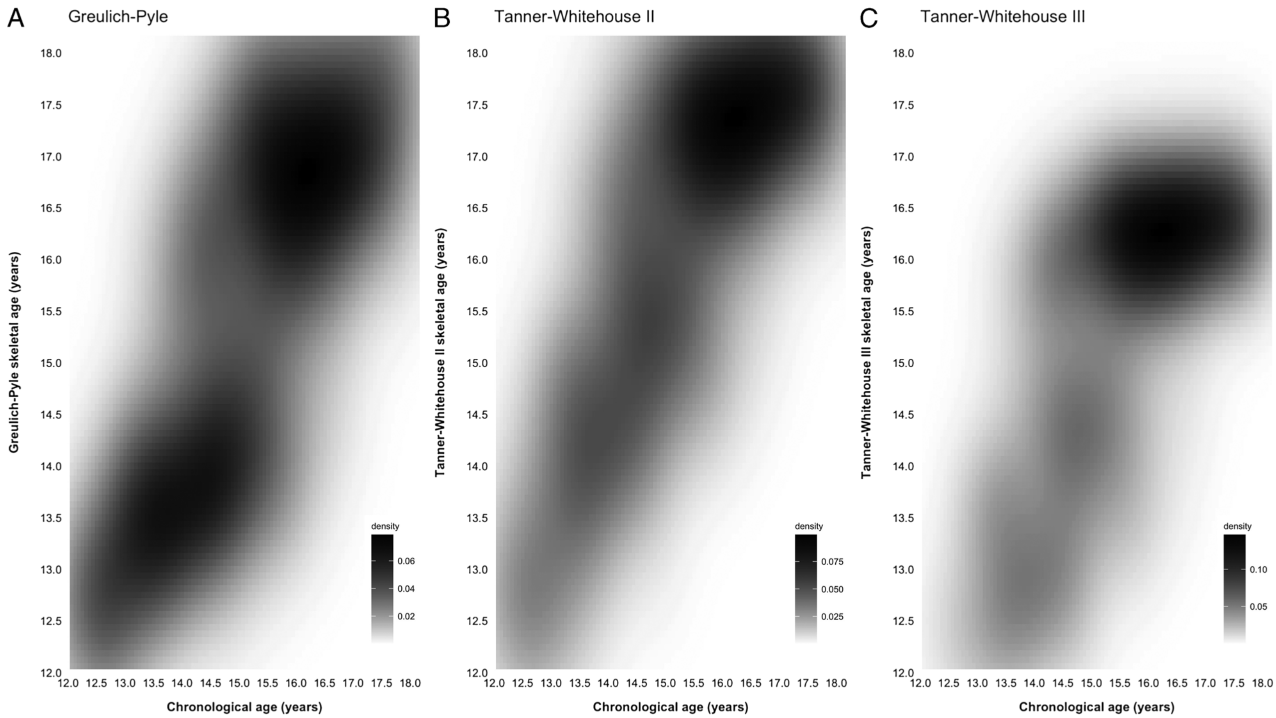

FIGURE 1—Density plot showing automated GP (A), TW-II (B), and TW-III (C) skeletal age distributions by chronological age.

timing and intensity random-effects (r = –0.65) suggested student-athletes entering puberty early showed the largest height velocity (see Figure, Supplemental Digital Content 4, Scatterplot matrix of random-effects correlations for size, timing, and intensity in the Superimposition by Translation and Rotation model, http://links.lww.com/MSS/C368).

For explorations by chronological age, the direction and width of the effects for actual *versus* predicted adult height as per the BoneXpert® method indicated a systematic overprediction in the range of 13.5 to approximately 14.5 yr (95% CI range, −1.9 to −1 cm). In the same age period, the estimated 95% PI were located around relatively more negative values (Table 1). Use of the TW-II method did not reveal substantial mean overprediction or underpredictions in adult height (95% CI range, −2.3 to 0.8 cm), with the estimated 95% PI located around relatively more negative values (Table 1). Systematic overpredictions resulted from the TW-III ratings until the attainment of 14.5 yr of age (Table 1). The uncertainty (95% CI) for the observed mean overpredictions ranged from approximately 1 to 4 cm, with the 95% PI in the range of −8.6 to 3.8 cm. The pooled residual SD estimate used to calculate the 95% PI for actual versus predicted adult height for BoneXpert®, TW-II, and TW-III protocols by chronological age was ±2.3 cm, ±2.5 cm, and ±2.8 cm, respectively.

For descriptions by reference skeletal age, the BoneXpert® protocol systematically overpredicted adult height at GP skeletal ages between approximately 13.5 and 15 yr (Table 2). In this particular age range, the degree of the uncertainty for the mean overprediction was approximately 1 to 3 cm, with the respective 95% PI ranging from −6.7 to 2.5 cm. The differences between actual *versus* TW-II predicted adult height were generally symmetric around zero (Table 2). The mean differences between methods corresponded to approximately 1 cm between 14 and 16 yr (95% CI range, −1.9 to −0.2 cm) yet located within

TABLE 1. Estimated marginal means for the difference between actual vs predicted adult height by chronological age.

	BoneXpert					TW-II					TW-III				
CA (yr)	Δ (cm)	95% CI		95% PI		Δ (cm)	95% CI		95% PI		Δ (cm)	95% CI		95% PI	
12.5	−0.9	−1.6	−0.2	−5.5	3.8	−1.6	−2.3	−0.8	−6.6	3.5	−2.8	−3.6	−1.9	−8.3	2.7
13	−1.1	−1.7	−0.6	−5.8	3.5	−1.3	−1.9	−0.7	−6.3	3.7	−3.1	−3.8	−2.4	−8.6	2.4
13.5	−1.3	−1.8	−0.9	−6.0	3.3	−1.0	−1.5	−0.4	−6.0	4.1	−3.1	−3.7	−2.4	−8.6	2.5
14	−1.4	−1.9	−1.0	−6.0	3.2	−0.6	−1.1	−0.1	−5.6	4.4	−2.5	−3.1	−1.9	−8.0	3.0
14.5	−1.3	−1.8	−0.9	−6.0	3.3	−0.2	−0.7	0.3	−5.2	4.8	−1.7	−2.3	−1.1	−7.2	3.8
15	−1.1	−1.5	−0.7	−5.7	3.6	0.1	−0.3	0.6	−4.9	5.2	−0.9	−1.4	−0.3	−6.4	4.6
15.5	−0.7	−1.2	−0.3	−5.4	3.9	0.3	−0.1	0.8	−4.7	5.4	−0.3	−0.8	0.2	−5.8	5.2
16	−0.4	−0.8	0.0	−5.0	4.2	0.3	−0.1	0.8	−4.7	5.4	0.0	−0.6	0.5	−5.5	5.4
16.5	−0.1	−0.6	0.4	−4.7	4.5	0.1	−0.4	0.6	−4.9	5.1	0.0	−0.6	0.6	−5.5	5.5
17	0.1	−0.4	0.6	−4.5	4.8	−0.3	−0.9	0.3	−5.3	4.7	−0.1	−0.7	0.6	−5.6	5.4
17.5	0.3	−0.5	1.0	−4.4	5.0	−0.8	−1.6	0.0	−5.9	4.2	−0.1	−1.1	0.8	−5.7	5.4

CA, chronological age.

TABLE 2. Estimated marginal means for the difference between actual vs predicted adult height by reference skeletal age.

	BoneXpert					TW-II					TW-III				
SA (yr)	Δ (cm)	95% CI		95% PI		Δ (cm)	95% CI		95% PI		Δ (cm)	95% CI		95% PI	
12.5	−0.1	−0.8	0.6	−4.3	4.0	−0.6	−1.4	0.3	−5.7	4.6	−2.9	−3.6	−2.3	−6.6	0.7
13	−0.8	−1.4	−0.2	−4.9	3.3	−0.5	−1.3	0.2	−5.6	4.6	−4.0	−4.6	−3.4	−7.7	−0.3
13.5	−1.9	−2.3	−1.4	−6.0	2.3	−0.6	−1.3	0.0	−5.8	4.5	−4.6	−5.2	−4.1	−8.3	−1.0
14	−2.6	−3.1	−2.1	−6.7	1.5	−0.9	−1.6	−0.3	−6.0	4.2	−4.4	−4.9	−3.8	−8.0	−0.7
14.5	−2.5	−3.1	−1.9	−6.7	1.6	−1.2	−1.8	−0.6	−6.3	3.9	−3.1	−3.7	−2.6	−6.8	0.5
15	−1.6	−2.2	−1.0	−5.8	2.5	−1.3	−1.9	−0.7	−6.4	3.8	−1.5	−2.1	−0.9	−5.2	2.2
15.5	−0.5	−1.1	0.1	−4.6	3.7	−1.1	−1.7	−0.5	−6.3	4.0	−0.1	−0.7	0.5	−3.8	3.6
16	0.3	−0.3	0.8	−3.9	4.4	−0.8	−1.4	−0.2	−5.9	4.3	0.6	0.2	1.0	−3.0	4.3
16.5	0.3	−0.1	0.8	−3.8	4.4	−0.3	−0.9	0.2	−5.4	4.8	0.8	0.4	1.3	−2.8	4.5
17	−0.1	−0.6	0.4	−4.2	4.0	0.1	−0.4	0.5	−5.0	5.2					
17.5	−0.3	−0.8	0.2	−4.4	3.8	0.3	−0.1	0.8	−4.8	5.4					

SA, reference skeletal age.

thresholds deemed of practical relevance (Δ = ± 1 cm). In this age period, the estimated 95% PI were located around relatively more negative values (Table 2). Adoption of the TW-III method systematically overpredicted adult height until a TW-III skeletal age of 15 yr (95% CI range, −5.2 to −1 cm), with the estimated 95% PI spanning relatively more negative values (Table 2). The pooled residual SD estimate used to calculate the 95% PI for actual *versus* predicted adult height for BoneXpert®, TW-II, and TW-III protocols by reference skeletal age was ±2.1 cm, ±2.6 cm, and ±1.8 cm, respectively.

DISCUSSION

Adult height prediction is a standard procedure in clinical pediatrics relevant to the understanding of the individual subject growth and maturation status for diagnostic and therapeutic purposes. In this context, we examined, for the first time, the application of common adult height prediction methods to a sample of elite youth soccer players from the Middle East. The main findings of our retrospective, mixed-longitudinal study revealed adult height predictions based on TW-II were relatively more precise than estimations from BoneXpert® and TW-III methods across both the chronological and skeletal age measurement ranges. Our line of evidence substantiates the notion that TW-II (RUS) can be considered the method of choice for adult height prediction in elite youth Arab soccer players.

The evaluation of an adult height prediction method constitutes, in practice, a formal validation of the reference skeletal age rating protocol given the population of interest (26). The conventional Bayley–Pinneau and TW protocols may perform reasonably in normal samples (19), and our study addressed a number of aspects relevant to accurate adult height prediction. Roche and colleagues (27) posited that existing skeletal age standards require adjustment when applied to nonreference populations. Studies conducted in Arab populations are, nevertheless, limited to explorations of manual versus automated method comparisons (29). In this context, our findings provide a meaningful contribution highlighting potential limitations of GP skeletal age as a method of choice in Arab youth male athletes. Specifically, automated skeletal age assignment as per the GP criteria potentially influenced the precision of adult height predictions observed at some development stages using the modified Bayley–Pinneau method. The mean effects for actual minus predicted adult height differences were not practically equivalent at chronological and GP skeletal ages ranging for 13.5 to 14.5 yr and 13.5 to 15.0 yr, respectively (Tables 1 and 2). The advances in digital radiography now permit reliable automated analysis of skeletal age independent of intrarater variability yet methods were validated on reference scales (14,15). Although BoneXpert® skeletal age was validated for White, African American, Hispanic, and Asian ethnicities on the GP system (16), data of children from the Brush Foundation Longitudinal Growth Study performed between 1931 and 1942 in Ohio informed this scale development (2). When applied to nonreference samples, the potential for the modified Bayley–Pinneau to provide imprecise adult height predictions relates to the properties of the reference atlas given the influence of environmental factors, ethnicity, secular changes, and socioeconomic status on skeletal age (2).

Conversely, TW scores are independent of geography and time (24). Irrespective of the rating method, the degree of imprecision we observed in TW-II predicted adult height suggested the existing protocols would probably remain valid when applied to this population but not for TW-III (Tables 1 and 2; see Tables, Supplemental Digital Content 5, Estimated marginal means for the difference between actual versus predicted adult height by chronological age [manual ratings], http://links.lww.com/MSS/C369; and Supplemental Digital Content 6, Estimated marginal means for the difference between actual versus predicted adult height by reference skeletal age [manual ratings], http://links.lww.com/MSS/C370). The substantial adult height overpredictions based on TW-III is an important and interesting finding of our study. Whereas counterintuitive, the number of differences between TW-II and TW-III are fundamentally practical and inherent to the nature of the prediction equations. First, the RUS, and not skeletal age, is included as an explanatory variable in the TW-III equation (11,12). Second, the TW-II protocol permits the inclusion of annual height increments over a broader chronological age range (minimum: 11 yr) than TW-III (minimum: 12 yr). Third, and accordingly, the parameter estimates from TW-II and TW-III prediction equations are, by definition, not practically equivalent (11,12). Tanner–Whitehouse-II (RUS) was developed on samples limited to British children from the Harpenden Growth Study and International Children's Centre London

Longitudinal Growth study (11). Tanner–Whitehouse-III (RUS) was based on a more heterogeneous pool of British, Belgian, Italian, Spanish, Argentinean, American (Texas), and Japanese children and adolescents surveyed between 1969 and 1995 (12). In practice, TW-II and TW-III methods provide qualitatively different descriptions of full maturity. The TW-III does not describe full maturity in the same way as TW-II. The final stages of radius and ulna define full maturity in TW-II, whereas the commencement of plate fusion, not full closure, is the final stage in TW-III (11,12). Tanner and colleagues (12) posited that actual maturity was represented more accurately by the actual maturity scores rather than skeletal age assumed to be prone to the influence of environmental factors. However, our study results indicated that TW-III is unlikely a meaningful advancement if applied for adult height prediction to this population of youth Arab athletes.

Our findings advanced current knowledge on the application of different adult height prediction protocols in normal and sports populations from practical and clinical standpoints. Nevertheless, formal comparisons of our results with existing studies are limited given the different outcome statistics reported in published research (16,26) and other methodological inconsistencies (13,17,21,26,30,33,34). Specifically, the calculation of height prediction errors as Δ (cm) = predicted *minus* actual adult height precludes understanding the true direction and sign of deviations in the predicted values. Bayley provided one of the very first explorations illustrating the degree of error in predicted adult height by chronological age and skeletal age using data of boys from the Harvard Growth Study (23). The mean effects for actual *minus* predicted adult height were practically equivalent in this study, with differences generally symmetric around zero across the GP skeletal age measurement range. This appears reasonable, and compared to our study, it reflects the more plausible contextual similarities between populations from Massachusetts and Ohio studied at the same time in the 19th century (23). Our findings also confirmed estimations of predicted adult height are prone to imprecision prior to the growth spurt period (19). The age at peak height velocity (see Figure, Supplemental Digital Content 3, Population mean height and velocity curves, http://links.lww.com/MSS/C367) in our population (13.62 yr; 95% CI, 13.55 to 13.70 yr) was similar to that observed in elite youth soccer players from Belgium (43) and United Kingdom (6) yet based on a larger study sample (n = 125) addressing methodological guidelines (41,42). Cameron et al. (31) applied the TW-II protocol to a sample of Canadian boys from the general population (11.6 ± 0.3 yr), with the observed a mean difference of −1.7 cm (95% CI, −2.6 to −0.9 cm) similar to our effects at 12.5 yr (Table 1). Likewise, Preece (19) found mean overpredictions for TW-II of −1.4 cm (95% CI, −2.5 to −0.3 cm) and −1.8 cm (95% CI, −2.8 to –0.9 cm) in a subsample of 24 boys age 11 to 13 yr from the Harpenden Growth Study. In sport, Ostojic (21) was the first to explore the application of the TW-II predicted adult height protocol to youth Serbian athletes from a wide range of disciplines including basketball (n = 71), soccer (n = 25), volleyball (n = 14), swimming (n = 12), and other sports (n = 13). Results suggested that TW-II provided reasonable adult height predictions, although effects are not comparable in terms of sign and direction with our study. Collectively, our study provided large-scale empirical evidence on the application of different adult height prediction methods to a population of elite youth soccer players from the Middle East.

From a real-world perspective, the estimation of predicted adult height for immature players can be an invaluable assessment to inform strategies relevant to talent identification, selection, and development (21). For example, in soccer, stature is a determinant for chances to play at elite level in the positions of goalkeeper and central defender (44). Notwithstanding the value of the available approaches, the determination of skeletal age in high-performance settings is not without limitations. The use of radiographs to assess skeletal age in sport represents a nonmedical human imaging procedure requiring careful consideration as in the case of assessing age in children and young people subject to immigration control (45). This assessment is generally requested as an adjunct to the diagnosis or the monitoring of treatment of disorders of growth or sexual maturation in clinical settings, and practitioners deem radiation exposure unacceptable when no benefit would accrue to the individual (46). In sports medicine, the use of imaging informed by clinical indications is well established because the outcome will influence patient management (7). In the context of talent identification and development processes (21), the International Atomic Energy Agency Safety Standards (no. SSG-55) indicated that nonmedical human imaging would require formal justification for benefit by authorities for sports organizations, players, medical professionals, and regulatory bodies (7).

Accordingly, radiation dosage is another aspect inherent to skeletal age assessments (47). The actual upper dosage for a typical hand-wrist x-ray for clinical purposes is reported as 0.001 mSv for pediatric patients, which falls well below the annual limit and constitutes an exposure equivalent to less than 20 min of natural background radiation or 2 min on a transatlantic flight (47). Crude estimations suggested the 40-yr mortality risk was 5.1×10^{-8} for a roentgenogram of the hand in teenagers based on a general dose of 0.00015 mSv (47). Practitioners and experts in this field deemed such doses minimal to prevent well-designed research projects from obtaining ethical approval (47) yet requiring formal justification for benefit from relevant authorities in sports performance and similar nonclinical settings (7). In this context, noninvasive methods were proposed as potential alternatives to the use of skeletal age (48). For example, in 1975, development of the Roche–Wainer–Thissen method aimed to address the limitations of existing protocols (22). This method included different explanatory variables as skeletal age, recumbent length, weight, and mid-parental stature to obtain more precise estimations (22). Khamis and Roche (48) later revised this protocol, including chronological age, weight, height, and mid-parental height as explanatory variables developing equations based on boys from Southwest Ohio enrolled in the Fels Longitudinal Study. Irrespective of the protocol, the mean errors in the prediction were approximately 5 cm (48) and substantially larger than what observed by researchers with other methods in normal populations (19,23,31). Importantly, accurate estimation of predicted adult height rests on the measurement of indicators such as growth increment, maturity, and maturity increment (22). Also, measuring midparental height may not always be feasible with the proposed alternatives (48) for this protocol equation rendering it potentially impractical with or without the inclusion of skeletal age in the equation as in the case of the present study. Recent investigations in youth soccer (6) adopted this approach to derive surrogate measures such as the percentage of predicted adult height (%). Whereas practical, this measure remains of limited utility for clinical purposes or grouping in sport if the error in the prediction is not accounted for. Nevertheless, our study is not without limitations, with a particular reference to the assessment of adult height predictions using different methods illustrated in the literature (16,22) also in other developmental phases within the growth process (e.g., 8 to 11 yr) relevant to professional coaches and practitioners (49). Furthermore, the availability of consistent skeletal age assessments over a large age range can be relevant to examine the onset of the growth spurt according to chronological and skeletal age, respectively (50). Given our data, exploration of these particular aspects was not, however, practically feasible.

CONCLUSIONS

The TW-II adult height prediction method provided the most consistent estimates in youth soccer players from the Middle East, with this protocol yielding adult height predictions with a precision (SD) of approximately ±2.6 cm. These findings supported the evaluation of maturation according to TW-II skeletal age assessment in this population. Automated skeletal age assessment using validated, computerized methods is a valid approach enabling reliable measurements independent of rater training and availability. Our findings extend knowledge on the precision of different adult height prediction methods applied to an underexplored population, providing an important contribution that may inform the digital implementation of other adult height prediction protocols based on automated image analysis.

The authors wish to thank Prof Robert M. Malina for his initial suggestions on this manuscript procedures and Ricardo Pinto to support the collection of the adult height data.

The authors declare no conflicts of interest. This study is not funded. The results of this study do not constitute endorsement by ACSM. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. Molinari L, Gasser T. The human growth curve: distance, velocity and acceleration. In: Molinari L, Cameron N, Hauspie RC, editors. *Methods in Human Growth Research*. Cambridge: Cambridge University Press; 2004. pp. 27–54.
2. Cameron N. Measuring maturity. In: Molinari L, Cameron N, Hauspie RC, editors. *Methods in Human Growth Research*. Cambridge: Cambridge University Press; 2004. pp. 108–40.
3. Kato S, Ashizawa K, Satoh K. An examination of the definition 'final height' for practical use. *Ann Hum Biol*. 1998;25(3):263–70.
4. Tanner JM. Sequence, tempo, and individual variation in the growth and development of boys and girls age twelve to sixteen. *Daedalus*. 1971;100(4):907–30.
5. Malina RM, Coelho ESMJ, Figueiredo AJ, et al. Tanner-Whitehouse skeletal ages in male youth soccer players: TW2 or TW3? *Sports Med*. 2018;48(4):991–1008.
6. Parr J, Winwood K, Hodson-Tole E, et al. Predicting the timing of the peak of the pubertal growth spurt in elite male youth soccer players: evaluation of methods. *Ann Hum Biol*. 2020;47(4):400–8.
7. International Atomic Energy Agency. Radiation Safety of X Ray Generators and Other Radiation Sources Used for Inspection Purposes and for Non-medical Human Imaging, IAEA Safety Standards Series No. SSG-55, IAEA, Vienna. 2020. Vienna: International Atomic Energy Agency; 2020.
8. Engebretsen L, Steffen K, Bahr R, et al. The International Olympic Committee Consensus statement on age determination in high-level young athletes. *Br J Sports Med*. 2010;44(7):476–84.
9. Greulich WW, Pyle SI. *Radiographic Atlas of the Skeletal Development of the Hand and Wrist*. 2nd ed. Palo Alto, CA: Stanford University Press; 1959, 238, 393.
10. Acheson RM. The Oxford method of assessing skeletal maturity. *Clin Orthop*. 1957;10:19–39.
11. Tanner JM, Whitehouse RH, Cameron N, Marshall WA, Healy M, Goldstein H. *Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method)*. 2nd ed. London: Academic Press; 1983.
12. Tanner JM, Healy M, Goldstein H, Cameron N. *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)*. 3rd ed. London: Academic Press; 2001.
13. Joss EE, Temperli R, Mullis PE. Adult height in constitutionally tall stature: accuracy of five different height prediction methods. *Arch Dis Child*. 1992;67(11):1357–62.
14. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans Med Imaging*. 2009;28(1):52–66.
15. Thodberg HH, Jenni OG, Ranke MB, Martin DD. Standardization of the Tanner-Whitehouse bone age method in the context of automated image analysis. *Ann Hum Biol*. 2012;39(1):68–75.
16. Thodberg HH, Juul A, Lomhot J, et al. *Adult Height Prediction Models*. New York: Springer; 2012.
17. Lenko HL. Prediction of adult height with various methods in Finnish children. *Acta Paediatr Scand*. 1979;68(1):85–92.
18. Tanner JM. Issues and advances in adolescent growth and development. *J Adolesc Health Care*. 1987;8(6):470–8.
19. Preece MA. Prediction of adult height: methods and problems. *Acta Paediatr Scand Suppl*. 1988;347:4–11.
20. Malina RM, Cumming SP, Rogol AD, et al. Bio-banding in youth sports: background, concept, and application. *Sports Med*. 2019; 49(11):1671–85.
21. Ostojic SM. Prediction of adult height by Tanner-Whitehouse method in young Caucasian male athletes. *QJM*. 2013;106(4):341–5.
22. Cameron N. Prediction. In: Molinari L, Cameron N, Hauspie RC, editors. *Methods in Human Growth Research*. Cambridge: Cambridge University Press; 2004. pp. 354–73.
23. Bayley N. Tables for predicting adult height from skeletal age and present height. *J Pediatr*. 1946;28:49–64.
24. Roemmich JN, Blizzard RM, Peddada SD, et al. Longitudinal assessment of hormonal and physical alterations during normal puberty in boys. IV: predictions of adult height by the Bayley–Pinneau, Roche–Wainer–Thissen, and Tanner–Whitehouse methods compared. *Am J Hum Biol*. 1997;9(3):371–80.
25. Alshamrani K, Offiah AC. Applicability of two commonly used bone age assessment methods to twenty-first century UK children. *Eur Radiol*. 2020;30(1):504–13.
26. Martin DD, Schittenhelm J, Thodberg HH. Validation of adult height prediction based on automated bone age determination in the Paris

Longitudinal Study of healthy children. *Pediatr Radiol*. 2016;46(2): 263–9.
27. Roche AF, Davila GH, Eyman SL. A comparison between Greulich–Pyle and Tanner–Whitehouse assessments of skeletal maturity. *Radiology*. 1971;98(2):273–80.
28. Alshamrani K, Messina F, Offiah AC. Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. *Eur Radiol*. 2019;29(6):2910–23.
29. Alshamrani K, Hewitt A, Offiah AC. Applicability of two bone age assessment methods to children from Saudi Arabia. *Clin Radiol*. 2020;75(2):156.e1–9.
30. Brämswig JH, Fasse M, Holthoff ML, von Lengerke HJ, von Petrykowski W, Schellong G. Adult height in boys and girls with untreated short stature and constitutional delay of growth and puberty: accuracy of five different methods of height prediction. *J Pediatr*. 1990;117(6):886–91.
31. Cameron N, Mirwald RL, Bailey DA, Davies PS. The application of new height-prediction equations (Tanner–Whitehouse mark 2) to a sample of Canadian boys. *Ann Hum Biol*. 1985;12(3):233–9.
32. de Waal WJ, Greyn-Fokker MH, Stijnen T, et al. Accuracy of final height prediction and effect of growth-reductive therapy in 362 constitutionally tall children. *J Clin Endocrinol Metab*. 1996;81(3): 1206–16.
33. Harris EF, Weinstein S, Weinstein L, Poole AE. Predicting adult stature: a comparison of methodologies. *Ann Hum Biol*. 1980;7(3): 225–34.
34. Zachmann M, Sobradillo B, Frank M, Frisch H, Prader A. Bayley–Pinneau, Roche–Wainer–Thissen, and Tanner height predictions in normal children and in patients with various pathologic conditions. *J Pediatr*. 1978;93(5):749–55.
35. Malina RM. Skeletal age and age verification in youth sport. *Sports Med*. 2011;41(11):925–47.
36. Wood SN. *Generalized Additive Models: An Introduction with R*. 2nd ed. Boca Raton: CRC Press; 2017.
37. Lenth RV, Buerkner P, Herve M, Love J, Riebl H, Singmann H. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.4. https://cran.r-project.org/web/packages/emmeans/index.html. In: 2021.
38. Marra G, Wood SN. Coverage properties of confidence intervals for generalized additive model components. *Scand J Stat*. 2012;39(1): 53–74.
39. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol*. 2003;22(1):85–93.
40. Francq BG, Lin D, Hoyer W. Confidence, prediction, and tolerance in linear mixed models. *Stat Med*. 2019;38(30):5603–22.
41. Cole TJ. Optimal design for longitudinal studies to estimate pubertal height growth in individuals. *Ann Hum Biol*. 2018;45(4):314–20.
42. Simpkin AJ, Sayers A, Gilthorpe MS, Heron J, Tilling K. Modelling height in adolescence: a comparison of methods for estimating the age at peak height velocity. *Ann Hum Biol*. 2017;44(8):715–22.
43. 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):221–30.
44. Deprez D, Fransen J, Boone J, Lenoir M, Philippaerts R, Vaeyens R. Characteristics of high-level youth soccer players: variation by playing position. *J Sports Sci*. 2015;33(3):243–54.
45. Aynsley-Green A, Cole TJ, Crawley H, Lessof N, Boag LR, Wallace RM. Medical, statistical, ethical and human rights considerations in the assessment of age in children and young people subject to immigration control. *Br Med Bull*. 2012;102:17–42.
46. Cole AJ, Webb L, Cole TJ. Bone age estimation: a comparison of methods. *Br J Radiol*. 1988;61(728):683–6.
47. Martin DD, Wit JM, Hochberg Z, et al. The use of bone age in clinical practice—part 1. *Horm Res Paediatr*. 2011;76(1):1–9.
48. Khamis HJ, Roche AF. Predicting adult stature without using skeletal age: the Khamis–Roche method. *Pediatrics*. 1994;94(4 Pt 1):504–7.
49. 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):1269–78.
50. Molinari L, Gasser T, Largo R. A comparison of skeletal maturity and growth. *Ann Hum Biol*. 2013;40(4):333–40.