1 2 9 0

UNIVERSIDADE D
COIMBRA

Paulo Moura Relvas de Sousa e Silva

ASSESSMENT OF BIOLOGICAL MATURATION IN MALE ADOLESCENT ATHLETES

PhD Thesis of the Doctorate Program in Sport Sciences, Branch of Sport Training, supervised by Professors Manuel João Cerdeira Coelho-e-Silva, Robert Marion Malina, and Sean P. Cumming, and submitted to the Faculty of Sport Sciences and Physical Education of the University of Coimbra

December 2022

University of Coimbra
Faculty of Sport Sciences and Physical Education

Assessment of Biological Maturation in Male Adolescent Athletes

Paulo Moura Relvas de Sousa e Silva

PhD Thesis of the Doctorate Program in Sport Sciences, Branch of Sport Training, supervised by Professors Manuel João Cerdeira Coelho-e-Silva, Robert Marion Malina, and Sean P. Cumming, and submitted to the Faculty of Sport Sciences and Physical Education of the University of Coimbra

December 2022

1 2 9 0

UNIVERSIDADE D COIMBRA

Sousa-e-Silva, P. (2022). *Assessment of biological maturation in male adolescent athletes*. PhD Thesis of the Doctorate Program in Sport Sciences, Branch of Sport Training. University of Coimbra. Coimbra, Portugal

“Per aspera ad astra.”
Lucius Annaeus Seneca

To my dear mother, Maria Paula de Melo Moura Relvas† (1953-2019).
Thank you for everything!

Para a minha querida mãe, Maria Paula de Melo Moura Relvas† (1953-2019).
Obrigado por tudo!

Acknowledgments

There were many people who directly or indirectly contributed to the achievement of another stage in my life.

I was lucky to have three mentors who, although with different roles, had a great influence on my life and academic journey: 1) Professor Manuel João Coelho-e-Silva, my academic mentor, I thank you for all your support, not only for help, availability, and constant feedback for the realization of this thesis, but also for all contribution to my professional and personal growth. Thank you, professor! 2) Miguel Guerreiro Silva (Migo), my mentor in sport training, I thank you mainly for the values transmitted over these last 20 years, as a coach and also as a friend. I highlight three: Loyalty, honesty and resilience.3) To my mother, to whom I dedicate this work. There are no words that express everything she has done for me, and I know how proud she would be to see me finish this step! Thank you, mother!

To my co-supervisors, Professor Robert Malina and Professor Sean Cumming, thank you for all the contribution, guidance, and expertise sharing to the realization of this thesis.

To my Journal Club colleagues João Pedro Duarte, Diogo Martinho, Daniela Costa, Raquel Costa, thank you for all the mutual help, sharing and support all over these years. I also highlight the true friendships I developed with some of them!

Thank my family, my friends and the friends who became family! A special thank you to my wife Joana, for all her understanding and unconditional support, especially in these last 4 years.

Finally, thank the most important people in my life because I fight for them every day, my children Maria and Manuel.

The work presented in this thesis was partially supported by the Portuguese Foundation for Science and Technology (MCTES)

Grant: SFRH/BD/138608/2018

Abstract

The present thesis is divided into eight chapters. Chapter I introduce the basic concepts related to growth and biological maturation, the historical perspective of skeletal maturation, the most common somatic indicators, and a review of the literature about the use of biological maturation in youth sports. Chapter II was dedicated to the general methodology used in the studies that integrated the present thesis. Studies that report data quality on the skeletal age assessment methods were lacking, mainly in the scoring methods, Tanner-Whitehouse (TW) and FELS, whereby the two subsequent chapters, III and IV, aimed to assess data quality on TW version 2 (TW2) and FELS methods. These studies were a novelty, and it was highlighted: (i) overall agreement is very acceptable, both for intra- and inter-observer agreement; (ii) in TW method it is important not to look only for the total disagreements but at the intraclass correlation coefficient (ICC), since different bones have different scores. For example, capitate was the bone with more agreement, but with a lower ICC, as it is a bone with great scoring; (iii) in FELS method trapezium and trapezoid appeared as most problematic bones for both observers. (iv) Although the error exists, especially in inter-observer agreement, the impact on skeletal age (SA) is negligible. After verifying data quality on SA assessment methods, next chapter (chapter V) aimed to assess the concordance between concurrent methods of SA assessment, considering a multisport sample of male adolescent athletes from different sports. Difference between SA and CA increased from 11.0-11.9 until 14.0-14.9, in all the three concurrent methods, GP, TW2 RUS, FELS. After classifying players by maturity status (late, on time, advanced), cross-tabulation showed FELS with a greater agreement with TW2 RUS than with GP. The next chapter (VI) was dedicated mainly to compare two methods that estimate parameters of the growth curve, fitting longitudinal height records of 58 male soccer players assessed during five seasons. Last study (chapter VII) compares estimated age at peak height velocity (APHV) assessed with two maturity offset equations (the original from Mirwald et al., 2002; the modified from More, 2015), with observed APHV. The sample were the same from the preceding chapter. Results confirmed that the use of maturity offset is limited to average matures, with CA near to the PHV in the moment of the assessment. Finally, last chapter (VIII) summarizes the main findings of the present thesis. Cross-sectional studies highlighted TW and FELS as reliable methods to assess SA, although some careful is needed as it was noted in study 3 that different methods can produce different results, especially at 14-14.9 years. Longitudinal data fitted by SITAR model showed that somatic indicator such as maturity offset is not a valid indicator of maturity timing and status in adolescent athletes.

Key words: Skeletal age, Greulich-Pyle; Tanner-Whitehouse; FELS; Somatic indicators; Peak height velocity; Youth athlete; SITAR.

Resumo

A presente tese está dividida em oito capítulos. O Capítulo I apresenta os conceitos básicos relacionados com o crescimento humano e maturação biológica, a perspetiva histórica da maturação esquelética, os indicadores somáticos mais comuns e uma revisão da literatura sobre a aplicação da maturação biológica em jovens atletas. O Capítulo II foi dedicado à metodologia geral utilizada nos estudos que integraram a presente tese. Tendo-se percebido que estavam em falta estudos que relatassem a qualidade dos dados sobre os métodos de avaliação da idade óssea, principalmente nos métodos descritivos, Tanner-Whitehouse (TW) e FELS, os dois capítulos subsequentes, III e IV, tiveram como objetivo avaliar a qualidade dos dados nos protocolos TW versão 2 (TW2) e FELS. Destacou-se o seguinte: (i) a concordância geral é bastante aceitável, tanto intra- quanto inter-observador; (ii) no método TW é importante não olhar apenas para os desacordos totais, mas sim para o coeficiente de correlação intraclasse (ICC), pois ossos diferentes têm pontuações diferentes. Por exemplo, o capitato (grande osso) foi o osso com maior concordância, porém com menor ICC, por ser um osso com uma grande contribuição na pontuação total; (iii) no método FELS o trapézio e o trapezoide foram os ossos mais problemáticos para ambos os observadores; (iv) Embora o erro exista, principalmente na concordância inter-observador, o impacto na idade esquelética (SA) é insignificante. Após verificar a qualidade dos dados sobre os métodos de avaliação da SA, o seguinte capítulo (capítulo V) teve como objetivo avaliar a concordância entre métodos concorrentes de avaliação da SA, considerando uma amostra substantiva de 1778 atletas adolescentes masculinos de diferentes modalidades desportivas. A diferença entre a SA e a idade cronológica (CA) aumentou gradualmente de 11.0-11.9 para 14.0-14.9, em todos os três métodos, Greulich-Pyle (GP), TW2 RUS, FELS. Depois de classificar os jogadores de acordo com o seu estado de maturação esquelética (atrasado, normomaturo, avançado), a tabulação cruzada mostrou o método de FELS com uma maior concordância com o método de TW2 RUS do que com o método de GP. O capítulo seguinte (VI) foi dedicado principalmente a comparar dois métodos que estimam parâmetros da curva de crescimento, numa amostra longitudinal de 58 jogadores de futebol masculino avaliados durante cinco épocas desportivas. O último estudo (capítulo VII) compara a idade estimada no pico de velocidade de crescimento (APHV) através de duas equações preditivas (a original - Mirwald et al., 2002; e uma modificada - Moore, 2015), com a APHV observada. A amostra foi a mesma do capítulo anterior. Os resultados confirmaram que o uso das equações preditivas é limitado a atletas normomaturos, ou com uma CA próxima ao PHV no momento da avaliação. Finalmente, o último capítulo (VIII) resume as principais conclusões da presente tese. Estudos transversais destacaram os métodos de TW e FELS como métodos fidedignos para avaliar a SA, embora alguns cuidados sejam necessários, já que diferentes métodos podem produzir diferentes resultados, especialmente nas idades compreendidas entre os 14.0-14.9 anos, tal como observado no estudo 3. Os dados longitudinais referentes à estatura dos jovens futebolistas, ajustados através do modelo SITAR, demonstraram que a equação preditiva para estimar o pico de velocidade crescimento apresenta sérias limitações.

Palavras-chave: Idade esquelética; Greulich-Pyle; Tanner-Whitehouse; FELS; Indicadores Somáticos; Pico de velocidade de crescimento; Jovem atleta; SITAR.

Contents

ACKNOWLEDGMENTS .... **V**
ABSTRACT .... **VIII**
RESUMO .... **IX**
CONTENTS .... **XI**
LIST OF TABLES .... **XIV**
LIST OF FIGURES .... **XVIII**
LIST OF ABBREVIATIONS .... **XX**

Chapter I
1. GENERAL INTRODUCTION **3**
1.1. Basic concepts 3
1.2. Growth 4
1.3. Growth spurt 6
1.4. Skeletal maturation 6
1.4.1. Greulich-Pyle method to assess skeletal age 8
1.4.2. Tanner-Whitehouse method to assess skeletal age 11
1.4.3. FELS method to assess skeletal age 18
1.5. The algorithm to predict maturity-offset 21
1.5.1. Validation studies 23
1.6. Prediction of adult size 25
1.7. Youth Sports 26
1.8. Objectives 28
1.9. References 28

Chapter II
2. METHODS **35**
2.1. Procedures and Sample 35
2.2. Anthropometry 35
2.3. Biological maturation 36
2.3.1. Skeletal maturation 36
2.3.2. Somatic maturation 38

2.4. Analyses....39
2.5. References....39

Chapter III: Study 1

3. Skeletal age assessed by TW2 using 20-bone, carpal and RUS score systems: Intra-observer and inter-observer agreement among male pubertal soccer players....43

3.1. Abstract....45
3.2. Introduction....56
3.3. Methods....48
3.4. Results....50
3.5. Discussion....58
3.6. Conclusions....61
3.7. References....61

Chapter IV: Study 2

4. Assessment of intra-observer reproducibility and inter-observer agreement of Fels protocol to determine skeletal age among male adolescent tennis players.67

4.1. Abstract....69
4.2. Introduction....70
4.3. Methods....71
4.4. Results....73
4.5. Discussion....77
4.6. Conclusions....80
4.7. References....81

Chapter V: Study 3

5. Concurrent assessment of skeletal age among male adolescents of different sports....87

5.1. Abstract....89
5.2. Introduction....90
5.3. Methods....92
5.4. Results....94
5.5. Discussion....102

5.6. References 104

Chapter VI: Study 4

6. Ages at peak height velocity in male soccer players 11-16 years: relationships with skeletal age and comparisons among longitudinal studies 109

6.1. Abstract 111

6.2. Introduction 112

6.3. Methods 114

6.4. Results 118

6.5. Discussion 127

6.6. Conclusions 130

6.7. References 130

Chapter VII: Study 5

7. Observed and predicted ages at peak height velocity in soccer players 137

7.1. Abstract 139

7.2. Introduction 140

7.3. Methods 141

7.4. Results 145

7.5. Discussion 153

7.6. Conclusions 158

7.7. References 159

Chapter VIII

8. GENERAL DISCUSSION 169

8.1. The importance of data quality in the determination of skeletal age 169

8.2. Interrelationship between skeletal and somatic maturity 174

8.3. Conclusions, practical application and future directions 176

8.4. References 177

List of Tables

Table 1.1. Differences between growth, maturation, and development (adapted from Malina et al., 2004)...3

Table 1.2. Sample in which the developmental stages from the Harpenden growth study were assigned...12

Table 1.3. Sample used in Tanner-Whitehouse method, second edition (1975) ...15

Table 1.4. Sample used in Tanner-Whitehouse method, third edition (2001)...17

Table 2.1. Characteristics of the five studies...35

Table 2.2. Analysis of the five studies...39

Table 3.1. Frequencies of developmental stages for each bone assessed with TW2 method assigned by observers A and B on two occasions (time moment 1 *versus* time moment 2), absolute and relative agreement rates, and intra-observer error in assessments of SA among 142 adolescent male soccer players ...52

Table 3.2. Descriptive statistics (mean ± standard deviation) for the scores of each bone in the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observer A on two occasions (time moment 1 *versus* time moment 2), paired t-tests, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients among the 142 adolescent male soccer players...53

Table 3.3. Descriptive statistics (mean ± standard deviation) for each bone score with the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observer B on two occasions (time moment 1 *versus* time moment 2), paired t-tests, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients in 142 adolescent male soccer players...54

Table 3.4. Descriptive statistics (mean ± standard deviation) for each bone score in the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observers A and B, paired t-test, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients in 142 adolescent male soccer players 55

Table 4.1. Descriptive statistics (means ± standard deviations) for skeletal ages of male tennis players (n=97) by time-moment (TM) separately for observer A and observer B; results of comparisons between time moments, technical errors of measurement (TEM) and coefficients of variation (%CV). Frequencies of intra-observer error by bone........75

Table 4.2. Cross-tabulations of maturity status classifications based on assessments of observers A and B among male adolescent tennis players for the total sample (n=97)..76

Table 5.1. Descriptive statistics for chronological age, stature, and body mass by age group among adolescents participating in competitive youth sports (n=1778)...............96

Table 5.2. Descriptive statistics for chronological age, skeletal age (FELS method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778)..97

Table 5.3. Descriptive statistics for chronological age, skeletal age (TW2 RUS method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778)..98

Table 5.4. Descriptive statistics for chronological age, skeletal age (Greulich-Pyle method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778)...99

Table 5.5. Cross-tabulation between frequencies of skeletal maturity categories derived from concurrent methods used to estimate skeletal age and statistics of agreement between protocols...100

Table 6.1. Means (M) and standard deviations (SD) for chronological age (CA), skeletal age (SA), height and weight for the longitudinal sample by observation 118

Table 6.2. Estimated parameters for take-off (TO) and peak height velocity (PHV) of the adolescent spurt based on the SITAR and FPCA methods applied to the longitudinal height records of 58 soccer players and differences between the respective estimates (SITAR minus FPCA) with the two methods (M, mean; SD, standard deviation) 119

Table 6.3. Frequencies and cross-tabulations of maturity status classifications (late, on time, early)[1] based on ages at PHV with the SITAR and FPCA models, percentage agreement, Chi square (χ^2) and Cohen's Kappa (κ); means and standard deviations for ages at PHV in the respective maturity groups are also indicated 120

Table 6.4. Frequencies and cross-tabulations of maturity status classifications based on ages at PHV with the SITAR and FPCA models and on Fels skeletal ages (SA – CA) at observations 1, 3 and 5, percentage agreement, Chi square (χ^2) and Cohen's Kappa (κ); means and standard deviations for SA – CA differences in the respective maturity groups are also indicated 122

Table 7.1. Sample sizes and descriptive statistics for chronological age (CA) at prediction, observed maturity offset and predicted maturity offset and ages at PHV, and the difference of predicted age at PHV minus observed ages at PHV (criterion) with the original (Mirwald) and modified (Moore) equations in soccer players at each observation.† 147

Table 7.2. Intercepts and slopes, and respective standard errors (SE) and 95% confidence limits based on the weighted Deming regression of predicted ages at PHV (y-axis) and observed (actual) ages at PHV (x-axis) for the Mirwald and Moore prediction equations at each observation (Obs) in youth soccer players.* 147

Table 7.3. Sample sizes and descriptive statistics (mean ± standard deviation) for chronological age (CA) at prediction, observed maturity offset and predicted maturity offset, predicted ages at PHV and the difference of predicted age at PHV minus observed ages at PHV (criterion) with the original (Mirwald) and modified (Moore) equations at each observation in players classified as advanced, average and delayed based on observed ages at PHV.† 148

Table 7.4. Number of participants by maturity status† according to predicted ages at PHV with the Mirwald equation and, separately, with the Moore equation who were within ±0.50 year of observed age at PHV (SITAR model) at each observation in youth soccer players..152

Table 7.5. Correlations at each observation between predicted maturity offset and predicted APHV (CA – predicted maturity offset) with chronological age (CA), height (Ht), sitting height (SitHt), estimated leg length (LegLt) and body weight (Wt) for the Mirwald equation and with CA and Ht for the Moore equation 154

Table 8.1. Data quality on Tanner-Whitehouse method: versions 1,2, 3......................172

Table 8.2. General characteristics of concurrent methods to assess SA: GP, TW2, and FELS...173

List of Figures

Figure 3.1. Interobserver agreement for the determination of skeletal age by two observers, considering concurrent systems (20-bone, Carpal and RUS) and indication of the mean for each observer, bivariate correlation coefficient between the series produced by observers A and B, intra-class correlation coefficient; complemented with Bland-Altman analysis to inspect intra-individual differences expressed against the mean values 57

Figure 4.1. SAs of individual tennis players assigned by observer A (x-axis) are plotted relative to those assigned by observer B (y-axis) with the respective means and standard deviations (panel A); Bland-Altman analysis of the intra-individual differences (y-axis) relative to the mean of both observers (x-axis) (panel B) 76

Figure 5.1. Mean stature and mean body mass of the competitive age-groups (11-12 years, 13-14 years and 15-16 years) plotted relative to U.S. percentiles (n = 1778), considering CA (Panels A-B), and SA (Panels C-H)...101

Figure 6.1. Aggregation of ages of PHV (years) in samples of male soccer players based on meta-analysis, including the subgroup analysis. Q: Cochran's Q-statistic (weighted sum of squares), QM: Cochran's Q-statistic for subgroups, I2: percentage of variability in effect sizes which is not due sampling error, τ2: between-study variance in a given set of samples (years squared); plots: means and 95% CIs of individual studies; diamonds: width represents 95% CI for each model aggregated by subsample and for all studies (below)...126

Figure 7.1. Predicted ages at PHV with the Mirwald equation plotted relative of observed age at PHV at each observation for individual soccer players classified as early, average and late maturing. The diagonal corresponds to the line of identity (x=y) 150

Figure 7.2. Predicted ages at PHV with the Moore equation plotted relative of observed age at PHV at each observation for individual soccer players classified as early, average and late maturing. The diagonal corresponds to the line of identity (x=y) 150

Figure 7.3. Means for predicted maturity offset (MO) plotted relative to means for chronological age for each year of observation in the current sample of players (filled diamonds) and in samples of male soccer players extracted from the literature...........156

Figure 7.4. Means for predicted maturity offset (MO) plotted relative to means for height for each year of observation in the current sample of players (filled diamonds) and in samples of male soccer players extracted from the literature..156

List of Abbreviations

%CV	Coefficient of variation
χ^2	Chi-square
A	Advanced
APHV	Age at peak heigh velocity
BMI	Body Mass Index
BMJ	British Medical Journal
B-spline	Basis spline
BTT	Bock-Thissen-du Toit
CA	Chronological age
CI	Confidence interval
CL	Confidence limits
d	d-cohen value
D	Delayed
f	Absolute frequencies
FDA	Functional Data Analysis
FFM	Fat free mass
FM	Fat mass
FPCA	Functional Principal Component Analysis
GP	Greulich-Pyle
Ht	Height
ICC	Intra-class correlation coefficients (ICC)
κ	Cohen's Kappa
KG	Khamis-Guo
KR	Khamis-Roche
LegLt	Estimated leg length
Lin CCC	Lin concordance correlation coefficient
LLTS	Leuven Longitudinal Twin Study
LLTS	Leuven Longitudinal Twin Study
LTAD	Long-term athlete development
M	Mean
MAD	Median absolute deviation

MO	Maturity offset
Mod	Modified
mSv	Millisievert
NR	Not reported
NY	New York
Obs	Observer
OT	On time
p	Significance value
PB-I	Preece-Baines model I
PBMAS	Pediatric Bone Mineral Accrual Study
PHV	Peak height velocity
PWV	Peak weigh velocity
rho	Spearman correlations
RUS	Radius-ulna-short bones
SA	Skeletal age
SD	Standard deviation
SE	Standard errors
SGDS	Saskatchewan 5 Growth and Development Study
SGDS	Saskatchewan Growth and Development Study
SITAR	Superimposition by Translation and Rotation
SitHt	Sitting height
t	T-value of paired t-test
TEM	Technical error of measurement
TM	Time-moment
TM1	Time-moment 1
TM2	Time-moment 2
TO	Take-off
TW	Tanner-Whitehouse method
TW1	TW method - version 1
TW2	TW method - version 2
TW3	TW method - version 3
U13	Under-13 players
Wt	Body weight
Yrs	Years

Chapter I

General introduction

1. GENERAL INTRODUCTION

1.1. Basic concepts

Growth, maturation, and development are three interacting processes that children, and adolescent experience. The terms are often synonymously used, although each has a specific meaning (Malina, Bouchard, & Bar-Or, 2004). In fact, growth is the dominant activity in the first two decades of human being life and it refers to an increase in the whole-size or specific parts of the body. As an infant growth, stature and body mass increase. Different parts of the body follow a specific growth curve. For example, legs growth earlier and attain adult size compared to the trunk. Meantime, biological maturation is defined as the process of achieving the adult state. Maturation corresponds to an underlying process while maturity refers to a state at a given time. Although biological maturation is an unique process, observable indicators can vary according to the biological system. For example, skeletal maturation can be defined as the process between the cartilaginous tissue of the fetus and the fully ossified skeleton of the adult. Development can be viewed in different contexts. It refers to the specialization of the cells when tissues are being formed. In parallel, in the context of motor development, it summarizes acquisition and refinement of motor patterns and skills. Note, however, that development can also be in terms of psychological or social domains. Table 1.1 summarize the main differences between these three concepts: growth, maturation and development.

Table 1.1. Differences between growth, maturation, and development (adapted from Malina et al., 2004).

GROWTH	MATURATION	DEVELOPMENT
Size	Skeletal	Cognitive
Proportions	Sexual	Emotional
Physique	Somatic	Social
Composition	Neuroendocrine	Motor
Systemic	Neuromuscular	Moral

Two terms are implicit in the concept of maturation. Timing refers to a specific date in which a maturational event occurs (e.g., age at peak height velocity). Tempo refers to the rate of the progression (e.g., how quickly or slowly an individual passes from the appearance of a bone center to a fully ossified bone). All individuals reach the adult state, but their timing (size attained at a given point) and tempo (rate in attaining the mature state) vary.

Mirwald et al. (2002) describes an example to explain the concept of timing and tempo: Two males (A and B) were tested at 11.4 years (A) and 11.3 years (B), showing difference in stature of 4.7 cm, and in weigh of 6.2 kg. The two were classified as pre-PHV. At 14.0 years, the heavier and taller individual already attained his peak height velocity and, consequently, is classified as post-PHV, while the shortest and lightest individual had not yet reached the PHV, classified as Pre-PHV. Inter-individual difference in stature increased to 25.8 cm. Corresponding value for body mass was 13.7 kg.   At 17.0 years, the shortest and lightest individual attained PHV and the difference between the two individuals was less than one centimeter and about two kilograms.   They were about the same size before and after adolescence but reached PHV at different timings (13.09 and 15.09 years). What differed was the tempo of growth, showing a differential progression to adult state.

1.2. Growth

The oldest record about human growth appeared in a Greek elegy (poetry) from the 6th Century B.C. Solon, a poet, lawmaker, and statesman from Atenas divided human growth into successive periods of seven years each, calling it “Hebdomads”. A translated version of the original elegy is presents on Tanner’s book (Tanner, 2010, pp. 1-2):

> *“A young boy acquires his first ring of teeth as an infant (literally while unable to speak) and sheds them before he reaches the age of 7 years. When the god brings to an end the next seven-year period, the boy shows the signs of beginning puberty (or: of beginning pubic hair). In the third hebdomad, the body enlarges, the chin becomes bearded and the bloom of the boy’s complexion is lost. In the fourth hebdomad physical strength is at its peak and is regarded as the criterion of manliness. In the fifth hebdomad a man should take thought of marriage and*

seek sons to succeed him. In the sixth hebdomad a man's mind is in all things disciplined by experience and he no longer feels the impulse to uncontrolled behavior. In the seventh he is at his prime in mind and tongue, as also in the eight, the two together making fourteen years. In the ninth hebdomad, though he still retains some strength, he is too feeble in mind and speech for the greatest excellence. If a man continues to the end of the tenth hebdomad, he has not encountered death before due time."

The term "Anthropometry" was developed by the German naturalist Johann Sigismund Elsholtz (1623-1688), the first man with the concerning of measuring the human body (Tanner, 2010). Elsholtz's graduation thesis, entitled "Anthropometria", was issued in 1654. The instrument used by Elsholtz was an anthropometer with a transverse rod, the *regula*, moving up and down the vertical rod. When it came to the representation of the growth in children, the artistic community was more well ahead of the medical one (Tanner, 2010), particularly because painters and sculptors needed instructions about body proportionalities. It was in 18th century, that appeared a book with the same name of Elsholtz thesis: "Anthropometria". The author was Johann Georg Bergmuller, an oil and fresco painter, professor of painting at Augsburg, Germany. The book deal specifically with changes of proportions during childhood (Tanner, 2010), and it gave a geometrical rule with the generation of a height curve which omitting any pubertal spurt.

The first ever longitudinal human growth study and the most famous of all records of Human Growth was made by Philippe Guéneau de Montbeillard, measuring his own son, between the date of his birth until he reached 18 years (1759 - 1777). Later, the experiment made by Montbeillard was published by Georges-Louis Leclerc, Comte de Buffon, in a supplement to the *Histoire Naturelle* (Leclerc, 1778, pp. 101-108). It was established the existence of the pubertal growth spurt, seasonal changes in growth spurt, and even the occurrence of shrinkage during the day. The presented measurements by Bufon are in *Pieds, peuces, and ligues*, as the meter *métre* was only introduced in 1795 by the French Delambre and Méchain. A reproduction of the growth curve with the data from Bufon, is presented in the book of Tanner: "Fetus into man" (Tanner, 1978).

1.3. Growth spurt

In boys, adolescent growth spurt begins at 11-12 years old, while in girls tend to occur slightly earlier, that is about 9-10 years. The growth rate increases until it reaches a peak called peak height velocity (PHV). It is an indicator of somatic maturation. Beside PHV, take-off (TO) corresponding to the initiation of the spurt, is the other primary parameter in the growth curve (Malina et al., 2004). Growth spurt in body mass begins slightly later that the growth spurt in height. Body weight can be divided in two major components: fat mass (FM) and fat free mass (skeletal muscle and bone mineral) (FFM). FFM has a growth pattern like body weight and FM tend to increase more gradually during childhood and adolescence (Malina et al., 2004).

After the PHV the rate of the growth spurt in height slows although growth continues until 16-18 years old and 18-20 years old, respectively in girls and boys. Growth curves for children and adolescent were published by CDC (Kuczmarski et al., 2002), and these data gives information about the size attained at a specific age, from 2 to 20 years. Data combined national surveys from 1959 to 2000, from different ethnic populations. Body mass from children from National Health and Nutrition Examination Survey II (1976-1980) and III (1988-1994) were not included in the development of body mass growth charts, due to the prevalence of obesity among youths from USA, which could result in inadequate percentile curves from a public health viewpoint. Body weight, FFM and muscle mass occur, on average, several months after the PHV (Malina et al., 2004).

1.4. Skeletal maturation

Wilhelm Roentgen discovered the x-rays in November 1895. The first radiograph that was published was in January 1896 in the *British Medical Journal (BMJ)*, by Alan Campbell-Swinton, who obtained an x-ray of his own hand (Campbell-Swinton, 1896), although the first radiograph was made by Roentgen to his own wife, Anna Bertha Ludwig. Meantime, Sidney Rowland, investigated Roentgen's discovery. His first report occurred in a new Journal: *Archives of Clinical Skiagraphy*, that became the actual *British Journal of Radiology* (Rowland, 1896). Rowland named *skiagram* to what was called "the new photography”, because *skia* is the Greek word for a shadow. The term *Roentgenogram* or *Roentgenograph* was popularized. Besides that, Wilhelm Roentgen suggested the term *x-rays* (x: algebraic symbol for an unknown quantity), radiation that generated the *radiography* (Eisenberg, 1995).

In a study of 500 healthy children at the Children’s Hospital of Boston, Dr. Arial Wellington George found that the most practical and reliable index of development was represented by the hand-wrist (Rotch, 1908). Through this, Rotch realized the importance of describing the developmental stages of skeletal maturation using the hand-wrist area, applying it to the many physical and educational problems that he considered at that time: school, athletics, and child labor. The sample was composed for 200 cases, perhaps 10 or 12 children of each CA (it is not fully detailed) and place them under their respective stages of anatomic development, designating these stages not by years but by letters: A, B, C, D, and so on until letter M. The purpose was grading for kindergarten, school, athletics, and child labor. These alphabetic grouping was made considering the appearance of the hand-wrist bones and their development (without specify the developmental differences: “much more advanced in development”). For example: A - cases which show *os magnum* (capitate) and *unciform* (hamate) bones, appearing in the first year of life; B – cases showing the presence of the lower epiphysis of radius, plus the bones present in group A; C – presence of *cuneiform* (triquetral), plus A and B. Heads of metacarpal bones and the epiphyses of the first phalanges also used as controls; M – “Very much more advanced in development than L and the pisiform bone almost as large as cuneiform . All the bones of the wrist are much more developed than in any previous group.”. In Rotch study, the author noted the following: 1) a large hand with larger bones does not necessarily imply advanced development; 2) Normal appearances shown in radiographs of boys and girls do not materially differ; 3) Left-hand didn’t show any material different in development compared with the right.

(Pryor, 1925) identified the time of ossification of the hand-wrist bones and the epiphyseal-diaphyseal fusion of the long bones, on a sample of 64 males (12.5 – 22.8 years) and 81 females (12.0 – 22.5 years). It was reported a distinct cartilaginous gap between the epiphyses and diaphysis in both hands of females and males under 13 years of age. After that, a rapidly increase in skeletal maturation was observed in females, comparing with males. Later, Hellman (1928) described the ossification of the hand-wrist bones, with emphasis on metacarpals and phalanges, among 60 girls from the Hebrew Orphan Asylum of New York (USA) with an initial aged 10.25-12.25 years, who were assessed for four consecutive years. Letter A indicated no epiphyseal-diaphyseal fusion, with an epiphysis distinctly separate from the diaphysis, until letter E, corresponding to complete fusion without any visible line. Hellman also reported the duration of each stage. Other findings by Hellman included: 1) is the first metacarpal a true metacarpal or a phalange? The epiphysis in the first metacarpal is at the proximal end, like in phalanges, while in the other four is at the distal end; 2) Regarding the ages of increments in the growth of the phalanges with growth in stature, it was found that the greatest increment in the bones of the hand preceded the greatest increment in the stature of the adolescents. The phalanges had the greatest increment at 13 years old, and one year after, at 14 years old, the subjects had the greatest increment in height (peak height velocity).

The preceding authors did not present systems to assess skeletal age. Todd (1937) was the first to publish an atlas of the hand-wrist skeletal maturation, on a process that lasted 10 years, from 1926 to 1936. The standards presented in the atlas cover a period that last from three months to 16.25 years in girls (35 standards), and 18.75 in boys (40 standards). First 15 months comprised an interval between standards of three months (3,6,9,12,15). After that, interval enlarged to six months (1.75 years to 16.25 years in girls, or 18.75 in boys). The sample were from Cleveland (Ohio, USA). They were all white children, mostly from European ancestry. The author reported a total of over 4000 children, with more than half been examined 12 or 13 times. In fact, 7060 radiographs were selected (3657 for males, 3403 for females), choosing one radiograph for each age group (40 for males and 35 for females). All films were organized in the order of their skeletal maturity progress. It was also common to find children two years advanced or retarded. A central group was selected with the exclusion of earlier and late children in order to represent a standard plate. Todd claimed that although this process appears to be simple and fast, it was the most monotonous and time-consuming tasks ever undertaken in his laboratory.

1.4.1. Greulich-Pyle method to assess skeletal age

Meanwhile, Greulich and Pyle (1950) emerged as the most well-known atlas of skeletal maturation of the hand-wrist (Greulich & Pyle, 1959). This method is considered a revision from Todd’s work. Although the first edition was released in 1950, it suffered few changes and a second edition is still being used (Greulich & Pyle, 1959). Briefly, four new male standards were added; a new female standard was added at age of 3 years; in some instances, a standard plate has been replaced;-the illustrations have been redrawn and the descriptions revised; supplementary material was added, particularly the tables by Bayley and Pinneau for adult height prediction from skeletal age.

In 1988, it was estimated that Greulich-Pyle method (GP) was used by about 76% of pediatricians (Roche, Cameron, & Thissen, 1988). It continues to be widely used today. The previous work for Todd (Bush Foundation collection at Western Reserve University School of Medicine) was used for the design of GP method. Greulich & Pyle (1959) stated that the use of a hand-wrist x-ray gives an objective measure about the progress of an individual biological maturation, being possible to compare with their peers from the same CA and sex. Nevertheless, authors make it clear that the hand-film assessment should be regarded as a complement to, and not as a substitute for, other valid methods of physical status of children. The standard plates presented in the Atlas, represent groups of healthy children. Each of the standards was selected from 100 films of children of the same age and sex. In most cases, the standard was representative of the central tendency. From the newborn to one year, the length between examinations was three months. From one year to five, was around six months, and from five to 18.0 (females) or 19.0 (males), was around one year, although in some ages the length is reduced to six months. A more detailed information is presented below:

- Female (27 plates): newborn, 3 months, 6 months, 9 months, 1 year, 1.25, 1.5, 2.0, 2.5, 3.0, 3.5, 4.2, 5.0, 5.75, 6.83, 7.83, 8.83, 10.0, 11.0, 12.0, 13.0, 13.5, 14.0, 15.0, 16.0, 17.0, 18.0 years.

- Male (31 plates): newborn, 3 months, 6 months, 9 months, 1 year, 1.25, 1.5, 2.0, 2.67, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 11.5, 12.5, 13.0, 13.5, 14.0, 15.0, 15.5, 16.0, 17.0, 18.0, 19.0 years.

The method presents drawings with the different stages of skeletal development of each one of the hand-wrist bones, accompanied with a brief description for each drawing. The authors recommend the following:

- To compare the film with the standard plate of the atlas, of the same sex and nearest CA;
- To consider subjacent standards, younger and older;
- To select the standard which is more closely to the film, using also the maturity indicators for a detailed assessment;
- After having the standard which appears more closely to the film, a more detailed comparison is made. It is recommended by the authors the habit of considering the bones assessment in a regular order, avoiding overlooking features that may be important in making the assessment: Radius, Ulna, Capitate, Hamate, Triquetral, Lunate, Scaphoid, Trapezium, Trapezoid, Pisiform, Metacarpals, Phalanges. The adductor sesamoid usually appears, after the pisiform ossification.
- If the bone is in the same stage of development of the corresponding standard in the atlas, it should be given the skeletal age that is in the standard. If not, it needs to be compared with the same bone from the adjacent standards. If the film does not correspond to any of the standards it should be given the mean age between the two more closely subsequent plates.

Acheson (1954) criticized the inspectional technique and proposed a different system: 1) every round bone and epiphysis can make its own contribution to the assessment, and so, assessment of a film can be made regardless of the pattern in which ossification is occurring; 2) small increases of maturity are recorded; 3) maturation is given a yardstick of its own, the units being Oxford Maturity units; 4) the same standards are used for both sexes, so that a direct comparison can be made between the unit status of a boy and a girl. In summary, the rationale was to create an unique unit of measure for skeletal maturity assessment. The sample was composed by an healthy group of children between six months and five years of age, participating in the Oxford Health Survey. That unit measurement, called oxford maturity units, was awarded into each bone. The author put the hypothesis that the carpus and epiphyseal ossification do not proceed at equal rates in all children, so probably it would not be appropriate to add the carpal and RUS bones scores together.

1.4.2. Tanner-Whitehouse method to assess skeletal age

Tanner and Whitehouse were noticed on Acheson article: “Tanner and his colleagues now are engaged in preparing standards for the hand and wrist from birth to maturity”(Acheson, 1957, p.25). In 1948, James Mourilyan Tanner, a well-known Pediatrician, along with his colleague Reginal Henry Whitehouse, from the Royal Army Medical Corps, started a government-funded research, in order to assess the effects of war on the nutritional status of children living in a children’s home in Harpenden, Hertfordshire, North London: The Harpenden Growth Study. Furthermore, they developed a longitudinal study that, over more than 20-year period, recorded the growth and photographed pubertal development of children until they reached their adult stature. The initial sample was composed by children with ages between 0.9 and 20 years, who were assessed between 1949 and 1969. They were predominantly children of manual workers or the lower middle class (Tanner, 2010). Whitehouse designed all the anthropometric instruments and became anthropometrist, making about 250.000 measurements between 1950 and 1975. Besides that, he rates all the x-rays from the Harpenden Growth Study.

A new system for assessing skeletal maturation of children emerged in 1959: Tanner-Whitehouse method or the “bone approach method”, with the descriptions and illustrations of the stages of development of 28 bones of the hand-wrist (radius, ulna, capitate, hamate, triquetral, lunate, trapezium, trapezoid, metacarpals (I to V), proximal phalanges (I to V), medial phalanges (II to V), and distal phalanges (I to V), together with radiographic instruction techniques. Stages of development were derived chiefly from the longitudinal material of the “Harpenden Growth study”, with x-rays available every 6 months or less, for 10 years. These stages had been assigned to each bone in the left hand of the following six groups (Table 1.2). They were all normal children, more or less healthy.

Table 1.2. Sample in which the developmental stages from the Harpenden growth study were assigned.

	Place	Boys	Girls	Study	Type of Data	Age covered	Years of the Study	Observations
1	London	91	82	Child Study Centre Group University of London	Mixed Longitudinal	1 month - 6 years	1952-?	/
2	Oxford	109	105	Oxford Child Health Survey (Dr. A. Stewart and Mr. D. Hewitt)	Mixed Longitudinal	6 months - 5 years	1946-1953	/
3	Ayr	462	444	State Schools (excluding Grammar Schools)	Cross-sectional	3-14 years	1956	Data from a Ministry of Health Survey.
4	Kilmarnock	450	470	State Schools (excluding Grammar Schools)	Cross-sectional	3-14 years	1956	Data from a Ministry of Health Survey.
5	Ganges (Sample are from all over UK)	212	/	Cadets at H.M.S. Ganges Training school for entering Royal Navy	Cross-sectional	15-16 years	1960	Sample from all over Britain, from much the same section of the population sampled in Kilmarnock and Ayr.
6	Mayfield	/	139	Mayfield Comprehensive School	Cross-sectional	14-16 years	1960	Approximately the same section of the population sampled in Kilmarnock and Ayr.
Total		1324	1240			1 month - 16 years	1952-1960	
		2564						

2177 x-rays taken from the same number of children in the cross-sectional survey (3 years - 16 years).
3000 x-rays taken from 387 children in the two mixed longitudinal survey (1 month - 5 years).

The scoring system was released in 1962 (Tanner, Whitehouse, & Healy, 1962). The stages were A, B, C, D, E, F, G, H, I, J instead of numeric stages 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, to avoid the impression that stages were equally spaced; no difference has been demonstrated between left and right hand; second and fourth digit were excluded since they did not add more additional information than the three or four digit: 20-bone system. Each stage corresponds to a certain score and after assessing the 20 bones, the total score had a corresponding skeletal age. Maximum score which can be achieved is 1000 points, corresponding to 18.9 years in boys and 16.9 years in girls. The scoring system use the rationale that long and round bones are controlled by different factors, and consequently a different score was desirable. A simple mean of all bones would give too much emphasis to the long bones, mainly the metacarpals and phalanges, whose maturation is closely linked. Adductor sesamoid and pisiform were omitted. The weight-score of the 20-bone system is the following: 200 points Radius and Ulna (10% each), 300 points metacarpals and phalanges (10% each digit, considering metacarpal and phalanges as a whole), 500 points for carpals (about 7% each one of the seven carpals).

Each stage has one, two, or three written criteria, marked as (i), (ii), or (iii), and that written criteria should be followed with caution, with illustrations serve only as a complement for the assessment. If a stage has only one written criteria, then it should be present for considering that stage; if a stage has two written criteria, then it is sufficient to have only one met; if a stage has three written criteria, then two of them must be met to consider that stage. Additionally, the written criteria (i) from the previous stage must be present.

The method suffered the first revision in 1975 (Tanner et al., 1975). This second revision (TW2) eliminated the final stages of the radius (stage J), ulna (stage I), and five of the seven carpals (stage I; with the exception of hamate and trapezium), due to the difficulty of rating. However, the criteria for the indicators were not modified. The scores attached to each stage were changed, differentiated for boys and girls. This revision provided alternative skeletal ages based, respectively, on three possible scoring systems: 20-bone (using the 20 bones of the hand-wrist like in TW1), CARPAL (only the carpal bones), and RUS (radius, ulna, and short bones). The authors considered that RUS maturity would be preferable. The sample was almost the same as the sample from the first edition, with a few differences: the cross-sectional samples were exactly the same (Ayr, Kilmarnock, Ganges, and Mayfield), longitudinal surveys pass from 3000 x-rays to 5500 x-rays (London – from 173 boys and girls to 200; the Harpenden sample which was used to design the stages of development of the method was used cross-sectionally in TW2 – 111 boys and girls), the Oxford longitudinal survey used cross-sectionally was not altered from TW1 to TW2. Detailed information is present in Table 1.3.

Table 1.3. Sample used in Tanner-Whitehouse method, second edition (1975).

	Place	Boys	Girls	Study	Type of Data	Age covered	Years of the Study	Observations
1	London	100	100	International Children's Centre Co-ordinated Growth Study Teams, London Study	Longitudinal data used cross-sectionally	1 month - 18 years	1952-1972	Approximate number (100 boys, 100 girls). Sample numbers decreasing to 50 of each sex at 18 years.
2	Harpenden	52	59	Children's Home (Harpenden Growth Study)	Mixed Longitudinal used cross-sectionally	5-21 years	1950-1970	/
3	Oxford	109	105	Oxford Child Health Survey Hewitt and Stewart (1952), Hewitt (1959)	Mixed Longitudinal used cross-sectionally	6 months - 5 years	1946-1953	/
4	Ayr	462	444	State Schools (excluding Grammar Schools)	Cross-sectional	3 years - 14 years	1956	Data from a Ministry of Health Survey.
5	Kilmarnock	450	470	State Schools (excluding Grammar Schools)	Cross-sectional	3 years - 14 years	1956	Data from a Ministry of Health Survey.
6	Ganges	212	/	Cadets at H.M.S. Ganges Training school for entering Royal Navy	Cross-sectional	15 years - 16 years	1960	Sample from all over Britain, from much the same section of the population sampled in Kilmarnock and Ayr.
7	Mayfield	/	139	Mayfield Comprehensive School	Cross-sectional	14 years - 16 years	1960	Approximately the same section of the population sampled in Kilmarnock and Ayr.
TOTAL		1385	1317			1 month - 21 years	1946-1972	
		2702						

In 2001 a third version (Tanner, Healy, Goldstein, & Cameron, 2001), originally called EA90 was adopted, although some authors prefer to use TW2 for assessing athletes (Malina et al., 2018). This last version retained the RUS and CARPAL systems, eliminating the 20-bone system. Criteria and assigned scores for each bone were kept and tables for converting the sum of maturity scores for the Carpal SA were not modified, but those for RUS SA were. To the British sample of 2702 children from the previous versions of TW (UK60), were added a Belgian sample (B70), a Spanish sample (S80), a Japanese sample (J85), an Italian sample (I90), an Argentinian sample (Arg70), and a North American sample (Tx90). A more detailed information is presented in Table 1.4.

Table 1.4. Sample used in Tanner-Whitehouse method, third edition (2001).

	Project Name	Place	Boys	Girls	Study	Age covered	Years of the Study	Observations
1	UK60	UK (London, Harpenden, Oxford, Ayr, Kilmarnock, Ganges, Mayfield)	1385	1317	Samples from TW1 and TW2 (Tanner et al., 1962, 1975)	1 month - 21 years	1946-1972	/
2	B70	Belgium (Leuven)	21174	9698	Leuven growth study of belgian boys; Leuven growth study of flemish girls (Beunen et al., 1990)	12 - 20 years; 6 - 19 years	1969-1974; 1979-1980	/
3	S80	Spain (Bilbao)	462	444	Atlas developed from children of Hospital de Basurto, Bilbao (Hernandez at al. 1991)	1 - 17.2 years	1978-1987	5266 x-rays; middle and low-middle class.
4	J85	Japan (Tokyo)	701	691	(Ashizawa et al., 1996)	3 - 18 years	1986	/
5	I90	Italy (Genoa)	952	879	(Vignoli et al., 1999)	8 - 16.8 years; 8 - 15.9 years	1975-1985	Children assessed at Instituto di Puericultura e Medicina Neonatale of the University of Genoa.
6	Arg70	Argentina (La Plata City)	388	387	(Lejarraga et al., 1997)	4 - 12 years	1971	/
7	Tx90	USA (Woodlands, Texas)	225	225	Project Heartbeat (Tanner et al., 1997)	8 - 16 years	1993-1996	1090 films; sample: generally above-average income; confirmatory study of 190 films from longitudinal series of 23 boys from Virginia.
Total			25287	13641		1 month - 21 years	1946-1996	
			38928					

1.4.3. FELS method to assess skeletal age

The *Fels Research Institute* was founded in 1929 with a single complex research project called, the Fels Longitudinal Study (Roche, 1992). In 1929, Arthur Morgan was the president of Antioch College in Yellow Springs and claimed a longitudinal study from conception to adulthood. He approached Mr. Samuel Fels, a Philadelphia businessman and philanthropist, who gave the enthusiastic and financial support for the project. Several top specialists in the analysis of data relating to growth, maturation, and body composition, worked at Fels Research Institute: Richard N. Baumgartner, Pamela Byard, William C. Chumlea, Christine E. Cronk, Frank Falkner, Stanley M. Gar, Shumei Guo, Harry Israel III, Arthur B. Lewis, Debrabata Mukherjee, Earle L. Reynolds, Meinhard Robinow, Alex F. Roche, Roger M. Siervoguel, Lester W. Sontag.

The Hand-wrist method emerged 13 years after the RWT method for the Knee (Roche et al., 1988). Hand-wrist area have advantages: 1) little radiation required; 2) facility in positioning; 3) number of bones in a relatively small area. Several authors contributed to the development of FELS hand-wrist method, including statistical advice, testing and measurements, and analysis of data. The 677 children evolved in Fels method (355 boys and 322 girls) were from southwestern Ohio: 35% lived in cites of medium size, about 50% in small cities and the remaining 15% on farms. Only 15 children enrolled in the Fels longitudinal study were black, but the serial radiographs of these children were not included in the developmental of the method. A total of 13.823 x-rays (7308 for boys and 6515 for girls) were used to develop Fels method. Ages were distributed from one month to 22 years: 1,3,6,9 and 12 months; from 12 months until 18 years with six-months interval; from 18 to 22 years with two years interval (18, 20, 22).

Roche and colleagues combined the data into a single estimate of skeletal age, providing the standard error which was not available in all the previous skeletal age methods. The procedure was developed first using radiographs of boys and then applied to girls, using the rationale that the maturation in boys occur slower than in girls and if some of the indicators were sex-specific:

- Step 1: compilation of the possible hand-wrist indicators, using previous literature, with Greulich-Pyle atlas (1959) being the most important source. Pisiform was omitted as provide very little information. Second and forth digit were also omitted as they provide redundant information in relation to the first, third and fight digit.

- Step 2: Rewritten of the indicators with standardization and an anatomically correct terminology.

- Step 3: Numerous objective possible grades were tested for each indicator. Then, they were combined or excluded as testing proceeded. Several assessors tested possible indicators and there were frequent discussions leading to changes in definitions of indicators. Intra and inter-observer differences were assessed.

- Step 4: Indicators were not considered useful were identified. A serial radiographs of 10 boys (primary test group) were studied at the ages for which an indicator was considered useful. Then, the radiographs were graded twice by two assessors, working independently. Each indicator needed to satisfy five criteria to be considered useful: Discrimination (variability within CA groups); Universality (each grade of every indicator must occur during the skeletal maturation process of the subject. Additionally, indicators with more than two grades do not need to reach 100% at a determined age, because children differ in their rates of maturation. However, the least and most mature grade need to be universal at some ages. If they were not universal by 18 years, the age was extended to 22. Very few were not all universal at 22, but in 98% of the cases the most mature grade was present); Reliability (intra and inter-observer differences less than 8% for grade indicators and less than 0.5mm for metric indicators); Validity (prevalence of grades of valid maturity indicators must change systematically with age, until the most mature grade became universal); Completeness (the extent to which each indicator could be recorded, restricting the age ranges during which some indicators were recorded. More changes in lesser time are more informative indicators).

- Step 5: Some indicators were combined (fewer indicators, more grades), others (with many grades) were separated into two indicators with two or more grades.

- Step 6: The retained indicators from the previous steps were assessed using a secondary test group, from a radiographic series of 20 boys (1,3,6,9, and 12 months; 12 to 18 years with half-year intervals). The radiographs were graded twice by two assessors, independently, generally the same professionals that assess the primary group. Reliability and completeness were tested with the second group and validity, discrimination and universality tested later using the all sample. Findings from the secondary group led to new modifications in the indicators descriptions as well as the age ranges during which they seemed to be valuable. New descriptions were tested when changes were made. As a result, from the 144 indicators described in the literature, 33 were excluded (23%), remaining 111.

- Step 7: All suitable x-rays from boys were graded for each useful indicator, except for the radiographs from the secondary test group that were not graded again.

- Step 8: Remaining data were analyzed for universality, validity, and completeness. For validity, the prevalence of reversals was calculated for each indicator from the ages when the least mature grade was present, until the most mature grade was present, in 75% of the children. Consequently, the ages where the indicator was frequently absent or present, were excluded. Among the final indicators, highest prevalence of reversals was 14.8%, both for boys and girls. Reversals were more prevalent for metric indicators.

- Step 9: Grades prevalence by age were recorded graphically on probability paper, estimating the approximate age at which grade reached 50%. When this didn't occur, three possible steps were made: (i) age range extended; (ii) adjacent grades were combined and the descriptions of indicators modified; (iii) they were retained if their retention increased the separation between the ages at which the less and more mature grades reached 50% prevalence levels.

Felshw 1.0 is the software of method, that requires the grade of each indicator that should be assessed, considering CA and sex of the child. The outputs are the estimated skeletal age in years and the standard error of this estimate. The program does not require the insertion of all grades or metric indictors for a specific age, but if some indicators are missing, estimated skeletal age is less precise, with a larger standard error of estimate. If the estimated skeletal age differs from the child's CA for more than two years, indicators from CA corresponding to the estimated skeletal age should be assessed and inserted in Felshw software, to obtain a revised estimated skeletal age with the that additional data.

When the estimated skeletal age is outside of what is considered "normal range", the assessor should consider the standard error of the estimate, which indicates the level of accuracy of the assessment. If its large (>90th percentile), this could indicate possible grading mistakes or incomplete set of indicators. The solution is to reassess the radiograph by the same or another assessor and then compare the two results. The method was applied to 500 radiographs taken in the First Health and Nutrition Examination Survey that was conducted by the National Center for Health Statistics (USA). The means of the assigned SA were very similar to the CAs, which indicate that the method is appropriate for the US population of children from that time.

1.5. The algorithm to predict maturity-offset

Although skeletal age assessment is considered the best method to assess skeletal maturation, it requires specialized equipment and assessors. Advancement in technology reduces the exposure to radiation to 0.001 millisievert (mSv), which is equivalent to 3 hours of television (Malina, 2011).

Non-invasive methods like estimation of peak at age velocity (PHV) require longitudinal data. Mirwald, Baxter-jones, Bailey, & Beunen (2002) hypothesized that the ratio between leg length and sitting height could contribute to estimate a maturity status. In fact, years from peak height velocity was possible to be predicted from anthropometry. The sex-specific predictive models were based on three samples. The predictive equations were developed using Saskatchewan Pediatric Bone Mineral Accrual Study (BMAS) (Bailey, 1997; Bailey, Mckay, Mirwald, Crocker, & Faulkner, 1999), a mixed-longitudinal study to assess the factors associated with bone mineral accrual in growing children. The children from the used sample were assessed 4 years from PHV and 3 years after PHV (113 boys and 115 girls). To verify the predictive equations, one sample from Saskatchewan Growth and Development Study (SGDS) (Mirwald, 1978) was used and another from the Leuven Longitudinal Twin Study (LLTS) (Beunen, Thomis, & Maes, 2000). The SGDS sample were measured annually from 1964 to 1973. The LLTS measured 190 twins from 1985 to 1999, at semiannual intervals from 10 to 16 years old, as also at the age of 18.

For the Canadian studies (BMAS and SGDS) height, sitting heigh and body mass were assessed (average of two measurements or the median of three measurements, in case of a difference of more than 4mm for weight and sitting height and 0.4kg for weight in the two first measurements). Leg length to sitting height ratio was calculated. A higher ratio was observed in boys closer to the age from PHV (0 years), indicating a greater leg length. This ratio increased until the age of PHV, then decreased after PHV. The aim of the study was the development of a gender-specific multiple regression equation, using leg length to sitting height ratio from a single measurement.

The maturity offset (MO) was the dependent variable in multiple regression analysis. Independent variables were CA and anthropometric variables (height, sitting height, subischial leg length, and weight). The interaction between CA and each one of the anthropometric variables were also assessed. Five ratios were calculated: weight/height; body mass index (BMI); sitting height/height; leg length/height; leg length/sitting height. From these 15 independent variables, sex-specific multiple regression equations, were developed through hierarchical entry with consideration given both biological and statistical significance of potential entry variables to predict MO. The final equation to predict MO was as follows:

Maturity offset (years)= -9.236+(0.0002708 × (Leg Length × Sitting Height)) + (- 0.001663 × (CA × Leg Length)) + (0.007216 × (CA × Sitting Height)) + (0.02292 × (Weight by Height Ratio × 100))
($R = 0.94$, $R^2 = 0.891$, $SEE = 0.592$)

Because the equation depends on anthropometric procedures, the authors advised that careful is recommended in the anthropometric assessment, standardizing the measurement procedures, with special attention on the sitting height variable which as a direct relationship with a great number of independent variables.

1.5.1. Validation studies

Validation studies were attempted in two distinct samples. Malina & Kozieł (2014) attempted to validate the predictive equation in a sample of Polish boys followed from 8 to 18 years. The data was retrieved from the Wroclaw Growth Study (1961-1972) that provided a reasonable cross-section of the Wroclaw population. Sample included 426 polish urban boys, observed in April/May 1961. All were born in 1953 and had a mean of 8 years at the time of observation (6.50 to 8.49 years). Then they were measured every year at the same date (April/May). Stature, body mass, and sitting height were measured and inter-observer technical error of measurement were 0.29 and 0.35cm, for stature and sitting weight, respectively. Serial stature records were fitted using Preece–Baines Model 1 to derive APHV. CA, mass, stature, sitting height and estimated leg length at each serial observation, were used to predict MO according to the equation proposed by Mirwald et al., (2002).

Maturity status was calculated for each boy (late, average, early) using the sample mean and standard deviation for actual APHV, 14.06 ± 1.11 years. Using a ± 1.0-year band, boys were classified as average between 13.1 and 15.1 years. The CA and number of subjects in each maturation group was: early (n=36; 12.57 ± 0.41 years); average (n=117, 13.97 ± 0.52 years); late (=40, 15.69 ± 0.56 years). MO predicted APHV and differences between predicted and actual APHV at each CA from 8 to 18 years were calculated for the three groups.

Mean predicted APHV at 12 years of age was almost similar to actual APHV (-0.01 years) but prior to 12 years were earlier than actual APHV (-1.47 years at 8 years, reducing gradually the difference until 12 years), and after 12 years predicted ages increased positively (0.18 years at 13 years until 0.82 years at 18 years). Results indicated limitations of the prediction equation since it underestimates APHV in earlier ages and overestimated at older ages. Consequently, authors recommended that maturity offset should be generalized with caution and uniquely around year of APHV among average maturing individuals.

Another study was aimed to validate the original maturity offset equation in a cohort study of American girls and boys followed longitudinally from childhood to adolescence, from the Fels Longitudinal study (Malina, Choh, Czerwinski, & Chumlea, 2016). The sample was of European ancestry (white) from Southwestern Ohio. The cohort include 137 youth, from which 74 boys, born in the 60s, 70s, and 80s, with all complete records from 8 to 18 years. The selected birth decades fit on the birth decades from the three studies where the MO equations were developed: SGDS (1966-1973), LLTS (1985-1999), and PBMAS (1991-1997). Briefly, ages at PHV did not differ by birth decades. The height records from early childhood to young adulthood for each children were modeled with the triple logistic Bock-Thissen-du Toit (BTT) model. Mean age for the age at PHV for the 74 boys was 13.70 ± 1.03 years (11.28 – 16.10 years). The sample was classified in different maturity groups, based on actual age at PHV ± 1 SD: early (PHV <1 SD), average (PHV ≥ -1SD and ≤ +1SD), late maturing (PHV > 1SD). Only 13 subjects were classified as early (12.15 ± 0.43 years); 50 as average (13.75 ± 0.54 years) and 11 as late (15.30 ± 0.41 years). MO approximates zero at 14 years in boys, and then becomes positive. At 11 and 12 years, the mean differences between predicted APVH and actual APHV were not significant, but in all other ages the mean differences were significant ($p < 0.03$). Predicted age at PHV was substantially lower than actual APHV at 8 years, but these differences decline with advancing CA until 12 years. Then they increase until 18 years, except at the range 13-15 years where they are stable and not statistically different. Mean differences between predicted APHV and actual APHV varied between maturity groups in all age groups, from 8 to 18 years. Results highlighted that predicted MO and predicted APHV were dependent upon CA at prediction. Standard deviations of predicted APHV were reduced compared to actual values. By comparing maturity groups, it was evident that predictions were affected by individual differences in actual APHV.

In summary, the available studies showed that predictions of age at PHV may be useful near the time of actual PHV in average maturing boys within a narrow CA range, but not in the other contrasting maturity groups, confirming the results obtained with the sample of polish boys. These results showed that the use of the original MO equation for the purpose of talent development or to adapt sport-specific programs may result in erroneous decisions due to its limitations. The studies regarding validation studies are less abundant in the youth sport literature.

1.6. Prediction of adult size

Methods for predicting adult stature have been widely used since 50s, but the first methods required skeletal age as a predictive variable (Bayley & Pinneau, 1952; Khamis & Guo, 1993; Roche, Wainer, & Thissen, 1975; Tanner et al., 1975, 1983). The statistical methods used in the development of the RWT method and Khamis-Guo (the modified RWT with Fels SA instead of GP), were applied to other method, Khamis-Roche (KR) (Khamis & Roche, 1994). This method was developed in the absence of skeletal age to simplify predictions of adult stature since SA requires radiation (although minimal) and experienced observers to read the films for assessing SA. The sample was composed by participants from the Fels longitudinal study and included 223 males and 210 females, with stature measured at 18 years. The KR method include three predictor variables: current stature, current weight, and midparent stature. All these variables were obtained in the Fels longitudinal study, with stature and weight being obtained with 6-month interval periods, from 3 to 18 years.

The derived regression equation for the method used stature, weight and mid-parent stature. The equation adopts coefficients for each CA group (from 4 to 17.5 years, with 6 months-interval). Authors calculated the median absolute deviation (MAD) for the differences between predicted and true adult stature as an indicator of accuracy. For example, at a certain chronological age, predicted adult stature was 178cm and the MAD for that age and gender is 3cm meaning that the majority of the individuals plot within the range 175-181cm. Although the average MAD for KR exceed the MAD for Khamis-Guo (KG) (Khamis & Guo, 1993), the differences were negligible. However, careful is needed when using the KR. The original authors recommended the use of Khamis-Guo in males of 14 years, instead of Khamis-Roche (p. 507):

> *"Although the differences in errors between the KR and the RWT method are small, except for the 90% error bounds at about 14 years in males…the latter is recommended when it is possible to obtain an accurate skeletal age…".*

In summary, adult stature can be predicted using three variables: stature, weight, and midparent stature. The intercept and regression coefficients for the three anthropometric variables are described in Tables 1 and 2 of Khamis-Roche paper (Khamis & Roche, 1994).

1.7. Youth Sports

The involvement of children in competitive sports is a social trend in current societies. In recent years, young athletes are being recruited for elite training centers progressively at younger ages in order to increment the volume of training sessions and the number of sessions per week. Some are totalizing 30 to 40h every week 12 to 12 months/ year (Armstrong, 2019). The early specialization raises several concerns about the negative effects of intensive training on developmental trajectories including increased risk of injuries, burnout, among others. Only a trivial percentage of youth athletes identified as talent and selected for training programs are reaching the adult elite level (Armstrong, 2019).

In the context of youth sports, Cacciari et al. (1990) assessed the effects of sport participation on growth, considering indicators of biological maturation, endocrine levels, and anthropometric measurements, in a sample of 399 males (175 athletes, 224 controls), aged 10-16 years. For the SA assessment the authors considered the "Carpal X-ray" suggesting an arbitrary adaptation of the original GP protocol. In fact, the authors of Fels method alerted that although it has been stated that the atlas method was easy, quick, and simple, the method should not be viewed as so, when applied to individual bones like it should be done (Roche et al., 1988). In other words, many authors seem adopt an inappropriate application of GP protocol considering the whole hand-wrist bones.

The literature suggested that among the age group 10-12 years late and maturing participants are equally represented. From that age, the number of early maturing boys are more likely to increase their prevalence while late maturing boys decline (Malina et al., 2000; Malina, Rogol, Cumming, Coelho-e-Silva, & Figueiredo, 2015). Although being advanced in biological maturation is associated to larger body size compared to late and average maturing peers, it seems that maturation does not influence all functional capacities and, in general, the sport-specific skills (Coelho-e-Silva, Figueiredo, Carvalho, & Malina, 2008; Figueiredo, Gonçalves, Coelho-e-Silva, & Malina, 2009a). With increments in CA, late maturing participants tend to have fewer opportunities for playing and the hypothesis of selective exclusion from sport is obvious. The preceding was already evidenced in soccer (Malina et al., 2000), basketball (Coelho-e-Silva et al., 2008), tennis (Myburgh, Cumming, Coelho-e-Silva, Cooke, & Malina, 2016), and table tennis.

Despite the concept of "Biobanding" is relatively new, in 1908 Rotch already questioned CA grouping in youth sports, considering what he called "anatomic age" (Rotch, 1908). Briefly, biobanding is a decision of grouping adolescent players within a given CA range into "bands" for specific training and/or competitions. The few attempts tend to adopt percent of predicted adult height as an indicator of biological maturation (Bradley et al., 2019; Cumming et al., 2018; Cumming, Lloyd, Oliver, Eisenmann, & Malina, 2017; MacMaster et al., 2021; Malina et al., 2019; Rogol, Cumming, & Malina, 2018).

Meantime, the literature demonstrated a modest agreement between maturity classifications based on SA-CA using Fels method and non-invasive methods (Khamis-Roche and Maturity offset) among male soccer players (Malina, Coelho-e-Silva, Figueiredo, Carling, & Beunen, 2012). The previous was also suggested in youth tennis (Myburgh, Cumming, & Malina, 2019). As a result, differences in the classification of players by biological maturation may emerged, when different methods are used.

Additionally, different methods for the determination of SA would be another source of variation. Malina, Chamorro, Serratosa, & Morate (2007) compared SAs of elite Spanish soccer players aged 12.5-16.1 years using TW3 RUS and Fels. The authors found a trend for higher SA values when players were assessed by Fels protocol. A participant can reach a maximum of 18.0 years when assessed by Fels, and no more than 16.5 years when the film is assessed using TW3.

More recently, Malina et al. (2018) examined the difference between the assessment of two different versions of TW method: TW2 and TW3, based on several samples of soccer players, aged 10.93-17.94 years, from eight countries, totalizing 1831 athletes. The results consistently suggested lower SA by the TW3 compared to TW2. The methods used by researchers and coaches seem to have implications for the distribution of players by maturity status which is implicit to long-term athletic development model (Jason Moran et al., 2020; Murtagh et al., 2018). In Portugal, the medical exam for the young athletes who require to compete at an older competitive level includes the assessment of SA using GP method. A related question refers to data quality due to the observer is not being systematically addressed in the literature using concurrent methods to assess SA.

1.8. Objectives

The present thesis is divided into six chapters. Chapter II describes the methods used in the five manuscripts, that is study design, samples, anthropometry, skeletal age, somatic indicators and analysis. Chapters III-VII correspond to five independent manuscripts. The first manuscript aimed to report intra and inter-observer agreement using the Tanner-Whitehouse method: 20-bone, RUS, CARPAL. Chapter IV assumes a similar approach to determine data quality using Fels protocol. Chapter V reported the agreement between concurrent methods to obtain SA using a multi-sport sample of 1778 male youth athletes aged 11-17 years old. Stature and body mass of the samples were plotted against the CDC reference data for CA and, in parallel, considering SA assessed by the GP, TW2 RUS, and FELS methods. Chapter VI compare two alternative methods for the estimation of the adolescent growth spurt: Superimposition by Translation and Rotation (SITAR) and Functional Principal Component Analysis (FPCA) in a longitudinal sample of 58 male adolescent soccer players followed by five seasons. It also assessed the relationships among maturity status assessed by invasive (skeletal age by FELS method) and non-invasive (age at PHV) maturity indicators. The final independent study (chapter VII) is devoted only to somatic indicators and used the same longitudinal sample of the previous study. Two maturity offset equations (Mirwald et al., 2002; Moore et al., 2015) were tested on the same longitudinal sample of the previous study whom ages at PHV were estimated using SITAR model. Chapter VIII summarizes the general discussion based on the main findings of the previous chapters.

1.9. References

Acheson, R. M. (1954). A method of assessing skeletal maturity from radiographs; a report from the Oxford child health survey. Journal of Anatomy, 88(4), 498–508.

Acheson, R. M. (1957). The Oxford method of assessing skeletal maturity. Clinical Orthopaedics, 10, 19–39.

Armstrong, N. (2019). Development of the youth athlete. London (UK): Routledge.

Bailey, D. A. (1997). The Saskatchewan Pediatric Bone Mineral Accrual Study: Bone Mineral Acquisition During the Growing Years. International Journal of Sports

Medicine, 18(S 3), S191–S194.

Bailey, D. A., Mckay, H. A., Mirwald, R. L., Crocker, P. R. E., & Faulkner, R. A. (1999). A Six-Year Longitudinal Study of the Relationship of Physical Activity to Bone Mineral Accrual in Growing Children: The University of Saskatchewan Bone Mineral Accrual Study. Journal of Bone and Mineral Research, 14(10), 1672–1679.

Bayley, N., & Pinneau, S. R. (1952). Tables for predicting adult height from skeletal age: Revised for use with the greulich-pyle hand standards. The Journal of Pediatrics, 40(4), 423–441.

Beunen, G., Thomis, M., & Maes, H. H. (2000). Genetic variance of adolescent growth in stature. Annals of Human Biology, 27(2), 173–186.

Bradley, B., Johnson, D., Hill, M., McGee, D., Kana-ah, A., Sharpin, C., … Malina, R. M. (2019). Bio-banding in academy football: player's perceptions of a maturity matched tournament. Annals of Human Biology, 46(5), 400–408.

Cacciari, E., Mazzanti, L., Tassinari, D., Bergamaschi, R., Magnani, C., Zappulla, F., … Tani, G. (1990). Effects of sport (football) on growth auxological, anthropometric and hormonal aspects. European Journal of Applied Physiology and Occupational Physiology, 61(1–2), 149–158.

Campbell-Swinton, A. A. (1896). The new kind of radiation. British Medical Journal.

Coelho-e-Silva, M. J., Figueiredo, A. J., Carvalho, H. M., & Malina, R. M. (2008). Functional capacities and sport-specific skills of 14- to 15-year-old male basketball players: Size and maturity effects. European Journal of Sport Science, 8(5), 277–285.

Cumming, S. P., Brown, D. J., Mitchell, S., Bunce, J., Hunt, D., Hedges, C., … Malina, R. M. (2018). Premier League academy soccer players' experiences of competing in a tournament bio-banded for biological maturation. Journal of Sports Sciences, 36(7), 757–765.

Cumming, S. P., Lloyd, R. S., Oliver, J. L., Eisenmann, J. C., & Malina, R. M. (2017). Bio-banding in sport: Applications to competition, talent identification, and strength and conditioning of youth athletes. Strength and Conditioning Journal, 39(2), 34–47.

Eisenberg, R. L. (1995). Radiology: An Illustrated History. Maryland height (Missouri, USA): Mosby Inc.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009). Youth soccer players, 11-14 years: maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Greulich, W., & Pyle, S. (1959). Radiographic atlas of skeletal development of the hand and wrist (2nd ed.). Stanford, CA: Stanford University Press.

Hellman, M. (1928). Ossification of epiphysial cartilages in the hand. American Journal of Physical Anthropology, 11(2), 223–257.

Khamis, H. J., & Guo, S. (1993). Improvement in the Roche-Wainer-Thissen stature prediction model: A comparative study. American Journal of Human Biology, 5(6), 669–679.

Khamis, H. J., & Roche, A. F. (1994). Predicting adult stature without using skeletal age: the Khamis-Roche method. Pediatrics, 94(4 Pt 1), 504–507.

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., … Johnson, C. L. (2002). 2000 CDC Growth Charts for the United States: Methods and Development. Vital and Health Statistics, Series 11,(246).

Leclerc, G. L. (1778). Histoire Naturelle, générale et particulière, servant de suite à l'histoire naturelle de l'homme.

MacMaster, C., Portas, M., Parkin, G., Cumming, S., Wilcox, C., & Towlson, C. (2021). The effect of bio-banding on the anthropometric, physical fitness and functional movement characteristics of academy soccer players. PLoS ONE, 16(11 November).

Malina, R. M. (2011). Skeletal Age and Age Verification in Youth Sport. Sports Medicine, 41(11), 925–947.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Chamorro, M., Serratosa, L., & Morate, F. (2007). TW3 and Fels skeletal ages in elite youth soccer players. Annals of Human Biology, 34(2), 265–272.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of maturity offset in the fels longitudinal study. Pediatric Exercise Science, 28(3), 439–455. https://doi.org/10.1123/pes.2015-0090

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Carling, C., & Beunen, G. P. (2012). Interrelationships among invasive and non- invasive indicators of biological maturation in adolescent male soccer players. Journal of Sports Science, 30(15), 1705–1717.

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Philippaerts, R. M., Hirose, N., Peña Reyes, M. E., … Buranarugsa, R. (2018). Tanner-Whitehouse Skeletal Ages in Male Youth Soccer Players: TW2 or TW3? Sports Medicine, 48(4), 991–1008.

Malina, R. M., Cumming, S. P., Rogol, A. D., Coelho-e-Silva, M. J., Figueiredo, A. J., Konarski, J. M., & Kozieł, S. M. (2019). Bio-Banding in Youth Sports: Background, Concept, and Application. Sports Medicine, 49(11), 1671–1685.

Malina, R. M., & Kozieł, S. M. (2014). Validation of maturity offset in a longitudinal sample of Polish boys. Journal of Sports Sciences, 32(5), 424–437.

Malina, R. M., Peña Reyes, M., Eisenmann, J., Horta, L., Rodrigues, J., & Miller, R. (2000). Height, mass and skeletal maturity of elite Portuguese soccer players aged 11–16 years. Journal of Sports Sciences, 18, 685–693.

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho-e-Silva, M. J., & Figueiredo, A. J. (2015). Biological maturation of youth athletes: Assessment and implications. British Journal of Sports Medicine, 49(13), 852–859.

Mirwald, R. L. (1978). The Saskatchewan Growth and Development Study. In M. Ostyn, G. Beunen, & J. Simons (Eds.), Kinanthropomety II (pp. 289–305). Baltimore, Maryland, USA: University Park Press.

Mirwald, R. L., Baxter-jones, A., Bailey, D., & Beunen, G. (2002). An assessment of maturity from anthropometric measurements. Medicine & Science in Sports & Exercise, 34(4), 689–694.

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D. G., Cameron, N., & Brasher, P. M. A. (2015). Enhancing a somatic maturity prediction model. Medicine and Science in Sports and Exercise, 47(8), 1755–1764.

Moran, J., Paxton, K., Jones, B., Granacher, U., Sandercock, G. R., Hope, E., & Ramirez-Campillo, R. (2020). Variable long-term developmental trajectories of short sprint speed and jumping height in English Premier League academy soccer players: An applied case study. Journal of Sports Sciences, 38(22), 2525–2531.

Murtagh, C. F., Brownlee, T. E., O'Boyle, A., Morgans, R., Drust, B., & Erskine, R. M. (2018). Importance of speed and power in elite youth soccer depends on maturation status. Journal of Strength and Conditioning Research, 32(2), 297–303.

Myburgh, G. K., Cumming, S. P., Coelho-e-Silva, M. J., Cooke, K., & Malina, R. M. (2016). Growth and maturity status of elite British junior tennis players. Journal of Sports Sciences, 34(20), 1957–1964.

Myburgh, G. K., Cumming, S. P., & Malina, R. M. (2019). Cross-Sectional Analysis Investigating the Concordance of Maturity Status Classifications in Elite Caucasian Youth Tennis Players. Sports Medicine - Open, 5(1), 27.

Pryor, J. W. (1925). Time of ossification of the bones of the hand of the male and female and union of epiphyses with the diaphyses. American Journal of Physical Anthropology, 8(4), 401–410.

Roche, A. F. (1992). Growth Maturation and Body Composition: The Fels Longitudinal

Study 1929-1991. New York (NY): Cambridge University Press.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Roche, A. F., Wainer, H., & Thissen, D. (1975). The RWT method for the prediction of adult stature. Pediatrics, 56(6), 1027–1033.

Rogol, A. D., Cumming, S. P., & Malina, R. M. (2018). Biobanding: A new paradigm for youth sports and training. Pediatrics, 142(5).

Rotch, T. M. (1908). Chronologic and anatomic age in early life. Journal of the American Medical Association, LI(15), 1197.

Rowland, S. D. (1896). Report on the Application of the New Photography to Medicine and Surgery: [Report to the British Medical Journal.]. BMJ, 2(1875), 1676–1677.

Tanner, J. M. (1978). Foetus into Man: Physical Growth from Conception to Maturity. London: Open Books Publishing.

Tanner, J. M. (2010). A History of the Study of Human Growth (1st ed.). Cambridge (UK): Cambridge University Press.

Tanner, J. M., Healy, M., Goldstein, H., & Cameron, N. (2001). Assessment of Skeletal Maturity and Prediction of Adult Height (TW3) Method (3rd ed.). Saunders Ltd.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1975). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London: Academic Press.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1983). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London (UK): Academic Press.

Tanner, J. M., Whitehouse, R. J., & Healy, M. J. R. (1962). A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2,600 healthy British children. Part II: The Scoring System. Paris: International Children's Centre.

Todd, T. (1937). Atlas of skeletal maturation (1st ed.). St. Louis, MO: V. Mosby Company.

Chapter II

Methods

2. Methods

2.1. Procedures and Sample

Portuguese Foundation for Science and Technology (SFRH/BD/138608/2018) supported the five studies of the present thesis and part of the data collection was achieved through a signed protocol between University of Coimbra and the Portuguese Institute of Sports and Youth (IPDJ/FCDEF.UC/2017-01). The investigation was performed in agreement with the ethical procedures from the *World Medical Association Declaration of Hensinki* (World Medical Association, 2013) as well as the last update of the *Ethical Standards in Sport and Exercise Science Research* (Harriss, Macsween, & Atkinson, 2019). Depending on the study, different samples were considered. Table 2.1 summarizes the characteristics of each study.

Table 2.1. Characteristics of the five studies.

number of study	type of data	sample	sports	age range (years)
1	Cross-sectional	142	soccer	11-15.3
2	Cross-sectional	97	tennis	8.7-16.8
3	Cross-sectional	1778	multisport	11.0-16.9
4	Longitudinal	58	soccer	11.9-15.9
5	Longitudinal	58	soccer	11.9-16.8

2.2. Anthropometry

Anthropometric variables were assessed according to the procedures proposed by (Lohman, Roche, & Martorell, 1988). Stature and sitting height were measured to the nearest 0.1cm, using an Harpenden portable stadiometer (model 98.603, Holtain Ltd, Crosswell, Crymych, UK) and a sitting-height table stadiometer (model 98.603, Holtain Ltd, Crosswell, Crymych, UK), respectively. Body mass was measured to the nearest 0.1kg, using a digital scale (SECA 770, Hanover, MD, USA). For Study 3, stature and body mass were plotted relative to U.S. reference data (Kuczmarski et al., 2002), and the process was repeated considering skeletal age.

2.3. Biological maturation

2.3.1. Skeletal maturation

A left hand-wrist radiograph was considered to assess skeletal maturation. Three methods were used: Greulich-Pyle (GP) (Greulich & Pyle, 1959) also known as the atlas or inspectional method; Tanner-Whitehouse (TW), more specifically the first revised version (Tanner et al., 1975), also recognized as the bone-specific approach method; and the FELS method, which has the similarity with TW of being a scoring method (Roche et al., 1988). They all consist in matching a youth hand-film with a set of criteria that can be pictorial, verbal, or both, although the scoring methods had the pictorial standards only as a complement for the verbal assessment (Malina et al., 2004). Final common aim is to assign a SA, in which a scale that will derive it vary according to the criteria and procedures considered in which one of the methods. Relatively to the sample characteristics, GP and FELS derived from children from Ohio, USA, although they were from European ancestry. TW derived from children from all over UK.

GP method (Greulich & Pyle, 1959) is based on the previous work from Todd (1937) and consists in matching the film of the child for assessment with the standard plate of the atlas, of the same sex and nearest CA. Then, a new comparison should be made with the subjacent standards. The closest standard plate should be selected, and the process should be repeated for the 30 bones of the hand-wrist, including radius and ulna, eight carpals including pisiform, five metacarpals, five proximal and five distal phalanges, four middle phalanges, and adductor sesamoid. Finally, the mean of the 30 bone skeletal ages corresponds to the global skeletal age of the participant. GP method was considered in Study 3.

TW2 method (Tanner et al., 1975) consider a maturity scale from letters A (no visible center of ossification) to I (complete fusion or adult form) with a punctuation for each letter, depending on the bone and stage selected. Each stage may have one to three written criteria (marked as (i), (ii), (iii) that should be carefully followed, with illustrations serving only as a complement for the assessment. If a stage has only one criteria, then that criteria must be met; if a stage has two criteria, only one criteria is sufficient to consider the stage; if a stage has three criteria, then at least two criteria should be met. Nevertheless, the criteria i) of the preceding stage must be always present. This first revised version of TW method has three possible scoring systems and the three were considered: 20-bone uses 20 bones of the hand-wrist, the 30 bones considered in the GP method without pisiform, adductor sesamoid and the bones of the 2nd and 4th digit as the authors considered they give redundant information's, since 1st, 3rd and 5th digit are already considered; CARPAL system consider only seven carpals of the hand-wrist: capitate, hamate, triquetral, lunate, scaphoid, trapezium, and trapezoid; RUS system consider only the 13 long bones, this is, the 20 bones without the seven carpals. All of three systems. All of three systems consider a total weight score of 1000 points, corresponding to the mature state, defined by Tanner et al. (1975) as ADULT, as a SA is not assigned. Thus 999 point correspond to 17.9 years in 20-bone system, 14.9 years in CARPAL system, and 18.1 years in RUS system. The three systems were considered in Study 1 and the RUS system in Study 3.

FELS method (Roche et al., 1988) is the more detailed and time consuming of the three methods. It considers 22 bones, the bones which are considered in TW method plus pisiform and adductor sesamoid. Each bone has one to eight maturity indicators, totalizing 98 indicators. From this total 58 are binary indicators, 27 ordinal indicators and 13 metric indicators. Nevertheless, only 20 to 66 indicators are really assessed, depending on the CA of the participant. Each one of the indicators may have two to five grades. Like TW method, written criteria should be the primary guideline for the assessment, with the pictorial standards serve only as a complementary resource. The assessed maturity indicators, according to the CA and sex of the participant, are entered into a computer software (Felshw 1.0), that provide the SA of the subject to a maximum of 18.0 years, in which the subject should be considered mature. Additionally, FELS method can provide a standard error of the estimated SA. FELS method was considered for studies 2, 3 and 4.

SA minus CA was used for maturity status classification (Malina et al., 2004): if the athlete had a SA younger than CA by more than 1.0 year, the athlete was classified as late (delayed) mature; if the athlete had an SA older than CA by more than 1.0 year, the athlete was classified as early (advanced) mature; if SA was within ±1.0 year of CA, then the athlete was classified as on time (average) mature. The band of 1.0 year is often used in youth sports for maturity status classification as approximates standard deviations for SAs in age groups 10-17 years, provides a wider range of players classified as on time, and allows for errors associated with the assessments (Malina et al., 2011).

2.3.2. Somatic maturation

Parameters of the growth curve were assessed with longitudinal data for studies 4 and 5. The 58 soccer players were assessed during five competitive seasons and the longitudinal height records were fitted with Superimposition by Translation and Rotation (SITAR) (studies 4 and 5) and Functional Principal Component Analysis (FPCA) (study 4).

SITAR model (Cole, 2020; Cole, Donaldson, & Ben-Shlomo, 2010) fits the raw data of all players with a basis spline (B-spline), a generalization of the Bézier curve. It lay over all curves, average them and back-projects the average curves into the original data as a growth model trough uniform transformations - translation and rotation.

A combination of the general Functional Data Analysis (FDA) with FCPA results in the FPCA growth model (Králík et al., 2021). Boys from the Brno Growth Study were the training set. The raw data of subjects were fit with B-spline curves that were modeled with FPCA. Six principal components for phase of the growth curves and six principal components for amplitude of the growth curves were used as a generative model to fit the newly analyzed data based on the Levenberg–Marquardt optimization algorithm. Detailed information about the model is available in the R package growthfd (Kíma & Králík, 2022; Králík et al., 2021).

Predicted age at peak heigh velocity (APHV), was obtained with maturity offset (MO) equations: the original MO equation (Mirwald et al., 2022) and one of the two modified MO equations (the one that doesn't require sitting height) (Moore et al., 2015). MO is an indicator of maturity timing and is noninvasive, as it requires only CA and anthropometric variables to predict the APHV, trough the MO (years from PHV). The original equation (equation 1) and the modified equation (equation 2) are the following:

Equation 1: Maturity offset (years)= -9.236+(0.0002708 × (Leg Length × Sitting Height))+(-0.001663 × (CA × Leg Length))+(0.007216 × (CA × Sitting Height))+(0.02292 × (Weight by Height Ratio × 100))

Equation 2: Maturity offset (years)= -7.999994+(0.0036124 × (CA × Height))

2.4. Analyses

Data analysis was performed according to the aim of each study (Table 2.2). The procedures were done mostly using SPSS (SPSS Inc., IBM Company, Armonk, NY, USA). R-software (R Core Team, 2019) was also used in studies 4 and 5 (R package version 0.0.2; 2018; R package version 1.1.2; 2020; package version 3.1–149; metafor package). Figures were made using GraphPad Prism (version 5 for Windows, GraphPad Software, San Diego, CA, USA). Significancy values were established at 0.05.

Table 2.2. Analysis of the five studies.

Analyses	Study				
	1	2	3	4	5
Agreements/ Disagreements	•	•			
Paired *t*-tests	•				
Techninal errors of measurement	•	•			
Coeficients of variation	•	•			
Intra-class correlation coeficient	•	•			
Cohen-*d*	•	•			
Magnitude effects for *d*-values	•	•			
Bivariate correlation coefficient	•	•			
Bland-Altman analysis	•	•			
Cross-tabulation analysis		•	•	•	
Cohen kappa coefficient			•	•	
SITAR model				•	•
FPCA model				•	
Chi-squared test				•	
Deming regression					•

2.5. References

Cole, T. J. (2020). SITAR: Super Imposition by Translation and Rotation Growth Curve Analysis. Retrieved from https://cran.r-project.org/package=sitar

Cole, T. J., Donaldson, M. D. C., & Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. International Journal of Epidemiology, 39(6), 1558–1566.

Greulich, W., & Pyle, S. (1959). Radiographic atlas of skeletal development of the hand and wrist (2nd ed.). Stanford, CA: Stanford University Press.

Harriss, D. J., Macsween, A., & Atkinson, G. (2019). Ethical Standards in Sport and Exercise Science Research: 2020 Update. International Journal of Sports Medicine, 40(13), 813–817.

Kíma, O., & Králík, M. (2022). R package growthfd for fitting FPCA-based growth curve models. Retrieved from https://ondrej-klima.github.io/growthfd/index.html

Králík, M., Klíma, O., Čuta, M., Malina, R. M., Kozieł, S., Polcerová, L., … Zemčík, P. (2021). Estimating Growth in Height from Limited Longitudinal Growth Data Using Full-Curves Training Dataset: A Comparison of Two Procedures of Curve Optimization-Functional Principal Component Analysis and SITAR. Children (Basel, Switzerland), 8(10).

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., … Johnson, C. L. (2002). 2000 CDC Growth Charts for the United States: Methods and Development. Vital and Health Statistics, Series 11,(246).

Lohman, T. G., Roche, A. F., & Martorell, R. (1988). Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1975). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London: Academic Press.

Todd, T. (1937). Atlas of skeletal maturation (1st ed.). St. Louis, MO: V. Mosby Company.

World Medical Association. World Medical Association Declaration of Helsinki, 310 JAMA (2013).

Chapter III

Study 1

Skeletal age assessed by TW2 using 20-bone, carpal and RUS score systems: Intra-observer and inter-observer agreement among male pubertal soccer players.

3. Skeletal age assessed by TW2 using 20-bone, carpal and RUS score systems: Intra-observer and inter-observer agreement among male pubertal soccer players.

3.1. Abstract

The purpose of this study was to determine intra- and inter-observer agreement for the three skeletal ages derived from the TW2 method among male pubertal soccer players. The sample included 142 participants aged 11.0-15.3 years. Films of the left hand-wrist were evaluated twice by each of two observers. Twenty bones were rated and three scoring systems used to determine SA adopting the TW2 version: 20-bone, CARPAL and RUS. Overall agreement rates were 95.1% and 93.8% for, respectively, Observer A and Observer B. Although, agreement rates between observers differed for 13 bones (5 carpals, metacarpal-I, metacarpal-III, metacarpal-V, proximal phalanges-I, III and V, distal phalanx-III), intra-class correlations were as follows: 0.990 (20-bone), 0.969 (CARPAL), and 0.988 (RUS). For the three SA protocols, BIAS was negligible: 0.02 years (20-bone), 0.04 years (CARPAL), and 0.03 years (RUS). Observer-associated error was not significant for 20-bone SA (TEM=0.25 years, %CV=1.86) neither RUS SA (TEM=0.31 years, %CV=2.22). Although the mean difference for CARPAL SAs between observers (observer A: 12.48±1.18 years; observer B: 12.29±1.24 years; t=4.662, p<0.01), the inter-observer disagreement had little impact (TEM: 0.34 years: %CV: 2.78). The concordance between bone-specific developmental stages seemed was somewhat more problematic for the carpals than for the long bones. Finally, when error due to the observer is not greater than one stage and the replicated assignments had equal probability for being lower or higher compared to initial assignments, the effect on SAs was trivial or small.

Keywords: Youth sports, Adolescence, Youth athletes, Tanner-Whitehouse, Bone age.

3.2. Introduction

Growth, maturation and development are central processes in long-term participation of children and youth in competitive sports. The preceding processes are not synonymous or interchangeable. Growth involves quantitative changes in body size, proportions, shape and composition (Malina et al., 2004). Development implies changes in behavioral domains: cognitive, emotional, social and motor. Finally, maturation marks progress towards the adult state which varies with the biological system: dental, sexual, somatic and skeletal. In the context of youth sports, the concepts are implicit to the long-term athletic development model (Jason Moran et al., 2020). Two indicators of biological maturation are commonly used in studies of adolescent athletes. In boys, sexual maturation includes genital and pubic hair development that are limited to pubertal years. Skeletal age (SA) requires standard radiographs of the hand and wrist and is generally considered the preferred indicator as it can be assessed through the first two decades of life. Several protocols are available to determine SA. They are similar in principle and require a radiograph of the hand-wrist. Advancements in technology have reduced exposure to radiation to about 0.001 millisievert (mSv), which is equivalent to three hours of television (Malina, 2011). The SA of a child or adolescent represents the chronological age (CA) at which a specific level of maturity of the hand-wrist bones is attained relative to the reference population upon which the method was created. The Greulich-Pyle (Greulich & Pyle, 1959) and Fels (Roche et al., 1988) methods were developed on American children and adolescents, while the Tanner-Whitehouse (Tanner et al., 1962) is originally based on British youth and was subsequently modified (Tanner et al., 2001, 1975).

Observations on Portuguese soccer players aged 11-12 years suggest that youth delayed and advanced in terms of skeletal maturity status based on Fels SAs were equally represented, whereas among players 13-14 years late maturing players were underrepresented whereas those classified as average and early were over-represented (Malina et al., 2000). The previous confirms soccer as highly selective and the literature suggests a gradient in body size among Portuguese male soccer players aged 13.0-14.1 years of age (Coelho-e-Silva et al., 2010): those selected for a regional team were taller, heavier, advanced in AS given by the Fels method and had more playing experience than teammates who were not selected. Meantime, among French youth players (Carling, Le Gall, & Malina, 2012), those who signed a professional contract and played at least one game as a professional were significantly taller and heavier and had a higher estimated aerobic power at baseline (~13 years) compared to peers who did not sign a professional contract, although the groups did not differ in skeletal maturity assessed by the Greulich Pyle protocol. Finally, a survey of the skeletal maturity status of Serbian soccer players aged 14 years using the Tanner-Whitehouse method, more precisely the radius-ulna-short bones (RUS), noted that late maturing players were more likely to attain a professional career compared to early maturing peers (Ostojic et al., 2014).

The above cited youth soccer literature produced different results which highlight the need to discuss the generalization of studies based on concurrent methods of SA assessment. In fact, results may reflect variation in the methods to determine SA and/or specific characteristics of youth soccer in Portugal, France and Serbia. The literature already examined the agreement of concurrent protocols for SA determination, particularly during pubertal years overlapping to selection, specialization into playing positions (Coelho-e-Silva et al., 2010) and vulnerability to sport injuries (van der Sluis et al., 2014). For example, the SAs of 40 male Spanish soccer players aged 12.5-16.1 years were assessed with the TW3 and Fels methods (Malina, Chamorro, et al., 2007). A consistent trend for lower SAs with the TW3 RUS compared to Fels was evident. More recently (Malina et al., 2018), two versions of TW RUS method (TW2 versus TW3) were compared in a large international sample of male soccer players aged 10.9-17.9 years. Across the CA range of the sample, TW3 RUS SAs were consistently lower than TW2 RUS SAs. The preceding studies have implications for the classification of youth players by maturity status. Advances in digital imaging technologies combined to research dealing with machine learning have led to the emergence of informatic applications that automatically estimate SA from digitalized radiographs (Thodberg, Kreiborg, Juul, & Pedersen, 2009). Meantime, sonography has been proposed as an alternative non-invasive method for determining SA (Mentzel et al., 2005). The preceding includes the operator, while obtaining the image, as an additional source of error.

Taking into account the preceding, error is a central issue in determination of SA. In the study of Spanish players (Malina, Chamorro, et al., 2007), intra-observer differences for Fels and TW3 SAs fluctuated 0.1 to 0.4 years, while technical errors of measurements were small: Fels SA was 0.04 year, TW3 SA was 0.06 year. The objective of the present study was to evaluate intra-observer and inter-observer agreements for the SAs derived by TW2 method among male adolescent soccer players: 20-bone protocol (TW2 20 bone SA), carpals (Carpal SA), and 13 long bones (RUS SA). It is hypothesized that even trained observers produce errors in the assessment of TW2 SAs.

3.3. Methods

Procedures

The present study is derived from the PRONTALSPORT Project (Growth, maturation and athletic performance in pubertal athletes). The project followed the ethical standards established for sports sciences (Harriss et al., 2019) and was approved by the Ethics Committee for Sports Sciences by the University of Coimbra (CE/FCDEF-UC/00122014). Participants were recruited from clubs of Portuguese Midlands having a written agreement with University of Coimbra. Parents of the players signed an informed consent, while the players provided assent. They were informed that their participation was voluntary and they could withdraw at any time. All data were collected within a 2-week period in the Coimbra University Stadium for anthropometry and posterior-anterior radiographs of the left hand-wrist were obtained on the same day at a certified clinic.

Sample

The sample included 142 male adolescent soccer players aged 11.0-15.3 years. All participants were registered in the Portuguese Football Federation as *infantiles* and *initiates*. The clubs competed in a 9-month tournament (from middle September until late May). In general, clubs trained 3-5 sessions per week (90-120 minutes) and competed once per week (usually on Saturdays or Sundays).

Chronological and skeletal ages

CA was calculated as the difference between birthdate and the date of the visit to the clinic. The films of the left hand-wrist were evaluated twice by each of the two observers. Observer A (first author) completed a 3-year Bsc in Sport Sciences in addition to a 2-year Msc in Youth Sports including a 27-hour course dealing with biological maturation. Subsequently, enrolled in the PhD programme and already complete a 45-hour training in the assessment of skeletal age that includes 100 assessments using concurrent methods to determine SA. Before assessing the x-rays of the current study, over the past four years determined SA of more than 1000 cases. The second author is Professor at the University of Coimbra over the past 27 years and was trained by the last author in the determination of SA more than 20 years ago and already assessed more than 5000 films using Greulich Pyle, Tanner-Whitehouse and Fels protocols. Repeated assessments by each of the two observers were obtained after one month.

The TW method - version 2 (TW2) was used to assess skeletal age (Tanner et al., 1975). The method is based on matching a specific bone on the radiograph with the verbally described criteria for specific stages for the bones. Twenty bones were rated: 13 long bones (radius, ulna, the metacarpals and the proximal, middle and distal phalanges of the first, third and fifth digits) and seven carpals (excluding the pisiform. Stages were essentially the same of the original TW version (Tanner et al., 1962).

The three scoring systems are specific to each SA in the TW2 version: 20-bone, CARPAL and RUS. A specific point score is assigned to each stage for each individual bone. The scores for each bone are summed to give a skeletal maturity score, which ranges from zero (immaturity) to 1000 (maturity). The CARPAL and the RUS bones were somewhat arbitrarily weighted so that each contributed 50% to the total skeletal maturity score and the overall differences between bones within each group were minimized. Finally, sex-specific tables convert the total score at a particular system (20-bone, RUS and CARPAL) into an individual SA. As noted, 1000 points indicates the skeletally mature state and an SA is not assigned for individuals who are skeletally mature (Malina, 2011).

Analyses

Frequencies for bone-specific developmental stages were presented separately for each occasion (time-moment 1; time-moment 2) for observer A and observer B. Rates of intra-observer agreement were calculated for each individual bone and for the total of observations (142 participants multiplied by 20 bones, 2840 observations). Discrepancies for stages between time-moments were noted as -2, -1, +1, +2 as time moment 2 minus time moment 1). Intra-observer mean differences were also calculated using paired t-tests, separately for bone-specific scores (points) and also for the three systems (20-bone, RUS, CARPAL) and for respective SAs. The preceding was done separately for observer A and observer B. Based on time-moment 2 for each observer, similar analyses were done to examine inter-observer variation in assessments. Technical errors of measurement (TEM), coefficients of variation (%CV) and intra-class correlation coefficients (ICC) were calculated. The magnitude effect was calculated using d-values (Cohen, 1988) and interpreted as follows (Hopkins, Marshall, Batterham, & Hanin, 2009): $d<0.20$ (trivial), $0.20<d<0.60$ (small), $0.60<d<1.20$ (moderate), $1.20<d<2.00$ (large), $2.00<d<4.00$ (very large), and >4.00 (nearly perfect). The analyses using SAs (20-bone SA, Carpal SA, RUS SA) as dependent variables were limited to participants who were not skeletally mature. Significance level was set at 5%. Analyses were performed using the Statistical Package for the Social Sciences version 26.0 (SPSS Inc., IBM Company, Armonk, NY, USA) and GraphPad Prism (version 5 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com).

3.4. Results

Developmental stages for each of the 20 bones at time moment 1 and time moment 2 for each observer are summarized in Table 3.1. Agreement was 95.1% and 93.8% for, respectively, Observer A and Observer B. Intra-observer error assessed as the difference between time-moment 2 minus time-moment 1 was equally distributed: 70 negative (2.5%) and 69 positive (2.4%) for observer A; and 91 negative (3.2%) and 85 positive (3.0%) for observer B. Technical errors of measurements, coefficients of variation and intra-class correlations for Observer A are summarized in Table 3.2. For the 20-bone system, mean difference between time moments was significant for the capitate ($t=2.022$, $p<0.05$), although the CV was less than 5% and ICC was 0.823. The ICC fluctuated between 0.823 and 0.993 for, respectively, the capitate and distal phalanx-I; the coefficient was 0.997 (TEM=8.01, %CV=0.95) for the 20-bone score. In the CARPAL protocol, the capitate was again the single bone presenting an intra-observer mean difference (t=2.022, p<0.05; TEM=6.41, %CV=3.03; ICC=0.834). In contrast, there was negligible variation in the CARPAL score (TEM=8.99, %CV=0.97; ICC=0.993). For the RUS protocol, mean differences were not significant and the ICC coefficient for the RUS score was 0.997 (TEM=13.92, %CV =2.60). Similarly, intra-observer agreement for observer B on the three scoring systems is summarized in Table 3.3. Overall, ICC scores were acceptable for each system: 20-bone (TEM=9.68, %CV=1.15; ICC=0.996), CARPAL (TEM=12.95, %CV=1.32; ICC=0.990), RUS (TEM=19.96, %CV=3.68; ICC=0.994). Significant intra-individual mean differences were noted for medial phalanx-V (t=-3.754, p<0.01; TEM=0.66, %CV=3.59; ICC=0.971) and distal phalanx-I (t=-4.488, p<0.01; TEM=2.49, %CV=9.68; ICC=0.934) in the 20-bone protocol, and only for the trapezium (t=-2.921, p<0.01; TEM=2.46, %CV=2.43; ICC=0.987) with the CARPAL protocol. Significant intra-individual differences were absent with the RUS protocol. Agreement rates for bone stages between observers A and B are presented in Table 3.4. Overall, they were greater than 80% with the 20-bone protocol, but significant differences were noted for 13 bones (5 carpals, metacarpal-I, metacarpal-III, metacarpal-V, proximal phalanges-I, III and V, distal phalanx-III). Bone-specific ICC coefficients ranged from 0.791 to 0.974. The lack of concordance between observers was similar for the CARPAL and RUS systems. Divergence between observers was noted for four of the seven CARPALS and for eight of 13 bones in the RUS system. However, the ICC coefficients for the total scores for each system were 0.990 (20-bone), 0.969 (CARPAL), and 0.988 (RUS).

Table 3.1. Frequencies of developmental stages for each bone assessed with TW2 method assigned by observers A and B on two occasions (time moment 1 *versus* time moment 2), absolute and relative agreement rates, and intra-observer error in assessments of SA among 142 adolescent male soccer players.

Bone	frequencies by stages according to TW2 method																Agreements	disagreements (stage differences)			
	time moment 1								time moment 2												
	B	C	D	E	F	G	H	I	B	C	D	E	F	G	H	I	f (%)	-2	-1	+1	+2
Intra-observer (A)																					
Radius					4	69	40	29					1	71	43	27	130(91.5%)		5	7	
Ulna			2	4	55	63	18				2	4	50	69	17		132(93.0%)		3	7	
Capitate					1	3	138						1	7	134		138(97.2%)		4		
Hamate					5	10	55	72					5	9	56	72	133(93.7%)		4	5	
Triquetral				14	20	42	66					14	21	41	66		137(96.5%)		3	2	
Lunate				1	15	36	90					1	16	31	94		137(96.5%)		1	4	
Scaphoid			2	7	18	41	74				1	8	19	41	73		133(93.7%)		5	4	
Trapezium			1	2	17	34	46	42			1	3	17	34	45	42	137(96.5%)		4	1	
Trapezoid			1	1	9	53	78				1	1	8	53	79		134(94.4%)		3	5	
Metacarpal I			3	32	21	60	15	11			2	29	27	58	15	11	135(95.1%)		2	5	
Metacarpal III				10	51	52	20	9				7	54	52	20	9	139(97.9%)			3	
Metacarpal V				32	23	62	19	6				32	22	62	19	7	137(96.5%)		1	4	
Proximal phalange I				12	56	59	7	8				11	58	58	7	8	138(97.2%)		2	2	
Proximal phalange III				17	58	40	18	9				15	61	41	17	8	134(94.4%)		5	3	
Proximal phalange V				26	48	47	13	8				26	50	46	12	8	137(96.5%)		4	1	
Medial phalange III				26	43	58	10	5				21	50	56	9	6	134(94.4%)		2	6	
Medial phalange V				46	43	38	10	5				43	52	32	8	7	131(92.3%)		6	5	
Distal phalange I				8	64	42	12	16				8	66	41	11	16	139(97.9%)		3		
Distal phalange III					64	56	13	9					69	52	10	11	134(94.4%)		6	2	
Distal phalange V				8	63	53	8	10				9	65	50	8	10	132(93.0%)		7	3	
All bones n																	2701		70	69	
%																	95.1%		2.5%	2.4%	
Intra-observer (B)																					
Radius					5	67	50	20					9	63	53	17	127(89.4%)		11	4	
Ulna				10	48	54	30					11	47	54	30		131(92.3%)		7	4	
Capitate					1	5	136							10	132		135(95.1%)		5	2	
Hamate					3	16	42	81					3	13	45	81	131(92.3%)		4	7	
Triquetral				16	30	38	58					15	27	44	56		133(93.7%)		3	6	
Lunate				2	18	39	83						21	43	78		126(88.7%)		10	6	
Scaphoid			1	16	28	26	71				1	14	26	31	70		131(92.3%)		3	8	
Trapezium			1	4	27	31	41	38			1	4	26	31	41	39	135(95.1%)		2	5	
Trapezoid				2	17	43	80					2	17	43	80		134(94.4%)		4	4	
Metacarpal I			2	7	53	53	16	11			2	6	53	56	14	11	136(95.8%)		3	3	
Metacarpal III				1	46	63	23	9				1	49	60	23	9	137(96.5%)		4	1	
Metacarpal V				17	34	70	15	6				16	34	71	15	6	136(95.8%)		2	4	
Proximal phalange I				7	57	62	7	9				3	67	54	8	10	129(90.8%)		6	7	
Proximal phalange III				7	61	46	18	10				7	57	52	16	10	134(94.4%)		3	5	
Proximal phalange V				20	50	44	17	11				20	53	42	15	12	133(93.7%)		6	3	
Medial phalange III				18	62	44	10	8				18	63	45	8	8	135(95.1%)		5	2	
Medial phalange V				42	64	23	7	9				41	64	23	5	9	133(93.7%)		6	3	
Distal phalange I				24	54	30	15	19				21	57	30	15	19	137(96.5%)		1	4	
Distal phalange III					65	51	11	15					64	53	10	15	136(95.8%)		3	3	
Distal phalange V				5	74	43	6	14				5	76	38	9	14	135(95.1%)		3	4	
All bones n																	2664		91	85	
%																	93.8%		3.2%	3.0%	

TW2 (Tanner-Whitehouse version 2); f (frequencies).

Table 3.2. Descriptive statistics (mean ± standard deviation) for the scores of each bone in the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observer A on two occasions (time moment 1 *versus* time moment 2), paired t-tests, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients among the 142 adolescent male soccer players.

Yi: dependent variable	descriptive statistics		Paired t-test		magnitude effect		TEM	%CV	ICC
	TM1	TM2	t	p	d	(qualitative)			
20-bone									
Capitate	115±5	114±7	2.022	0.045	0.131	trivial	3.20	2.79	0.823
Hamate	98±10	98±10	-0.312	0.756	-0.008	trivial	2.09	2.14	0.976
Triquetral	50±12	50±12	0.259	0.796	0.004	trivial	1.60	3.18	0.991
Lunate	53±9	54±9	-1.709	0.090	-0.042	trivial	1.96	3.67	0.978
Scaphoid	49±10	48±10	0.422	0.674	0.011	trivial	2.24	4.62	0.976
Trapezium	47±9	46±10	0.935	0.352	0.014	trivial	1.21	2.59	0.992
Trapezoid	48±10	48±10	-0.600	0.549	-0.018	trivial	2.37	4.96	0.968
Radius	87±14	88±12	-1.189	0.237	-0.046	trivial	4.30	4.90	0.943
Ulna	66±12	67±12	-1.423	0.157	-0.045	trivial	3.10	4.66	0.964
Metacarpal-I	26±5	26±4	-1.434	0.154	-0.031	trivial	0.83	3.24	0.984
Metacarpal_III	19±4	19±4	-1.744	0.083	-0.032	trivial	0.62	3.18	0.988
Metacarpal-V	19±4	19±4	-0.930	0.354	-0.014	trivial	0.51	2.72	0.992
Proximal phalange-I	25±4	25±4	-0.203	0.839	-0.005	trivial	0.87	3.46	0.978
Proximal phalange-III	21±3	21±3	-0.610	0.543	-0.015	trivial	0.68	3.25	0.979
Proximal phalange-V	20±4	20±4	0.733	0.465	0.014	trivial	0.57	2.89	0.987
Medial phalange-III	20±4	20±3	-1.845	0.067	-0.054	trivial	0.84	4.28	0.968
Medial phalange-V	18±3	18±3	-0.496	0.621	-0.011	trivial	0.60	3.26	0.981
Distal phalange-I	26±5	26±5	1.514	0.132	0.021	trivial	0.59	2.24	0.993
Distal phalange-III	19±3	19±3	1.918	0.057	0.054	trivial	0.69	3.62	0.971
Distal phalange-V	18±3	18±3	1.417	0.159	0.041	trivial	0.71	3.99	0.969
20-bone score	843±105	844±103	-0.941	0.348	-0.009	trivial	8.01	0.95	0.997
Carpal									
Capitate	212±10	211±13	2.022	0.045	0.128	trivial	6.41	3.03	0.834
Hamate	184±14	185±14	-0.333	0.740	-0.011	trivial	3.91	2.12	0.958
Triquetral	105±22	105±22	0.407	0.684	0.006	trivial	2.61	2.48	0.993
Lunate	111±13	111±13	-1.422	0.157	-0.031	trivial	2.47	2.22	0.982
Scaphoid	104±15	104±14	0.252	0.802	0.006	trivial	2.82	2.70	0.981
Trapezium	104±13	103±13	1.527	0.129	0.022	trivial	1.64	1.59	0.992
Trapezoid	103±14	104±14	-0.672	0.503	-0.019	trivial	3.35	3.24	0.972
Carpal score	924±78	923±78	1.117	0.266	0.015	trivial	8.99	0.97	0.993
RUS									
Radius	126±50	126±48	0.646	0.520	0.017	trivial	10.55	8.38	0.976
Ulna	94±40	95±39	-0.851	0.396	-0.022	trivial	8.50	9.01	0.977
Metacarpal-I	34±13	35±13	-0.374	0.709	-0.004	trivial	1.11	3.22	0.996
Metacarpal_III	28±11	29±11	-1.744	0.083	-0.014	trivial	0.72	2.53	0.998
Metacarpal-V	27±11	27±11	-1.635	0.104	-0.016	trivial	0.88	3.27	0.997
Proximal phalange-I	34±12	34±12	0.685	0.495	0.007	trivial	1.04	3.07	0.996
Proximal phalange-III	28±10	28±9	0.668	0.505	0.012	trivial	1.42	5.02	0.988
Proximal phalange-V	26±9	26±9	1.178	0.241	0.014	trivial	0.91	3.47	0.995
Medial phalange-III	27±9	28±9	-1.657	0.100	-0.026	trivial	1.22	4.46	0.991
Medial phalange-V	25±9	25±9	1.206	0.230	0.023	trivial	1.53	6.12	0.986
Distal phalange-I	35±13	35±13	1.717	0.088	0.017	trivial	1.11	3.17	0.996
Distal phalange-III	25±8	25±9	0.536	0.593	0.013	trivial	1.77	6.98	0.978
Distal phalange-V	24±9	24±9	1.196	0.234	0.032	trivial	1.94	8.09	0.974
RUS score	535±182	535±180	0.302	0.763	0.003	trivial	13.92	2.60	0.997

TW2 (Tanner-Whitehouse version 2); t (t-value of paired t-test); p (significance value); d (d-cohen value); TEM (technical error of measurement); %CV (coefficient of variation); ICC (intra-class correlation coefficient).

Table 3.3. Descriptive statistics (mean ± standard deviation) for each bone score with the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observer B on two occasions (time moment 1 *versus* time moment 2), paired t-tests, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients in 142 adolescent male soccer players.

Yi: dependent variable	descriptive statistics		Paired t-test		magnitude effect		TEM	%CV	ICC
	TM1	TM2	t	p	d	(qualitative)			
20-bone									
Capitate	115±6	114±7	1.471	0.143	0.108	trivial	3.97	3.47	0.764
Hamate	98±10	99±10	-0.781	0.436	-0.024	trivial	2.50	2.54	0.966
Triquetral	49±12	49±12	-0.105	0.917	-0.002	trivial	2.26	4.66	0.982
Lunate	52±10	52±10	1.330	0.186	0.051	trivial	3.13	6.04	0.946
Scaphoid	47±11	47±11	-0.673	0.502	-0.014	trivial	1.94	4.10	0.985
Trapezium	45±10	45±10	-0.878	0.381	-0.018	trivial	1.69	3.72	0.985
Trapezoid	48±10	48±10	0.000	1.000	0.000	trivial	2.18	4.59	0.976
Radius	87±13	86±15	1.696	0.092	0.073	trivial	5.14	5.96	0.928
Ulna	67±13	67±13	0.364	0.717	0.009	trivial	2.76	4.12	0.977
Metacarpal-I	26±4	26±4	-0.661	0.510	-0.010	trivial	0.45	1.70	0.992
Metacarpal_III	20±3	20±3	1.728	0.086	0.039	trivial	0.62	3.07	0.982
Metacarpal-V	19±3	19±3	-1.000	0.319	-0.019	trivial	0.53	2.76	0.988
Proximal phalange-I	26±4	26±4	-0.243	0.808	-0.010	trivial	1.22	4.73	0.941
Proximal phalange-III	22±3	22±3	-1.590	0.114	-0.026	trivial	0.38	1.74	0.991
Proximal phalange-V	20±3	20±3	0.602	0.548	0.017	trivial	0.79	3.94	0.973
Medial phalange-III	20±3	20±3	0.728	0.468	0.011	trivial	0.41	2.06	0.992
Medial phalange-V	18±3	19±3	-3.754	<0.001	-0.102	trivial	0.66	3.59	0.971
Distal phalange-I	25±6	26±8	-4.488	<0.001	-0.178	trivial	2.49	9.68	0.934
Distal phalange-III	19±3	19±3	-0.355	0.723	-0.009	trivial	0.67	3.48	0.974
Distal phalange-V	18±3	18±3	0.609	0.543	0.013	trivial	0.49	2.71	0.985
20-bone score	841±106	841±105	0.208	0.836	0.002	trivial	9.68	1.15	0.996
Carpal									
Capitate	212±12	210±14	1.405	0.162	0.103	trivial	8.01	3.80	0.765
Hamate	185±14	185±14	-1.127	0.262	-0.034	trivial	3.48	1.88	0.967
Triquetral	102±22	103±22	-0.873	0.384	-0.018	trivial	3.73	3.65	0.985
Lunate	109±14	109±13	0.771	0.442	0.031	trivial	4.61	4.23	0.941
Scaphoid	101±17	102±16	-1.460	0.147	-0.031	trivial	2.94	2.89	0.984
Trapezium	101±16	101±14	-2.921	0.004	-0.055	trivial	2.46	2.43	0.987
Trapezoid	103±15	103±15	0.000	1.000	0.000	trivial	3.23	3.14	0.977
Carpal score	912±86	912±83	-0.339	0.735	-0.006	trivial	12.05	1.32	0.990
RUS									
Radius	122±45	119±43	1.760	0.080	0.054	trivial	11.44	9.49	0.965
Ulna	101±47	101±47	0.087	0.931	0.003	trivial	12.17	12.18	0.965
Metacarpal-I	35±12	35±12	0.103	0.918	0.002	trivial	1.72	4.90	0.990
Metacarpal_III	30±10	30±10	1.345	0.181	0.025	trivial	1.59	5.29	0.987
Metacarpal-V	27±10	27±10	-0.539	0.591	-0.011	trivial	1.65	6.09	0.986
Proximal phalange-I	35±12	35±12	-0.373	0.710	-0.010	trivial	2.54	7.31	0.975
Proximal phalange-III	29±9	30±9	-0.580	0.563	-0.011	trivial	1.43	4.84	0.987
Proximal phalange-V	27±10	27±10	0.970	0.334	0.017	trivial	1.47	5.37	0.989
Medial phalange-III	27±10	27±9	1.176	0.241	0.024	trivial	1.62	5.93	0.985
Medial phalange-V	25±10	24±9	1.122	0.264	0.023	trivial	1.59	6.46	0.985
Distal phalange-I	34±15	35±15	-1.407	0.162	-0.013	trivial	1.14	3.31	0.997
Distal phalange-III	26±10	26±10	-0.172	0.864	-0.003	trivial	1.37	5.26	0.990
Distal phalange-V	24±9	24±9	-0.138	0.891	-0.002	trivial	1.29	5.32	0.991
RUS score	543±186	540±183	1.194	0.234	0.015	trivial	19.96	3.68	0.994

TW2 (Tanner-Whitehouse version 2); t (t-value of paired t-test); p (significance value); d (d-cohen value); TEM (technical error of measurement); %CV (coefficient of variation); ICC (intra-class correlation coefficient).

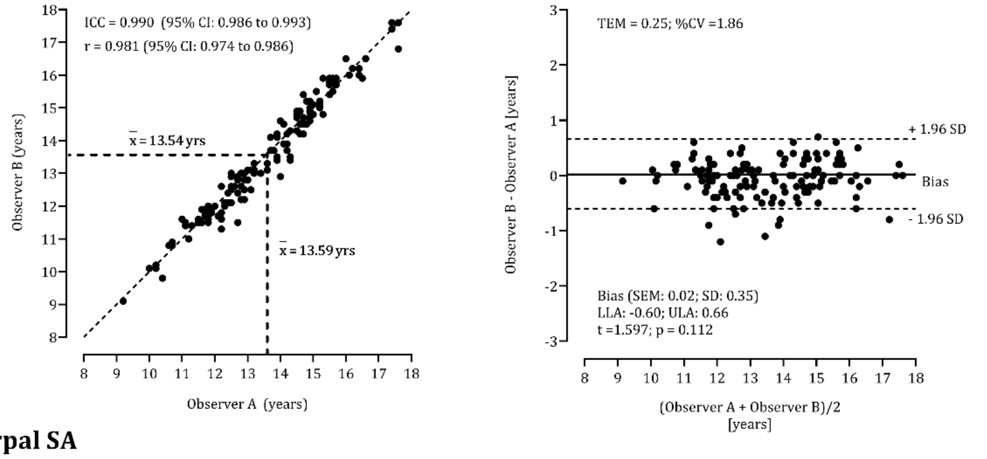
Table 3.4. Descriptive statistics (mean ± standard deviation) for each bone score in the three scoring systems (TW2 20-bone, Carpal, RUS) assigned by observers A and B, paired t-test, effect sizes, technical errors of measurement, coefficients of variation and intra-class correlation coefficients in 142 adolescent male soccer players.

Variable	%	descriptive statistics		Paired t-test		magnitude effect		TEM	%CV	ICC
	agr.	Obs A	Obs B	t	p	d	(qualitative)			
20-bone										
Capitate	95	114±7	114±7	0.656	0.513	0.046	trivial	3.97	3.47	0.791
Hamate	87	98±10	99±10	-2.347	0.020	0.093	trivial	3.24	3.30	0.940
Triquetral	86	50±12	49±12	4.124	0.000	0.135	trivial	2.26	4.66	0.956
Lunate	83	54±9	52±10	3.968	0.000	0.207	small	4.38	8.34	0.884
Scaphoid	83	48±10	47±11	4.420	0.000	0.111	trivial	2.41	5.04	0.974
Trapezium	82	46±10	45±10	3.982	0.000	0.108	trivial	2.32	5.04	0.971
Trapezoid	81	48±10	48±10	0.647	0.519	0.031	trivial	3.94	8.26	0.912
Radius	80	88±12	86±15	3.219	0.002	0.173	trivial	6.32	7.27	0.880
Ulna	80	67±12	67±13	-0.553	0.581	0.024	trivial	4.49	6.72	0.930
Metacarpal-I	80	26±4	26±4	-4.153	0.000	0.177	trivial	1.52	5.85	0.925
Metacarpal_III	82	19±4	20±3	-4.464	0.000	0.205	small	1.49	7.49	0.910
Metacarpal-V	80	19±4	19±3	-3.962	0.000	0.156	trivial	1.32	6.91	0.937
Proximal phalange-I	87	25±4	26±4	-2.622	0.010	0.122	trivial	1.52	5.98	0.915
Proximal phalange-III	81	21±3	22±3	-4.573	0.000	0.226	small	1.34	6.31	0.894
Proximal phalange-V	80	20±4	20±3	-3.122	0.002	0.122	trivial	1.17	5.95	0.940
Medial phalange-III	80	20±3	20±3	0.346	0.730	0.015	trivial	1.20	6.05	0.925
Medial phalange-V	82	18±3	19±3	-1.890	0.061	0.081	trivial	1.01	5.51	0.930
Distal phalange-I	80	26±5	26±8	-0.327	0.744	0.020	trivial	3.25	12.36	0.854
Distal phalange-III	87	19±3	19±3	-2.183	0.031	0.089	trivial	1.02	5.36	0.936
Distal phalange-V	80	18±3	18±3	-0.262	0.793	0.012	trivial	1.13	6.30	0.917
20-bone score		844±103	841±105	2.027	0.045	0.034	trivial	14.79	1.76	0.990
Carpal										
Capitate		211±13	210±14	0.599	0.550	0.042	trivial	8.01	3.81	0.792
Hamate		185±14	185±14	-0.967	0.335	0.054	trivial	6.38	3.45	0.877
Triquetral		105±22	103±22	4.290	0.000	0.122	trivial	5.47	5.27	0.967
Lunate		111±13	109±13	3.577	0.000	0.183	trivial	5.97	5.42	0.891
Scaphoid		104±14	102±16	4.685	0.000	0.150	trivial	4.44	4.30	0.957
Trapezium		103±13	101±14	4.334	0.000	0.129	trivial	3.70	3.62	0.964
Trapezoid		104±14	103±15	1.149	0.253	0.054	trivial	5.84	5.66	0.915
Carpal score		923±78	912±83	4.871	0.000	0.133	trivial	19.78	2.16	0.969
RUS										
Radius		126±48	119±43	3.024	0.003	0.134	trivial	17.64	14.41	0.921
Ulna		95±39	101±47	-2.637	0.009	0.136	trivial	19.27	19.72	0.892
Metacarpal-I		35±13	35±12	-2.065	0.041	0.041	trivial	2.09	6.02	0.986
Metacarpal_III		29±11	30±10	-4.000	0.000	0.142	trivial	3.28	11.20	0.949
Metacarpal-V		27±11	27±10	-0.876	0.383	0.027	trivial	2.64	9.79	0.966
Proximal phalange-I		34±12	35±12	-3.080	0.002	0.091	trivial	2.98	8.68	0.966
Proximal phalange-III		28±9	30±9	-4.529	0.000	0.147	trivial	2.63	9.09	0.957
Proximal phalange-V		26±9	27±10	-3.785	0.000	0.120	trivial	2.71	10.14	0.960
Medial phalange-III		28±9	27±9	1.186	0.238	0.046	trivial	3.01	10.99	0.944
Medial phalange-V		25±9	24±9	1.182	0.239	0.042	trivial	2.77	11.22	0.954
Distal phalange-I		35±13	35±15	0.820	0.414	0.024	trivial	3.54	10.17	0.968
Distal phalange-III		25±9	26±10	-2.808	0.006	0.095	trivial	2.68	10.44	0.956
Distal phalange-V		24±9	24±9	-1.082	0.281	0.043	trivial	3.02	12.57	0.941
RUS score		535±180	540±183	-1.751	0.082	0.031	trivial	27.45	5.11	0.988

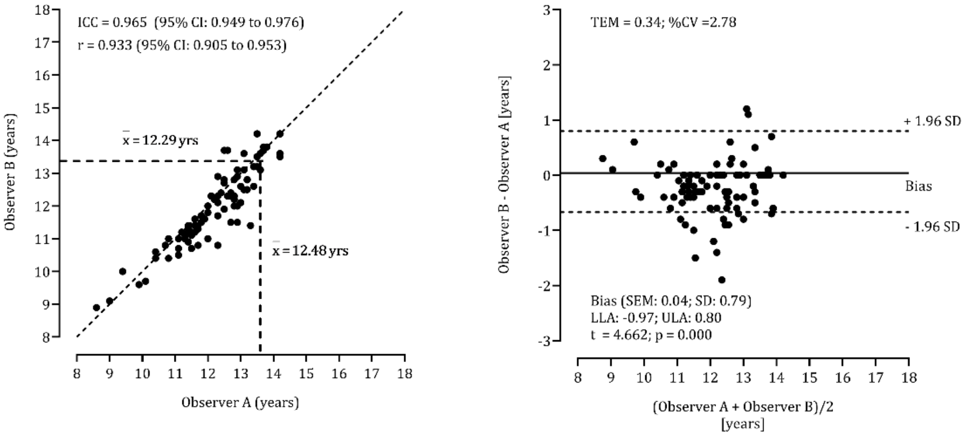
TW2 (Tanner-Whitehouse version 2); t (t-value of paired t-test); p (significance value); d (d-cohen value); TEM (technical error of measurement); %CV (coefficient of variation); ICC (intra-class correlation coefficient); Obs (observer).

The impact of observer-associated variation in the point scores for each protocol on the respective SAs is illustrated in Figure 3.1. Comparisons between observers were not significant for 20-bone SAs (TEM=0.25 years, %CV=1.86; ICC=0.990, 95%CI: 0.986 to 0.993) neither the RUS SAs (TEM=0.31 years, %CV=2.22; ICC=0.984, 95%CI: 0.978 to 0.989). Although the difference between mean CARPAL SAs of the two observers was small, 12.48±1.18 years and 12.29±1.24 years for observers A and B, respectively (t=4.662, $p<0.01$), the difference between observers had little impact as shown in the respective panel: ICC=0.965, 95%CI: 0.949 to 0.976, TEM=0.34 years, %CV=2.78). For the three SA protocols, BIAS was negligible: 0.02 years (20-bone), 0.04 years (CARPAL), and 0.03 years (RUS).

20-bone SA



Carpal SA



RUS SA

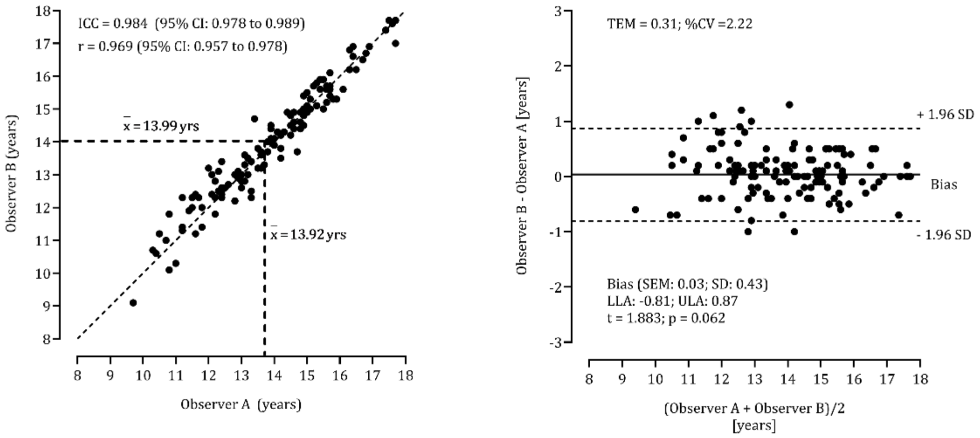


Figure 3.1. Interobserver agreement for the determination of skeletal age by two observers, considering concurrent systems (20-bone, Carpal and RUS) and indication of the mean for each observer, bivariate correlation coefficient between the series produced by observers A and B, intra-class correlation coefficient; complemented with Bland-Altman analysis to inspect intra-individual differences expressed against the mean values.

3.5. Discussion

The present study evaluated intra-observer agreement for SA assessments on two independent occasions using the TW2 20-bone, CARPAL and RUS protocols among male soccer players 11-15 years of age. Overall agreement between the two time-moments was acceptable for the three systems. Discrepancies did not exceed one stage and there was no specific trend for the replicate assessment to exceed or fall below that for the initial assessment. With the 20-bone protocol, bone-specific technical errors of measurement were always < 5% of one observer and exceeded 5% for only three bones by the other observer. Disagreement seemed slightly higher for the CARPAL and RUS protocols which are based on smaller number of bones; this likely reflected the scoring system as the 20 bone, CARPAL and RUS protocols were based on a 1000-point scale. Nevertheless, allowing for several problematic bones, intra-observer agreement for the respective SAs were acceptable both in terms of scores and assigned SAs.

TW2 protocol has been updated (TW3) and has been used in the sports sciences (Romann & Fuchslocher, 2016). The original version (TW1) was developed on a British sample of average socioeconomic status (Tanner et al., 1962). The scores were designed to represent biological weights for each of 20 bones bone and the overall score was obtained by summing the scores of the 20 bones. Specific tables were used to convert the 20-bone score into a SA (20-bone TW1-SA). The first revision of the method (TW2) retained the verbal criteria for the respective stages of the 20 bones with few refinements (Tanner et al., 1975): radius (stage J was deleted), ulna (stages I was deleted) and for five carpals (capitate, triquetral, lunate, scaphoid, trapezoid) the final stage I was deleted. This initial revision included changes in the scores associated with each stage. Three maturity scores were separately developed for boys and girls to derive an SA with each protocol: carpals (CARPAL TW2 -SA), radius, ulna and short bones (RUS TW2 -SA) in addition to the 20-bone TW2SA.

The most recent revision for the TW protocol (TW3) incorporated several additional samples of children and adolescents in revising the tables for converting the CARPAL and RUSs into SAs (Tanner et al., 2001). The British samples of the initial study dated from 1950s was retained while samples from Belgium (Leuven Growth Study in the 1970s)

(Beunen et al., 1990), Spain (Bilbao in the 1980s) (Hernandez, Sanchez, Sobradillo, & Rincon, 1991), Japan (Tokyo in 1986) (Ashizawa et al., 1996), Italy (north of Italy) (Vignolo et al., 1999), Argentina (LaPlata in the 1970s) (Lejarraga, Guimarey, & Orazi, 1997), and the U.S. (Texas, European-American ancestry) (Tanner, Oshman, Bahhage, & Healy, 1997) were added.

The specific stages and corresponding scores were the same as in TW2, but the TW3 revision deleted the 20-bone SA. As such, the TW3 revision includes only sex-specific tables CARPAL TW3-SA and RUS TW3-SA. In addition, skeletal maturity for the RUS TW3 protocol is attained at 16.5 years for males and 15.0 years for females. In the preceding versions of the TW method, the pre-mature state (999 points) for males corresponded to an SA of 17.9 years with the TW2 20-bone, 18.1 years with the TW2-RUS and 14.9 years with the TW2-CARPAL scoring protocols.

Early studies reporting intra-observer agreement of the TW2 method date to 1970s. In a sample of Swedish 122 boys and 90 girls 1 month to 7 years of age, replicate assessments had an overall agreement rate of about 80% (Taranger et al., 1976). Among 3817 Danish school children 7 to 18 years TW 20 bone scores largely matched the British reference (Helm, 1979). In the preceding study, 90 radiographs were rated twice and agreement rates were 88-89% for the long bones and 84-96% for the short bones. Since the carpals attained the final stages at earlier ages compared to long bones, the Danish study decided to examine x-rays from 7-13 years old boys and 7-11 years old girls, in a total of 60 cases, to obtain an agreement rate ranging 82-93%.

Meantime, TW2 assessments was previously carried out using three observers (Beunen & Cameron, 1980). Two sets of x-rays in a random order obtained from the Harpenden Longitudinal Growth Study and from the Leuven Longitudinal Study of Belgian Boys were used to test the agreement rates between observers. Significant differences were found in mean SA between observers for 20-bone SA and CARPAL SA. In contrast, no significant differences in mean SA between observers were found for RUS SA. In the present study, after converting scores to SAs, inter-observer mean differences were not significant for the TW2-20bone and for TW2-RUS SAs. In contrast, the inter-observer difference with the TW2-CARPAL protocol was a source of error with 15 cases exceeding the limits of agreement in the present study. Among 110 Danish children and adolescents aged 6-16 years (Wenzel & Melsen, 1982), intra-observer agreement fluctuated between 82% to 100%, and consistent with the current study, disagreements did not exceed more than one stage with capitate diagnosed as the most critical bone for disagreement.

Inter-observer agreement rates TW SA assessment are less frequently reported in the literature compared to intra-observer differences. In a study of Dutch children (van Venrooij-Ysselmuiden & van Ipenburg, 1978), 60 radiographs of boys 10 through 16 years were rated with the TW protocol by an expert and a Dutch author. The percentage of agreement for the ulna was 83% and that for the radius 66% with a systematic disagreement that was concentrated in the assessment of stage F. This prompted the authors to hypothesize a differential impact of observer expertise among youth 10-12 years of age. Meantime, in the present study of soccer players, disagreements between observers that exceeded the limits of agreement were concentrated between 11-13 years for TW2 Carpal SA and between 11.5-14.0 years for TW2 20-bone SA (see Figure 3.1).

The literature on the skeletal maturity status of youth soccer players has consistently shown that the sport tends to favor early maturing players as they transition into the adolescent years (Malina, 2011; Malina et al., 2000). A band of plus/minus 1.0 year is commonly used to classify players as late, average or early maturing. In the present study and based on assessments of observer A (first author), early maturing players represented 36% at time moment 1 (TM1) and 37% (TM2) while using the TW2 20-bone SA, thus suggesting that intra-individual error marginally impacted the frequencies of maturity status. Corresponding estimates of maturity classifications with TW2-RUS SA classified 49% and 50% of the participants as advanced in TM1 and TM2, respectively. In contrast, percentages of players classified as advanced with TW2-CARPAL SA were, respectively, 8% and 11%. By inference, intra-observer error in assessments did not appear to influence maturity status classifications.

The present study highlights the expertise of SA assessments with the TW2 protocol among adolescent soccer players. The study is novel as it considers intra-observer analyses for each bone in addition to the three protocols (20-bone, CARPALS, RUS) both using scores and assigned SAs as the dependent variable. Nevertheless, few limitations should be considered. First, the study was focused on the ability of two observers and inter-examiner agreement is essential in research projects using more than two examiners. Additionally, the results are limited to a sample of 142 male soccer players 11-15 years. Given the CA range, it was not possible to evaluate early stages for specific boys, e.g., stages B-E for the radius, capitate, hamate and distal phalange III; B-D for the triquetral, lunate, metacarpals II-V, proximal phalanges I-V, and distal phalanges I and V; and for stages B-C of the ulna, scaphoid, trapezium, trapezoid, and metacarpal I. By inference, there is a need for additional research on pre-teens, especially for CARPAL protocol. Note, the age interval of the current sample included 25 participants who were classified as skeletally mature and as such were not included in the calculations illustrated in Figure 3.1. Nevertheless, the literature generally considers descriptors of the stages for round bones (carpals) more difficult to evaluate compared to long bones and as noted, the capitate has been previously indicated as problematic (Beunen & Cameron, 1980; Medicus, Grøn, & Moorrees, 1971; Tiisala, Kantero, & Tamminen, 1971). The carpals are more difficult to evaluate because they involve assessments of shape and radiopaque lines or zones, whereas assessments of the long bones tend to concentrate on the centers of ossification and epiphyseo-diaphysial relationships and fusion (Roche, 1989).

3.6. Conclusions

In summary, the assignment of developmental stages is specific for each bone and is somewhat more problematic for the round (carpals) than for the long bones. Examiners should be encouraged to evaluate their expertise on perhaps 100 images spanning a broad range of CAs. Data quality using adolescent samples should not be generalized to early ages. Finally, if disagreements between replicate assessments are not greater than one stage and shows equal probability for the replicates to be lower or higher compared to initial assignments, the effect on assigned SAs appears to be trivial or small.

3.7. References

Ashizawa, K., Asami, T., Anzo, M., Matsuo, N., Matsuoka, H., Murata, M., … Tsukagoshi, K. (1996). Standard RUS skeletal maturation of Tokyo children. Annals of Human Biology, 23(6), 457–469.

Beunen, G., & Cameron, N. (1980). The reproducibility of TW2 skeletal age assessments by a self-taught assessor. Annals of Human Biology, 7(2), 155–162.

Beunen, G., Lefevre, J., Ostyn, M., Renson, R., Simons, J., & Van Gerven, D. (1990). Skeletal maturity in Belgian youths assessed by the Tanner-Whitehouse method (TW2). Annals of Human Biology, 17(5), 355–376.

Carling, C., Le Gall, F., & Malina, R. M. (2012). Body size, skeletal maturity, and functional characteristics of elite academy soccer players on entry between 1992 and 2003. Journal of Sports Sciences, 30(15), 1683–1693.

Coelho-e-Silva, M. J., Figueiredo, A. J., Simões, F., Seabra, A., Natal, A., Vaeyens, R., … Malina, R. M. (2010). Discrimination of U-14 soccer players by level and position. International Journal of Sports Medicine, 31(11), 790–796.

Cohen, J. (1988). Statistical power analysis for behavioral sciences (2nd ed.). Mahaw (NJ): Lawrence Erlbaum Associates.

Greulich, W., & Pyle, S. (1959). Radiographic atlas of skeletal development of the hand and wrist (2nd ed.). Stanford, CA: Stanford University Press.

Harriss, D. J., Macsween, A., & Atkinson, G. (2019). Ethical Standards in Sport and Exercise Science Research: 2020 Update. International Journal of Sports Medicine, 40(13), 813–817.

Helm, S. (1979). Skeletal maturity in Danish schoolchildren assessed by the TW2 method. American Journal of Physical Anthropology, 51(3), 345–351.

Hernandez, M., Sanchez, E., Sobradillo, B., & Rincon, J. M. (1991). Skeletal Maturation and Height Prediction: Atlas and Scoring Methods. Madrid: Diaz de Santos.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise, 41(1), 3–12.

Lejarraga, H., Guimarey, L., & Orazi, V. (1997). Skeletal maturity of the hand and wrist of healthy Argentinian children aged 4-12 years, assessed by the TWII method. Annals of Human Biology, 24(3), 257–261.

Malina, R. M. (2011). Skeletal Age and Age Verification in Youth Sport. Sports Medicine, 41(11), 925–947. https://doi.org/10.1016/j.yspm.2012.01.014

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Chamorro, M., Serratosa, L., & Morate, F. (2007). TW3 and Fels skeletal ages in elite youth soccer players. Annals of Human Biology, 34(2), 265–272.

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Philippaerts, R. M., Hirose, N., Peña Reyes, M. E., … Buranarugsa, R. (2018). Tanner-Whitehouse Skeletal Ages in Male Youth Soccer Players: TW2 or TW3? Sports Medicine, 48(4), 991–1008.

Malina, R. M., Peña Reyes, M., Eisenmann, J., Horta, L., Rodrigues, J., & Miller, R. (2000). Height, mass and skeletal maturity of elite Portuguese soccer players aged 11–16 years. Journal of Sports Sciences, 18, 685–693.

Medicus, H., Grøn, A. -M, & Moorrees, C. F. A. (1971). Reproducibility of rating stages of osseous development. (Tanner-Whitehouse system). American Journal of Physical Anthropology, 35(3), 359–371.

Mentzel, H.-J., Vilser, C., Eulenstein, M., Schwartz, T., Vogt, S., Böttcher, J., … Kaiser, W. A. (2005). Assessment of skeletal age at the wrist in children with a new ultrasound device. Pediatric Radiology, 35(4), 429–433.

Moran, J., Paxton, K., Jones, B., Granacher, U., Sandercock, G. R., Hope, E., & Ramirez-Campillo, R. (2020). Variable long-term developmental trajectories of short sprint speed and jumping height in English Premier League academy soccer players: An applied case study. Journal of Sports Sciences, 38(22), 2525–2531.

Ostojic, S. M., Castagna, C., Calleja-González, J., Jukic, I., Idrizovic, K., & Stojanovic, M. (2014). The biological age of 14-year-old boys and success in adult soccer: Do early maturers predominate in the top-level game? Research in Sports Medicine, 22(4), 398–407.

Roche, A. F. (1989). Relative utility of carpal skeletal ages. American Journal of Human Biology, 1(4), 479–482.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Romann, M., & Fuchslocher, J. (2016). Assessment of skeletal age on the basis of DXA-derived hand scans in elite youth soccer. Research in Sports Medicine, 24(3), 200–211.

Tanner, J. M., Healy, M., Goldstein, H., & Cameron, N. (2001). Assessment of Skeletal Maturity and Prediction of Adult Height (TW3) Method (3rd ed.). Saunders Ltd.

Tanner, J. M., Oshman, D., Bahhage, F., & Healy, M. (1997). Tanner-Whitehouse bone age reference values for North American children. Journal of Pediatrics, Vol. 131, pp. 34–40.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1975). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London: Academic Press.

Tanner, J. M., Whitehouse, R. J., & Healy, M. J. R. (1962). A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2,600 healthy British children. Part II: The Scoring System. Paris: International Children's Centre.

Taranger, J., Bruning, B., Claesson, I., Karlberg, P., Landström, T., & Lindström, B. (1976). Skeletal development from birth to 7 years. Acta Pædiatrica, 65(15), 98–108.

Thodberg, H. H., Kreiborg, S., Juul, A., & Pedersen, K. D. (2009). The BoneXpert method for automated determination of skeletal maturity. IEEE Transactions on Medical Imaging, 28(1), 52–66.

Tiisala, R., Kantero, R. L., & Tamminen, T. (1971). A mixed longitudinal study on skeletal maturation in healthy Finnish children aged 5 to 10 years. Human Biology, 43(2), 224–236.

van der Sluis, A., Elferink-Gemser, M. T., Coelho-e-Silva, M. J., Nijboer, J. A., Brink, M. S., & Visscher, C. (2014). Sport injuries aligned to peak height velocity in talented pubertal soccer players. International Journal of Sports Medicine, 35(4), 351–355.

van Venrooij-Ysselmuiden, M. E., & van Ipenburg, A. (1978). Mixed longitudinal data on skeletal age from a group of Dutch children living in Utrecht and surroundings. Annals of Human Biology, 5(4), 359–380.

Vignolo, M., Naselli, A., Magliano, P., Di Battista, E., Aicardi, M., & Aicardi, G. (1999). Use of the new US90 standards for TW-RUS skeletal maturity scores in youths from the Italian population. Hormone Research, 51(4), 168–172.

Wenzel, A., & Melsen, B. (1982). Replicability of assessing radiographs by the Tanner and Whitehouse-2 method. Human Biology, 54(3), 575–581.

Chapter IV

Study 2

Assessment of intra-observer reproducibility and inter-observer agreement of Fels protocol to determine skeletal age among male adolescent tennis players.

4. Assessment of intra-observer reproducibility and inter-observer agreement of Fels protocol to determine skeletal age among male adolescent tennis players.

4.1. Abstract

Background: Skeletal age (SA) is used for estimating biological maturation and is often employed as part of the medical examination to readiness to participate in competitive sports. This study aimed to determine intra-observer (reproducibility) and inter-observer agreement of SA values adopting Fels method among male adolescent tennis players.

Methods: The sample included 97 male tennis players aged 8.7-16.8 years. Radiographs were observed twice by two independent trained observers. Participants were classified into the following maturity categories: late, average, early, and mature based on the differences between SA and chronological age (CA).

Results: The intra-individual differences between repeat estimations were (observer A: d=0.008; observer B: d=0.001). Coefficient of variation was respectively 1.11% and 1.75% for observer A and observer B. Inter-observer mean differences were negligible (t=1.252, p=0.210) and the associated intra-class correlation coefficient was nearly perfect (ICC=0.995). When tennis players were classified according to maturity categories, concordance between observers was 90%, indicating a strong degree of agreement.

Conclusion: The Fels method appeared reproducible and showed an acceptable level of inter-observer agreement between trained examiners. Although the SA scores derived from the Fels protocol presented errors of trivial magnitude, classification of adolescent male tennis players by maturity groups should not be considered as 100 percent reliable. Consequently, experienced observers are crucial to perform medical examinations.

Keywords: Youth sports; Biological maturation; Bone Age; Skeletal maturation; Maturity status.

4.2. Introduction

Biological maturation describes the process of progress toward the mature or adult state. Skeletal age (SA) is a commonly used indicator of maturity status in studies of youth athletes (Coelho-e-Silva et al., 2022; Martinho et al., 2022). Estimates of SA are derived from reference standards pertaining to the mean chronological age (CA) at which a specific state of maturity of the hand-wrist bones is attained relative to the general population upon which the method of assessment was developed. Three methods are commonly used to estimate SA (Malina et al., 2004): Greulich-Pyle (GP) (Greulich & Pyle, 1959), Tanner-Whitehouse (TW) (Tanner et al., 2001, 1975, 1962) and Fels (Roche et al., 1988). Although the three methods vary, each requires a radiograph of the hand-wrist and is based on the universal and invariant sequences of development of each bone from initial calcification to the mature state.

Although automated protocols for the assessment of radiographs have also been implemented (Booz et al., 2019; Grendstad et al., 2020) assessment of SA is generally done by experienced examiners. Studies do not ordinarily report intra-examiner and inter-examiner variability in assessments. The latter may reflect the broad range of normal variability that is commonly accepted in the clinical context. Nevertheless, studies reporting errors associated to the examiner in assessments of SA commonly use relatively small samples (Figueiredo et al., 2009a; Malina, Chamorro, et al., 2007). In fact, a previous study reported an intra-observer mean difference of -0.02±0.09 years between repeated assessments and a TEM of 0.06 years (Malina, Dompier, Powell, Barron, & Moore, 2007) when using Fels SAs. The preceding study was based on 15 American football players aged 9-14 years (10% of the sample). Similarly, among soccer players 11-17 years, assessments of Fels SA were repeated in 15 players (11% of the sample) to determine a mean difference of 0.08±0.17 years (Malina et al., 2000). There is a need to investigate the degree to which variation in estimations of SA exist within and between observers in relatively large samples. Additionally, the preceding studies were performed in sports which tend to exclude late maturing boys and to favour average and early maturing participants as chronological age and sport specialization increase (Malina et al., 2000). Doing so would allow scholars to better document the proportion of error or variation in assessments that can be attributed to the examiner as a source of error; and by doing it in youth tennis, it is an opportunity to confirm previous conclusions obtained from studies in youth soccer.

In the context of the preceding discussion, the purpose of this study is to determine intra- and inter-observer reproducibility of SA assessments of approximately 100 radiographs of male tennis players evaluated by two observers within an interval of one month. Inter-observer agreement in bone-specific indicators was initially evaluated. Differences between the two observers were addressed as inter-individual mean differences of concurrent SA estimates based on the second assessments. Finally, each player was classified by each of the two observers as delayed (late), on time (average), advanced (early) or skeletally mature based on the assessments of each observer and the agreement between classifications was evaluated.

4.3. Methods

Study design and participants

The present study followed ethical standards established for sports sciences (Harriss et al., 2019). The project was approved by the *Ethics Committee for Sports Sciences* by the *University of Coimbra* (CE/FCDEF-UC/00122014). Parents of the tennis players signed an informed consent. Participants were informed about the objectives, procedures, benefits, risks and also that they could withdraw from the study at any time. The sample included 97 male tennis players aged 8.7 through 16.8 years. Previous studies assessed intra-observer reproducibility in 15-participant subsamples (Malina, Dompier, et al., 2007; Malina et al., 2000). In contrast, a recent publication considered intra-observer and inter-observer agreement among 142 male pubertal soccer players with skeletal age assessed using TW2 score systems (Sousa-e-Silva et al., 2022). All players trained for at least two seasons under the supervision of a certified coach in a tennis club. They also competed in official tournaments organized by the national tennis federation. CA was calculated as the difference between birthdate and the date of the visit to the clinic for the radiograph.

Determination of skeletal age

A posterior-anterior radiographs of the left hand-wrist was obtained following the medical exams adopted by the Portuguese Institute of Sports (Law 204/2006; act 11/2012). SA was estimated with the Fels method (Roche et al., 1988). The protocol requires the evaluation of 22 bones: radius, ulna, carpals (capitate, hamate, triquetral, lunate, scaphoid, trapezium, trapezoid, pisiform), three metacarpals (I, III, V), proximal and distal phalanges of three digits (I, III, V), middle phalanges of digits III and V in addition to the absence or presence of the adductor sesamoid. Evaluations are based on specific criteria for each bone, i.e., the presence or absence of the ossification centre, changes in shape, radiopaque lines, and epiphyseal-diaphyseal fusion of the long bones and attainment of adult morphology for the carpals. Measurements of epiphyseal and diaphyseal widths of the long bones are also required. The grades assigned for each bone and the epiphyseal and metaphyseal widths are entered into a computer program (Felshw 1.0) that calculates the SA and its standard error of estimate for the individual. CA was subtracted from the SA for each player (SA minus CA). Players were classified as late (SA<CA by more than 1.0 year), on time or average (SA within ±1.0 year of CA), or early (SA>CA by more than 1.0 year); if the player had attained skeletal maturity, he was simply indicated as mature, i.e., an SA is not assigned. The band ±1.0 year is commonly used in samples of youth, both non-athletes and athletes (Malina et al., 2004). The band accommodates the in SA *per se* and also the variation associated with error in SA assessments.

Analyses

All radiographs were independently evaluated by two experienced individuals (observer A and observer B) on two occasions (labelled time-moment 1 and time-moment 2). Both observers were experienced in the estimation of SA using the Fels method; each had completed more than 1000 examinations over the past few years. The frequencies of intra-observer errors by bone were calculated. Technical error of measurement (TEM) and coefficient of variation (%CV) were determined separately for each observer in addition to intra-observer agreement tested using paired t-test. Subsequently, a paired t-student test was also performed to determine inter-observer agreement for time-moment 2 in parallel to determination of intra-class correlation coefficient (ICC). The magnitude of the effect was calculated using Cohen's d-values (Cohen, 1988) which were interpreted as follows: $d<0.20$ (trivial), $0.20<d<0.60$ (small), $0.60<d<1.20$ (moderate), $1.20<d<2.00$ (large), $2.00<d<4.00$ (very large), and >4.00 (nearly perfect) (Hopkins et al., 2009). Finally, the limits of agreement between observers were assessed using Bland–Altman analysis [16]. The preceding analyses excluded participants classified as skeletally mature. The concordance of maturity status classifications based on SA – CA between observers was determined for the total sample. Statistical Package for the Social Sciences version 26.0 (SPSS Inc., IBM Company, Armonk, NY, USA) and GraphPad Prism (version 5 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com) were used in the analyses. Significance level was set at $p<0.05$.

4.4. Results

Descriptive statistics for the estimations of SA for observer A and observer B at each time moment are summarized in Table 4.1. Intra-individual mean differences were of trivial in magnitude, while %CV were residual: 1.11% and 1.75%. Overall, the two observers accumulated 605 errors (observer A: 234 errors; observer B: 371 errors) between replicate assessments. Based on the summed frequencies of the two observers, the ordered list of problematic indicators was as follows: 29 errors in metacarpal (MET III-4: proximo-medial projection of the epiphysis of metacarpal III), 28 errors in trapezoid (TPD-3: shape of the medial margin of the trapezoid), 22 errors in triquetral (TRI-2: shape of the lateral margin of the triquetral), 22 errors in trapezium (TPM-4: radiopaque line or zone within the proximal margin of the trapezium). SAs of individual tennis players assigned by observer A (x-axis) are plotted relative to those assigned by observer B (y-axis) in Figure 4.1 (panel A) which also includes the respective means and standard deviations. The mean difference between observers was not significant ($t=1.252$, $p=0.210$) while the ICC was nearly perfect (ICC= 0.995). The Bland-Altman analysis of the intra-individual differences (y-axis) relative to the mean of both observers (x-axis) is illustrated in Figure 4.1 (panel B). Only four cases exceeded the limits of agreement (lower limit of agreement: -0.457 year; upper limit of agreement: 0.552 year). Results of the cross-tabulations of maturity status classifications based on assessments of observers A and B for the total sample are summarized in Table

4.2. The overall agreement was 90%, i.e., 87 of the 97 players were classified in the same maturity category by the two observers.

Table 4.1. Descriptive statistics (means ± standard deviations) for skeletal ages of male tennis players (n=97) by time-moment (TM) separately for observer A and observer B; results of comparisons between time moments, technical errors of measurement (TEM) and coefficients of variation (%CV). Frequencies of intra-observer error by bone.

Variable		Intra-observer agreement: Observer A†		Observer B‡	
Skeletal Ages					
TM1	Mean±SD	13.18±2.47		13.07±2.39	
TM2	Mean±SD	13.18±2.45		13.09±2.40	
Paired t-test	t	0.098		0.563	
	p	0.922		0.575	
Magnitude effect	d	0.008		0.001	
	(qualitative)	(trivial)		(trivial)	
TEM		0.15		0.23	
%CV		1.11		1.75	
Error		Bone	(f)	Bone	(f)
		Trapezoid	32	Trapezium	45
		Trapezium	28	Metacarpal III	40
		Metacarpal V	21	Trapezoid	37
		Proximal phalange I	21	Radius	32
		Metacarpal III	18	Metacarpal V	31
		Proximal phalange III	18	Triquetral	24
		Triquetral	16	Proximal phalange I	19
		Radius	11	Proximal phalange III	17
		Medial phalange III	11	Metacarpal I	15
		Metacarpal I	10	Lunate	14
		Adductor sesamoid	8	Medial phalange III	12
		Scaphoid	7	Scaphoid	12
		Medial phalange V	6	Proximal phalange V	11
		Distal phalange I	6	Ulna	11
		Distal phalange III	6	Distal phalange V	8
		Proximal phalange V	4	Medial phalange V	7
		Lunate	3	Pisiform	7
		Pisiform	3	Distal phalange I	6
		Distal phalange V	2	Distal phalange III	6
		Capitate	2	Capitate	6
		Hamate	1	Hamate	6
		Ulna	0	Adductor sesamoid	5
			234		371

† observer A assigned one participant as skeletally mature and were excluded; ‡ observer B assigned four participants as skeletally mature and were excluded; TM1 (time-moment 1); TM2 (time-moment 2); SD (standard deviation); d (cohen d-value); p (significance level); TEM (technical error of measurement); %CV (coefficient of variation); f (absolute frequencies).

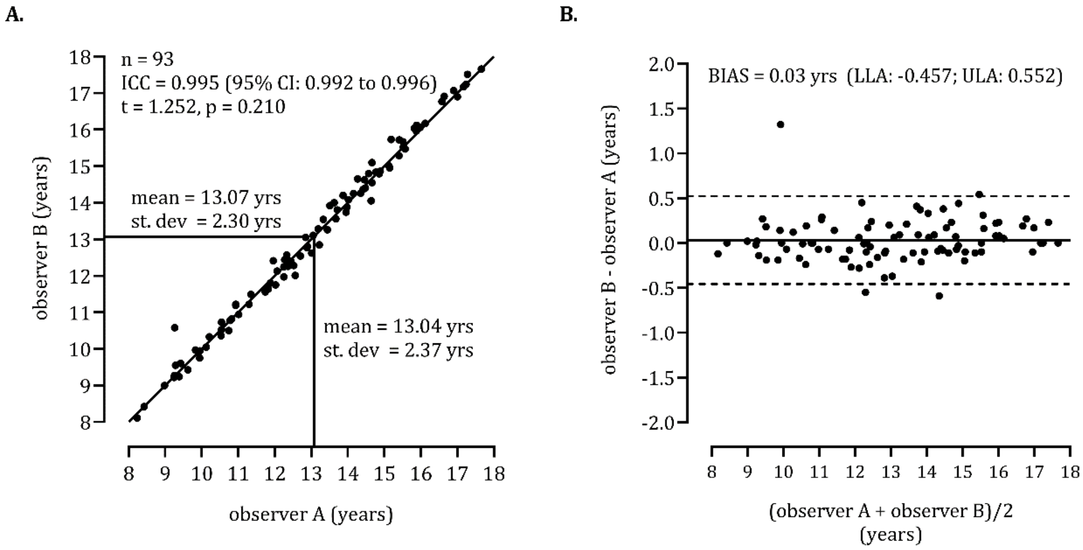


Figure 4.1. SAs of individual tennis players assigned by observer A (x-axis) are plotted relative to those assigned by observer B (y-axis) with the respective means and standard deviations (panel A); Bland-Altman analysis of the intra-individual differences (y-axis) relative to the mean of both observers (x-axis) (panel B).

Table 4.2. Cross-tabulations of maturity status classifications based on assessments of observers A and B among male adolescent tennis players for the total sample (n=97).

Maturity Status (Observer A)	Maturity Status (Observer B)				Total
	Late	Average	Early	Mature	
Late	**10**	2	0	0	12
Average	1	**45**	1	2	49
Early	0	3	**31**	1	35
Mature	0	0	0	**1**	1
Total	11	50	32	4	97

Bold values indicate the same maturity status classification with each method of SA assessment.

4.5. Discussion

SA age is generally accepted as the best method for the estimation of biological maturity status (Malina et al., 2004). The present study evaluated intra-observer the reproducibility of SA assessments among 97 male adolescent tennis players 8.7 through 16.8 years of age. Although both observers were well-experienced in the Fels SA assessment protocol, 234 errors were noted between assessments of observer A and 371 between assessments of observer B. By inference, error is part of SA assessment process. Nevertheless, the intra-observer variability had a trivial effect on intra-individual differences in mean SAs. The error associated with each individual observer was negligible in terms of the %CV. Among the 22 bones assessed by the two observers, five bones – two carpals (trapezium, trapezoid), two metacarpals (III and V) and the radius accounted for 49% of the differences between observations. Nevertheless, the agreement between the two assessors was in the present study acceptable. The plot of the SA assessments of each assessor approached the line of identify (Figure 4.1). The differences between observers were not significant and the ICC approached nearly perfection. On the other hand, only 87 of the 97 players were classified as having the same skeletal maturity status (delayed, one time, advanced, mature).

The assessment protocols assume a universal and invariant sequence of qualitative changes in each of the bones comprising the hand-wrist complex (Roche et al., 1988). The rate at which bones progress from the cartilage model to the mature state varies among bones and also within individuals. More importantly, the progress in the maturation of the skeleton can be monitored with standardized radiographs of the hand and wrist, which is comprised of two types of bones: long (distal radius and ulna, metacarpals, and phalanges) and round (carpals, adductor sesamoid). Corresponding protocols have been developed for other regions of the body - knee, foot and ankle, as summarized by Roche et al. (Roche et al., 1988). In addition, maturity status based on SA assessment spans approximately the first two decades of life, while indicators of sexual maturity status span the pubertal years.

Among 250 girls and 339 boys aged 2-15 years, intra-individual variability of SAs obtained from hand-wrist and knee radiographs was examined (Aicardi et al., 2000). Knee SA was derived using Roche-Wainer-Thissen (RWT) protocol. Results evidenced SA closer to CA while using the knee method rather than the hand-wrist protocols. In fact, among late maturing individuals, GP, TW 20-bone, TW radius, ulna, short-bones (RUS), and FELS SAs tended to be lower than RWT knee estimates. Conversely, among early maturing youth, the hand-wrist estimates tended to be higher than RWT knee bone ages. Previously, the same research group (Vignolo et al., 1999) addressed the accuracy and precision of FELS SA assessments relative to GP and TW SAs in a sample of Italian children and adolescents 1 through 17 years (171 males, 156 females). Two observers rated the radiographs and one of them re-assessed the radiographs after 6 months. The gradient of inter-observer error expressed as standard errors was 0.165, 0.203 and 0.293 for, respectively, FELS, GP and TW2-20 bone SAs.

In the context of youth sports, few studies have addressed the reproducibility of FELS SAs in small samples. Among 15 American football players aged 9-14 years, intra-observer mean difference was -0.02 (Malina, Dompier, et al., 2007). Among male adolescent soccer players, replicate assessments of Fels SA in 15 players indicated a mean difference of 0.08 years (Malina et al., 2000). Finally, repeated assessments of radiographs of 10 soccer players (25% of the sample) by the same observer indicated small differences for Fels SA (0.01 year) and TW3 RUS SA (0.04 year) with corresponding TEM about 0.04 year and 0.06 year, respectively for each protocol (Malina, Chamorro, et al., 2007). In the current study, the mean differences between assessments were -0.01 year and 0.02 year, respectively for observer A and observer B corresponding to TEMs as follows: 0.15 year and 0.23 year. Finally, among U12 and U14 soccer players, assessments of Fels SA for 18 players by two observers indicated a mean difference of 0.03 (Malina et al., 2012) which corresponds to bias obtained by the Bland-Altman in Figure 4.1.

In the current study, the same protocol used by two observers, and 10.3% of the players were misclassified in the skeletal maturity groups. The present study demonstrates that error is also implicit in SA determination and impacts the distribution into skeletal maturity groups. The classification criteria used in the present study is consistent with other studies of the general population of youth (Malina, Katzmarzyk, & Beunen, 1999) and of athletes (Coelho-e-Silva et al., 2010, 2022; Myburgh et al., 2016), although a narrower span to define early and late maturity has been reported (H. Kemper, Verschuur, & Ritmeester, 1986). The preceding band is within the range of standard errors of assessment.

The present study has limitations that should be recognized. The Fels method is characterized by age and sex-specific bone indicators (Roche et al., 1988) and consequently the findings in the present study about the most critical bones should not be generalized. By inference, future studies aimed to examine Fels protocol should include younger ages and also samples from the general population with equivalent distribution of skeletal maturity status (delayed, average, advanced) at all ages across adolescence. The current findings are limited to 97 male tennis players aged 8.7-16.8 years. To determine the influence of training frequency in bone inter-arm asymmetry in youth tennis, 24 male players aged 10.6 years were grouped depending on the number of weekly sessions (Sanchis-Moysi, Dorado, Olmedillas, Serrano-Sanchez, & Calbet, 2010). It was concluded that the asymmetry of bone mineral content and lean mass in dominant arm depends on the number of weekly hours devoted to youth tennis. The literature suggests that the effect of sports practice has repercussions on bone mineral content, and this is totally independent of skeletal maturation. One study compared asymmetry of upper arms among 24 tennis players of both sexes across circumpubertal years (7-13 years) and concluded that training time was the best predictor of inter-arm differences (Palaiothodorou, Antoniou, & Vagenas, 2020). It would have been more adequate to compare players of contrasting stages of sexual maturation within the same age, separately for males and females. Moreover, the study assessed biological maturation as “Tanner scale” and it was not informed if the evaluation followed pubic hair development or genitalia criteria. Future research needs to examine intra-observer reproducibility and inter-observer agreement in the determination of SA among females and test the impact of data quality in the distribution of skeletal maturity categories by age groups in the general population and perhaps among female athletes. Although the concordance between observers regarding the classification by skeletal maturity categories was acceptable in the present study, the sample is limited to youth tennis. Selective practices associated with different sports likely influence the distribution of players by maturity status. For example, late maturing females tend to persist in artistic gymnastics, while early maturing boys tend to persist in ice hockey, soccer and swimming.

Biological maturation is implicit in models of talent identification, selection, and development. Estimates of age at peak height velocity is central to the long-term athlete development (LTAD) model. According to LTAD, the identification of early, average, and late maturing participants in youth sports is essential for optimal trainability and readiness, that is, selection of appropriate training methodologies and organization of competitions without increasing risk of injuries (Balyi, Cardinal, Higgs, Norris, & Way, 2005). Among male adolescent soccer players, it was evident that non-invasive and invasive indicators of biological maturation were only moderately correlated (Malina et al., 2012). Considering the limitations of sexual maturation across the two decades of life and recognizing limitations of predicted age at peak height velocity from anthropometry (Kozieł & Malina, 2018; Malina et al., 2016; Malina & Kozieł, 2014a), a valid and reliable assessment of biological maturation is given by skeletal age (Malina et al., 2004).

Error associated to the observer seems to be modest and concentrated in carpals and metacarpals. The preceding information is relevant for training potential examiners. An arbitrary number of 100 observations is recommended as a training session before SA assessments, preferably using radiographs of similar age groups since bone-specific indicators depends on sex and CA. Inter-observer data quality is recommended for large studies involving more than one examiner. Finally, for the purpose of age verification in youth sports, decision-makers need to be aware of the error due to examiners and method selected to assess SA.

4.6. Conclusions

The present study confirms the Fels method as an objective and reliable method to determine SA. The protocol requires experienced observer and although this requisite, variance in longitudinal and cross-sectional studies tend to include a small portion of error due to the observer. In general, the impact of agreement within a single observer or between observers is negligible on SA and consequently does not substantially affect the classification of youth sport participant as delayed, average, or advanced, particularly when maturity categories consider a ± 1.0-year band between CA and SA.

4.7. References

Aicardi, G., Vignolo, M., Milani, S., Naselli, A., Magliano, P., & Garzia, P. (2000). Assessment of skeletal maturity of the hand-wrist and knee: A comparison among methods. American Journal of Human Biology, 12(5), 610–615.

Balyi, I., Cardinal, C., Higgs, C., Norris, S., & Way, R. (2005). Canadian Sport for Life: Long-term athlete development resource paper V2. Vancouver, BC: Canadian Sport Centres.

Booz, C., Wichmann, J. L., Boettger, S., Al Kamali, A., Martin, S. S., Lenga, L., … Kaltenbach, B. (2019). Evaluation of a Computer-Aided Diagnosis System for Automated Bone Age Assessment in Comparison to the Greulich-Pyle Atlas Method: A Multireader Study. Journal of Computer Assisted Tomography, 43(1), 39–45.

Coelho-e-Silva, M. J., Figueiredo, A. J., Simões, F., Seabra, A., Natal, A., Vaeyens, R., … Malina, R. M. (2010). Discrimination of U-14 soccer players by level and position. International Journal of Sports Medicine, 31(11), 790–796.

Coelho-e-Silva, M. J., Konarski, J. M., Krzykała, M., Galas, S., Beata, P., Żurek, P., … Malina, R. M. (2022). Growth and maturity status of young male table tennis players. Research in Sports Medicine, 30(1), 61–79.

Cohen, J. (1988). Statistical power analysis for behavioral sciences (2nd ed.). Mahaw (NJ): Lawrence Erlbaum Associates.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009). Youth soccer players, 11-14 years: maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Grendstad, H., Nilsen, A., Rygh, C. B., Hafstad, A., Kristoffersen, M., Iversen, V. V., … Gundersen, H. (2020). Physical capacity, not skeletal maturity, distinguishes competitive levels in male Norwegian U14 soccer players. Scandinavian Journal of Medicine & Science in Sports, 30(2), 254–263.

Greulich, W., & Pyle, S. (1959). Radiographic atlas of skeletal development of the hand and wrist (2nd ed.). Stanford, CA: Stanford University Press.

Harriss, D. J., Macsween, A., & Atkinson, G. (2019). Ethical Standards in Sport and Exercise Science Research: 2020 Update. International Journal of Sports Medicine, 40(13), 813–817.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. Medicine and Science in Sports and Exercise, 41(1), 3–12.

Kemper, H., Verschuur, R., & Ritmeester, J. (1986). Maximal aerobic power in earlyand late-maturing teenagers. In J. Rutenfranz, R. Mocellin, & F. Kliimt (Eds.), Children and exercise XII (pp. 213–225). Champaign, IL: Human Kinetics.

Kozieł, S. M., & Malina, R. M. (2018). Modified Maturity Offset Prediction Equations: Validation in Independent Longitudinal Samples of Boys and Girls. Sports Medicine, 48(1), 221–236.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Chamorro, M., Serratosa, L., & Morate, F. (2007). TW3 and Fels skeletal ages in elite youth soccer players. Annals of Human Biology, 34(2), 265–272.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of maturity offset in the fels longitudinal study. Pediatric Exercise Science, 28(3), 439–455.

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Carling, C., & Beunen, G. P. (2012). Interrelationships among invasive and non- invasive indicators of biological maturation in adolescent male soccer players. Journal of Sports Science, 30(15), 1705–1717.

Malina, R. M., Dompier, T. P., Powell, J. W., Barron, M. J., & Moore, M. T. (2007). Validation of a noninvasive maturity estimate relative to skeletal age in youth football players. Clinical Journal of Sport Medicine, 17(5), 362–368.

Malina, R. M., Katzmarzyk, P. T., & Beunen, G. P. (1999). Relation between birth weight at term and growth rate, skeletal age, and cortical bone at 6-11 years. American Journal of Human Biology, 11(4), 505–511.

Malina, R. M., & Kozieł, S. M. (2014). Validation of maturity offset in a longitudinal sample of Polish boys. Journal of Sports Sciences, 32(5), 424–437.

Malina, R. M., Peña Reyes, M., Eisenmann, J., Horta, L., Rodrigues, J., & Miller, R. (2000). Height, mass and skeletal maturity of elite Portuguese soccer players aged 11–16 years. Journal of Sports Sciences, 18, 685–693.

Martinho, D. V., Coelho-e-Silva, M. J., Valente-dos-Santos, J., Minderico, C., Oliveira, T. G., Rodrigues, I., … Malina, R. M. (2022). Assessment of skeletal age in youth female soccer players: Agreement between Greulich-Pyle and Fels protocols. American Journal of Human Biology, 34(1), 1–11.

Myburgh, G. K., Cumming, S. P., Coelho-e-Silva, M. J., Cooke, K., & Malina, R. M. (2016). Growth and maturity status of elite British junior tennis players. Journal of Sports Sciences, 34(20), 1957–1964.

Palaiothodorou, D., Antoniou, T., & Vagenas, G. (2020). Bone asymmetries in the limbs of children tennis players: testing the combined effects of age, sex, training time, and maturity status. Journal of Sports Sciences, 38(20), 2298–2306.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Sanchis-Moysi, J., Dorado, C., Olmedillas, H., Serrano-Sanchez, J. A., & Calbet, J. A. L. (2010). Bone and lean mass inter-arm asymmetries in young male tennis players depend on training frequency. European Journal of Applied Physiology, 110(1), 83–90.

Sousa-e-Silva, P., Coelho-e-Silva, M. J., Seabra, A., Costa, D. C., Martinho, D. V., Duarte, J. P., … Malina, R. M. (2022). Skeletal age assessed by TW2 using 20-bone, carpal and RUS score systems: Intra-observer and inter-observer agreement among male pubertal soccer players. PLoS ONE, 17(8 August), 1–15.

Tanner, J. M., Healy, M., Goldstein, H., & Cameron, N. (2001). Assessment of Skeletal Maturity and Prediction of Adult Height (TW3) Method (3rd ed.). Saunders Ltd.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1975). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London: Academic Press.

Tanner, J. M., Whitehouse, R. J., & Healy, M. J. R. (1962). A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2,600 healthy British children. Part II: The Scoring System. Paris: International Children's Centre.

Vignolo, M., Naselli, A., Magliano, P., Di Battista, E., Aicardi, M., & Aicardi, G. (1999). Use of the new US90 standards for TW-RUS skeletal maturity scores in youths from the Italian population. Hormone Research, 51(4), 168–172.

Chapter V

Study 3

Concurrent assessment of skeletal age among male adolescents of different sports.

5. Concurrent assessment of skeletal age among male adolescents of different sports.

5.1. Abstract

Introduction: Skeletal age (SA) is perhaps the best maturity indicator to assess biological maturation, since it is the only method that spans infancy through the young adulthood. Three methods of assessment are commonly used: Greulich-Pyle (GP), Tanner-Whitehouse version 2 (TW2), and FELS. Present study aimed to evaluate the agreement among concurrent methods to assess SA.

Methods: Sample was composed by young athletes aged 11.0-16.9 years from several different sports (n=1778). SA was assessed by the three most common used methods: GP, TW2 RUS, and FELS. Body mass and stature were obtained and plotted relative to U.S. reference data, with the process being repeated considering SA assessed by the three concurrent methods.

Results: Differences between SA and chronological age (CA) (SA minus CA) increase positively and gradually in the three methods, from the age group 11.0–12.9 years (0.1, 0.2, and 0.9 years, for GP, FELS, and TW2, respectively) until 14.0-14.9 years (0.5, 1.3, 1.4 years). Regarding to the maturity status classification (delayed, on-time, advanced) FELS and TW2-RUS showed a moderate agreement of 66% (kappa= 0.467).

Discussion: Results showed a considerable variation between methods. SA increased on the three protocols, relatively to CA, being more evident at 14.0 to 14.9 years. 15.4% of participants were already mature at 15 years of age but the respective percentage for TW2-RUS were higher, 39.8%. When mean values of stature and body mass were plotted against the CDC references considering SA, mean values were close to the median for GP at 11-12 and 13-14 age groups, but below the median when FELS and TW2 were considered.

Keywords: Greulich-Pyle; TW2; FELS; Skeletal age; Biological maturation.

5.2. Introduction

Biological maturation is defined as the process from immaturity towards adult state. Moreover, maturation can be assessed essentially in two different aspects: maturity status (level of maturation at the chronological age of observation) and maturity timing (chronological age at which a specific maturational event occurs). The two are related but are not equivalent (Malina et al., 2004). Indicators of maturation vary according to the biological system. For example, skeletal maturation is the process from the cartilaginous tissue of the fetus, until the fully ossified skeleton of the adult (Acheson, 1954). In fact, skeletal age (SA) is perhaps the best indicator since it spans infancy through adolescence (Malina et al., 2015).

The ossification of the bones in the hand-wrist were originally described by Todd (1937) as an atlas. It was subsequently revised (Greulich & Pyle, 1959) and is still being used today. The protocol involves 31 standard plates from newborn to 19.0 years old in males, with intervals of about one-year, although in some ages the length was reduced to six months. Meantime, a new system for assessing skeletal age published in 1962 as Tanner-Whitehouse method (TW) (Tanner et al., 1962). It requires the assessment of 20 bones of the hand-wrist. Each bone is evaluated following verbal stages starting from letter A (no center of ossification) to letter I or J (complete epiphyseal-diaphyseal fusion or adult form). The stages correspond to a certain score and the sum of all scores is converted into a SA. The preceding corresponds to TW – version 1 that suffered a revision in 1975 (Tanner et al., 1975). Briefly, although criteria were not modified, the final stages of the radius (stage J), ulna (stage I), and five of the seven carpals (stage I) were excluded. In addition, scores corresponding to stages were modified and three alternative score systems emerged: 20-bone, CARPAL, and RUS (radius, ulna, and short bones). TW2 is still used (Malina et al., 2018). More recently, second revision was produced (TW3). After combining samples from the previous versions with others from Belgium, Spain, Japan, Italy, Argentina and USA, it was decided to retain RUS and CARPAL systems and the 20-bone SA was eliminated (Tanner et al., 2001). The number of stages and respective verbal description were kept with modification concentrated in the modification of the conversions of RUS scores and CARPAL scores into corresponding SAs. The preceding version, TW3, considered mature state at 16.5 years for boys. Finally, FELS protocol (Roche et al., 1988) was obtained from the Fels longitudinal study that assessed a total of 13.823 films from 677 children. They were evaluated from one month to 22 years with about 6-month interval. The technique includes 98 grades as maturity indicators and 13 metrics indicators. The individual SA is derived by using a software.

SAs assessed obtained from the different methods are not supposed to equivalent. Actually, SA represents the chronological age (CA) at which a specific level of maturity was attained by the reference sample. SA minus CA is often calculated for the purpose of classification participants in youth sports by maturity status (Coelho-e-Silva et al., 2022; Figueiredo, Coelho-e-Silva, Sarmento, Moya, & Malina, 2019; Figueiredo, Gonçalves, Coelho-e-Silva, & Malina, 2009b): late, on time, advanced. Individuals classified as mature should not be assigned to any SA (Malina, 2011). Among youth soccer players aged 10-12 years, late and maturing participants were equally represented and, with increasing age, early and skeletally mature players tended to be overrepresented (Malina et al., 2000). The trend of maturity-associated selection of adolescent athletes was also discussed in youth tennis (Myburgh et al., 2016) and youth table tennis (Coelho-e-Silva et al., 2022).

Agreement between protocols for determination of SA are still scarce in youth sports literature or based on limited samples, that is, mainly in soccer. Malina et al. (2007) assessed 40 elite Spanish players aged 12.5-16.1 years using FELS method and TW3. Meantime, a sample of 1831 male soccer players aged 10.93-17.94 years from eight countries were simultaneously assessed by TW2 and TW3 (Malina et al., 2018). More recently the agreement between FELS and Greulich-Pyle (GP) was tested among 441 Portuguese female soccer players (Martinho et al., 2022). The concept of biological maturation is implicit to talent identification and long-term athletic development (Jason Moran et al., 2020; Murtagh et al., 2018). By inference, differences between concurrent methods used to assess SA have obviously implications in the distribution of players by maturity status.

Taking into account the preceding, the present study was aimed to examine the intra-individual variation associated to different protocols that allow the determination of SA among a large sample of male adolescents from several sports who were assessed using GP, TW2 RUS, and FELS. In parallel, this study classifies each individual participant as mature, advanced, on time or delayed according to each of the concurrent SAs. Finally, mean values of stature for age and body mass for age reporting specific competitive age groups (11-12 years, 13-14 years, 15-16) is plotted using US reference data considering body size descriptors for CA and, in parallel, body size descriptors for concurrent SAs.

5.3. Methods

Procedures and sample

The current study examined a database including male youth athletes from different sports (basketball, handball, roller hockey, soccer, swimming, table tennis, tennis). Participants aged 11.0 and 17.0 years and combined chronological age, stature, body mass and a x-ray of the wrist and left hand to allow the determination of skeletal age The sample included 1778 male athletes grouped by age: 11.0-11.9 years, 12.0-12.9 years, 13.0-13.9 years, 14.0-14.9 years, 15.0-15.9 years, 16.0-16.9 years.

Anthropometry

Stature and body mass were assessed according to the procedures proposed by Lohman, Roche, & Martorell (1988). Stature was measured to the nearest 0.1cm, using an Harpenden portable stadiometer (model 98.603, Holtain Ltd, Crosswell, Crymych, UK). Body mass was measured to the nearest 0.1kg, using a digital scale (SECA 770, Hanover, MD, USA). The two anthropometric variables were plotted relative to U.S. reference data (Kuczmarski et al., 2002). and the process was repeated considering skeletal age assessed by the three concurrent methods: GP, TW2 RUS, and FELS.

Skeletal Age

Skeletal age was assessed by three methodologies: Greulich-Pyle (GP) (Greulich & Pyle, 1959), Tanner-Whitehouse II RUS system (TW2 RUS) (Tanner et al., 1975), and FELS (Roche et al., 1988). Greulich-Pyle consists in matching the individual bone under assessment, with the most closely plate presented in the atlas. The procedure is repeated for the 30 bones of the hand wrist: radius, ulna, capitate, hamate, triquetral, lunate, scaphoid, trapezium, trapezoid, pisiform, metacarpals from 1st to 5th digit, proximal phalanges from 1st to 5th digit, middle phalanges from 2nd to 4th digit, and distal phalanges from 1st to 5th digit, adductor sesamoid. The assigned SA for the subject by GP was the mean of the SA's of all rated bones. TW2 RUS, considered the assessment of 13 long bones of the hand-wrist: Radius, ulna, metacarpals and phalanges (omitting the 2nd and 4th digit). The rationale for the assessment is common for the 13 bones: appearance of the secondary centers of ossification (epiphyses) (stage A) to the epiphyseal-diaphyseal fusion (stage I), except for the ulna where the last stage is the beginning of ossification (stage H). The assessment is based on the written criteria attained to each stage, with the images serving only as a complementary support to the assessment. A specific point score was assigned to each stage and for each individual bone. The scores for the 13 bones were summed and a maturity score was given, ranging from 0 (without any secondary center of ossification) to 1000 points (mature, adult state). A sex-specific table for boys was consulted on the book, with the conversions of the total score into an individual SA. 999 points corresponds to 18.1 years, where for 1000 points no SA is assigned (the subject is considered ADULT: mature state). Fels protocol requires the evaluation of 22 bones: radius, ulna, eight carpals (capitate, hamate, triquetral, lunate, scaphoid, trapezium, trapezoid, pisiform), metacarpals I, III and V, proximal and distal phalanges of I, III and V, middle phalanges of the third and fifth digits, adductor sesamoid. Each bone considers one to eight maturity indicators, and each indicator range from two to five grades. A total of 111 possible maturity indicators are considered in the protocol, 13 of them metric indicators: epiphyseal/diaphyseal width ratio of the long bones. Written criteria are provided and images in the book could be used as a complementary support. The maturity indicators necessary to evaluate, considering the age and sex of the subject, were assigned and the grades were entered into a computer software (Felsw 1.0), that provided the SA of the subject. At 18.0 years of SA by Fels, the subject is considered mature.

Analyses

After determination of descriptive statistics for stature and body mass, SA, CA and SA minus CA for each of the methods (GP, TW2-RUS, and FELS) players were classified into four groups: average (on time), SA ± 1.0 year from CA; late (delayed), SA younger than CA by >1.0 year; early (advanced), SA older than CA by <1.0 year; skeletally mature (only applicable to TW2 and FELS). The band of ± 1.0 year approximated standard deviations of SA within one-year CA groups of youth males aged 11-17 years (Malina, 2011; Malina et al., 2018). Concordance of maturity status classification based on the three concurrent assessed methods was estimated with Cohen kappa coefficient (Viera & Garrett, 2005) in the total sample. The six age groups were reduced to the following: 11.0-12.9 years, 13.0-14.9 years, 15.0-16.9years. Finally, stature and body mass were plotted relative to U.S. reference data (Kuczmarski et al., 2002) and the illustrations were repeated considering the skeletal age for the three SA assessments. All analyses were performed with the use of IBM SPSS Statistics, version 26 (SPSS Inc., IBM Company, Armonk, NY, USA) and GraphPad Prism (version 8 for Windows, GraphPad Software, San Diego, California, USA, www.graphpad.com).

5.4. Results

Descriptive statistics for chronological age and body size descriptors are presented in Table 5.1 by single age groups. Table 5.2 summarizes descriptive statistics for SA assessed by FELS method, calculated differences between CA and SA and, finally, frequencies of skeletal maturity groups by single age groups. Mature players were observed from 14.0-14.9 years and upwards. Corresponding analyses adopting TW2-RUS and GP correspond to Tables 5.3-5.4, respectively. Of interest, individuals classified as skeletally mature appeared among 13-year-old participants and older age groups. The GP protocol did not diagnose any participant characterized by the mature state on all 30 bones and, consequently, none were classified as fully mature (see Table 5.4). A cross-tabulation between the frequencies of skeletal maturity status derived by the concurrent methods of SA assessment was calculated in Table 5.5. Briefly, FELS and TW2-RUS showed a moderate agreement of 66% (kappa= 0.467). Statistics were slightly more modest while combining GP and FELS (63% agreement, kappa= 0.376) or GP and TW2-RUS (63% agreement, kappa= 0.373).

Finally, mean values of stature and body mass considering the three age groups (11-12 years; 13-14 years; 15-16 years) were plotted against the CDC references as recommended, that is, according to CA. The illustrations (Figure 5.1) were separately repeated using SAs derived from the above-mentioned methods of SA determination. In fact, mean stature is consistently close to the median when X-axis corresponds to CA values. The trend is similar when using SA-GP in the X-axis. However, the mean values of each groups plotted below the median at younger categories (11-12 years, 13-14 years) when using FELS and TW2-RUS. Regarding the graphics of body mass for CA, soccer players tended to plot at percentile 50% among groups aged 11-12 years and 13-14 years while older players (15-16 years) approached the percentile 75%. The pattern was similar adopting SA-GP in x-Axis (panel H). Consistently to stature for SA, mean values of body mass for SA for participants aged 11-12 years and 13-14 years, when adopting TW2-RUS and FELS in X-axis plotted below the percentile 50% and at the median among boys aged 15-16 years.

Table 5.1. Descriptive statistics for chronological age, stature, and body mass by age group among adolescents participating in competitive youth sports (n=1778).

Age groups	n	CA (yrs)	Stature (cm)			Body mass (kg)		
			(range)	mean	SD	(range)	mean	SD
11.0-11.9 years	213	11.5±0.3	(130.1 – 172.4)	146.5	7.6	(23.6 – 65.0)	38.3	7.1
12.0-12.9 years	308	12.5±0.3	(133.0 – 182.8)	152.8	8.6	(23.0 - 83.1)	43.4	9.3
13.0-13.9 years	387	13.5±0.3	(134.8 – 187.0)	161.2	9.0	(23.0 – 92.9)	49.8	9.7
14.0-14,9 years	423	14.5±0.3	(143.0 – 196.7)	169.5	8.9	(24.0 – 94.5)	58.5	10.8
15.0-15.9 years	357	15.4±0.3	(153.0 – 208.6)	174.3	8.4	(43.8 – 127.3)	65.5	10.2
16.0-16.9 years	090	16.3±0.3	(150.2 – 194.5)	174.2	6.8	(41.6 – 90.3)	67.2	8.0

CA (chronological age); SD (standard deviation).

Table 5.2. Descriptive statistics for chronological age, skeletal age (FELS method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778).

Age groups (years)	N	not skeletlly mature						skeletally mature			Maturity status		
		n	CA (years)	SA (years)		SA - CA (years)		n	CA (years)		D	OT	A
				Range	M±SD	Range	M±SD		Range	M±SD			
11.0-11.9	214	214	11.5±0.3	7.0-17.1	11.7±1.5	-4.4-6.0	0.2±1.5	0	-	-	38	114	62
12.0-12.9	308	308	12.5±0.3	8.3-17.9	12.9±1.5	-4.2-5.2	0.5±1.4	0	-	-	46	147	115
13.0-13.9	389	389	13.5±0.3	10.2-	14.4±1.3	-3.5-4.1	0.9±1.3	0	-	-	31	178	180
14.0-14.9	423	409	14.5±0.3	10.3-	15.7±1.3	-4.4-3.7	1.3±1.3	14	14.1-14.9	14.6±0.3	13	168	228
15.0-15.9	357	302	15.4±0.3	13.3-	16.6±1.1	-2.1-3.0	1.2±1.1	55	15.0-15.9	15.6±0.3	6	105	191
16.0-16.9	091	069	16.3±0.3	14.6-	17.1±0.8	-1.8-2.0	0.8±0.8	22	16.0-16.9	16.4±0.3	3	33	33

M (mean); SD (standard deviation); CA (chronological age); SA (skeletal age); D (delayed); OT (on time); A (advanced).

Table 5.3. Descriptive statistics for chronological age, skeletal age (TW2 RUS method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778).

Age groups (years)	N	not skeletlly mature						skeletally mature			Maturity status (n)		
		n	CA (years)	SA (years)		SA - CA (years)		n	CA (years)		D	OT	A
				Range	M±SD	Range	M±SD		Range	M±SD			
11.0-11.9	214	214	11.5±0.3	7.5-16.2	12.4±1.5	-3.8-4.3	0.9±1.5	0	-	-	27	90	97
12.0-12.9	308	307	12.5±0.3	9.3-16.6	13.5±1.3	-3.0-4.0	1.1±1.3	1	-	-	24	106	177
13.0-13.9	389	385	13.5±0.3	11.2-17.7	14.8±1.2	-2.4-4.7	1.3±1.2	4	13.4-13.9	13.6±0.2	4	108	259
14.0-14.9	423	358	14.4±0.3	10.4-17.7	15.8±1.1	-4.3-3.7	1.4±1.1	65	14.1-14.9	14.6±0.3	10	123	225
15.0-15.9	357	214	15.4±0.3	13.7-17.7	16.4±0.8	-1.7-2.6	1.0±0.8	142	15.0-15.9	15.5±0.3	2	107	106
16.0-16.9	091	037	16.3±0.3	15.1-17.8	17.0±0.7	-1.2-1.6	0.7±0.7	54	16.0-16.9	16.3±0.3	1	23	13

M (mean); SD (standard deviation); CA (chronological age); SA (skeletal age); D (delayed); OT (on time); A (advanced).

Table 5.4. Descriptive statistics for chronological age, skeletal age (Greulich-Pyle method), differences between CA and SA in addition to absolute frequencies by skeletal maturity groups according to age groups among male adolescent participating in competitive sports (n=1778).

Age groups (years)	N	not skeletlly mature						Maturity status (n)		
		n	CA (years)	SA (years)		SA - CA (years)				
				Range	M±SD	Range	M±SD	D	OT	A
11.0-11.9	214	214	11.5±0.3	6.2-14.4	11.6±1.2	-4.8-3.1	0.1±1.2	29	142	43
12.0-12.9	308	308	12.5±0.3	7.8-15.4	12.7±1.1	-4.5-2.9	0.2±1.1	32	211	65
13.0-13.9	389	389	13.5±0.3	9.4-16.2	13.9±1.0	-3.6-2.9	0.4±1.0	25	266	98
14.0-14.9	423	423	14.5±0.3	9.7-18.5	15.0±1.0	-5.0-3.7	0.5±0.9	17	259	147
15.0-15.9	357	357	15.4±0.3	13.1-16.2	15.7±0.6	-2.3-1.2	0.2±0.6	17	328	12
16.0-16.9	091	091	16.3±0.3	14.3-16.2	16.0±0.4	-2.1-3.0	-0.3±0.5	6	85	0

M (mean); SD (standard deviation); CA (chronological age); SA (skeletal age); D (delayed); OT (on time); A (advanced).

Table 5.5. Cross-tabulation between frequencies of skeletal maturity categories derived from concurrent methods used to estimate skeletal age and statistics of agreement between protocols.

		FELS				TW2 (RUS)			
		D	OT	A	M	D	OT	A	M
TW2 (RUS)	D	70	12	0	0				
	OT	64	435	54	1				
	A	3	280	587	6				
	M	0	15	167	84				
	% agreement	66%							
	Kappa (p)	0.467							
GP	D	88	37	1	0	76	48	1	1
	OT	48	693	474	74	14	753	395	127
	A	1	12	333	17	0	11	296	56
	M	0	0	0	0	0	0	0	0
	% agreement	63%				63%			
	Kappa (p)	0.376				0.373			

TW2 (Tanner-Whitehouse method, version 2); GP (Greulich-Pyle method); D (delayed); OT (on time); A (advanced); M (mature); p (significance value); Lin CCC (Lin concordance correlation coefficient); 95% CI (95% confidence interval)

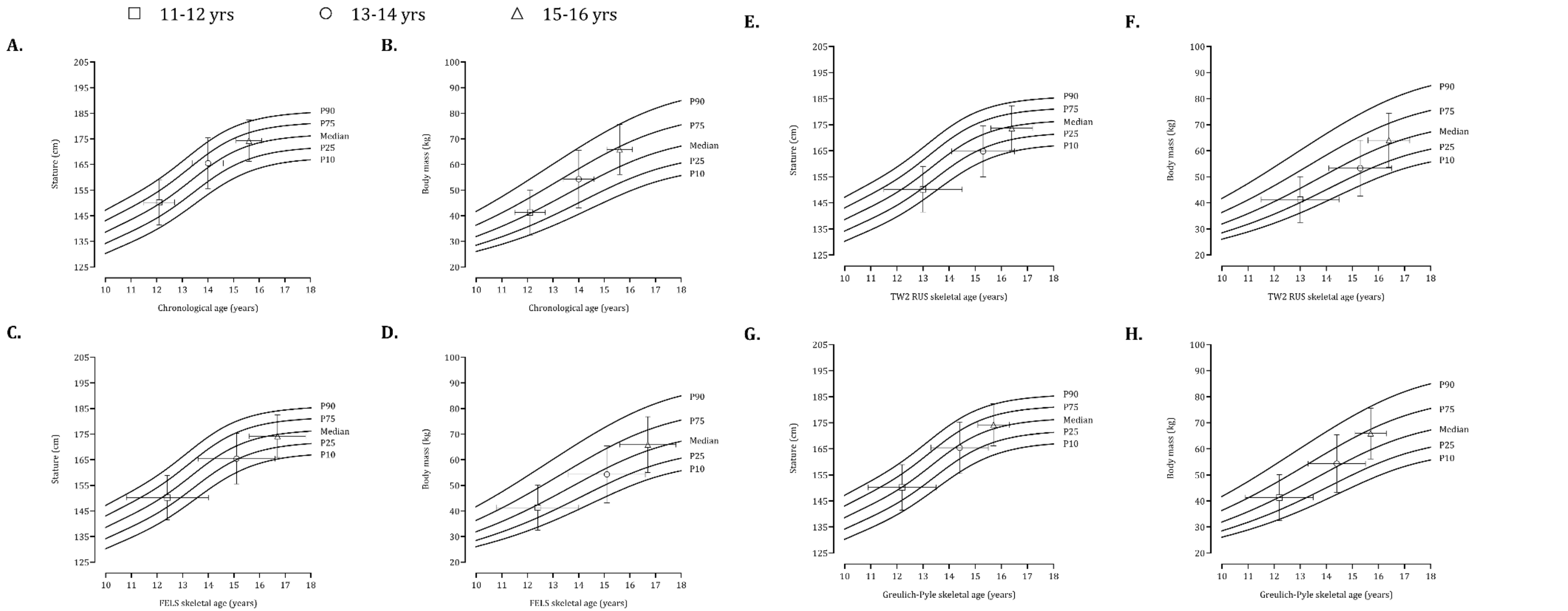


Figure 5.1. Mean stature and mean body mass of the competitive age-groups (11-12 years, 13-14 years and 15-16 years) plotted relative to U.S. percentiles (n = 1778), considering CA (Panels A-B), and SA (Panels C-H).

5.5. Discussion

The present study assessed the SAs through the three most common methods used in youth sports. Greulich-Pyle may be viewed as an inspectional method while TW2-RUS is based on scores converted into SA. Finally, FELS requires an informatic application to calculate SA and associated error from grades and transverse measurements. Nevertheless, SA values derived from each of the concurrent methods would not be supposed to be exactly the same, results showed a substantial variation between methods. At younger ages (11-11.9 years), mean SA_{GP} and SA_{FELS} plotted close to mean CA (0.1 and 0.2 years, respectively). In contrast, mean SA_{TW2} presented a 0.9-year difference to mean CA. With the advancing age groups, SA naturally increased on the three protocols, particularly until 14.0-to 14.9-year-olds, where the difference is more evident. Subsequently, when using FELS protocol 15.4% of participants were already mature at 15 years of age and the percentage increases to 24.2% among 16-year-old male adolescents of the current study. Respective percentages for TW2-RUS were 39.8% and 59.3%. In the current sample, GP method did not classify any participant as skeletally mature. The agreement between categories derived from concurrent SA protocols was moderate, that is, kappa statistics ranged 0.373 to 0.467.

A previous study (Van Lenthe, Kemper, & Van Mechelen, 1998) determined the SA of 30 boys from the *Amsterdam Growth and Health Study* according to TW2 and FELS. The age ranged 12-18 years with mean differences between SAs about 0.26 years at 12-year-olds (12.0-12.9 years) and slightly higher among 14-year-olds (0.47 years). Afterwards, intra-individual mean differences of late adolescents aged 16 years were residual, 0.05 years, in the 16-16.9-year group. The present study confirmed that the discrepancy between SA and CA is accentuated during late adolescent years.

Among a large sample of non-athletes, SA was simultaneously determined using GP, TW2, and FELS (Santoro, Marini, Fuzio, Introna, & de Donno, 2019). The study combined 204 subjects aged 4-19 years from Southern Italy and Benin (sub-Saharan Africa). No significant differences between the CA and the SA were found, independently of the assessment methods. Unfortunately, authors did not address the ethnic variation of the agreement between methods. In fact, the literature is still lacking regarding ethnic-associated variation of the agreement among methods. It is not of irrelevance to consider the limitation to generalize FELS and GP to other population besides middle-class white North American youth. The current sample was plotted against CDC percentiles (Kuczmarski et al., 2002) as summarized in Figure 5.1 (panels A-B). Younger sport participants tended to plot at the median on stature for CA and body mass for CA. With increasing age, mean values of the current sample at 13-14 years and 15-16 years tended to plot above the median, more pronounced in body mass for age which means that male adolescent athletes tended to carry more weight for height when compared to sex-and age standards. The preceding conclusions was also extracted from a study of Portuguese soccer players aged 11-16 years (Malina et al., 2000). Changes in the body size of athletes over time tend already received attention in the available literature (Malina, Figueiredo, & Coelho-e-Silva, 2017). In the early Olympic editions, 1896-1960, track and field 100m sprinters were characterized by a low BMI, with a weight values fluctuating between 56-75kg and a stature 167-186 cm. More recently, over the period 1964-2022, Olympic medalists presented larger body size: 176-195 cm, 75-94kg.

In the context of youth soccer, skeletal age (SA) tends to be advanced for chronological age (CA) in adolescent male soccer players (Coelho-e-Silva et al., 2010; Figueiredo et al., 2009a; Malina et al., 2000). Among 40 elite youth soccer players 12.5–16.1 years of age (Malina, Chamorro, et al., 2007), SA assessments with the TW3 and Fels methods were compared. Consistently with the preceding, in general, SA was in advance of CA at all age groups. Additionally, among fourteen players aged >15.0 years, two were mature with the Fels method, while 11 were already skeletally mature with the TW3 method with authors concluding that "*observations have implications for international age group competitions as well as for medico-legal circumstances that require CA verification*". Clinical staff and coaches in youth sports should be aware that the methods of assessing SA have a modest agreement and discrepancies between in biological age and CA are particularly evident after years mid-puberty.

When mean values of stature by age group (11-12 years, 13-14 years, 15-16 years) were plotted for SA, instead of CA, adopting CDC percentiles (Kuczmarski et al., 2002), the trend of progressive more weight for age with increasing values in x-axis was similar using SA_{GP} as previously noted for CA. By using both SA_{FELS} SA_{TW2RUS} in X-axis, mean values did not approach percentile 75% neither for stature not body mass (see Panels C-G in Figure 1).

5.6. References

Acheson, R. M. (1954). A method of assessing skeletal maturity from radiographs; a report from the Oxford child health survey. Journal of Anatomy, 88(4), 498–508.

Coelho-e-Silva, M. J., Figueiredo, A. J., Simões, F., Seabra, A., Natal, A., Vaeyens, R., … Malina, R. M. (2010). Discrimination of U-14 soccer players by level and position. International Journal of Sports Medicine, 31(11), 790–796.

Coelho-e-Silva, M. J., Konarski, J. M., Krzykała, M., Galas, S., Beata, P., Żurek, P., … Malina, R. M. (2022). Growth and maturity status of young male table tennis players. Research in Sports Medicine, 30(1), 61–79.

Figueiredo, A. J., Coelho-e-Silva, M. J., Sarmento, H., Moya, J., & Malina, R. M. (2019). Adolescent characteristics of youth soccer players: do they vary with playing status in young adulthood? Research in Sports Medicine (Print), 1–12.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009a). Youth soccer players, 11-14 years: maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009b). Youth soccer players, 11-14 years: Maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Greulich, W., & Pyle, S. (1959). Radiographic atlas of skeletal development of the hand and wrist (2nd ed.). Stanford, CA: Stanford University Press.

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., … Johnson, C. L. (2002). 2000 CDC Growth Charts for the United States: Methods and Development. Vital and Health Statistics, Series 11,(246).

Lohman, T. G., Roche, A. F., & Martorell, R. (1988). Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics.

Malina, R. M. (2011). Skeletal Age and Age Verification in Youth Sport. Sports Medicine, 41(11), 925–947.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Chamorro, M., Serratosa, L., & Morate, F. (2007). TW3 and Fels skeletal ages in elite youth soccer players. Annals of Human Biology, 34(2), 265–272.

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Philippaerts, R. M., Hirose, N., Peña Reyes, M. E., … Buranarugsa, R. (2018). Tanner-Whitehouse Skeletal Ages in Male Youth Soccer Players: TW2 or TW3? Sports Medicine, 48(4), 991–1008.

Malina, R. M., Figueiredo, A. J., & Coelho-e-Silva, M. J. (2017). Body Size of Male Youth Soccer Players: 1978–2015. Sports Medicine, 47(10), 1983–1992.

Malina, R. M., Peña Reyes, M., Eisenmann, J., Horta, L., Rodrigues, J., & Miller, R. (2000). Height, mass and skeletal maturity of elite Portuguese soccer players aged 11–16 years. Journal of Sports Sciences, 18, 685–693.

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho-e-Silva, M. J., & Figueiredo, A. J. (2015). Biological maturation of youth athletes: Assessment and implications. British Journal of Sports Medicine, 49(13), 852–859.

Martinho, D. V., Coelho-e-Silva, M. J., Valente-dos-Santos, J., Minderico, C., Oliveira, T. G., Rodrigues, I., … Malina, R. M. (2022). Assessment of skeletal age in youth female soccer players: Agreement between Greulich-Pyle and Fels protocols. American Journal of Human Biology, 34(1), 1–11.

Moran, J., Paxton, K., Jones, B., Granacher, U., Sandercock, G. R., Hope, E., & Ramirez-Campillo, R. (2020). Variable long-term developmental trajectories of short sprint speed and jumping height in English Premier League academy soccer players: An applied case study. Journal of Sports Sciences, 38(22), 2525–2531.

Murtagh, C. F., Brownlee, T. E., O’Boyle, A., Morgans, R., Drust, B., & Erskine, R. M. (2018). Importance of speed and power in elite youth soccer depends on maturation status. Journal of Strength and Conditioning Research, 32(2), 297–303.

Myburgh, G. K., Cumming, S. P., Coelho-e-Silva, M. J., Cooke, K., & Malina, R. M. (2016). Growth and maturity status of elite British junior tennis players. Journal of Sports Sciences, 34(20), 1957–1964.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Santoro, V., Marini, C., Fuzio, G., Introna, F., & de Donno, A. (2019). A comparison of 3 established skeletal age estimation methods in an African group from Benin and an Italian group from southern Italy. American Journal of Forensic Medicine and Pathology, 40(2), 125–128.

Tanner, J. M., Healy, M., Goldstein, H., & Cameron, N. (2001). Assessment of Skeletal Maturity and Prediction of Adult Height (TW3) Method (3rd ed.). Saunders Ltd.

Tanner, J. M., Whitehouse, R., Cameron, N., Marshall, W., Healy, M., & Goldstein, H. (1975). Assessment of skeletal maturity and the prediction of adult height (TW2 Method). London: Academic Press.

Tanner, J. M., Whitehouse, R. J., & Healy, M. J. R. (1962). A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2,600 healthy British children. Part II: The Scoring System. Paris: International Children's Centre.

Todd, T. (1937). Atlas of skeletal maturation (1st ed.). St. Louis, MO: V. Mosby Company.

Van Lenthe, F. J., Kemper, H. C. G., & Van Mechelen, W. (1998). Skeletal maturation in adolescence: A comparison between the Tanner- Whitehouse II and the Fels method. European Journal of Pediatrics, 157(10), 798–801.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Family Medicine, 37(5), 360–363.

Chapter VI

Study 4

Ages at peak height velocity in male soccer players 11-16 years: relationships with skeletal age and comparisons among longitudinal studies.

6. Ages at peak height velocity in male soccer players 11-16 years: relationships with skeletal age and comparisons among longitudinal studies.

6.1. Abstract

Background: To compare two estimates of ages at take-off (TO) and at peak height velocity (PHV) n a longitudinal sample of male soccer players and to evaluate maturity status based upon ages at PHV and skeletal age (SA), and to compare estimated ages at PHV in 13 longitudinal samples of soccer players.

Methods: Heights of 58 soccer players of European ancestry followed longitudinally across five seasons (11-16 years) were modeled with Superimposition by Translation and Rotation (SITAR) and Functional Principal Component Analysis (FPCA) to estimate ages at TO and PHV. SAs at observations 1, 3 and 5 were assessed with the Fels method. Ages at PHV in 13 longitudinal samples of soccer players (Europe 7, Japan 6) were evaluated with meta-analysis.

Results: SITAR and FPCA estimates were, respectively, 11.2±0.8 and 11.0±0.8 years for ages at TO, and 13.62±0.90 and 13.66±0.88 years for ages at PHV. An earlier age at PHV was associated with advanced skeletal maturity status. The systematic analysis indicated a north (later) - south (earlier) gradient in ages at PHV among players in Europe, which were later than ages at PHV among players in Japan.

Conclusions: Ages at TO and PHV were similar with SITAR and FPCA, and ages at PHV were most strongly correlated with SA at ~14 years. Ages at PHV showed a north-south geographic gradient in European studies and were later compared to Japanese studies.

Keywords: Adolescent spurt; Maturity timing; Youth athletes; Skeletal age.

6.2. Introduction

Acceleration in rate of growth in height in late childhood/early adolescence marks the onset or take-off (TO) of the adolescent growth spurt. The rate of growth accelerates until it reaches a peak (peak height velocity, PHV) and then decelerates until growth in height ceases in late adolescence or young adulthood. Ages at TO and at PHV are estimated from longitudinal height records. Estimates of the parameters of the adolescent spurt in earlier studies were based on graphic analysis of heights attained at a given CA and estimated velocities of growth (increments between measurements). Mathematical modeling of longitudinal height records for individuals facilitated estimation of the parameters. Estimates of age at PHV (years), PHV (cm/year) and height at PHV (cm) are provided with most models, while some provide estimates of age, velocity of growth and height at TO, and an estimate of adult height. The procedures provide a convenient means of comparing individual and/or group differences in parameters of the adolescent spurt in height (Hauspie & Chrząstek-Spruch, 1999; Marubini & Milani, 1986; Sanders et al., 2017).

Ages at PHV derived from longitudinal growth records of individual youth should not be confused with estimates based on predicted maturity offset defined as the time before PHV, and predicted age at PHV estimated as CA minus predicted maturity offset (Mirwald et al., 2002; Moore et al., 2015). Although predicted maturity offset and age at PHV are increasingly used as an indicator of maturity status in studies of youth athletes, the estimates are not equivalent with those based on longitudinal observations (Malina et al., 2004). Relative to observed maturity offset (CA at observation minus observed age at PHV) and ages at PHV in three longitudinal samples of boys followed 8 to 18 years, predicted maturity offset and ages at PHV varied with CA and body size at prediction, and both predictions were influenced by maturity status defined by observed age at PHV (Kozieł & Malina, 2018; Malina et al., 2016; Malina & Kozieł, 2014a; Malina, Kozieł, Králik, Chrzanowska, & Suder, 2021). Predicted maturity offset and ages at PHV were consistently later than observed offset and ages at PHV in early maturing boys, and generally earlier than observed offset and ages at PHV in late maturing boys.

Ages at PHV for youth athletes based on longitudinal data spanning adolescence are not extensive (Malina, 2021; Malina et al., 2015). This is in part a function of difficulties inherent in longitudinal studies *per se*, and the selectivity of sport, differential persistence and cessation of participation in a sport (drop out, injury, motivation, changing interests), changes in teams or clubs, among other considerations. Nevertheless, coaches and trainers are increasingly interested in monitoring growth rates in height and weight of youth players over relatively short intervals in an effort to evaluate growth status in an effort to individualize training and to reduce the risk of injury during the adolescent growth spurt(Cumming et al., 2017; Johnson, Cumming, Bradley, & Williams, 2022). In this context, further information on variation in the timing and intensity of growth at TO and PHV among youth athletes is important.

The purposes of this study are twofold. First, it compares two methods for estimating parameters of the adolescent growth spurt (TO, PHV) in a longitudinal sample of male youth soccer players 11-16 years of age; and evaluates maturity classification based on age at PHV, an indicator of maturity timing, relative SA, an indicator of maturity status at the time of observation. Second, it systematically compares estimated ages at PHV reported for longitudinal samples of soccer players.

6.3. Methods

Sample

Data for the present study were part of the *Coimbra Soccer Longitudinal Project*, which followed the guidelines established by the declaration of Helsinki (Harriss et al., 2019). Formal approval was obtained from the *University of Coimbra Sports Sciences and Physical Education Board* and included agreements with the Presidents of the respective soccer clubs. Written consent was obtained from parents or legal guardians of the players, and players were informed that participation was voluntary and that they could withdraw from the study at any time.

The baseline sample included 87 U13 players (11-12 years) from five clubs in the midlands of Portugal; players were classified as *infantiles* in the *Portuguese Soccer Federation*. All players except one were of European ancestry. At baseline, the sample had 1-6 years of experience in soccer (median 3 years) and participated in 3-5 training sessions (~90 minutes) and one game per week, usually on Saturday.

Heights and weights of the players, among other anthropometric dimensions, were measured within a two-week interval in December at baseline; players who persisted at the respective clubs were subsequently measured within the same two-week interval in December over the next five seasons. All measurements were taken by a single observer (MJC) at the University of Coimbra. Heights, with shoes removed, were measured to the nearest 0.1 cm using a stadiometer (Harpenden 98.603, Holtain Ltd, Croswell, UK). Weight was measured to the neared 0.1 kg using a SECA scale (model 770, Hanover, MD, US). Intra-observer technical errors of measurement were 0.27 cm for height and 0.47 kg for weight. CA at each observation was calculated as the difference between date of birth and date of a hand-wrist radiograph (see below) for observations one, three and five, and between date of birth and date of measurement for observations two and four.

Across the five years, 59 players had four or five annual height measurements. The longitudinal sample did not differ significantly from their 28 teammates at baseline: respectively, CA, 11.9±0.5 and 11.7±0.5 years; SA, 12.0±1.4 and 11.8±1.6 years; height, 144.8±6.9 and 144.3±6.5 cm; and weight, 37.6±6.0 and 38.8±7.0 kg; the distribution of players by pubic hair status also did not significantly differ.

Parameters of the Adolescent Spurt

The longitudinal height records of 58 players of European ancestry were successfully modeled with two methods to estimate parameters of the adolescent spurt: Superimposition by Translation and Rotation (SITAR) and Functional Principal Component Analysis (FPCA). The heights of one player of non-European ancestry (four observations) were not successfully modeled.

The SITAR procedure (Cole, 2020; Cole et al., 2010)[16,17] available in the R package *sitar* (R Core Team, 2019) was used. The model fits the raw data for all players with a curve (defined as a B-spline), superposes the curves of all cases, averages the curves and then back-projects the average curve into the original data as a growth model through uniform transformations: translation and rotation. A total of 269 measurements were available for the 58 players. Visual inspection of the model based on running plots with the raw data showed that the model fit the data very well. The mean residual was 0.0 cm by definition; the standard deviation of the residuals was 0.47 cm, and the mean absolute value of the residuals was 0.36 cm.

The FPCA growth model (Králík et al., 2021) is based on a combination of general Functional Data Analysis (FDA) and FCPA (Ramsay & Silverman, 2002, 2005). The complete postnatal growth curves of individual boys in the Brno Growth Study were the training set, which was fit by the B-spline curves of the raw data for all soccer players. The splines were modeled with the FPCA procedure; 12 Principal Components (6 for phase and 6 for amplitude of the growth curves) were then used as a generative model to fit the newly analyzed data based on the Levenberg–Marquardt optimization algorithm. Details of the specific calculations and functions of the model are available in the R package *growthfd* (Kíma & Králík, 2022; Králík et al., 2021). The mean of the 269 model residuals was 0.04 cm and the standard deviation of the residuals was 0.44 cm; the mean absolute value of the residuals was 0.33 cm.

Both methods provided estimates of age, velocity of growth and height at TO and at PHV for each player. Descriptive statistics for each variable were calculated. Each player was also classified as late (delayed), on time (average) or early (advanced) maturing based on estimated ages at PHV with the SITAR and FCPA models. A band of plus/minus 1.0 year of the respective mean ages at PHV for the total sample defined on time or average maturity status. An age at PHV greater than +1.0 year of the respective means indicated late maturity status, while an age at PHV less than –1.0 year of the respective means indicated early maturity status. A band of plus/minus one year was used as standard deviations for ages at PHV generally approximate about 1 year; the band also reflects variation in ages at PHV *per se* and that associated with the different methods of estimation (Malina et al., 2004).

Skeletal Age

Posterior-anterior radiographs of the left hand-wrist of players were taken at observations one, three and five. Skeletal age (SA) was evaluated with the Fels method (Roche et al., 1988), which utilizes specific criteria for the radius, ulna, carpals, the adductor sesamoid of the first metacarpal, the first, third and fifth metacarpals and phalanges, and ratios of linear measurements of epiphyseal and metaphyseal widths the first, third and fifth metacarpals and phalanges. The ratings were entered into the Felshw 1.0 Software program to derive an SA and its standard error, an indicator of maturity status at the time of observation. The mean difference between SAs by two independent assessors of 20 radiographs and the inter-observer technical error of measurement were, respectively, 0.03±0.04 years and 0.12 years; the inter-observer intra-class correlation was 0.99. Standard errors for SA assessments at observations one, three and five ranged, respectively, from 0.27 to 0.30 year (median 0.29), from 0.29 to 0.49 year (median 0.35), and from 0.30 to 0.48 (median 0.37) year. Four players were skeletally mature at observation five; an SA is not assigned to individuals who have attained skeletal maturity.

Based on SA and CA at each observations, the skeletal maturity status of each player was classified as average (on time), SA within ± 1.0 year of CA; delayed (late), SA younger than CA >1.0 year; or advanced (early), SA older than CA by >1.0year. The ±1.0 year band approximates standard deviations for SAs within specific CA groups, allows for error associated with assessments, and provides for broad range of youth who are classified as average in maturity status (Malina, 2011). As noted, an SA is not assigned to skeletally mature players. At observation five, 42 players had radiographs, but three did not have a measure of height and weight; of the 40 players with a measure of height and weight, one player did not have a radiograph.

Statistics

Descriptive statistics (means and standard deviations) at each observation for the longitudinal sample were calculated for CA, height and weight, and for SA and SA minus CA at the three observations. Corresponding statistics were calculated for estimates ages at TO and PHV (years), velocities of growth at TO and at PHV (cm/year), and heights at TO and at PHV (cm) based on the SITAR and FPCA methods. The differences between parameters of the growth spurt with the two methods were evaluated with paired sample t-tests and tests of equivalence using 90% equivalence boundaries representative of a moderate effect (±0.5 of Cohen's d). Correlations between parameters of the growth spurt and the difference of SA minus CA at observations one and three with ages at PHV were also calculated, and the concordance of maturity status classifications based on the two estimates of age at PHV and on skeletal maturity status at observations one, three and five were evaluated with chi square and unweighted Cohen's Kappa coefficients.

Systematic Comparisons among Samples of Soccer Players

Ages at PHV based on a variety of methods for 12 longitudinal samples of soccer players from Europe and Japan were compiled from the literature. Estimates were available for 6 samples of players from Europe: Wales (Bell, 1993), Denmark (Froberg, Anderson, & Lammert, 1991), Belgium (Philippaerts et al., 2006), Spain (Carvalho, Lekue, Gil, & Bidaurrazaga-Letona, 2017; Monasterio, Gil, & Bidaurrazaga-Letona, 2022; Monasterio et al., 2021), England (Parr et al., 2020), and Netherlands (Teunissen et al., 2020), and six samples of players from central Japan (Chuman, Hoshikawa, Iida, & Nichijima, 2013, 2014; Nariyama, Hauspie, & Mino, 2001; Saeki et al., 2021; Takei, Taketomi, Tanaka, & Torii, 2020). Several studies including the present study reported estimated ages at PHV based on two or three methods, while three studies provided estimates for subsamples of players from the same club. Excluding estimates based on graphic and incremental methods (Bell, 1993; Parr et al., 2020), and the FPCA method (present study), and limiting the estimate for Spanish players to the one based on the largest sample (Monasterio et al., 2022), ages at PHV in 13 samples of soccer players from Europe and Japan were subjected to a meta-analysis using methods available in the R-Package *metafor* (Viechtbauer, 2010). Sample size, mean and standard deviation for age at APHV of the 13 samples were used as estimates of effect size. The Random Effect Model was used because it can be reasonably assumed that the population with the same grand mean age at PHV was not sampled in the different studies of soccer players (i.e., the populations actually differed in ages at PHV). The restricted maximum likelihood method (REML estimator) was used to estimate the between-sample variance (τ^2, tau-squared).

6.4. Results

Descriptive statistics for CA, height and weight at each observation and for SA at observations 1, 3 and 5 in the longitudinal sample are summarized in Table 6.1. Corresponding statistics for parameters of TO and PHV are summarized in Table 6.2.

Table 6.1. Means (M) and standard deviations (SD) for chronological age (CA), skeletal age (SA), height and weight for the longitudinal sample by observation.

Obs	N	CA, yrs		SA, yrs		Height, cm		Weight, kg	
		M	SD	M	SD	M	SD	M	SD
1	58	11.9	0.5	12.0	1.4	144.9	6.9	37.6	6.0
2	58	12.9	0.5			151.7	7.9	42.3	7.3
3	58	13.9	0.5	14.2	1.1	159.2	7.7	48.7	8.4
4	55	14.9	0.5			165.5	6.7	54.9	7.9
5[a]	40	15.9	0.5			169.3	5.3	60.1	6.3
5[b]	35	15.8	0.5	16.3	1.1	169.2	5.4	59.7	6.2
5[c]	4	16.8	0.2			171.5	5.0	64.9	6.1

Obs (observations); CA (chronological age); SA (skeletal age); [a]Total sample of players with measures of height and weight at observation five; one player did not have a radiograph; [b]Not skeletally mature; [c]Skeletally mature.

Table 6.2. Estimated parameters for take-off (TO) and peak height velocity (PHV) of the adolescent spurt based on the SITAR and FPCA methods applied to the longitudinal height records of 58 soccer players and differences between the respective estimates (SITAR minus FPCA) with the two methods (M, mean; SD, standard deviation).

Parameters	SITAR			FPCA			SITAR - FPCA	
	M	SD	Range	M	SD	Range	M	SD
Age at TO, yrs	11.24	0.79	9.94-13.00	10.99	0.82	8.55-12.81	0.25	0.70
TO, cm/yr	4.62	0.52	3.37-6.06	4.58	0.36	3.85-5.67	0.03	0.47
Height at TO, cm	141.1	5.7	130.0-153.8	140.2	6.0	125.0-153.0	0.90	3.32
Age at PHV, yrs	13.62	0.90	11.92-15.59	13.66	0.88	11.90-15.49	-0.04	0.12
PHV, cm/yr	9.71	1.26	6.69-13.56	9.81	1.33	6.97-14.53	-0.10	0.87
Height at PHV, cm	157.1	5.7	145.9-169.7	157.9	5.6	147.7-170.1	-0.51	1.02

Ages at PHV

Estimated mean ages based on SITAR and FPCA are statistically different and not equivalent with each other for age at TO, 11.2±0.8 years and 11.0±0.8 years, respectively (t=2.73, p<0.01; Cohen's d=0.36), and for age at PHV, 13.6±0.9 years and 13.7±0.9 years, respectively (t=2.60, p=0.01; Cohen's d=0.34). Mean estimated heights at TO, 141.1±5.7 cm (SITAR) and 140.2±6.0 cm (FPCA), differ significantly (t=2.06, p<0.05) but are equivalent (Cohen's d=0.27), while mean estimated heights at PHV, 157.9±5.6 cm (FPCA) and 157.1±5.7 cm (SITAR), also differ significantly but are not equivalent (t=3.80, p<0.01, Cohen's d=0.50). Velocities of growth in height at TO, 4.6±0.5 cm/year (SITAR) and 4.6±0.4 cm/year (FPCA), do not differ statistically (t=0.56, p>0.05) and are equivalent (Cohen's d=.07). Velocities of growth in height at PHV, 9.7±1.3 cm/year, (SITAR) and 9.8±1.3 cm/year (FPCA), also do not differ statistically between methods and can be considered equivalent (Cohen's d=0.11).

Although the differences between estimated ages and heights at TO and at PHV with the SITAR and FCPA methods are statistically significant, they are quite small. The estimated ages at PHV and heights at PHV with the two methods are highly correlated, 0.99 and 0.98 (p<0.001), respectively, while the correlation for estimated PHVs with the two methods is slightly lower, 0.77 (p<0.001). Estimated ages and heights at TO with the two methods are correlated to a lesser extent, 0.62 and 0.84 (p<0.001), while the correlation for estimate velocity of growth at TO with the two methods is lower, 0.48 (p<0.001).

The cross-tabulation of maturity classifications based on the two estimates of age at PHV is summarized in Table 6.3. Overall, 90% of the players are classified as having the same maturity status based on ages at PHV with the SITAR and FPCA models. Four of the six players who are misclassified have estimated ages at PHV close to the plus/minus one year cut-offs, and the differences in ages at PHV with the two methods (SITAR minus FPCA) are negligible, 0.08, 0.07, 0.05 and 0.08 year. The differences between ages at PHV for the other two players are somewhat larger, 0.35 and -0.35 year.

Table 6.3. Frequencies and cross-tabulations of maturity status classifications (late, on time, early)[1] based on ages at PHV with the SITAR and FPCA models, percentage agreement, Chi square (χ^2) and Cohen's Kappa (κ); means and standard deviations for ages at PHV in the respective maturity groups are also indicated.

Age at PHV: FPCA, yrs			**Age at PHV: SITAR**, yrs			
			Late	On time	Early	Total
			14.92±0.35	13.54±0.53	12.37±0.20	
Late	14.98±0.27		**8**	1	0	9
On Time	13.69±0.52		3	**36**	0	39
Early	12.36±0.26		0	2	**8**	10
		Total	11	39	8	58

Agreement 90% χ^2 = 77.29* κ = 0.79*

* (p<0.01); [1]On time (average) is an age at PHV within ±1.0 year of the mean age at PHV for the total sample of 58 players with SITAR (13.62±0.90 years); late is a PHV >14.56 years; early is a PHV <12.72 years; corresponding on time classification with FPCA is (13.66±0.88 years); late is a PHV >14.54 years; early is a PHV <12.78 years.

Maturity Based on SA and Age at PHV

Spearman correlations (rho) between skeletal maturity status expressed as the difference of SA minus CA and age at PHV are moderate in early adolescence (~12 years, observation one), -0.53 (SITAR, p<0.01) and -0.54 (FPCA, p<0.001), but are higher in mid-adolescence (~14 years, observation three), -0.76 (SITAR, p<0.001) and -0.77 (FPCA, p<0.001). The negative correlations indicate an earlier age at PHV among players with a positive difference of SA minus CA, i.e., an SA in advance of CA.

The concordance of maturity status classifications (late, average or early) based on ages at PHV with SITAR and FPCA and on the difference of SA minus CA at each observation is summarized in Table 6.4. Maturity classifications at observations one and three, respectively, are concordant in 59% and 71% of the players for SITAR ages at PHV and in 62% and 74% of the players for FPCA ages at PHV. The Kappa coefficients are relatively low at observation one and moderate at observation three. Allowing for small numbers at observation five (~16 years), maturity classifications are concordant in 57% (SITAR) and 60% (FPCA) of the players, and the Kappa coefficient is moderate. Two of the four skeletally mature players at observation five are classified as on time and two as early maturing based on the ages at PHV. Mean ages at PHV for the four skeletally mature players are similar with SITAR (12.57±0.38 years) and FPCA (12.65±0.37 years), and the respective means are earlier than mean ages at PHV among early maturing CA peers who are not skeletally mature, SITAR (12.90±0.65 years) and FPCA (12.94±0.66 years), respectively.

Table 6.4. Frequencies and cross-tabulations of maturity status classifications based on ages at PHV with the SITAR and FPCA models and on Fels skeletal ages (SA – CA) at observations 1, 3 and 5, percentage agreement, Chi square (χ^2) and Cohen's Kappa (κ); means and standard deviations for SA – CA differences in the respective maturity groups are also indicated.

Skeletal Maturity Groups	Skeletal Age[2]	Maturity Groups: Age at PHV SITAR[1]				Age at PHV FPCA[1]			
		Late	On time	Early	Total	Late	On time	Early	Total
Observation 1									
Late	(-1.98±0.98 yrs)	**3**	6	0	9	**3**	6	0	9
On Time	(-0.02±0.59 yrs)	6	**25**	2	33	4	**26**	3	33
Early	(1.76±0.62 yrs)	2	8	**6**	16	2	7	**7**	16
Total	(0.16±1.38 yrs)	11	39	8	58	9	39	10	58
		Agreement 59%, χ^2=11.60**, κ = 0.25*				Agreement 62%, χ^2=13.49**, κ = 0.31*			
Observation 3									
Late	(-1.48±0.33 yrs)	**5**	3	0	8	**4**	4	0	8
On Time	(0.18±0.46 yrs)	5	**29**	1	35	4	**30**	1	35
Early	(1.66±0.54 yrs)	1	7	**7**	15	1	5	**9**	15
Total	(0.33±1.07 yrs)	11	39	8	58	9	39	10	58
		Agreement 71%, χ^2 = 28.75*, κ = 0.45*				Agreement 74%, χ^2=33.45**, κ = 0.51*			
Observation 5									
Late	(-1.76±0.38 yrs)	**5**	0	0	5	**3**	2	0	5
On Time	(0.00±0.72 yrs)	2	**14**	0	16	1	**15**	0	16
Early	(1.71±0.41 yrs)	0	12	**5**	17	0	10	**7**	17
Mature		0	2	2	4	0	2	2	4
Total	(0.53±1.33 yrs)	7	28	7	42	4	29	9	42
		Agreement 57%, χ^2 = 36.90*, κ = 0.45*				Agreement 60%, χ^2 = 27.17*, κ = 0.37*			

*(p<0.01); [1]See Table 6.3 for ages at PHV in the respective maturity groups; On time - SA within ±1.0 year of CA; late - SA behind CA by >1.0 year; early - SA advance of CA by > 1.0 year.

Ages at PHV in Studies of Soccer Players

Ages at PHV reported in the 13 studies of soccer players are summarized in Table 6.5. A variety of methods were used to estimate ages at PHV. Several studies estimated age at PHV with two or three estimates, but mean ages based on different methods within the same samples do not differ. Mean ages at PHV for the 58 Portuguese soccer players estimated with SITAR and FPCA are within the range of mean ages at PHV in the six longitudinal samples of soccer players in Europe, 12.9 to 14.2 years. The earliest estimated mean age at PHV, 12.9 years (standard deviation not reported), is for a sample of 33 Spanish players 10+ years measured on four to six occasions between 2009 and 2016 (Carvalho et al., 2017). Two estimates for larger samples from the same soccer club (n=110 and 124), 10-11+ years of age with 10 or more observations between 2000 and 2020 are later, 13.4±0.8 years (Monasterio et al., 2021) and 13.5±0.9 years (Monasterio et al., 2022). The analysis of the small sample used a two level polynomial, while the later analyses used the SITAR model. The differences likely reflect both sampling and methodological variation; the more recent estimates are also based on a larger number of height measurements for each player.

Table 6.5. Ages at PHV (years) and PHV (cm/year) in longitudinal samples of adolescent male soccer players in Europe (including the present study) and Japan.

Country	Study	Method	Competitive Level	Observations (obs)	N	APHV yrs M	SD	Range, yrs	PHV cm/yr M	SD
EUROPE										
Portugal	Present study	Sitar	prof clubs	11-12/15-16 yrs, 4-5	58	13.6	0.9	11.9-15.6	9.7	1.3
		FPCA		annual obs, 2003-2008	58	13.7	0.9	11.9-15.5	9.8	1.3
Wales	Bell (1993)	Graphic	school	12-15 yrs, 4 annual	32	14.2	0.8		9.6	1.8
		Moving incr		obs, 1981-1984		14.1	0.8		9.3	1.5
		Polynomials				14.2	0.9		9.5	1.5
Denmark	Froberg et al. (1991)	PB 1	local club	11-16 yrs, semi-annual obs over 6 yrs, 1980s	8	14.2	0.9	12.6-15.7		
Belgium	Philippaerts et al. (2006)	Polynomials	prof clubs	10-13/14-17yrs, annual obs 1996-2000	33	13.8	0.8			
Spain (same club)	Carvalho et al. (2017)	Polynomials	prof club	10-16 yrs, 4 obs 2009-2014	33	12.9		11.8-15.5*	8.1	
	Monasterio et a. (2021, 2022)	Sitar	prof club	10-11 yrs-16-18 yrs,	110	13.4	0.8		9.9	1.8
				≥10 obs, 2000-2020	124	13.5	0.9		10.1	2.0
England	Parr et al. (2020)	Sitar	prof club	5 seasons 12.4±0.6 yrs	27	14.1	0.8	12.6-15.5	9.8	
		Graphic		baseline, 17-20 obs, 2013-2017	27	14.2	0.8			
Netherlands	Teunissen et al. (2020)	PB 1	prof club	4 seasons 11.9±0.8 yrs baseline, 16-25 obs, 2008-2012	17	13.8	0.7	12.6-15.2		
JAPAN										
Fukui Prefecture	Nariyama et al. (2001)	PB 1	school	school records, 1970-1987, 6-18 yrs	83	13.7	1.1		8.8	1.1
Saitama	Saeki et al. (2021)	Auxal	school	school records 7-12 yrs + obs JHS	88	13.3	0.9			
Tokyo	Takei et al. (2020)	Auxal	rec league	school records + 6 obs over 2 yrs 2011-2016	201	13.4	0.9			
Shizouka	Chuman et al. (2013)	Triple logistic	prof club	school record: sub-elite	48	12.9	0.9			
				7-12yrs, club elite obs at 13 yrs	16	12.6	1.0			
Shizouka	Chuman et al. (2014)	Triple logistic	prof club	school records, club 6 obs, 7-15 yrs, 2008-2010	29	12.9	1.0			

*(estimated 95% credible interval)

The estimated mean ages at PHV for players from professional clubs in Europe are somewhat earlier than those for players from a school and local club. With the exception of one subsample of club players in Spain (Carvalho et al., 2017), means ages at PHV for club soccer players in Europe tend to be later than estimates for players at a professional club in Japan, 12.6 to 12.9 years (Table 6.5). The latter, in turn, are earlier than estimates for school and recreational league players in Japan.

Results of the meta-analysis of ages at PHV in the 13 samples of soccer players from Europe and Japan indicate significant heterogeneity (Q [df, 12]) = 103.45, $p < 0.0001$) while I^2 for the model of all studies is relatively high (93.3%), By inference, a systematic effect among samples is suggested. The effect of geographic location (Japan and Southern, Northern and Western Europe) as a moderator of age at PHV was then evaluated. The Mixed Effect Model indicates a statistically significant moderator effect (Q_M [df, 2] = 22.33, $p < 0.0001$); ages at PHV differ significantly among the geographic groups (Figure 6.1). Age at PHV is latest for players from Northern and Western Europe, earlier for players from Southern Europe, and earliest for players from Japan. Heterogeneity among samples in Northern and Western Europe (professional and local clubs and schools combined) is not significant (Q [df, 4] = 5.46, p = 0.2433), while heterogeneity among the samples of professional clubs in Europe (Southern, Northern and Western Europe together) is significant (Q [df, 4] = 15.4260, p = 0.0039). By inference, geographical distribution appears to be a more significant factor than level of competition; note, however, sample sizes in Northern and Western Europe are relatively small which may have reduced the statistical significance of differences among samples.

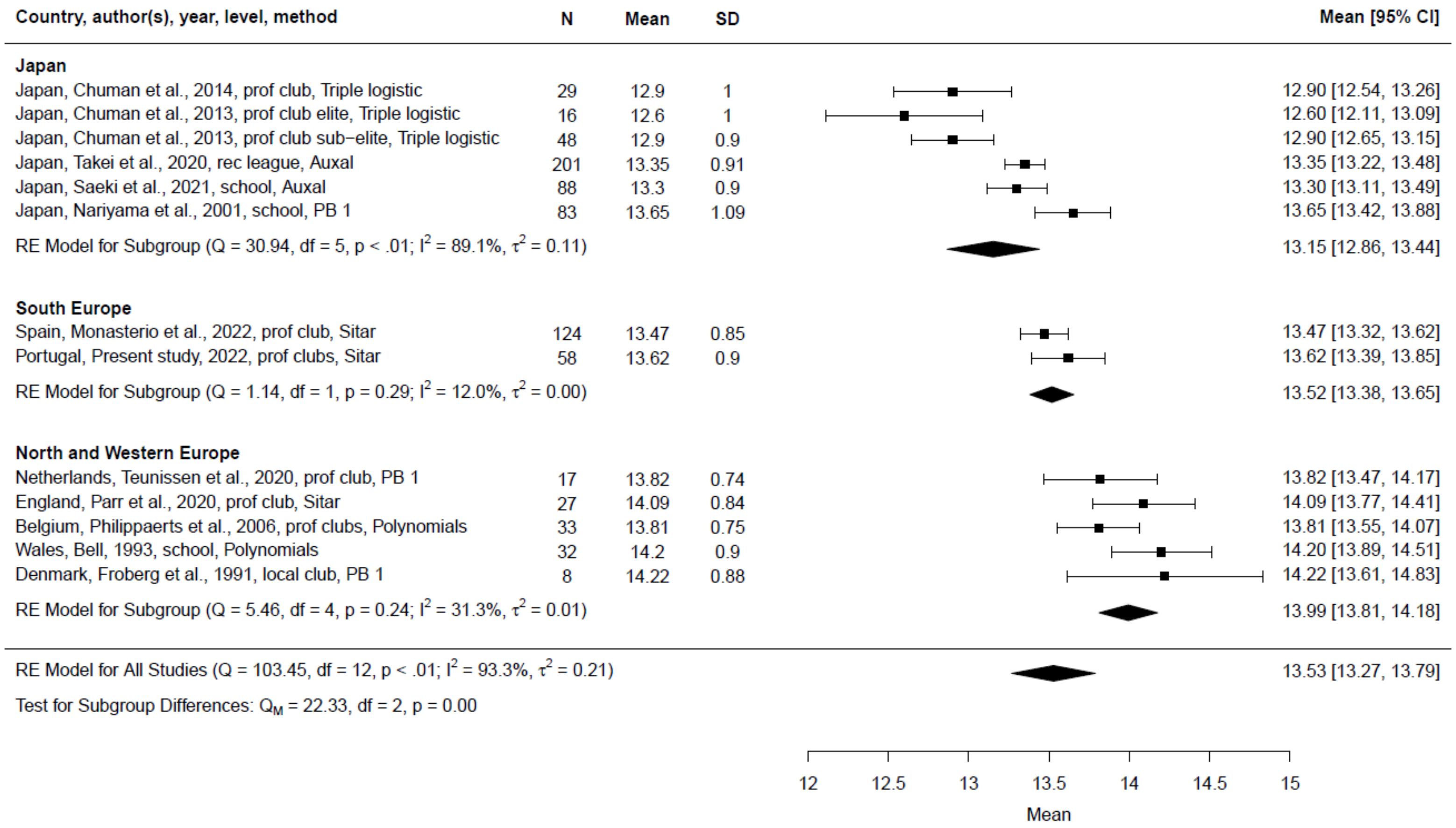


Figure 6.1. Aggregation of ages of PHV (years) in samples of male soccer players based on meta-analysis, including the subgroup analysis. Q: Cochran's Q-statistic (weighted sum of squares), QM: Cochran's Q-statistic for subgroups, I2: percentage of variability in effect sizes which is not due sampling error, τ2: between-study variance in a given set of samples (years squared); plots: means and 95% CIs of individual studies; diamonds: width represents 95% CI for each model aggregated by subsample and for all studies (below).

6.5. Discussion

Differences in estimated mean ages at TO and PHV and estimated mean PHVs based on the SITAR and FPCA models in the sample of 58 Portuguese soccer players (Table 6.2), though statistically significant, were very small in practical terms. Estimated mean ages at TO (11.2 and 11.0 years) and at PHV (13.6 and 13.7 years) among the soccer players were also within the ranges of reported mean ages in longitudinal samples of European boys spanning the 1970s through the present: 10.4 to 11.8 years for 21 estimates of mean age at TO, and 13.0 to 14.5 years for 64 estimates of mean age at PHV. Mean PHVs among the soccer players (9.7 and 9.8 cm/year) were also within the range of estimates in the general population: 27 estimates ranged from 7.8 to 11.5 cm/year (Malina et al., 2004) (the reference includes the citations for the studies reporting the respective means).

Estimated mean ages at PHV for the 58 Portuguese soccer players based on SITAR and FPCA were also within the range of mean ages at PHV estimated with different methods in six longitudinal samples of soccer players in Europe (Table 6.5). Though limited to a relatively small number of studies, results of the systematic analysis of ages at PHV in 13 samples of soccer players from Europe and Japan suggested earlier ages among players in Southern compared to Northern and Western Europe, while mean ages at PHV for Japanese club players, largely from Central Japan, tended to be earlier than corresponding ages in European players (Figure 6.1). The trend towards earlier ages at PHV among soccer players in Japan compared to players in Europe was consistent that with noted for ages at PHV in the general populations of youth in both regions (Malina et al., 2004). Studies in Japan are unique in that they commonly use serial height records of players measured annually at their respective schools. Beginning at 7 years of age, heights of school children in Japan are measured annually in April (Nariyama et al., 2001). In several instances, school records were complemented by measurements taken at the leagues and soccer clubs.

The preceding discussion was largely focused on mean ages at PHV among samples of soccer players. Variation in ages at PHV *among individual players* also merits attention. Ages at PHV among players in five of the European clubs ranged from 11.8 to 15.7 years (Table 6.5) were within that noted in several longitudinal samples of European boys, 11.3 to 17.3 years (Malina et al., 2004; Malina & Kozieł, 2014a; Malina et al., 2021). Of potential relevance, the range among soccer players may have been somewhat restricted by relatively the late CAs at initial observation and limited duration of the studies. The study of Belgian players (Philippaerts et al., 2006), for example, followed 76 players for 4 to 5 years; CAs at initial observation ranged from 10.4 to 13.7 years. Ages at PHV based on polynomials were successfully estimated for only 33 players (CA 12.1±0.7 years and SA 12.4±1.3 years at initial observation). Among the 43 players for whom ages at PHV could not be estimated, 25 were older (CA=12.6±0.5 years) and advanced in SA (13.5±1.2 years), while 18 were younger (CA=11.6±0.8 years) and delayed in SA (11.1±1.1 years) at initial observation. It is likely that PHV in players advanced in SA occurred before or at initial observation, while PHV in players slightly delayed in SA occurred at or after final observation, and as such could not be estimated.

Estimated ages at TO of the adolescent spurt among the 58 Portuguese players spanned 8.6 to 13.0 years (Table 6.2) and were largely in the range of ages at TO in three longitudinal samples of European boys, 9.0 to 15.0 years (Kozieł & Malina, 2018; Largo, Gasser, Prader, Stuetzle, & Huber, 1978; Preece & Baines, 1978). Corresponding estimates are lacking for the other longitudinal samples of soccer players. Nevertheless, the range in estimated ages at TO in soccer players highlights the importance monitoring the growth status of players from an earlier age. The Growth and Maturation Screening Programme implemented by the English Premier League, for example, now measures the heights and weights of all registered academy players 9 years and older every three to four months (Cumming et al., 2017). Along with corresponding observations for fitness and an academy-wide injury audit, the data provide a potentially unique opportunity to better understand the impact of the adolescent growth spurt upon fitness and performance and also on the incidence and burden of injury. Note, however, taking height measurements at relatively close intervals requires attention to inter- and intra-examiner measurement variability, diurnal variation in measurements, and seasonal variation in growth (Malina et al., 2004). Heights also should not be measured after a period of physical activity as in training programs and scrimmages.

Concordance of maturity classifications (late, average or early) based on ages at PHV (maturity timing) and on SA minus CA (maturity status at the time of observation) was modest at initial observation, 11.9±0.5 years, but was higher at the third observation, 13.9±0.5 years (Table 6.4). The observations were consistent with relationships among indicators of maturity timing close to the time of PHV among 111 boys in the Wrocław Growth Study (Bielicki, Koniarek, & Malina, 1984). Correlations between estimated ages at attaining SAs of 12.0 and 14.0 years and age at PHV were, respectively, 0.42 and 0.81. Correlations between the two estimates of age at PHV and the difference of SA minus CA in the 58 soccer players were similar at observations one (-0.54) and three (-0.77), i.e., advanced skeletal maturity status at 12 and 14 years was related with an earlier age at PHV, and the association was stronger closer to the time of PHV.

Inter-individual variation in estimated rates of growth in height (cm/year) at TO and at PHV in the longitudinal series of soccer players has implications for those working with youth athletes. Ranges of estimated rates of growth at PHV (Table 6.2) spanned 6.7 to 13.6 cm/year (SITAR) and 7.0 to 14.5 cm/year (FCPA). Based on monthly measurements of heights and weights of soccer players 11-19 years during the course of a season (September through April), estimated monthly increments of $\geq$0.6 cm/month in height and of $\geq$0.3 kg/m^2/month in the BMI, and an estimated monthly decline of $\geq$0.4 kg/m^2/month in the BMI were associated with an increased risk of injury (Kemper et al., 2015). Extending the monthly increments in height through a year, it was suggested that an estimated velocity of growth in height ≥7.2 cm/year was indicative that a player was within his growth spurt (Johnson et al., 2022; Kemper et al., 2015). The range of estimated PHVs in the sample of 58 soccer players (Table 6.2) suggests that some players with rates of growth <7.2 cm/year were in their growth spurts.

Epidemiological data suggest enhanced susceptibility to injury during the interval of the growth spurt, especially conditions associated with rapid growth, i.e., Osgood-Schlatter and Sever's disease (Belikan et al., 2022; Price, Hawkins, Hulse, & Hodson, 2004), and overuse (DiFiori et al., 2014). Management of training load and use of developmentally appropriate training protocols (activities emphasizing core strength, balance, coordination, mobility, and limiting accelerations and decelerations) may serve to mitigate injury risk during this interval of rapid growth (Cumming, 2018). Some athletes may also experience temporary disruptions or regressions in motor performances during the interval of the growth spurt, commonly labeled as adolescent awkwardness (Malina et al., 2004). Of potential relevance, recent evidence suggests that coach evaluations of match performances of youth soccer players tend to decline through the growth spurt, but return to pre-spurt levels at the cessation of the spurt (Hill, Scott, McGee, & Cumming, 2021). By inference, it is essential that coaches and others working with youth athletes are aware of the details of growth and maturation during the interval of the adolescent spurt, specifically individual differences in timing and tempo, when evaluating youth athletes, for example, delaying decisions until after the growth spurt, reviewing game film/player performance metrics prior to the onset and during the spurt, and/or allowing a player to play down an age group while they adjust to changes associated with the adolescent spurt (Hill, John, McGee, & Cumming, 2022).

6.6. Conclusions

Mean ages at PHV based on two methods of estimation in a longitudinal sample of 58 Portuguese soccer players were 13.6±0.9 years and 13.7±0.9 years and were within the range of observed means and standard deviations for ages at PHV in studies of the general population and of soccer players in Europe. Concordance of maturity classifications based on age at PHV, and the difference of SA minus CA was moderate, but was strongest close the age at PHV (observation 3, about 14 years). Systematic analysis of ages at PHV in 13 longitudinal samples of European and Japanese soccer players suggested a geographic gradient: northern Europe > southern Europe > Japan. The gradient was consistent with that observed in the general population of adolescents in Europe and Japan.

6.7. References

Belikan, P., Färber, L.-C., Abel, F., Nowak, T. E., Drees, P., & Mattyasovszky, S. G. (2022). Incidence of calcaneal apophysitis (Sever’s disease) and return-to-play in adolescent athletes of a German youth soccer academy: a retrospective study of 10 years. Journal of Orthopaedic Surgery and Research, 17(1), 83.

Bell, W. (1993). Body size and shape: A longitudinal investigation of active and sedentary boys during adolescence. Journal of Sports Sciences, 11(2), 127–138.

Bielicki, T., Koniarek, J., & Malina, R. M. (1984). Interrelationships among certain measures of growth and maturation rate in boys during adolescence. Annals of Human Biology, 11(3), 201–210.

Carvalho, H. M., Lekue, J. A., Gil, S. M., & Bidaurrazaga-Letona, I. (2017). Pubertal development of body size and soccer-specific functional capacities in adolescent players. Research in Sports Medicine, 25(4), 421–436.

Chuman, K., Hoshikawa, Y., Iida, T., & Nichijima, T. (2013). Relationship between sprint ability and maturity in elite and sub-elite pubescent male soccer players. Football Sci, (10), 10–17.

Chuman, K., Hoshikawa, Y., Iida, T., & Nichijima, T. (2014). Maturity and intermittent endurance in male soccer players during the adolescent growth spurt: A longitudinal study. Football Sci, (11), 39–47.

Cole, T. J. (2020). SITAR: Super Imposition by Translation and Rotation Growth Curve Analysis. Retrieved from https://cran.r-project.org/package=sitar

Cole, T. J., Donaldson, M. D. C., & Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. International Journal of Epidemiology, 39(6), 1558–1566. https://doi.org/10.1093/ije/dyq115

Cumming, S. P. (2018). A game plan for growth: how football is leading the way in the consideration of biological maturation in young male athletes. Annals of Human Biology, 45(5), 373–375.

Cumming, S. P., Lloyd, R. S., Oliver, J. L., Eisenmann, J. C., & Malina, R. M. (2017). Bio-banding in sport: Applications to competition, talent identification, and strength and conditioning of youth athletes. Strength and Conditioning Journal, 39(2), 34–47.

DiFiori, J. P., Benjamin, H. J., Brenner, J. S., Gregory, A., Jayanthi, N., Landry, G. L., & Luke, A. (2014). Overuse injuries and burnout in youth sports: a position statement from the American Medical Society for Sports Medicine. British Journal of Sports Medicine, 48(4), 287–288.

Froberg, K., Anderson, B., & Lammert, O. (1991). Maximal oxygen uptake and respiratory functions during puberty in boy groups of different physical activity. In F. Frenkl & I. Szmodis (Eds.), Children and Exercise: Pediatric Work Physiology XV (pp. 65–80).

Harriss, D. J., Macsween, A., & Atkinson, G. (2019). Ethical Standards in Sport and Exercise Science Research: 2020 Update. International Journal of Sports Medicine, 40(13), 813–817.

Hauspie, R., & Chrząstek-Spruch, H. (1999). Growth models: Possibilities and limitations. In F. E. Johnston, P. B. Eveleth, & B. Zemel (Eds.), Human Growth in Context (pp. 15–24). London (UK): Smith-Gordon.

Hill, M., John, T., McGee, D., & Cumming, S. P. (2022). ‘He’s got growth’: Coaches understanding and management of the growth spurt in male academy football. International Journal of Sports Science & Coaching, 174795412211224.

Hill, M., Scott, S., McGee, D., & Cumming, S. P. (2021). Are relative age and biological ages associated with coaches’ evaluations of match performance in male academy soccer players? International Journal of Sports Science & Coaching, 16(2), 227–235.

Johnson, D. M., Cumming, S. P., Bradley, B., & Williams, S. (2022). The influence of exposure, growth and maturation on injury risk in male academy football players. Journal of Sports Sciences, 40(10), 1127–1136.

Kemper, G., van der Sluis, A., Brink, M., Visscher, C., Frencken, W., & Elferink-Gemser, M. (2015). Anthropometric Injury Risk Factors in Elite-standard Youth Soccer. International Journal of Sports Medicine, 36(13), 1112–1117.

Kíma, O., & Králík, M. (2022). R package growthfd for fitting FPCA-based growth curve models. Retrieved from https://ondrej-klima.github.io/growthfd/index.html

Kozieł, S. M., & Malina, R. M. (2018). Modified Maturity Offset Prediction Equations: Validation in Independent Longitudinal Samples of Boys and Girls. Sports Medicine, 48(1), 221–236.

Králík, M., Klíma, O., Čuta, M., Malina, R. M., Kozieł, S., Polcerová, L., … Zemčík, P. (2021). Estimating Growth in Height from Limited Longitudinal Growth Data Using Full-Curves Training Dataset: A Comparison of Two Procedures of Curve Optimization-Functional Principal Component Analysis and SITAR. Children (Basel, Switzerland), 8(10).

Largo, R. H., Gasser, T., Prader, A., Stuetzle, W., & Huber, P. J. (1978). Analysis of the adolescent growth spurt using smoothing spline functions. Annals of Human Biology, 5(5), 421–434.

Malina, R. M. (2011). Skeletal Age and Age Verification in Youth Sport. Sports Medicine, 41(11), 925–947.

Malina, R. M. (2021). L’accelerazione (“spurt”) di crescita adolescenziale nei giovani atleti. Atletica Studi: Trimestrale Di Ricerca Scientifica e Tecnica Applicata All’Atletica Leggera, 52(1), 1–15.

Malina, R. M., Bouchard, C., & Bar-Or, O. (n.d.). Growth and Maturation: Physical Activity and Sport (3rd ed.). Champaign, IL (USA): Human Kinetics.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of maturity offset in the fels longitudinal study. Pediatric Exercise Science, 28(3), 439–455.

Malina, R. M., & Kozieł, S. M. (2014). Validation of maturity offset in a longitudinal sample of Polish boys. Journal of Sports Sciences, 32(5), 424–437.

Malina, R. M., Kozieł, S. M., Králik, M., Chrzanowska, M., & Suder, A. (2021). Prediction of maturity offset and age at peak height velocity in a longitudinal series of boys and girls. American Journal of Human Biology, 33(6).

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho-e-Silva, M. J., & Figueiredo, A. J. (2015). Biological maturation of youth athletes: Assessment and implications. British Journal of Sports Medicine, 49(13), 852–859.

Marubini, E., & Milani, S. (1986). Approaches to the analysis of longitudinal data. In F. Falkner & J. M. Tanner (Eds.), Human Growth: A Comprehensive Treatise, Vol 3 (2nd ed., pp. 79–94). New York: Plenum.

Mirwald, R. L., Baxter-jones, A., Bailey, D., & Beunen, G. (2002). An assessment of maturity from anthropometric measurements. Medicine & Science in Sports & Exercise, 34(4), 689–694.

Monasterio, X., Gil, S. M., Bidaurrazaga-Letona, I., et al. (2022). Estimating maturity status in elite youth soccer players: Evaluation of methods. under review.

Monasterio, X., Gil, S. M., Bidaurrazaga-Letona, I., Lekue, J. A., Santisteban, J. M., Diaz-Beitia, G., … Larruskain, J. (2021). The burden of injuries according to maturity status and timing: A two-decade study with 110 growth curves in an elite football academy. European Journal of Sport Science, 1–11.

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D. G., Cameron, N., & Brasher, P. M. A. (2015). Enhancing a somatic maturity prediction model. Medicine and Science in Sports and Exercise, 47(8), 1755–1764.

Nariyama, K., Hauspie, R. C., & Mino, T. (2001). Longitudinal growth study of male Japanese junior high school athletes. American Journal of Human Biology, 13(3), 356–364.

Parr, J., Winwood, K., Hodson-Tole, E., Deconinck, F. J. A., Parry, L., Hill, J. P., … Cumming, S. P. (2020). Predicting the timing of the peak of the pubertal growth spurt in elite male youth soccer players: evaluation of methods. Annals of Human Biology, 47(4), 400–408.

Philippaerts, R. M., Vaeyens, R., Janssens, M., Van Renterghem, B., Matthys, D., Craen, R., … Malina, R. M. (2006). The relationship between peak height velocity and physical performance in youth soccer players. Journal of Sports Sciences, 24(3), 221–230.

Preece, M. A., & Baines, M. J. (1978). A new family of mathematical models describing the human growth curve. Annals of Human Biology, 5(1), 1–24.

Price, R. J., Hawkins, R. D., Hulse, M. A., & Hodson, A. (2004). The Football Association medical research programme: an audit of injuries in academy youth football. British Journal of Sports Medicine, 38(4), 466–471.

R Core Team. (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Ramsay, J. O., & Silverman, B. W. (2002). Applied Functional Data Analysis: Methods and Case Studies. New York (USA): Springer-Verlag.

Ramsay, J. O., & Silverman, B. W. (2005). Functional Data Analysis (2nd ed.). New York (USA): Science + Business Media.

Roche, A. F., Cameron, W., & Thissen, D. (1988). Assessing the skeletal maturity of the hand-wrist: Fels method. Springfield, Illinois: Charles C. Thomas.

Saeki, J., Iizuka, S., Sekino, H., Suzuki, A., Maemichi, T., & Torii, S. (2021). Optimum Angle of Force Production Temporarily Changes Due to Growth in Male Adolescence. Children, 8(1), 20.

Sanders, J. O., Qiu, X., Lu, X., Duren, D. L., Liu, R. W., Dang, D., … Cooperman, D. R. (2017). The Uniform Pattern of Growth and Skeletal Maturation during the Human Adolescent Growth Spurt. Scientific Reports, 7(1), 1–9.

Takei, S., Taketomi, S., Tanaka, S., & Torii, S. (2020). Growth pattern of lumbar bone mineral content and trunk muscles in adolescent male soccer players. Journal of Bone and Mineral Metabolism, 38(3), 338–345.

Teunissen, J. W., Rommers, N., Pion, J., Cumming, S. P., Rössler, R., D’Hondt, E., … Malina, R. M. (2020). Accuracy of maturity prediction equations in individual elite male football players. Annals of Human Biology, 47(4), 409–416.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software, 36(3).

Chapter VII

Study 5

Observed and predicted ages at peak height velocity in soccer players.

7. Observed and predicted ages at peak height velocity in soccer players.

7.1. Abstract

The purpose of the study was to evaluate predicted maturity offset (time before age at PHV) and age at PHV (chronological age [CA] minus maturity offset) in a longitudinal sample of 58 under-13 club level soccer players in central Portugal for whom ages at PHV were estimated with the SITAR model. Two maturity offset prediction equations were applied: the original equation which requires CA sitting height, estimated leg length, height and weight, and a modified equation which requires CA and height. Predicted maturity offset increased, on average, with CA at prediction throughout the age range considered, while variation in predicted maturity offset and ages at PHV within CA groups was considerably reduced compared to variation in observed ages at offset and at PHV. Predicted maturity offset and ages at PHV were consistently later than observed maturity offset and age at PHV among early maturing players, and earlier than observed in late maturing players. Both predicted offset and ages at PHV with the two equations were, on average, later than observed among players maturing on time. Intra-individual variation in predicted ages at PHV with each equation was considerable. The results for soccer players were consistent with similar studies in the general population and two recent longitudinal studies of soccer players. The results question the utility of predicted maturity offset and age at PHV as valid indicators of maturity timing and status.

Key words: Adolescence; Maturation; Growth spurt, Youth athletes; Training.

7.2. Introduction

Predicted maturity offset, defined as time before peak height velocity (PHV) (Mirwald et al., 2002; Moore et al., 2015), and estimated age at PHV, i.e., chronological age (CA) minus predicted offset, are widely used as estimates of maturity status (state of maturation at the time of observation) and/or timing (age at which a specific maturational event occurs) in studies of youth athletes and to a lesser extent in studies youth physical activity and fitness (Malina, 2014; Moran et al., 2017; Moran, Sandercock, Rumpf, & Parry, 2016). The original sex-specific equations require CA, sitting height, estimated leg length, height and weight (Mirwald et al., 2002), while the modified equations (Moore et al., 2015) require CA and height (both sexes) or sitting height (boys). Validation studies of the original equations in three independent longitudinal series, the Wroclaw Growth Study (Malina & Kozieł, 2014a, 2014b), the Fels Longitudinal Study (Malina et al., 2016) and the Cracow Growth Study (Malina et al., 2021), and of the modified equations in two of the samples (Kozieł & Malina, 2018; Malina et al., 2021) have indicated major limitations of the predictions in both males and females. The validity of the prediction equations has also been questioned in longitudinal samples of female artistic gymnasts (Malina et al., 2006) and soccer players (Parr et al., 2020; Teunissen et al., 2020), but sample sizes in longitudinal samples tend to be limited and to some extent select as they are limited to athletes who have persisted in the respective sports (Malina et al., 2015). Similarly, cross-sectional studies of tennis (Myburgh et al., 2019) and soccer (Malina et al., 2012) players have questioned maturity status classifications based on the original prediction equations (Mirwald et al., 2002) relative to classifications based on skeletal age.

Current interest in the application of the maturity offset prediction equations in samples of youth athletes is considerable (Malina, 2014; Moran et al., 2017, 2016). Predictions based on the original and modified equations, however, depend upon CA and body size at prediction, have reduced variation relative to observed ages at PHV, and have major limitations with early and late maturing youth as defined by observed ages at PHV (Kozieł & Malina, 2018; Malina et al., 2016; Malina & Kozieł, 2014a, 2014b; Malina et al., 2021). The latter are problematic as advanced (early for CA) or delayed (late for CA) maturity status is often of major concern in developmental studies of youth athletes (Malina et al., 2015).

In the context of the preceding, the purpose of the present study is to evaluate predicted maturity offset and derived ages at PHV with the original (Mirwald et al., 2002) and modified (Moore et al., 2015) equations in a sample of 58 soccer players for whom age at PHV was determined from longitudinal height records. The study specifically considers variation in the predictions in three contexts: (i) relative to actual maturity offset and observed age at PHV at each observation, (ii) among players differing in the timing of observed age at PHV, and (iii) within individual players.

7.3. Methods

Research design and procedures

The data set for the present study was extracted from the Coimbra Soccer Longitudinal Project (Valente-dos-Santos et al., 2012). This project was conducted according to the standards established by the declaration of Helsinki (Harriss et al., 2019), and formal approval was obtained from the University of Coimbra Sports Sciences and Physical Education Board (FCDEFUC/AAC/2003; FCDEFUC/ADCA/2003; FCDEFUC/CFM/2003; FCDEFUC/CFUC/2003; FCDEFUC/GRVM/2003). Signed institutional agreements were also obtained from the Presidents of the respective clubs. All players were registered with the Portuguese Soccer Federation. Male players 11-14 years of age were recruited from five clubs in the Midlands of Portugal; the initial sample included 159 players (Figueiredo et al., 2009a). Written consent was obtained from parents or legal guardians of the participants who were informed that contribution to the study was voluntary and that they could withdraw from the study at any time. All observations were completed at the Biokinetics Laboratory of the Coimbra University Stadium.

Sample

According to the Portuguese Soccer Federation, male soccer players were grouped as infantiles (aged 11-12 years, n=87) and initiates (aged 13-14 years, n=72). The analysis in the current study is limited to under-13 players (U13) who were measured annually in

December for four or five years (n=59). CAs at baseline ranged from 10.98 to 12.94 years. The heights of one player were not successfully modeled; consequently, the final sample was composed of 58 players. All players were of European ancestry, except one. Participants trained and competed September through May. They had a median of 3 years of soccer experience at baseline (range: 2-6 years). The clubs had 3-5 training sessions per week (each about 90-120 minutes) and usually one game, mainly on Saturdays.

Anthropometry

Participants wore shorts and a t-shirt; shoes were removed. Height and sitting height were measured to the nearest 0.1cm using, respectively, a stadiometer (Harpenden 98.603, Holtain Ltd, Croswell, UK) and a table (Harpenden sitting height table, model 98.607, Holtain Ltd, Crosswell, UK). Body weight was measured to the nearest 0.1 kg using a scale (SECA 770, Hanover, MD, USA). The heights, sitting heights and weights of players who continued at the respective clubs were subsequently measured on an annual basis. Measurements were made by a single observer (MJCS). Intra-observer technical errors of measurement for height, sitting height and weight were 0.27 cm, 0.31 cm and 0.47 kg, respectively (Harriss et al., 2019).

Age at PHV

The longitudinal height records were fit with the Superimposition by Translation and Rotation (SITAR) model (Cao, Hui, & Wong, 2018; Cole, 2020; Cole et al., 2010; R Core Team, 2019) to derive an age at peak height velocity for each player. As noted, the heights of one player were not successfully modeled; the estimated age at PHV for the player was outside of the empirical data range and was inconsistent with his advanced skeletal maturity status at observations one and three. Mean age at PHV for the remaining 58 players was 13.60±0.85 years, with a range from 11.89 to 15.49 years.

Predicted Maturity Offset

Maturity offset, defined as time before PHV, was predicted at each observation for the 58 players with the original equation for boys (Mirwald et al., 2002):

Equation 1:

Maturity offset (years)= -9.236+(0.0002708 × (Leg Length × Sitting Height))+(-0.001663 × (CA × Leg Length))+(0.007216 × (CA × Sitting Height))+(0.02292 × (Weight by Height Ratio × 100))

Leg length was estimated at each observation as standing height minus sitting height. The need to multiply the weight by height ratio by 100 was overlooked in the original report (Mirwald et al., 2002); in some publications using the equation, it is not clear if the adjustment was applied. Maturity offset was also predicted at each observation with a modified equation for boys that incorporated age and height (Moore et al., 2015):

Equation 2:

Maturity offset (years)= -7.999994+(0.0036124 × (CA × Height))

The equation with age and height was selected for evaluation as it is increasingly used [26-31]. Standard errors for the original (Mirwald et al., 2002) and modified (Moore et al., 2015) equations were, respectively, 0.592 and 0.542 year. Predicted maturity offset and predicted age at PHV with the respective equations are subsequently labelled in the text, tables and figures as Mirwald and Moore, respectively.

Predicted Age at PHV

Predicted age at PHV was estimated as CA minus predicted maturity offset at each observation for individual players with the respective equations.

Observed Maturity Offset

Observed or actual maturity offset at each observation was estimated as CA at prediction minus observed age at PHV based on the SITAR model.

Analyses

Descriptive statistics were calculated at each observation for CA and actual offset, for predicted maturity offset and age at PHV, and for the difference of predicted age at PHV minus observed age at PHV with the two prediction equations. The players were also classified as advanced (early), average (on time) or delayed (late) maturing relative to the mean and standard deviation for age at PHV (SITAR) in the total sample. Average was defined as an age at PHV within ±1 SD of the mean age at PHV for the total sample of 58 players (13.60±0.85 years); delayed was a PHV >14.45 years and advanced was a PHV <12.75 years.

Linear mixed-effect models with the data grouped by subjects (random effects) were used with a maximum-likelihood estimator to evaluate the variance structure of the dependent variable, i.e., the differences between observed and predicted ages at PHV. Separate analyses were done for predictions with the Mirwald and Moore equations. The difference of the dependent variable from zero (so-called unconditional means model) was initially tested. In the second step, the effect of predictions at observations 1-5 (as a time variable) on the dependent variable at the population level (fixed effect) and at the intra-individual level (i.e., random slopes model, or unconditional growth model) were tested. Note that the term growth in the statistical context refers to the general change in a dependent variable with a time variable; it does not refer to growth in the biological sense. Finally, maturity status was added as a fixed factor to test the effect of variation in maturity timing on the predictions. The model was run separately for predictions with the Mirwald and Moore equations in the R-software (Cole, 2020) with the help of the nlme statistical package (Pinheiro, Bates, DebRoy, & Sarkar, 2020).

Weighted Deming regression using the jackknife procedure (Cornbleet & Gochman, 1979) was used to compare observed age at PHV estimated with the SITAR model (i.e., the reference) with predicted ages at PHV at each observation based on the two equations. The weighted Deming procedure considers both x (observed age at PHV) and y (predicted age at PHV) as subject to measurement error whereas simple regression permits only the y variable to have an associated error. With the weighted Deming regression, systematic differences between x and y are indicated by the intercept, while proportional differences are indicated by the slope.

Many applications of the equations use predicted maturity offset to classify youth as pre-PHV, at/circa/mid-PHV, or post-PHV using a band of -0.5 to +0.5 year to define the interval at PHV(Jakovljevic et al., 2016; López-Plaza, Alacid, Muyor, & López-Miñarro, 2017; Machado, Bonfim, & Costa, 2009; Peña-González, Fernández-Fernández, Cervelló, & Moya-Ramón, 2019; Read, Oliver, De Ste Croix, Myer, & Lloyd, 2017, 2018; Read, Oliver, Myer, De Ste Croix, & Lloyd, 2018; Rodríguez-Rosell et al., 2016; Živković et al., 2019); a band of -1.0 to +1.0 year is used less often (Buchheit & Mendez-Villanueva, 2013; Cripps, Hopper, & Joyce, 2016; Hammami, Chaouachi, Makhlouf, Granacher, & Behm, 2016; Mendez-Villanueva et al., 2010; Morris et al., 2020). On the other hand, some studies do not report the specific cut-offs that were used (Asadi, Ramirez-Campillo, Arazi, & Sáez de Villarreal, 2018; Brownstein, Ball, Micklewright, & Gibson, 2018; Lloyd et al., 2020). The standard errors of the prediction equations, 0.592 and 0.542 year, also approximate the narrow cut-offs used in many studies. Thus, the number and percentage of predicted ages at PHV with each equation within ±0.5 year of observed age at PHV at each observation were estimated for players of contrasting maturing status and for the total sample.

7.4. Results

Descriptive statistics for CA at prediction, observed and predicted maturity offset, predicted ages at PHV and the difference of predicted age at PHV minus observed age at PHV (the reference for comparison) with the original (Mirwald) and modified (Moore) equations in the soccer players are summarized in Table 7.1. Observed and predicted maturity offset increase linearly, on average, across the five observations. Predicted offset with the Moore equation is similar to actual offset at observation one but is then less than actual offset at subsequent observations. Predicted maturity offset with the Mirwald equation is less than actual offset and predicted offset with the Moore equation across the five observations.

Corresponding trends in predicted ages at PHV and the difference of predicted minus observed ages at PHV parallel those for maturity offset. Standard deviations for predicted maturity offset and ages at PHV with both equations are consistently lower than corresponding standard deviations for observed offset and age at PHV across the five observations. Variability is reduced more so with the Moore compared to the Mirwald equation.

Results of the mixed effects model indicate significant differences between observed and predicted ages at PHV with the Mirwald equation ($F=40.95$; $p<0.001$) and also with the Moore equation ($F=9.39$; $p<0.01$). The differences between observed and predicted ages at PHV with the respective equations at each observation increase significantly with subsequent observations with the Mirwald ($F=22.81$; $p<0.001$) and the Moore ($F=172.97$; $p<0.001$) equations, although the 95% confidence intervals indicate that the difference at observation 1 for the Moore equation is not different from zero.

Intercepts based on weighted Deming regressions for the two prediction equations are well above zero at each observation, indicating that the methods differ significantly by a constant error (Table 7.2). Estimated slopes for each regression are <1.0, indicating significant proportional differences between predicted ages at PHV with each equation and observed age at PHV. Overall, the results indicate systematic error for predicted ages at PHV.

Descriptive statistics for the three maturity groups are summarized in Table 7.3. Sample sizes and ages at PHV of players in each maturity timing group were as follows: advanced, n=8, 12.30±0.27 years; average, n=38, 13.50±0.51 years; delayed, n=12, 14.76±0.27 years.

Table 7.1. Sample sizes and descriptive statistics for chronological age (CA) at prediction, observed maturity offset and predicted maturity offset and ages at PHV, and the difference of predicted age at PHV minus observed ages at PHV (criterion) with the original (Mirwald) and modified (Moore) equations in soccer players at each observation.†

Observations	n	CA	Maturity Offset (yrs) Observed	Maturity Offset (yrs) Predicted Mirwald	Maturity Offset (yrs) Predicted Moore	Predicted age at PHV (yrs) Mirwald	Predicted age at PHV (yrs) Moore	Predicted minus Observed age at PHV Mirwald	Predicted minus Observed age at PHV Moore
1	58	11.9±0.5	-1.72±0.86	-2.09±0.51	-1.78±0.49	13.97±0.36	13.66±0.27	0.37±0.71	0.06±0.74
2	58	12.9±0.5	-0.70±0.86	-1.26±0.61	-0.93±0.56	14.15±0.45	13.82±0.35	0.56±0.63	0.23±0.68
3	58	13.9±0.5	0.29±0.86	-0.26±0.71	-0.01±0.57	14.14±0.51	13.90±0.37	0.55±0.56	0.30±0.64
4	55	14.9±0.5	1.33±0.86	0.71±0.66	0.89±0.55	14.17±0.52	13.99±0.36	0.62±0.59	0.44±0.72
5	40	15.9±0.5	2.34±0.87	1.63±0.52	1.72±0.50	14.25±0.43	14.16±0.32	0.71±0.82	0.65±0.92

†Observed (actual) maturity offset was calculated as CA at prediction minus observed age at PHV for each player, see text for details.

Table 7.2. Intercepts and slopes, and respective standard errors (SE) and 95% confidence limits based on the weighted Deming regression of predicted ages at PHV (y-axis) and observed (actual) ages at PHV (x-axis) for the Mirwald and Moore prediction equations at each observation (Obs) in youth soccer players.*

Obs	Mirwald						Moore					
	Intercept			Slope			Intercept			Slope		
	value	SE	(95% CL)	value	SE	95% CL)	value	SE	(95% CL)	value	SE	(95% CL)
1	10.26	0.77	(8.75; 11.81)	0.27	0.06	(0.15; 0.39)	11.17	0.53	(10.10; 12.24)	0.18	0.04	(0.10; 0.26)
2	8.35	0.94	(6.46; 10.24)	0.43	0.07	(0.28; 0.57)	9.86	0.67	(8.52; 11.20)	0.29	0.05	(0.19; 0.39)
3	7.01	0.77	(5.45; 8.57)	0.52	0.06	(0.41; 0.64)	9.36	0.62	(8.11; 1.62)	0.33	0.05	(0.24; 0.43)
4	7.26	0.77	(5.72; 8.79)	0.51	0.06	(0.40; 0.62)	10.49	0.76	(8.98; 12.01)	0.26	0.06	(0.15; 0.37)
5	11.18	1.07	(9.01; 13.35)	0.23	0.08	(0.06; 0.39)	13.58	0.97	(11.62; 15.54)	0.04	0.07	(-0.10; 0.19)

*All intercepts and slopes are significant in showing, respectively, systematic and proportional differences between predicted and observed ages at PHV with each equation.

Table 7.3. Sample sizes and descriptive statistics (mean ± standard deviation) for chronological age (CA) at prediction, observed maturity offset and predicted maturity offset, predicted ages at PHV and the difference of predicted age at PHV minus observed ages at PHV (criterion) with the original (Mirwald) and modified (Moore) equations at each observation in players classified as advanced, average and delayed based on observed ages at PHV.†

Obs	n	CA (yrs)	Maturity Offset (years)			Predicted APHV (years)		Predicted minus Observed APHV (years)	
			Observed	Predicted		Mirwald	Moore	Mirwald	Moore
				Mirwald	Moore				
Early									
1	8	11.4±0.5	-0.90±0.44	-2.14±0.63	-1.94±0.64	13.54±0.29	13.34±0.22	1.24±0.26	1.04±0.34
2	8	12.4±0.5	0.02±0.44	-1.14±0.70	-0.98±0.72	13.55±0.36	13.40±0.33	1.25±0.35	1.10±0.42
3	8	13.4±0.5	1.11±0.44	-0.13±0.68	-0.04±0.68	13.54±0.37	13.45±0.31	1.24±0.32	1.15±0.40
4	8	14.4±0.5	2.11±0.44	0.81±0.55	0.76±0.65	13.59±0.26	13.64±0.32	1.29±0.23	1.34±0.43
5	7	15.4±0.5	3.10±0.47	1.45±0.55	1.35±0.64	13.92±0.25	14.02±0.29	1.65±0.17	1.75±0.38
Average									
1	38	11.9±0.5	-1.57±0.76	-2.05±0.54	-1.73±0.49	13.98±0.32	13.66±0.24	0.48±0.53	0.16±0.53
2	38	12.9±0.5	-0.56±0.76	-1.21±0.64	-0.87±0.57	14.16±0.37	13.82±0.27	0.65±0.50	0.32±0.48
3	38	13.9±0.5	0.43±0.76	-0.15±0.72	0.08±0.57	14.08±0.39	13.86±0.27	0.58±0.46	0.36±0.48
4	37	14.9±0.5	1.46±0.75	0.85±0.63	1.00±0.50	14.09±0.35	13.94±0.28	0.61±0.55	0.46±0.58
5	26	15.9±0.5	2.39±0.87	1.70±0.49	1.77±0.40	14.23±0.37	14.16±0.31	0.69±0.72	0.62±0.72
Late									
1	12	12.0±0.4	-2.73±0.26	-2.18±0.38	-1.83±0.38	14.21±0.29	13.85±0.22	-0.55±0.34	-0.90±0.30
2	12	13.0±0.4	-1.72±0.26	-1.51±0.39	-1.09±0.39	14.55±0.30	14.13±0.23	-0.21±0.36	-0.63±0.31
3	12	14.0±0.4	-0.73±0.26	-0.70±0.50	-0.28±0.46	14.73±0.32	14.31±0.24	-0.03±0.38	-0.45±0.34
4	10	15.0±0.4	0.22±0.22	0.08±0.53	0.57±0.54	14.91±0.35	14.42±0.28	0.14±0.44	-0.35±0.43
5	7	16.2±0.4	1.40±0.26	1.58±0.62	1.88±0.57	14.63±0.51	14.33±0.37	-0.17±0.49	-0.48±0.47

Obs (observations); †Players were classified as late, average or early maturing on the basis of their observed age at PHV (SITAR model) relative to age at PHV for the total sample of soccer players- see text for details. Observed (actual) maturity offset was calculated as CA at prediction minus observed age at PHV for each player.

Results of the mixed-effects model comparing the three maturity groups indicate that maturity status as a fixed factor has a significant effect on the dependent variable (predicted ages at PHV) with both the Mirwald ($F=36.85$, $p<0.001$) and the Moore ($F=51.28$, $p<0.001$) equations. By inference, predicted ages at PHV differ relative to observed age at PHV in each group. However, the interaction between observation and maturity group is significant for the Mirwald equation ($F=5.39$, $p=0.005$) and indicates different slopes of change in predicted ages at PHV with consecutive observations in the maturity groups. The latter reflects the trend in differences between predicted and observed ages at PHV for the Mirwald equation which are not significant except at observation 1. In contrast, the interaction term is not significant for the Moore equation.

In the context of the results of the mixed-effects model, differences between predicted and observed ages at PHV are significant and positive across the five observations among players advanced in maturity timing (i.e., early ages at PHV). The predicted ages at PHV are consistently later observed ages at PHV.

Results are similar for players maturing on time (average), i.e., predicted ages at PHV with the two equations are later than observed age at PHV. Across the five observations, the differences between predicted and observed ages at PHV with each equation are significant, although predicted age at PHV with the Moore equation at observation one approaches that for observed age at PHV.

Among late maturing players, in contrast, differences between predicted and observed ages at PHV for the Mirwald equation are not significant except at observation one, while differences between predicted and observed ages at PHV for the Moore equation are significant at each observation. The differences between predicted and observed ages at PHV with the Mirwald equation are also smaller than corresponding differences with the Moore equation.

Predicted ages at PHV (y-axis) for individual players with the Mirwald and Moore equations are illustrated relative to their respective observed ages at PHV (x-axis) in Figure 7.1 and 7.2, respectively. Intra-individual variation in predicted ages at PHV is considerable and ranges of predicted ages are reduced with the Moore equation. Relatively few predicted ages approximate observed ages at PHV in early and late maturing players.

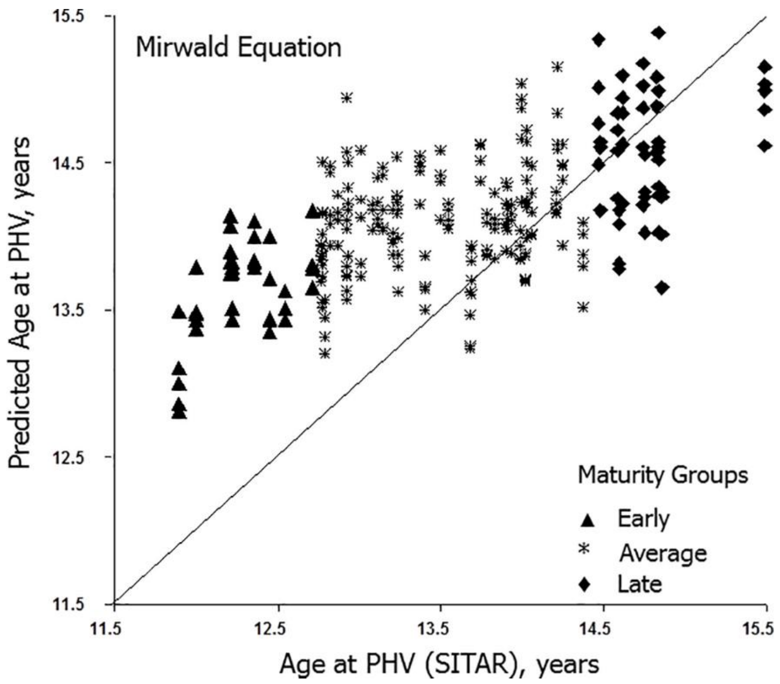


Figure 7.1. Predicted ages at PHV with the Mirwald equation plotted relative of observed age at PHV at each observation for individual soccer players classified as early, average and late maturing. The diagonal corresponds to the line of identity (x=y).

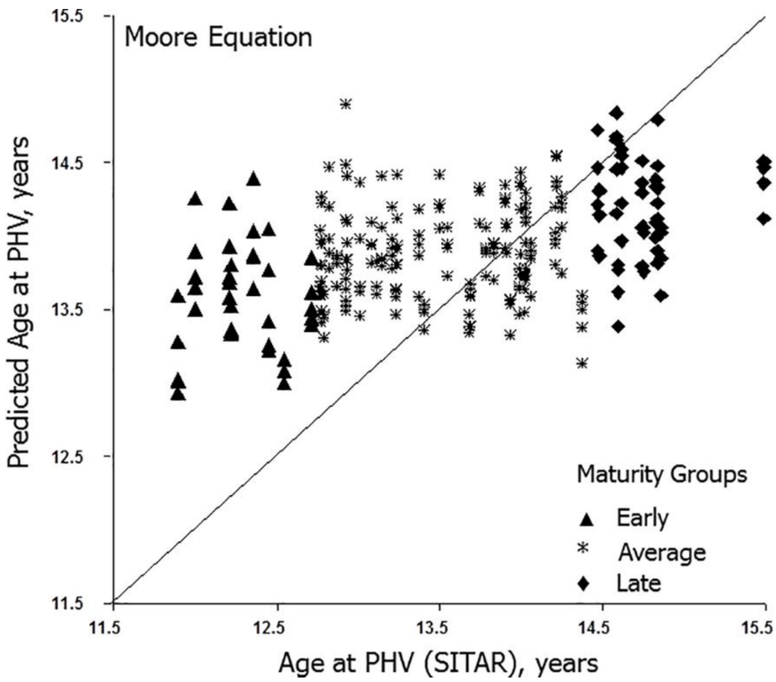


Figure 7.2. Predicted ages at PHV with the Moore equation plotted relative of observed age at PHV at each observation for individual soccer players classified as early, average and late maturing. The diagonal corresponds to the line of identity (x=y).

Across the five observations (Table 7.4), no predicted ages at PHV with the Mirwald equation are within ±0.5 year of observed age at PHV (SITAR) among the eight early maturing players (0 of 39). Corresponding estimates for predicted ages at PHV with the Mirwald equation within ±0.5 year of observed age at PHV across the five observations are 35 of 53 (66%) among late and 76 of 177 (43%) among average maturing players. For the total sample, 111 of 269 (41%) of predicted ages at PHV with the Mirwald equation are within ±0.5 year of observed age at PHV.

For the Moore equation, only 1 of 39 predicted ages at PHV (3%) is within ±0.5 year of observed age at PHV among early maturing players. On the other hand, 88 of 177 predicted ages at PHV (50%) among average and 23 of 53 predicted ages at PHV (43%) among late maturing players are within ±0.5 year of observed age at PHV. For the total sample, 112 of 269 (42%) of predicted ages at PHV with the Moore equation are within ±0.5 year of observed age at PHV.

Table 7.4. Number of participants by maturity status[†] according to predicted ages at PHV with the Mirwald equation and, separately, with the Moore equation who were within ±0.50 year of observed age at PHV (SITAR model) at each observation in youth soccer players.

Equation	Observations	Maturity Groups							
		Advanced		Average		Delayed		Total	
		N	n	N	n	N	n	N	n
Mirwald	1	8	0	38	17	12	6	58	23
	2	8	0	38	15	12	8	58	23
	3	8	0	38	18	12	10	58	28
	4	8	0	37	15	10	7	55	22
	5	7	0	26	11	7	4	40	15
	Total	39	0	177	76	53	35	269	111
		0%		43%		66%		41%	
Moore	1	8	0	38	18	12	1	58	19
	2	8	1	38	21	12	5	58	27
	3	8	0	38	19	12	7	58	26
	4	8	0	37	18	10	7	55	25
	5	7	0	26	12	7	3	40	15
	Total	39	1	177	88	53	23	269	112
	(%)	3%		50%		43%		42%	

[†]Players were classified as advanced, average or delayed on the basis of their observed age at PHV (SITAR model) relative to age at PHV for the total sample - see text for details.

7.5. Discussion

The 58 players comprising the present study was larger than samples in five other longitudinal studies of European youth soccer players, 8 to 33 players (Study 4 of the present thesis). Results of the application of the maturity offset prediction equations in the longitudinal series of Portuguese youth soccer players were consistent with recent studies of English (Parr et al., 2020) and Dutch (Teunissen et al., 2020) soccer players. The study of English players was limited to the Mirwald (Mirwald et al., 2002) equation, while that of the Dutch players considered the original and modified (Moore et al., 2015) equations in addition to an equation which predicted a maturity ratio (Fransen et al., 2018). Although the three studies varied in design, scope and focus, the results were consistent in highlighting major limitations of predicted maturity offset and predicted age at PHV in longitudinal samples of soccer players.

Applications of the original and modified equations in the longitudinal series of Portuguese youth soccer players were also consistent with validation studies of the maturity offset prediction equations in three longitudinal series of youth spanning late childhood through adolescence, one in the U.S. (Malina et al., 2016) and two in Poland (Kozieł & Malina, 2018; Malina et al., 2021). The three studies and the present study of soccer players used similar analytical methods and noted several major limitations of the prediction equations.

First, predicted maturity offset increased, on average, with CA at prediction throughout the age range considered in each study. In the study of soccer players, mean predicted ages at PHV based on the Moore equation increased, on average, with CA, while those based on the Mirwald equation increased from observation one to two, changed negligibly through observation four and then increased to observation five. The age-related trend probably reflects the predictors comprising each equation which increase, on average, with CA. The preceding is apparent in the correlations between predicted maturity offset and predicted age at PHV with CA, height, sitting height, estimated leg length and weight at each observation for the Mirwald equation and with CA and height with the Moore equation (Table 7.5). For predicted maturity offset, correlations within each CA group range from moderately high to high; correlations are highest for sitting height and tend to be lowest for estimated leg length. For predicted age at PHV, correlations are relatively low and positive for CA, but are negative and moderate to high for the anthropometric variables. Within an age group, taller and heavier players tended to have an earlier predicted age at PHV.

Table 7.5. Correlations at each observation between predicted maturity offset and predicted APHV (CA – predicted maturity offset) with chronological age (CA), height (Ht), sitting height (SitHt), estimated leg length (LegLt) and body weight (Wt) for the Mirwald equation and with CA and Ht for the Moore equation.

		Predicted Maturity Offset							Predicted APHV						
		Mirwald					Moore		Mirwald					Moore	
Observat	n	C	Ht	Sit	Leg	W	C	Ht	CA	Ht	Sit	Leg	W	CA	Ht
1	5	0.	0.	0.9	0.6	0.	0.	0.	0.3	-	-	-	-	0.3	-
2	5	0.	0.	0.9	0.7	0.	0.	0.	0.1	-	-	-	-	0.1	-
3	5	0.	0.	0.9	0.6	0.	0.	0.	0.0	-	-	-	-	0.1	-
4	5	0.	0.	0.9	0.5	0.	0.	0.	0.1	-	-	-	-	0.2	-
5	4	0.	0.	0.8	0.3	0.	0.	0.	0.4	-	-	-	-	0.4	-

‡Not significant; all other correlations are significant.

Second, variation in predicted maturity offset and ages at PHV within CA groups was consistently reduced compared to variation in observed ages at offset and at PHV. Variation was reduced more so with the Moore compared to the Mirwald equation.

Third, predictions varied with maturity status defined by observed age at PHV. Predicted maturity offset and ages at PHV were consistently later than observed maturity offset and age at PHV among early maturing boys, and earlier than observed in late maturing boys. By inference, maturity status defined by observed age at PHV influenced predicted ages at PHV in both early and late maturing boys. It should be noted, however, that Moore et al. [Moore et al., 2015, p. 1761] cautioned that "Our sample was not large enough to rigorously assess variation in prediction error due to early- and late-maturing children". This caution, however, is overlooked in applications of the equations. In contrast to early and late maturing youth, predicted ages at PHV appeared to be reasonably accurate for average maturing boys within approximately ±1 year of observed PHV. Unfortunately, the maturity status and/or timing of individuals is not ordinarily known in studies applying the prediction protocols.

Fourth, intra-individual variation in predicted ages at PHV with each equation was considerable in the present sample of soccer players and in each of the longitudinal studies. Ranges of predicted ages at PHV were reduced with the Moore compared to the Mirwald equation.

The dependency of predicted maturity offset upon CA and body size at prediction merits attention. Means for predicted maturity offset are plotted relative to means for CA and height in the present study and in samples of male soccer players extracted from the literature (Agostinete et al., 2020; Aquino et al., 2017; Borges et al., 2018; Campa et al., 2019; Deprez et al., 2015; Deprez, Fransen, Lenoir, Philippaerts, & Vaeyens, 2015; Doncaster, Iga, & Unnithan, 2018; Gibson, McCunn, MacNay, Mullen, & Twist, 2018; Gibson, Henning, & Twist, 2018; Gil et al., 2014; Gonçalves, Severino, Silva, & Figueiredo, 2011; Hernandez Camacho, Huelva Leal, Martinez-Sanz, Lahoz Ruano, & Vázquez Carrión, 2018; Laas et al., 2020; Lloyd et al., 2015; Lovell, Bocking, Fransen, & Coutts, 2018; Lovell, Bocking, Fransen, Kempton, & Coutts, 2018; Moreira et al., 2013; Parr et al., 2020; Read, Oliver, De Ste Croix, et al., 2018; Seabra et al., 2012; Trecoci, Longo, Perri, Iaia, & Alberti, 2019) are illustrated in Figure 7.3 and 7.4. The means plotted in the figures were limited to studies using the Mirwald equation, as it was more widely used in studies of soccer players. Predicted maturity offset increased linearly with CA and with height at prediction. The plotted means were largely based on one- or two-year CA groups, although several were based on players spanning age ranges of three or more years.

Studies reporting maturity offset by relative age categories within an age group and studies classifying players across variable age ranges as pre-, at/circa- or post-PHV based on predicted offset were not included in the graphs. In the former, players born early in the year were, on average, older and taller than those born later in the year, while in the latter, CA and height systematically increased from pre-, to at/circa- to post-PHV groups (see above). Given the trends, studies applying predicted maturity offset as a maturity indicator beg the following question: Is predicted maturity offset an indicator of time before age at PHV or is it an indicator of CA and size at prediction? By inference, the validity of predicted maturity offset as an indicator of the time before or after PHV can be questioned.

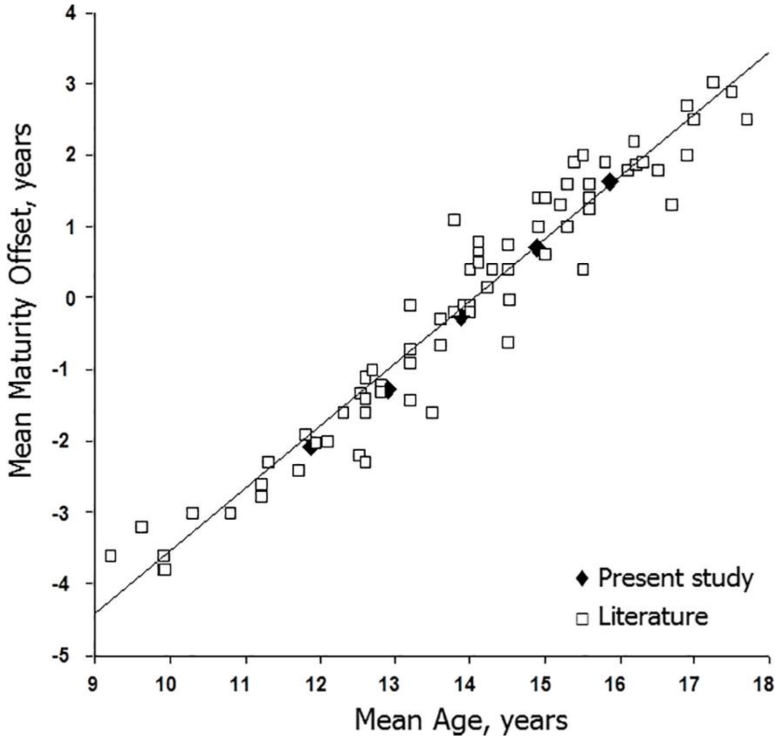


Figure 7.3. Means for predicted maturity offset (MO) plotted relative to means for chronological age for each year of observation in the current sample of players (filled diamonds) and in samples of male soccer players extracted from the literature.

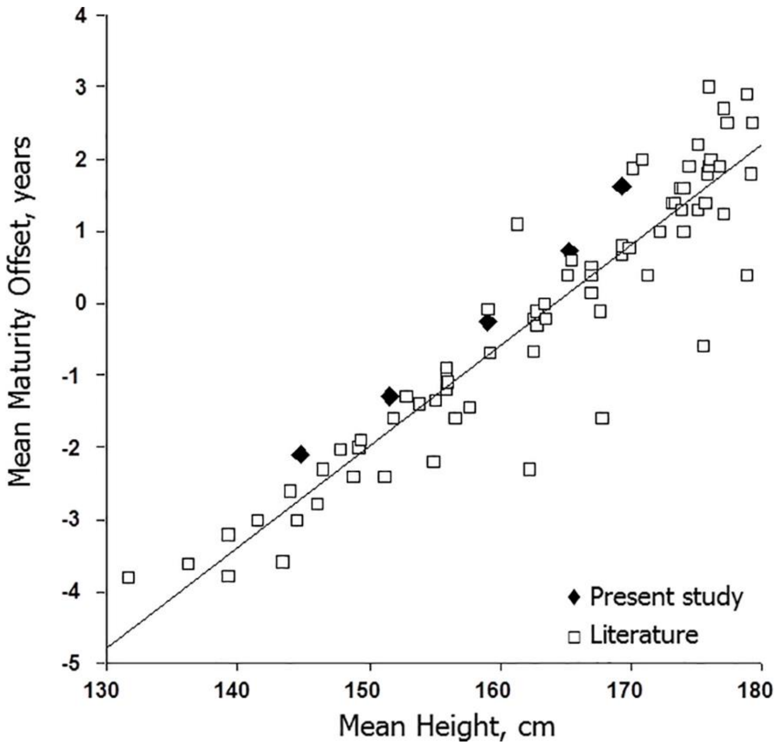


Figure 7.4. Means for predicted maturity offset (MO) plotted relative to means for height for each year of observation in the current sample of players (filled diamonds) and in samples of male soccer players extracted from the literature.

The preceding has implications for studies using predicted maturity offset as an indicator of maturity status among youth athletes in soccer and other sports. Predicted maturity offset is used most often to classify youth as pre-, at-/circa- and post- PHV, although mean CAs, heights and weights show, on average, a clear gradient across the respective maturity groups. Many studies simply compare the three groups with analysis of variance without controlling for the variation in CA and body size among groups (Jakovljevic et al., 2016; López-Plaza et al., 2017; Peña-González et al., 2019; Till & Jones, 2015; Zago et al., 2020). It is also unclear as to how CA-related variation in predicted offset or ages at PHV was addressed in studies applying the prediction equations in short-term longitudinal studies (Matthys et al., 2013; Till, Cobley, O'Hara, Chapman, & Cooke, 2013; Zuber, Zibung, & Conzelmann, 2016).

Although studies of youth athletes do not ordinarily indicate the ethnic composition of samples, the issue of ethnic variation is relevant as the maturity offset prediction equations were developed and validated on samples of European ancestry. The original Mirwald equation (Mirwald et al., 2002) requires sitting height and estimated leg length, while one of the Moore equations (Moore et al., 2015) for boys requires age and sitting height. Of potential relevance, population variability in the proportions of sitting height and estimated leg length to standing height is reasonably well established (Eveleth & Tanner, 1976, 1990). American youth of European (White), African (Black) and Hispanic ancestry, for example, vary in the proportions of sitting height and estimated leg length (Malina, Brown, & Zavaleta, 1987; Martorell, Malina, Castillo, Mendoza, & Pawson, 1988). The proportions of the Portuguese youth soccer players in the present study, as reflected in the sitting height-height ratio, were, on average, generally similar to those for American White and Hispanic youth, but different from American Black youth who have proportionally longer lower extremities. This trend was also noted in a recent study of soccer players in which players of non-European ancestry were taller with a lower sitting height/height ratio, i.e., proportionally longer legs, than players of European ancestry (Parr et al., 2020).

Results of the current study also have practical implications for those working with youth athletes. The interval of PHV is central to the Long-Term Athlete Development (LTAD) model for youth athletes (Balyi et al., 2005; Balyi & Hamilton, 2004), which calls for identifying youth of contrasting maturity status, i.e., early, average or late maturing. The LTAD, however, does not specify the method for doing so other than suggesting the monitoring of estimated growth velocities; the latter, however, have limitations over the short term. Estimated increments over short intervals (3-4 months), however, must be interpreted with care; they must be adjusted for the interval between measurements and evaluated relative to factors which influence short term estimates of growth rate – specifically measurement errors (both inter- and intra-observer), diurnal variation and perhaps seasonal variation (Malina et al., 2004). Nevertheless, application of predicted maturity offset in this context has the potential for misclassification and thus implications for player development. Those using predicted maturity offset per se or variations of the method to identify when players enter and exit the interval of the adolescent growth spurt should employ these methods with caution. If predicted offset is used to inform training design and prescription, it is essential that variation in chronological age at prediction and error associated with the prediction equations be considered. Perhaps additional or alternative methods might be used as a complement, for example, percentage of predicted adult stature attained at the time of observation. As noted above, the utility of estimated velocities of growth in height based on short term height increments has limitations.

The inability of the maturity offset prediction methods to effectively differentiate between early and late maturing youth implies that they should not be used to group players by maturity status as in bio-banding (Malina et al., 2019), or to adjust fitness and performance scores to accommodate individual differences in maturation (Abbott et al., 2021; Till & Jones, 2015). As age at PHV is over-estimated in early and under-estimated in late maturing youth, the majority of these players will likely be categorized as being on time and some will be grouped in equivalent bands. Similarly, maturity associated adjustments to performance or fitness scores in early and late maturing boys will, by virtue of these biases, be attenuated and regress towards a common mean.

7.6. Conclusions

Results of the study of applying maturity offset prediction equations to the sample of Portuguese soccer players were consistent with similar studies of soccer players and of the general population. Predicted maturity offset increased, on average, with CA at prediction throughout the age range considered. Variation in predicted maturity offset and ages at PHV within CA groups was consistently reduced compared to variation in observed ages at offset and at PHV. Predictions also varied with maturity status defined by observed age at PHV; predicted maturity offset and ages at PHV were consistently later than observed maturity offset and age at PHV among early maturing boys, and earlier than observed in late maturing boys. And, intra-individual variation in predicted ages at PHV with each equation was considerable and ranges of predicted ages at PHV were reduced more with the Moore compared to the Mirwald equations.

7.7. References

Abbott, S., Hogan, C., Castiglioni, M. T., Yamauchi, G., Mitchell, L. J. G., Salter, J., … Cobley, S. (2021). Maturity-related developmental inequalities in age-group swimming: The testing of “Mat-CAPs” for their removal. Journal of Science and Medicine in Sport, 24(4), 397–404.

Agostinete, R. R., Fernandes, R. A., Narciso, P. H., Maillane-Vanegas, S., Werneck, A. O., & Vlachopoulos, D. (2020). Categorizing 10 Sports According to Bone and Soft Tissue Profiles in Adolescents. Medicine and Science in Sports and Exercise, 52(12), 2673–2681.

Aquino, R., Alves, I. S., Padilha, M. B., Casanova, F., Puggina, E. F., & Maia, J. (2017). Multivariate Profiles of Selected versus Non-Selected Elite Youth Brazilian Soccer Players. Journal of Human Kinetics, 60, 113–121.

Asadi, A., Ramirez-Campillo, R., Arazi, H., & Sáez de Villarreal, E. (2018). The effects of maturation on jumping ability and sprint adaptations to plyometric training in youth soccer players. Journal of Sports Sciences, 36(21), 2405–2411.

Balyi, I., Cardinal, C., Higgs, C., Norris, S., & Way, R. (2005). Canadian Sport for Life: Long-term athlete development resource paper V2. Vancouver, BC: Canadian Sport Centres.

Balyi, I., & Hamilton, A. (2004). Long-Term Athlete Development: Trainability in Childhood and Adolescence–Windows of Opportunity, Optimal Trainability. Victoria, BC: National Coaching Institute British Columbia and Advanced Training and Performance Ltd.

Borges, P. H., Cumming, S., Ronque, E. R. V, Cardoso, F., Avelar, A., Rechenchosky, L., … Rinaldi, W. (2018). Relationship Between Tactical Performance, Somatic Maturity and Functional Capabilities in Young Soccer Players. Journal of Human Kinetics, 64, 160–169.

Brownstein, C. G., Ball, D., Micklewright, D., & Gibson, N. V. (2018). The Effect of Maturation on Performance During Repeated Sprints With Self-Selected Versus Standardized Recovery Intervals in Youth Footballers. Pediatric Exercise Science, 30(4), 500–505.

Buchheit, M., & Mendez-Villanueva, A. (2013). Reliability and stability of anthropometric and performance measures in highly-trained young soccer players: effect of age and maturation. Journal of Sports Sciences, 31(12), 1332–1343.

Campa, F., Silva, A. M., Iannuzzi, V., Mascherini, G., Benedetti, L., & Toselli, S. (2019). The Role of Somatic Maturation on Bioimpedance Patterns and Body Composition in Male Elite Youth Soccer Players. International Journal of Environmental Research and Public Health, 16(23), 1–11.

Cao, Z., Hui, L. L., & Wong, M. Y. (2018). IAPVBS: Individual Age at Peak Velocity Based on SITAR. Retrieved from https://rdrr.io/github/Zhiqiangcao/iapvbs

Cole, T. J. (2020). SITAR: Super Imposition by Translation and Rotation Growth Curve Analysis. Retrieved from https://cran.r-project.org/package=sitar

Cole, T. J., Donaldson, M. D. C., & Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. International Journal of Epidemiology, 39(6), 1558–1566.

Cornbleet, P. J., & Gochman, N. (1979). Incorrect least-squares regression coefficients in method-comparison analysis. Clinical Chemistry, 25(3), 432–438.

Cripps, A. J., Hopper, L., & Joyce, C. (2016). Maturity, Physical Ability, Technical Skill and Coaches' Perception of Semi-Elite Adolescent Australian Footballers. Pediatric Exercise Science, 28(4), 535–541.

Deprez, D., Fransen, J., Boone, J., Lenoir, M., Philippaerts, R., & Vaeyens, R. (2015). Characteristics of high-level youth soccer players: variation by playing position. Journal of Sports Sciences, 33(3), 243–254.

Deprez, D. N., Fransen, J., Lenoir, M., Philippaerts, R. M., & Vaeyens, R. (2015). A retrospective study on anthropometrical, physical fitness, and motor coordination characteristics that influence dropout, contract status, and first-team playing time in high-level soccer players aged eight to eighteen years. Journal of Strength and Conditioning Research, 29(6), 1692–1704.

Doncaster, G., Iga, J., & Unnithan, V. (2018). Assessing Differences in Cardiorespiratory Fitness With Respect to Maturity Status in Highly Trained Youth Soccer Players. Pediatric Exercise Science, 30(2), 216–228.

Eveleth, P. B., & Tanner, J. M. (1976). Worldwide Variation in Human Growth. Cambridge: Cambridge University Press.

Eveleth, P. B., & Tanner, J. M. (1990). Worldwide Variation in Human Growth (2nd ed.). Cambridge: Cambridge University Press.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009). Youth soccer players, 11-14 years: maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Fransen, J., Bush, S., Woodcock, S., Novak, A., Deprez, D., Baxter-Jones, A. D. G., … Lenoir, M. (2018). Improving the Prediction of Maturity From Anthropometric Variables Using a Maturity Ratio. Pediatric Exercise Science, 30(2), 296–307.

Gibson, NV, McCunn, R., MacNay, S., Mullen, T., & Twist, C. (2018). Playing exposure does not affect movement characteristics or physiological responses of elite youth footballers during an intensified period of competition. Science and Medicine in Football, 1–6.

Gibson, Neil V., Henning, G., & Twist, C. (2018). Movement characteristics, physiological and perceptual responses of elite standard youth football players to different high intensity running drills. Science and Medicine in Football, 2(4), 281–287.

Gil, S. M., Zabala-Lili, J., Bidaurrazaga-Letona, I., Aduna, B., Lekue, J. A., Santos-Concejero, J., & Granados, C. (2014). Talent identification and selection process of outfield players and goalkeepers in a professional soccer club. Journal of Sports Sciences, 32(20), 1931–1939.

Gonçalves, R. R., Severino, V., Silva, M. J. C. e, & Figueiredo, A. J. (2011). Age-related variation in anthropometric and maturity characteristics of soccer goalkeepers aged 11-14 years. Annals of Research in Sport and Physical Activity, (1), 70–81.

Hammami, R., Chaouachi, A., Makhlouf, I., Granacher, U., & Behm, D. G. (2016). Associations Between Balance and Muscle Strength, Power Performance in Male Youth Athletes of Different Maturity Status. Pediatric Exercise Science, 28(4), 521–534.

Harriss, D. J., Macsween, A., & Atkinson, G. (2019). Ethical Standards in Sport and Exercise Science Research: 2020 Update. International Journal of Sports Medicine, 40(13), 813–817.

Hernandez Camacho, J., Huelva Leal, A., Martinez-Sanz, J., Lahoz Ruano, M., & Vázquez Carrión, J. (2018). Peak height velocity and muscle mass in young soccer players. Revista Española de Nutrición Humana y Dietética, 22, 219–226.

Jakovljevic, S., Macura, M., Radivoj, M., Jankovic, N., Pajic, Z., & Erculj, F. (2016). Biological Maturity Status and Motor Performance in Fourteen-Year-old Basketball Players. International Journal of Morphology, 34(2), 637–643.

Kozieł, S. M., & Malina, R. M. (2018). Modified Maturity Offset Prediction Equations: Validation in Independent Longitudinal Samples of Boys and Girls. Sports Medicine, 48(1), 221–236.

Laas, M. M., Wright, M. D., McLaren, S. J., Eaves, D. L., Parkin, G., & Portas, M. D. (2020). Motion tracking in young male football players: a preliminary study of within-session movement reliability. Science and Medicine in Football, 4(3), 203–210.

Lloyd, R. S., Oliver, J. L., Myer, G. D., De Ste Croix, M. B., Wass, J., & Read, P. J. (2020). Comparison of Drop Jump and Tuck Jump Knee Joint Kinematics in Elite Male Youth Soccer Players: Implications for Injury Risk Screening. Journal of Sport Rehabilitation, 29(6), 760–765.

Lloyd, R. S., Oliver, J. L., Radnor, J. M., Rhodes, B. C., Faigenbaum, A. D., & Myer, G. D. (2015). Relationships between functional movement screen scores, maturation and physical performance in young soccer players. Journal of Sports Sciences, 33(1), 11–19.

López-Plaza, D., Alacid, F., Muyor, J. M., & López-Miñarro, P. Á. (2017). Sprint kayaking and canoeing performance prediction based on the relationship between maturity status, anthropometry and physical fitness in young elite paddlers. Journal of Sports Sciences, 35(11), 1083–1090.

Lovell, T. W. J., Bocking, C. J., Fransen, J., & Coutts, A. J. (2018). A multidimensional approach to factors influencing playing level and position in a school-based soccer programme. Science and Medicine in Football, 2(3), 237–245.

Lovell, T. W. J., Bocking, C. J., Fransen, J., Kempton, T., & Coutts, A. J. (2018). Factors affecting physical match activity and skill involvement in youth soccer. Science and Medicine in Football, 2(1), 58–65.

Machado, D. R. L., Bonfim, M. R. B., & Costa, L. T. (2009). Pico de velocidade de crescimento como alternativa para classificação maturacional associada ao desempenho motor. Rev. Bras. Cineantropom. Desempenho Hum, 11(1), 14–21.

Malina, R M, Brown, K. H., & Zavaleta, A. N. (1987). Relative lower extremity length in Mexican American and in American black and white youth. American Journal of Physical Anthropology, 72(1), 89–94.

Malina, R. M.. (2014). Top 10 research questions related to growth and maturation of relevance to physical activity, performance, and fitness. Research Quarterly for Exercise and Sport, 85(2), 157–173.

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, Maturation, and Physical Activity (2nd ed.). Ilinois: Human Kinetics.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of maturity offset in the fels longitudinal study. Pediatric Exercise Science, 28(3), 439–455.

Malina, R. M., Claessens, A. L., Van Aken, K., Thomis, M., Lefevre, J., Philippaerts, R. M., & Beunen, G. P. (2006). Maturity Offset in Gymnasts. Medicine & Science in Sports & Exercise, 38(7), 1342–1347.

Malina, R. M., Coelho-e-Silva, M. J., Figueiredo, A. J., Carling, C., & Beunen, G. P. (2012). Interrelationships among invasive and non- invasive indicators of biological maturation in adolescent male soccer players. Journal of Sports Science, 30(15), 1705–1717.

Malina, R. M., Cumming, S. P., Rogol, A. D., Coelho-e-Silva, M. J., Figueiredo, A. J., Konarski, J. M., & Kozieł, S. M. (2019). Bio-Banding in Youth Sports: Background, Concept, and Application. Sports Medicine, 49(11), 1671–1685.

Malina, R. M., & Kozieł, S. M. (2014a). Validation of maturity offset in a longitudinal sample of Polish boys. Journal of Sports Sciences, 32(5), 424–437.

Malina, R. M., & Kozieł, S. M. (2014b). Validation of maturity offset in a longitudinal sample of Polish girls. Journal of Sports Sciences, 32(14), 1374–1382.

Malina, R. M., Kozieł, S. M., Králik, M., Chrzanowska, M., & Suder, A. (2021). Prediction of maturity offset and age at peak height velocity in a longitudinal series of boys and girls. American Journal of Human Biology, 33(6).

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho-e-Silva, M. J., & Figueiredo, A. J. (2015). Biological maturation of youth athletes: Assessment and implications. British Journal of Sports Medicine, 49(13), 852–859.

Martorell, R., Malina, R. M., Castillo, R. O., Mendoza, F. S., & Pawson, I. G. (1988). Body proportions in three ethnic groups: children and youths 2-17 years in NHANES II and HHANES. Human Biology, 60(2), 205–222.

Matthys, S. P. J., Vaeyens, R., Fransen, J., Deprez, D., Pion, J., Vandendriessche, J., … Philippaerts, R. (2013). A longitudinal study of multidimensional performance characteristics related to physical capacities in youth handball. Journal of Sports Sciences, 31(3), 325–334.

Mendez-Villanueva, A., Buchheit, M., Kuitunen, S., Poon, T. K., Simpson, B., & Peltola, E. (2010). Is the Relationship Between Sprinting and Maximal Aerobic Speeds in Young Soccer Players Affected by Maturation? Pediatric Exercise Science, 22(4), 497–510.

Mirwald, R. L., Baxter-jones, A., Bailey, D., & Beunen, G. (2002). An assessment of maturity from anthropometric measurements. Medicine & Science in Sports & Exercise, 34(4), 689–694.

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D. G., Cameron, N., & Brasher, P. M. A. (2015). Enhancing a somatic maturity prediction model. Medicine and Science in Sports and Exercise, 47(8), 1755–1764.

Moran, J., Sandercock, G. R. H., Ramírez-Campillo, R., Meylan, C., Collison, J., & Parry, D. A. (2017). A meta-analysis of maturation-related variation in adolescent boy athletes' adaptations to short-term resistance training. Journal of Sports Sciences, 35(11), 1041–1051.

Moran, J., Sandercock, G., Rumpf, M., & Parry, D. (2016). Variation in Responses to Sprint Training in Male Youth Athletes: A Meta-analysis. International Journal of Sports Medicine, 38(01), 1–11.

Moreira, A., Mortatti, A., Aoki, M., Arruda, A., Freitas, C., & Carling, C. (2013). Role of free testosterone in interpreting physical performance in elite young Brazilian soccer players. Pediatric Exercise Science, 25(2), 186–197.

Morris, R. O., Jones, B., Myers, T., Lake, J., Emmonds, S., Clarke, N. D., … Till, K. (2020). Isometric Midthigh Pull Characteristics in Elite Youth Male Soccer Players: Comparisons by Age and Maturity Offset. Journal of Strength and Conditioning Research, 34(10), 2947–2955.

Myburgh, G. K., Cumming, S. P., & Malina, R. M. (2019). Cross-Sectional Analysis Investigating the Concordance of Maturity Status Classifications in Elite Caucasian Youth Tennis Players. Sports Medicine - Open, 5(1), 27.

Parr, J., Winwood, K., Hodson-Tole, E., Deconinck, F. J. A., Parry, L., Hill, J. P., … Cumming, S. P. (2020). Predicting the timing of the peak of the pubertal growth spurt in elite male youth soccer players: evaluation of methods. Annals of Human Biology, 47(4), 400–408.

Peña-González, I., Fernández-Fernández, J., Cervelló, E., & Moya-Ramón, M. (2019). Effect of biological maturation on strength-related adaptations in young soccer players. PLOS ONE, 14(7), e0219355.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2020). nlme: Linear and Nonlinear Mixed Effects Models. Retrieved from https://cran.r-project.org/package=nlme.

R Core Team. (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Read, P. J., Oliver, J., De Ste Croix, M., Myer, G., & Lloyd, R. (2017). Hopping and Landing Performance in Male Youth Soccer Players: Effects of Age and Maturation. International Journal of Sports Medicine, 38(12), 902–908.

Read, P. J., Oliver, J. L., De Ste Croix, M. B. A., Myer, G. D., & Lloyd, R. S. (2018). Landing Kinematics in Elite Male Youth Soccer Players of Different Chronologic Ages and Stages of Maturation. Journal of Athletic Training, 53(4), 372–378.

Read, P. J., Oliver, J. L., Myer, G. D., De Ste Croix, M. B. A., & Lloyd, R. S. (2018). The Effects of Maturation on Measures of Asymmetry During Neuromuscular Control Tests in Elite Male Youth Soccer Players. Pediatric Exercise Science, 30(1), 168–175.

Rodríguez-Rosell, D., Franco-Márquez, F., Pareja-Blanco, F., Mora-Custodio, R., Yáñez-García, J. M., González-Suárez, J. M., & González-Badillo, J. J. (2016). Effects of 6 Weeks Resistance Training Combined With Plyometric and Speed Exercises on Physical Performance of Pre-Peak-Height-Velocity Soccer Players. International Journal of Sports Physiology and Performance, 11(2), 240–246.

Seabra, A., Marques, E., Brito, J., Krustrup, P., Abreu, S., Oliveira, J., … Rebelo, A. (2012). Muscle strength and soccer practice as major determinants of bone mineral density in adolescents. Joint Bone Spine, 79(4), 403–408.

Teunissen, J. W., Rommers, N., Pion, J., Cumming, S. P., Rössler, R., D'Hondt, E., … Malina, R. M. (2020). Accuracy of maturity prediction equations in individual elite male football players. Annals of Human Biology, 47(4), 409–416.

Till, K., Cobley, S., O'Hara, J., Chapman, C., & Cooke, C. (2013). A longitudinal evaluation of anthropometric and fitness characteristics in junior rugby league players considering playing position and selection level. Journal of Science and Medicine in Sport, 16(5), 438–443.

Till, K., & Jones, B. (2015). Monitoring anthropometry and fitness using maturity groups within youth rugby league. Journal of Strength and Conditioning Research, 29(3), 730–736.

Trecoci, A., Longo, S., Perri, E., Iaia, F., & Alberti, G. (2019). Field-based physical performance of elite and sub-elite middle-adolescent soccer players. Research in Sports Medicine, 33, 1232–1236.

Valente-dos-Santos, J., Coelho-e-Silva, M. J., Simões, F., Figueiredo, A. J., Leite, N., Elferink-Gemser, M. T., … Sherar, L. (2012). Modeling Developmental Changes in Functional Capacities and Soccer-Specific Skills in Male Players Aged 11-17 Years. Pediatric Exercise Science, 24(4), 603–621.

Zago, M., Moorhead, A. P., Bertozzi, F., Sforza, C., Tarabini, M., & Galli, M. (2020). Maturity offset affects standing postural control in youth male soccer players. Journal of Biomechanics, 99, 109523.

Živković, M., Stojiljković, N., Antić, V., Pavlović, L., Stanković, N., & Jorgić, B. (2019). The motor abilities of handball players of different biological maturation. Facta Universitatis, Series: Physical Education and Sport, 17(1), 125–133.

Zuber, C., Zibung, M., & Conzelmann, A. (2016). Holistic Patterns as an Instrument for Predicting the Performance of Promising Young Soccer Players - A 3-Years Longitudinal Study. Frontiers in Psychology, 7, 1088.

Chapter VIII

General Discussion

8. GENERAL DISCUSSION

8.1. The importance of data quality in the determination of skeletal age

Advances in digital imaging technologies and machine learning led to the development of artifacts, such as BoneXpert, that automatically estimate skeletal age (SA) from digitalized skeletal hand-wrist radiographs (Thodberg et al., 2009). BoneXpert uses a three-layer imaging process to reconstruct the bone borders and architecture. It adjusts SA to the Greulich-Pyle method and transform the values to the Tanner-Whitehouse (TW) version 3 (TW3) or FELS stages and estimates of SA. Briefly, BoneXpert provides a standardized, cost effective for estimating SA.

Literature on data quality regarding SA assessment is scarce and is still lacking specifically for the concurrent protocols and score systems as in TW version 2 (TW2). The first two studies in the current thesis assessed intra and interobserver agreement of these two protocols. Table 8.1 refers the studies which considered data quality on TW protocol, including Study 1, and six notes are possible to be summarized as follows:

(i) Number of studies which considered TW1 are greater, but with a significant lower agreement then TW2, due to the changes that occurred with the revised version;

(ii) None of the samples were composed specifically by youth athletes. In fact, the majority represented the general non-athletic population;

(iii) The literature did not systematically consider intra and inter-observer agreement;

(iv) Agreement level in Study 1 was, in general, in accordance with the results from the other studies using TW2, presented in Table 8.1;

(vi) The literature did not examine the effect of disagreements on the concurrent scoring systems and the impact on the final SA.

In the 20-bone system and in the CARPAL system, the bone with the greater scoring is the Capitate. By inference, an error in the above-mentioned bone had a larger impact on intra-class correlation than any other bone. The preceding was observed in the present study with an agreement between examiners for capitate close to 95% combined to ICC equal to 0.791. Conversely, distal phalange V showed an agreement of 80% between the two observers and an ICC of 0.917.

In addition, an error between two examinations at the most advanced stages may have a different impact while converting scores into SA, compared to an error at the early stages that correspond to a modest score increment. The TW2 protocol eliminated the last stage of Radius (J – complete fusion), whereby the last stage in TW-2 became the beginning of fusion (stage I). The difference between stage I and H is very wide (75 points), which can translate into a huge difference in SA, with implications on maturity classification.

Reproducibility in FELS protocol is even more scarce in the literature. Existing studies focused on the variability in overall SAs, within and between observers (Malina, Chamorro, et al., 2007; M. Vignolo et al., 1992) and little is known about the difficulty of maturity indicators and respective grades. In that particular, study 2 should be viewed as a novelty. With few exceptions, an overall intra-observer agreement over 80% was found for most of the indicators. As expected, interobserver agreement was not perfect. Nevertheless, the impact on SAs was negligible in the current study. Mean difference in SA was 0.03 years and with only three cases in 97 films plotted beyond the limits of agreement. These results confirmed the validity of verbal information of FELS to determine SA. However, younger ages would require the assessment of a different set of bones, indicators and grades corresponding to different verbal and visual instructions.

After examining data quality on both protocols (TW in study 1; FELS in study 2), agreement among the three most used methods to assess SA and skeletal maturity was tested in study 3 (GP, TW2 RUS and FELS) using a multisport sample of adolescent athletes (about 1800 boys divided by one-year age groups from 11 to 17 years). The three methods have the same basic assumptions:

(i) They all use a left hand-wrist radiograph;

(ii) An invariable sequence of the maturity indicators is global to the three methods;

(iii) The final aim is always to get a SA of the individual, comparing with the reference sample in which the method was developed.

Table 8.2 presented the general characteristics of each method. It seems that TW2 tends to classify more youth athletes as mature and is the method that is furthest from CA from younger ages. Results from Study 3 indicated that at 11.0-11.9 years, FELS plotted very close to mean CA. The same occurs regarding age-specific mean SA by GP. Reference samples from the two methods, although from different studies, were exactly from the same region, state of Ohio, USA. The discrepancy between SA and CA increase with advancing CA, reaching a largest value at 14-0-14.9 years. The preceding is common to the three methods. The preceding age it is coincident with many changes that occur during the period of pubescence, like PHV or peak weigh velocity (PWV).

Table 8.1. Data quality on Tanner-Whitehouse method: versions 1,2, 3.

Authors (year)	TW version	Sample/films	Age (years)	Population	Intra-agreement	Inter-agreement
Acheson, Vicinus, & Fowler (1966)	TW1	50 (boys and girls)	2-18	Grossly undernourished or had suffered a chronic disease	95% CL: ± 0.61 years to 0.63 years (2 obs)	/
		11 (from the 50 above)			95% CL: 0.38 years to ± 0.94 years (4 obs)	/
Tiisala et al. (1969)	TW1	Total: 513 (219 boys, 294 girls); For Inter-observer:328 (140 boys, 188 girls)	1-5	General (Longitudinal "Model Child" Study, Finland)	/	72.2%
Tiisala et al. (1971)	TW1	Total: 505 (200 boys, 305 girls) For Inter-observer: 296 (117 boys, 179 girls)	6-10	General (Longitudinal "Model Child" Study, Finland)	/	75.4%
Medicus, Gron & Moorrees (1971)	TW1 (mod)	300 (85 boys and girls)	4-18	General (Forsyth Longitudinal Study of Twins, USA)	300 x-rays:80.9%	280 films: 77.6%
Ashizawa (1974)	TW1	770 (418 boys and 352 girls)	6-14	General (Tokyo, Japan)	71.9% (3 readings)	/
Taranger et al. (1976)	TW2	2191 (212 total: 122 boys, 90 girls)	1 month - 7	Urban Children (Longitudinal, Sweden)	about 80%	/
Van Venrooij-Ysselmuiden & Ipenburg (1978)	TW1	120 films /10 to 20 in each age group)	8-16 (girls)	Growth Study at the University Dental School of Utrecht (The Netherlands)	0.01 to 0.09 years	/
	TW1	2264 (628 girls and 504 boys x 2 moments: 10%)	8 - 17			<0.15 years
	TW1	60 films from the total sample	10-16 (boys)		/	71.3%
	TW2				/	75.8%
Helm (1979)	TW2	1880 boys and 1937 girls 90 films randomly selected	7-16 (boys) 7-14 (girls)	School Children (Copenhagen, Denmark)	Long Bones: 88-89% Short Bones: 91% (84-96%) Carpal Bones: 86% (82-93%)	/
Beunen & Cameron (1980)	TW2	112 (12 films: 10 per child)	1-19	Harpenden Longitudinal Growth Study (England)	/	83.6%
		50	12.06 - 19.69	Leuven Growth Study of Belgian Boys (Belgium)	All bones: 89.1%, 92.2% RUS: 87.4%, 91.4 Carpals: 92.3%, 93.7%	83.2%
Wenzel & Melsen (1982)	TW2	110 (55 boys, 55 girls): films 10%	6-16	General (Denmark)	82 to 100%	/
Ye, Wang, & Cao (1992)	TW2	2122 (1123 boys, 999 girls)	2-20	School Children and adolescents (Changsha, South China)	90.5%	/
Silva, Freitas, Beunen, & Maia (2010)	TW3	46 x-rays from "Leuven Growth Study of Belgian Boys"	NR	Leuven Growth Study of Belgian Boys (Belgium)	81.6%	68%
		40 x-rays from "Harpenden Growth Study"	NR	Harpenden Longitudinal Growth Study (England)	/	81.3%
		40 films from “Crescer com Saúde no Cariri”	NR	Project “Crescer com Saúde no Cariri” (Brazil)	87.9%	/
Sousa-e-Silva et al. (2022)	TW2	142 males	11.0-15.3	Young soccer players	95.1%, 93.8%	82.8%

TW (Tanner-Whitehouse); mod (modified); NR (not reported); CL (confidence limits); obs (observer); RUS (Radius, ulna and short bones).

Table 8.2. General characteristics of concurrent methods to assess SA: GP, TW2, and FELS.

Method	Year	Type of method	Number of bones	Place	Sample (n)	Number of films	Age covered	Years covered
GP	1959	Visual inspection method (Atlas:31 plates for boys, 27 plates for girls)	30 bones	Cleveland, Ohio, United States of America	*	27 plates for boys, 31 plates for girls*	newborn - 19.0 years	1931-1942
TW2	1975	Scoring method (Stages from A to I)	20 bones (20-bone) 13 bones (RUS) 7 bones (CARPALS)	United Kingdom	2702 (1385 boys, 1317 girls)	7677	1 month - 21 years	1946-1972
FELS	1988	Scoring method (98 maturity indicators, each indicator may have 2 to 5 grades; 13 metric indicators)	22 bones	Southwestern Ohio, United States of America	677 (355 boys, 322 girls)	13.823	1 month - 22 years	1932-1977

*Each of the standards was chosen from 100 films of children of the same sex and age.

8.2. Interrelationship between skeletal and somatic maturity

Non-invasive protocols provide an easiest artifact to assess biological maturation. The maturity offset (MO) equation (Mirwald et al., 2002) emerged from anthropometric variables (stature, sitting height, leg length, and body mass) and is being widely used. Two validation studies were realized, one using a European sample from Poland (Malina & Kozieł, 2014a), the other using a North-American sample with European ancestry from the FELS longitudinal study. Briefly, the two validation studies indicated limitations on the MO equation with age at peak heigh velocity (APHV) except for average maturing individuals around years of maximal velocity in stature. The preceding findings showed that the use of the original MO equation for talent development purposes or to some sport-specific adaptation training programs may result in inaccurate decisions.

Nevertheless, Moore et al. (2015) provided two additional equations to the original MO, one of them without the need of using sitting height as dependent variable. Sample included 79 boys from Pediatric Bone Mineral Accrual Study (PBMAS), 42 from University of British Columbia's (HBS-III), and 38 from the Harpenden Growth Study (HGS). The authors founded that the prediction error continue to be slightly higher in late and early matures, increasing farther a child is away from expected APHV. In parallel, Fransen et al. (2018) aimed to develop an equation that supposedly mitigates error among individuals far from APHV. Two databases were considered: PBMAS (the Canadian sample used to develop the original MO equation) and a sample of 1330 Belgian soccer players, aged 8.0-17.0 years. However, validation studies to support the algorithms are lacking and available literature is not promising.

Parr et al. (2020), assessed a sample of 23 boys from professional academies of the English Premier League during five competitive seasons. Longitudinal records were fitted with SITAR model (Cole et al., 2010). The mean predicted age at PHV at 13.0 years using MO equation was 15.1 ± 0.5 years (14.0-16.0 years) and the observed age at PHV was 14.2 ± 0.8 years (12.6 – 15.5 years). It was evident by using predicted APHV from maturity-offset equations. Meantime, Teunissen et al. (2020) assessed the accuracy of the original (Mirwald) and modified (Moore 1, Moore 2, Fransen) equations on a sample of 17 boys from a professional youth soccer Dutch academies. The youths aged 11.9±0.8 years in the baseline and were assessed during four competitive seasons with anthropometric measurements at least three-months interval. Age at PHV were estimated using Preece-Baines model I and was correspondent to 13.8±0.7 years (almost similar to the mean of the three samples from the original MO protocol: 13.9±0.8 years). Nonetheless, the range of the observed age was wider (12.6-15.2 years) than any of the four predictive equations (13.2–15.5 years: Mirwald; 13.3–15.3 years: Moore 1; 12.9-4.8 years: Moore 2; 13.2–15.1: Fransen). The preceding was already noted in Polish (Kozieł & Malina, 2018; Malina & Kozieł, 2014a; Malina et al., 2021) and American (Malina et al., 2016) samples. By using Mirwald equation none of the players presented a mean predicted APHV equal to the observed APHV. In most of the players the observed APHV occurred earlier than the predicted. Nevertheless, Mirwald and Fransen equations apparently had more stable values over time than the two equations by Moore. As noted by the authors, ethnic variation in size, proportionality, shape and composition corresponds to limitations in the generalization from original samples to other populations. Other limitation in the preceding study (Teunissen et al., 2020) is the 17-individual sample.

Study 4 compared estimates of APHV and age at takeoff (TO) on a sample of male soccer players aged 11-16 years, using SITAR and Functional Principal Component Analysis (FPCA). This was a novelty, as any of the preceding studies used Preece-Baines model I (PB-I) which was considered a limitation (Teunissen et al., 2020). Meantime, concordance between maturity classification based on estimated age at PHV and skeletal age by FELS method was assessed. Systematic analysis of ages at PHV in 13 longitudinal samples of European and Japanese soccer players was made, suggesting a population gradient: ages at PHV earlier for Japanese boys and later for the north-European boys, with the south-Europeans in the middle of the Japanese and north-Europeans. Estimated ages at PHV occurred earlier than the ages at PHV of Mirwald, Parr, and Teunissen, 13.6±0.9 years and 13.7±0.9 years, for SITAR and FPCA, respectively. Differences in ages at TO and PHV, based on SITAR and FPCA were negligible suggesting that any of the two can be used for monitoring the growth spurt of young athletes.

During the interval of growth spurt, the individuals are more susceptible to injuries and in many sports years of peak velocity in stature overlap with sport specialization in swimming as well as in track and fields while, in other sports, correspond to sport selection such as in soccer (Coelho-e-Silva et al., 2010). In other words, being informed about inter-individual variation on age at take-off and at peak height velocity is obviously relevant for coaches.

The last study confirmed the literature regarding major limitations of the predicted MO. In a sample of 58 male adolescent soccer players, SITAR was considered for the calculation of PHV instead of PB-I that was previously used in the literature (Teunissen et al., 2020). It was concluded, as in the validation studies (American and Polish samples), that mean predicted APHV increased with CA. Moreover, variation in predicted MO and APHV within a single CA groups was reduced compared to variation in observed ages at offset and PHV. Additionally, predictions varied according to maturity status. Finally, intra-individual variation in predicted ages at PHV was considerable. Based on the preceding, MO should not be predicted for the selection of adequate training methods in youth sports. It does not correspond to an adequate protocol for assessing the timing and tempo of adolescent athletes.

8.3. Conclusions, practical application and future directions

Problems with accurate chronological age (CA) occur on relative regular basis in youth sports (Malina, 2011). As a result, there is increasing discussion of 'bone age' or SA for the purpose of verifying CA. The discussion has been previously used in medicolegal contexts for many years. The present thesis demonstrated that SA from different protocols is not identical and, additionally, true inter-individual variability in skeletal age occurs within players without any need for age verification:

A previous study diagnosed a gradient of elite > club > dropout clearly defined among 13- to 14-year-old players at male soccer players (Figueiredo et al., 2009b). Elite players were older chronologically and skeletally, larger in body size and performed better in functional capacities and three skill tests than club players and dropouts. By inference, the assessment of biological maturation assumes special interest in the interpretation of performance level among adolescent athletes.

Inter-individual differences in size, maturity status, function, and behavior among youth of the same chronological age (CA) have long been a concern in grouping for sport.

Biobanding is an interesting attempt to accommodate maturity-associated variation among youth in sport. Several applications of biobanding in youth soccer have indicated positive responses from players and coaches (Malina et al., 2019). The potential utility of biobanding for appropriate training loads, injury prevention, and fitness assessment merits closer attention, specifically during the interval of pubertal growth.

Currently, Portuguese Sports Medicine units are considering medical examinations to approve overclassification in youth sports, that is, moving early matures to forward age groups (DL 345/99, 27 August; DL 255/2012, 29th November). The medical protocol includes the evaluation of left ventricular mass, estimates of body composition derived from anthropometry, and skeletal age using Greulich-Pyle method. From the previous, it is of obvious interest to examine intra- and inter-observer agreement in the determination of SA, agreement between concurrent methods and maturity-associated variation of body composition and left ventricular mass.

8.4. References

Coelho-e-Silva, M. J., Figueiredo, A. J., Simões, F., Seabra, A., Natal, A., Vaeyens, R., … Malina, R. M. (2010). Discrimination of U-14 soccer players by level and position. International Journal of Sports Medicine, 31(11), 790–796.

Cole, T. J., Donaldson, M. D. C., & Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. International Journal of Epidemiology, 39(6), 1558–1566.

Decree Law no. 345/99 of 27th August, from Presidency of the Council of Ministers. Republic Diary No. 345/99, Series 1 of 1999-08-27. Accessed on 1 Dec. 2022. Available at www.dre.pt.

Decree Law no. 255/2012 of 29th November, from Presidency of the Council of Ministers. Republic Diary No. 255/2012, Series 1 of 1999-11-29. Accessed on 1 Dec. 2022. Available at www.dre.pt.

Figueiredo, A. J., Gonçalves, C. E., Coelho-e-Silva, M. J., & Malina, R. M. (2009). Youth soccer players, 11-14 years: Maturity, size, function, skill and goal orientation. Annals of Human Biology, 36(1), 60–73.

Fransen, J., Bush, S., Woodcock, S., Novak, A., Deprez, D., Baxter-Jones, A. D. G., … Lenoir, M. (2018). Improving the Prediction of Maturity From Anthropometric Variables Using a Maturity Ratio. Pediatric Exercise Science, 30(2), 296–307.

Kozieł, S. M., & Malina, R. M. (2018). Modified Maturity Offset Prediction Equations: Validation in Independent Longitudinal Samples of Boys and Girls. Sports Medicine, 48(1), 221–236.

Malina, R. M. (2011). Skeletal Age and Age Verification in Youth Sport. Sports Medicine, 41(11), 925–947.

Malina, R. M., Chamorro, M., Serratosa, L., & Morate, F. (2007). TW3 and Fels skeletal ages in elite youth soccer players. Annals of Human Biology, 34(2), 265–272.

Malina, R. M., Choh, A. C., Czerwinski, S. A., & Chumlea, W. C. (2016). Validation of maturity offset in the fels longitudinal study. Pediatric Exercise Science, 28(3), 439–455.

Malina, R. M., Cumming, S. P., Rogol, A. D., Coelho-e-Silva, M. J., Figueiredo, A. J., Konarski, J. M., & Kozieł, S. M. (2019). Bio-Banding in Youth Sports: Background, Concept, and Application. Sports Medicine, 49(11), 1671–1685.

Malina, R. M., & Kozieł, S. M. (2014). Validation of maturity offset in a longitudinal sample of Polish boys. Journal of Sports Sciences, 32(5), 424–437.

Malina, R. M., Kozieł, S. M., Králik, M., Chrzanowska, M., & Suder, A. (2021). Prediction of maturity offset and age at peak height velocity in a longitudinal series of boys and girls. American Journal of Human Biology, 33(6).

Mirwald, R. L., Baxter-jones, A., Bailey, D., & Beunen, G. (2002). An assessment of maturity from anthropometric measurements. Medicine & Science in Sports & Exercise, 34(4), 689–694.

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D. G., Cameron, N., & Brasher, P. M. A. (2015). Enhancing a somatic maturity prediction model. Medicine and Science in Sports and Exercise, 47(8), 1755–1764.

Parr, J., Winwood, K., Hodson-Tole, E., Deconinck, F. J. A., Parry, L., Hill, J. P., … Cumming, S. P. (2020). Predicting the timing of the peak of the pubertal growth spurt in elite male youth soccer players: evaluation of methods. Annals of Human Biology, 47(4), 400–408.

Teunissen, J. W., Rommers, N., Pion, J., Cumming, S. P., Rössler, R., D'Hondt, E., … Malina, R. M. (2020). Accuracy of maturity prediction equations in individual elite male football players. Annals of Human Biology, 47(4), 409–416.

Thodberg, H. H., Kreiborg, S., Juul, A., & Pedersen, K. D. (2009). The BoneXpert method for automated determination of skeletal maturity. IEEE Transactions on Medical Imaging, 28(1), 52–66.

Vignolo, M., Milani, S., Cerbello, G., Coroli, P., Di Battista, E., & Aicardi, G. (1992). FELS, Greulich-Pyle, and Tanner-Whitehouse bone age assessments in a group of Italian children and adolescents. American Journal of Human Biology, 4(4), 493–500.