Evaluation of the BAUSport System (By Sonicbone) for Assessing Skeletal Age in Youth

By Prof Sean Cumming, PhD

.

- Reader in Sport and Exercise Sciences
- Researcher in the area of Growth and Maturation in Youth Sports
- Professor at the University of Bath, Bath, UK
- Consultant to professional Sports' institutions & clubs, including the British Premier League & Football Association, the US soccer Federation and more.

March 18, 2019

Introduction

The purpose of this evaluation was to determine the validity and accuracy of the BAUSport method for assessing skeletal age (i.e., biological maturation) in youth. To determine the validity and accuracy of the method, BAUSport values obtained from twenty individuals were compared against a known standard for assessing biological maturation via skeletal age of the hand-wrist (i.e., Fels).

The Fels method for estimating skeletal age was developed from over 13,000 serial radiographs of 355 boys and 322 girls' that participated in the Fels longitudinal Study of Growth and Development. In comparison to other methods for assessing skeletal age (i.e., TW2 & TW3, Greulich-Pyle), the Fels method uses a greater, more diverse and more specific set of criteria (appearance, fusion, shape, ratios) to estimate skeletal age. It also considers fusion of the growth plate on the radius across its lateral, medial, and central segments. In, contrast to other methods of assessment, the Fels method also provides an associated standard error. The Fels method is an ideal standard to compare the BAUSport against as both methods consider not only the contour of bones or fusion of the growth plates, but also the degree of the ossification process (i.e., transformation from cartilage to bone tissues.

Procedures

Sample

SonicBone Medical provided digital images radiographs of the left hand-wrist for a total of twenty individuals. The chronological age at assessment, sex, and BAUSport skeletal age values for each individual connected to each digital image was also provided by **SonicBone** Medical. No information pertaining to the identity, nationality, or ethnicity of the individuals was provided. The total sample included 15 boys and 5 girls aged between 11 and 14 years for age. A comparatively narrow age group was requested in order to provide a more conservative test of the BAUSport system's ability to differentiate between children of a similar chronological age, yet varying maturity status.

Methods

Each radiograph was assessed by a single assessor experienced in using the Fels method for estimating skeletal age. The Fels method is an established index of biological maturation and is considered an objective measure of skeletal age. The Fels method has been previously employed as a standard upon which to compare and validate alternative estimates of biological maturation, including the Khamis Roche method for predicting adult stature, the Mirwald maturity offset method, and the Tanner Whitehouse 2 method for estimating skeletal age.

The Fels method for estimating skeletal age requires the assessor to make a series of judgements relative to the appearance, characteristics, and dimensions of specific bones on the hand and wrist. Using the Fels method, the total number of indicators assessed varies relative to the age and sex of the individual. Grades and measurements were entered into the Felshw 1.0 software (Felshw 1.0, Software Lifespan Health and Research Center, Departments of Community Health and Pediatrics, Booshoft School Medicine, Wright State University Dayton Ohio, USA) to derive an estimate of SA and its associated standard error. BAUSport scores were not accessed by the investigator until all of the Fels assessments had been fully completed. The individuals Fels ratings in addition to the calculated skeletal age and associated standard error for each individual can be found in the accompanying appendix section.

As the Fels and BAUSport assess maturation of the skeletal system one would expect a good degree of concordance between these methods. It should be noted, however, that whereas the BAUSport methods infers maturation from bone density at three sites (radius, carpals, phalanges); the Fels methods assesses maturation in terms **of** the appearance, shape, and dimensions of a broader range of bones within the hand-wrist. Accordingly, it is likely that some degree of discrepancy in two estimates of skeletal age will be observed.

For both the BAUSport and Fels methods the maturity timing for each individual (i.e., early, on-time, delayed) is defined by the difference between their skeletal and chronological age at the point of assessment, and represented in years. An individual with a skeletal age that is greater than their chronological age is considered to be advanced in maturation for their sex and age. In contrast, an individual who present a skeletal age that is less than their chronological age is considered delayed in maturation. When skeletal and chronological ages are equivalent, the individual is considered to be on time.

<u>Statistical Analyses</u>

A range of analyses was used to determine the validity and reliability of the BAUSport system; with the Fels skeletal age assessments serving as the standard upon which these qualities were determined. Descriptive analyses were first conducted to determine the individual, mean, maximum, and minimum differences in skeletal age as determined by the Fels and BAUSport methods. Absolute and relative Technical Errors of Measurement were also calculated to determine the absolute and relative error margin between the two methods.

A Bland-Altman Plot was then conducted to analyse the level of agreement between the two measures of skeletal age (Fels & BAUS). A one sample means t-test was then conducted to determine the presence of a fixed bias between the measurements of skeletal age. Similarly, 95% limits of agreement were also computed to estimate how far apart measurements use the two methods (Fels and BAUS) would fall in a typical population. Scatterplots and intra-class correlations were also conducted to examine the magnitude and direction of the associations between (1) the two measurements of skeletal age (Fels & BAUS) and (2) the differential between skeletal and chronological age using each method (Fels & BAUS). Finally, a concordance analysis was also conducted to examine the degree to which the two methods agreed in terms of being able to categorize players as being within or outside a two-year band (+-1 year) of the mean age of Peak Height Velocity (13.8 years in males, 11.8 years in females). All analyses were conducted using IBM SPSS version 24.0.

Results

Descriptive Analyses

The descriptive analyses for each of the individual is presented in Table 1. The mean age of the participants was 12.76 (SD= 0.92) years with a minimum and maximum age of 11 and 14 years, respectively. The mean skeletal age for the Fels and BAUSport methods were and 13.32 (SD= 1.39) and 13.22 (SD= 1.46) years respectively, with minimum and maximum values of 10.61 and 16.01 years for the Fels method and 10.5 and 16 years for the BAUSport method. The mean differences between the Fels and BAUSport estimates of skeletal age across all measurements was 0.11 (SD= 0.41) years with a minimum and maximum difference between the methods of +0.89 and -0.89 years.

Table 1. Descriptive statistics associated with each individual assessment of skeletal age (years) using the Fels and BAUSport methods.

Individual	Sex	Age	Fels SA	Fels SE	BAUS SA	Fels SA - BAUSport SA	Fels SA-CA	BAUSport SA-CA
M1	m	12	13.39	.3	12.5	0.89	1.39	0.5
M2	m	14	13.53	.36	13.25	0.28	-0.47	-0.75
M3	m	13.5	13.32	.33	13	0.32	-0.18	-0.5
M4	m	13	13.19	.31	13.5	-0.31	0.19	0.5
M5	m	12	10.61	.3	11.5	-0.89	-1.39	-0.5
M6	m	13.5	14.2	.32	14	0.2	0.7	0.5
M7	m	14	13.33	.36	13.25	0.08	-0.67	-0.75
M9	m	12.5	13.77	.3	13.5	0.27	1.27	1
M10	m	11.5	12.88	.29	12.75	0.13	1.38	1.25
M11	m	12.5	13.09	.3	13.25	-0.16	0.59	0.75
M12	m	14	16.01	.28	16	0.01	2.01	2
M13	m	13	15.32	.25	15.75	-0.43	2.32	2.75
M14	m	13	15.11	.25	14.75	0.36	2.11	1.75
M15	m	13	13.04	.31	12.5	0.54	0.04	-0.5
F1	f	11	11.89	.28	11.5	0.39	0.89	0.5
M20	m	13.1	12.77		12.64	0.13	-0.33	-0.46
F3	f	14	13.98	.31	13.75	0.23	-0.02	-0.25
F4	f	11.5	11.06	.29	10.5	0.56	-0.44	-1
F5	f	12	11.35	.29	11.25	0.1	-0.65	-0.75
F6	f	12.1	14.74	.34	15.25	-0.51	2.64	3.15

<u>Absolute and Relative Technical Errors of Measurement</u>

The absolute and relative technical errors of measurement for the differences in biological age assessed by the Fels and BAUSport method were 0.29 years and 2.5%, respectively. In this instance, the absolute TEM represents the standard deviation between repeated measures of the same individuals across the different methods. That is, the standard difference between the measures of skeletal age between the Fels and BAUSport methods is approximately 0.3 years. The relative TEM represents the standard differences between repeated measures of the same individual across the different methods, as a percentage of the total average of the variable (i.e., skeletal age) assessed. A relative TEM value of 2.5% can be considered acceptable for the purpose of repeat measures of an individual across methods (i.e., inter-method reliability).

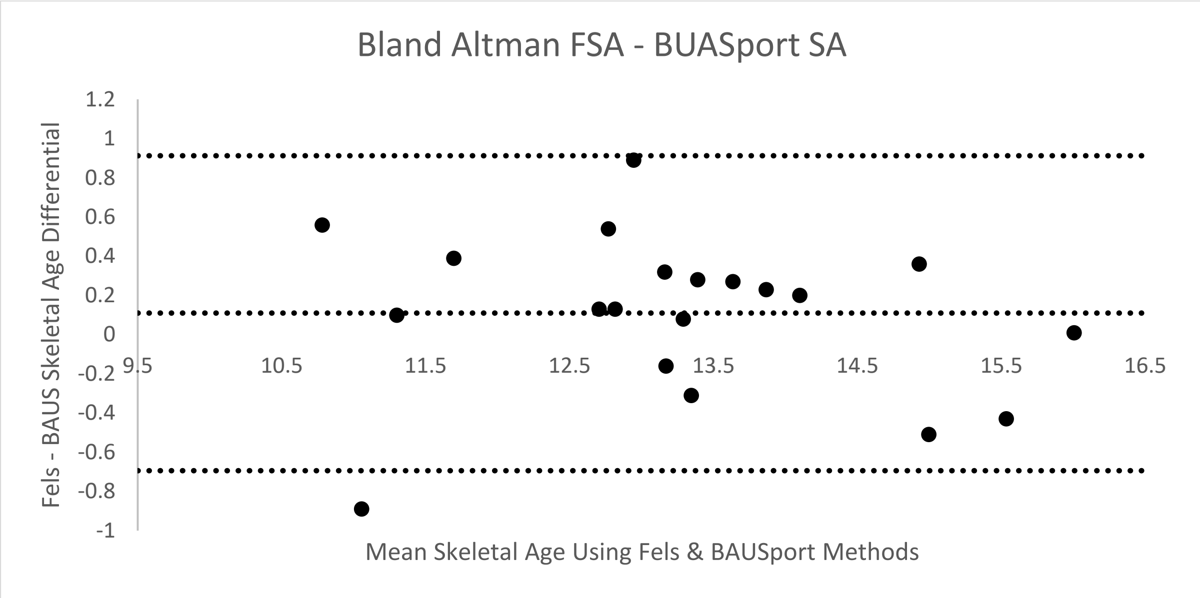
Bland Altman Plot and Estimated Bias

A Bland Atman Plot was generated to represent the differences between the Fels and BAUSport estimates of skeletal age across each individual and is presented in Figure 1. The estimated bias and upper and lower levels of agreements were also included in this analysis and are represented in Figure 1. The estimated bias across the assessments was 0.11 years (SD = 0.41). A one sample means t-test failed to identify the presence of a fixed effects (i.e., proportional) bias. As the estimated bias between the Fels and BAUSport methods did not different significantly from zero, t (19) =1.19, p=.25, this finding would indicate no need to adjust the skeletal age values produced by the BAUSport method when using it as an alternative to the Fels method. That said, it should be noted that the effect size associated with this t-value, though not statistically significant, is moderate in magnitude (Cohen's D =.55) and thus, this finding should be interpreted with some degree of caution.

The Upper and Lower Level of Agreements calculated at +-1.96 SD methods were 0.91 and -.69 years respectively, representing a band of +-.80 years from the estimated bias (See Figure 1). These values indicate that the agreement between the Fels and BAUSport methods, in the current evaluation, that falls with a range of 1.6 years in 95% of all assessments. The range of the levels of agreement, though broad, is notably lower (20%) than those observed for both boys and girls in an equivalent study comparing the Fels and Tanner Whitehouse (TW2) methods in Dutch and Youth (Van Lenthe et al., 1998) (1): ULOA-LLOA = 1.32 to -.68 in males, ULOA-LLOA = 1.18 to -.58 in females). It is also markedly lower than the equivalent ranges observed by Roche and colleagues (Roche 1980) (2) in a comparison of the Fels and TW2 methods. This suggests that while the Fels and BAUSport methods should not be considered as directly comparable or interchangeable, the BAUSport method presents superior levels of agreement with the Fels method than other established indices of skeletal age, such as the TW2.

The minimal detectable change for skeletal age assessment at 95% CI was .55 years. The MDC95 represents the minimal amount of difference skeletal age across measurements that we can be confident (at 95% probability) is not a result of measurement error. Accordingly, using the BAUSport method we can be confident that any observed differences in skeletal age at/or beyond .55 years, are not a result of measurement error.

Figure 2. Bland Altman plot representing differences in skeletal age values calculated by the Fels and BAUSport methods, with estimated bias and 95% upper and lower levels of agreement.



<u>Intraclass Correlations and Scatter Plots</u>

The intraclass correlation (average measures) between that Fels and BAUSport measures of skeletal age was positive, large in magnitude and achieved statistical significance (ICC = .98, df=19, p<.001). The intraclass correlation (average measures) for skeletal-chronological age differences between the Fels and BAUSport methods was also positive, large in magnitude and achieved statistical significance (ICC = .97, df=19, p<.001). These values are high, indicating a substantial level of agreement between the Fels and BAUSport method, whether it be for assessing skeletal age and/or the differential between skeletal and chronological age. The intraclass correlation values observed for the Fels and BAUSport methods are equal to, and in some instances higher, then the equivalent correlations observed among the many different methods for assessing skeletal age via hand-wrist x-rays. (i.e., Fels, Greulich-Pyle, Tanner-Whitehouse)

The scatterplots for each of the intraclass correlations are presented in figures 2 and 3. It should be noted, that in only one (M15) of the twenty individuals did the Fels and BAUSport values for skeletal age produce contrasting positive and negative values (Figure 3). In this instance the difference between the BAUSport and Fels values for skeletal age was 0.54 years.

Figure 2. Scatterplot and correlation between Fels and BAUSport estimates of skeletal age in boys and girls

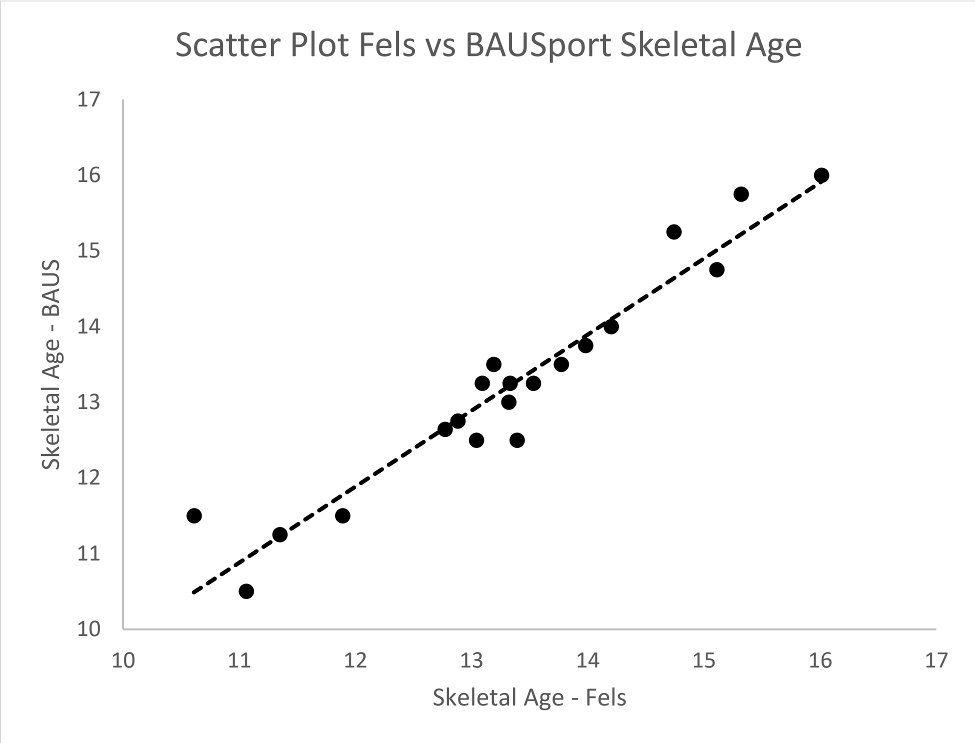
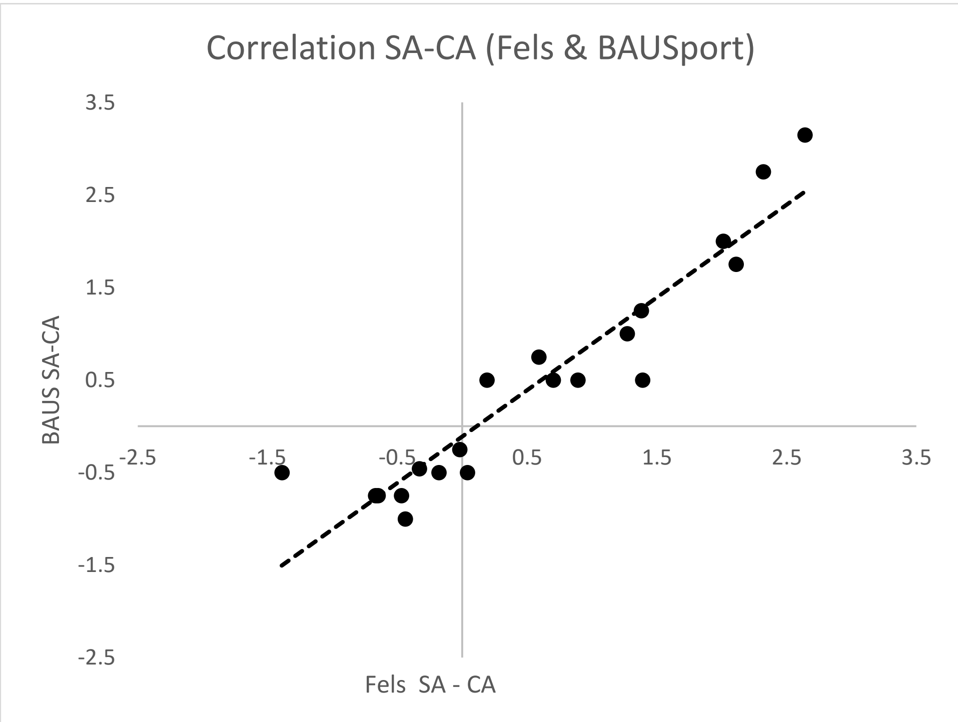


Figure 3. Scatterplot and correlation for Fels and BAUSport differential values for skeletal age (SA) and chronological age (CA).



<u>Fels and BUAS Concordance at Predicting PHV Status</u>

To examine the degree to which the Fels and BAUSport methods agreed in terms of terms of categorizing individuals as being within or outside +- 1 year band from the mean age of Peak Height

Velocity (13.8 Years of age for males; 11.8 years of age for females), Kappa coefficients were calculated. The level of agreement between the methods in categorising the individuals was 85% percent and statistically significant, demonstrating a good standard of concordance (Kappa = .70, p<.01). This suggest that the Fels and BAUSport methods can be demonstrate a high and statistically significant degree of concordance when employed for the purpose of classifying individuals as being within or outside a +- 1 year band from the mean ages of PHV.

Summary

The purpose of the evaluation was to examine the validity and accuracy of the BAUSport method as an index of skeletal age in youth. In the sample provided, the BAUSport demonstrated reasonably good convergent and concurrent validity in against an established index of biological maturation (i.e., Fels method). While the values produced by the BAUSport method may not be directly comparable to those produced by the Fels method, the level of concordance between these methods was as good, if not better, than equivalent levels of agreement observed among established measures of skeletal age (i.e., Fels, TW2-3, Greulich Pyle) (Malina, Bouchard, Bar-Or, 2004). (3)

The intraclass correlations between the methods for measures of skeletal age and the differential between skeletal and chronological age were positive, large in magnitude (i.e., >.50) and statistically significant. This suggests a very strong association between the Fels and BAUSport methods. It should be noted, however, that strong intraclass correlations, though indicative of agreement between methods, can be misleading. For example, it is possible for two measures of a constructs to present a high correlation yet vary markedly in their estimated values.

Technical Error of Measurement (TEM) is a common method for representing the standard error between agreement between measures in anthropometry. TEM values can be determined for both within (intra-) and between (inter) investigator reliability and can be considered in terms of absolute and relative values. The absolute TEM between the measures of skeletal age across the methods in the current evaluation was small in magnitude (0.3 years) and the relative Technical Error of Measurement (RTEM = 2.5%) was close the values considered acceptable (2%) for repeated measures across investigators using the same methodology. As the Fels and BAUSport methods do not employ the same methodology it could be argued that a RTEM value of 2.5% is representative of an acceptable level of reliability.

The Bland Altman plot demonstrated a comparatively good level of agreement between the methods and no evidence of an estimation bias. It should be noted, however, that the upper and lower levels of agreement and minimal detectable change were, however, moderate, indicating that that individual assessments of skeletal age using the BAUSport method, and differences across assessments, should be considered and interpreted within these bounds. The range of the level of agreements in the current evaluation, though broad, was more than 20% lower than the equivalent ranges observed in research examining levels of agreement between the Fels and TW2 method (an established and commonly applied indices of skeletal in youth).

In comparison to existing methods for estimating skeletal age and maturation in youth (all of which are X-ray based methods), the BAUSport system presents a number of potential advantages. First, as it is ultrasound based, it poses no risk in terms of exposure to radiation. As the readings are non-operator-dependent, it reduces the likelihood for error associated with intra or inter-investigator reliability. It's non-invasive and objective nature may also make it ideal for monitoring change to the pubertal growth spurt where more frequent testing of maturation status may be of interest. The method also provides useful indicator of 'biological age' which is of value in terms of identifying players who are advanced, on-time, and delayed in maturation. This information could also be used to help identify phases of accelerated growth and adjusting training load and content accordingly. Similarly, information pertaining to skeletal age could be used to group players by maturation status for the purpose of both training competition (i.e., bio-banding), optimising player safety and development. The BAUSport system offers many potential benefits for practitioners working with young athletes and could be used as part of a comprehensive athlete growth and maturation screening program coupled with traditional methods for assessing growth and maturation, (i.e., measure of height, weight, velocity curves) it could provide further valuable insight into the growth and development of young athletes (Malina 2007) (4).

BAUSport enables practitioners to perform routine, non-invasive tests on Youth for monitoring the processes growth and maturation. In Sport this has a tremendous value, mainly: for adjusting individual training programs to young athletes, for preventing injury risks and for talent identification and selection.

BAUSport complements is has the potential to support new worldwide trend in sports – such as the practice of Bio-Banding, for both: training and competition, alike.

Limitations of the consultancy project should also be noted. First, the result of this project are limited to a sample of twenty children between the ages of 11 and 14 years and may not generalise to other samples. Although the sample was large enough to produce statistically significant results, an analysis that included a greater number of individuals and had greater statistical power would provide a more stringent test of both validity and accuracy. Further, it should be noted that the current project was cross sectional in nature and only involved single assessment of skeletal age using the BAUSport system. As such, it was not possible to determine the intra-test reliability of the BAUSport system (i.e. repeat measures across the same individual). That said, it should be noted that the BAUS's demonstrated high reproducibility and repeatability in a previous clinical study (Rachmiel et al., 2017) (5). While the ethnicity of the participants was not provided, it has been argued that differences in bone maturity across races are more likely to result from environmental differences, rather than ethnicity, per se (Hochberg 2015) (6).

In summary, the results of this project suggest that the BAUSport system can serve as valid and reliable predictor of biological maturation in youth. In the absence of a true gold standard for the assessment of skeletal age, it is difficult to ascertain the degree to which a specific method is, or is not valid, or reliable. However, in comparison to other established method for determining skeletal age in youth (i.e., TW2-3 Greulich-Pyle), the BAUSport method performed as well, if not better. Accordingly, it can be argued that in comparison to existing methods, the BAUSport method serves as an equally valid and reliable (if not more) measure of skeletal age (as determined by the Fels method). While these results are encouraging, assessment of skeletal age using the BAUSport system should be considered within the limitations of this method (i.e., absolute and relative technical errors of measurement, upper and lower levels of agreement & minimal detectable change).

References

1. 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.
2. Roche, A. F. (1980). The measurement of skeletal maturation. In *Human physical growth and maturation* (pp. 61-82). Springer, Boston, MA.
3. Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). *Growth, maturation, and physical activity*. Champaign, IL. Human Kinetics.
4. 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.
5. Rachmiel M. et al., (2017) Bone age assessments by Quantitative Ultrasound (SonicBone) and by Hand X-ray Based Methods are Comparable. IMAJ, VOL19, September 2017; pp 533, 535
6. Hochberg, Z. (2015). On the need for National-, Racial- or Ethnic-Specific Standards for the Assessment of Bone Maturation. *European Journal of Endocrinology, 174(2),* 65-70.