

Comparison of Bone Age Assessments by Gruelich-Pyle, Gilsanz-Ratib, and Tanner Whitehouse Methods in Healthy Indian Children

Nikhil Shah, Vaman Khadilkar, Nikhil Lohiya¹, Hemchand K. Prasad², Prashant Patil³, Ketan Gondhalekar, Anuradha Khadilkar

Department of Growth and Pediatric Endocrinology, Hirabai Cowasji Jehangir Medical Research Institute, Jehangir Hospital, ¹Department of Pediatrics, Dr. D. Y. Patil Medical College, Hospital and Research Centre, Pimpri, Pune, Maharashtra, ²Department of Paediatric Endocrinology, Mehta Multispeciality Hospitals India Pvt Ltd, Chennai, Tamil Nadu, ³Department of Pediatric Endocrinology, NH SRCC Children's Hospital, Mumbai, Maharashtra, India

Abstract

Background: There are several methods of bone age (BA) assessment, which include Gruelich-Pyle (GP), Gilsanz-Ratib (GR), and Tanner Whitehouse-3 (TW-3) methods. Although GP atlas is the most widely used, there are concerns about its accuracy in children of different ethnicities, making the use of the TW-3 method an attractive option in Indian children. **Objectives:** 1) To assess the relationship of BA with chronological age (CA) as assessed by different methods (GP, GR, and TW-3) in healthy Indian children 2) To assess which of the three methods of BA assessment is more suitable in Indian children. **Methodology:** X-rays of 851 children (438 boys and 413 girls, aged 2–16.5 years) were analyzed by four independent observers using three different methods of BA estimation (GP, GR, and TW-3). Mean BAs were converted to Z-scores. For purpose of deciding which method of BA was most suitable in our cohort, a test of proportions and root mean square (RMS) deviations were computed. **Results:** Using the test of proportions, the TW-3 method was most suitable overall ($P < 0.05$). TW-3 method was again most applicable in prepubertal boys ($P < 0.05$), in prepubertal girls (although not significant, $P > 0.1$), and pubertal girls ($P < 0.05$). However, in pubertal boys, the GR atlas method was most suitable ($P < 0.05$). The same results were obtained when root mean square (RMS) deviations were computed. Interestingly, BA was underestimated in Indian boys irrespective of the method used. In Indian girls, however, the BA was underestimated till the pubertal growth spurt, after which there was rapid advancement of BA. **Conclusions:** Among the three methods (GP, GR, and TW-3), the BAs estimated by the TW-3 method were closest to CAs. Hence, it seems reasonable to recommend the use of the TW-3 method for BA estimation in the Indian population till an Indian standard bone age atlas is developed.

Keywords: Bone age, children, Gilsanz-Ratib, Gruelich-Pyle, India, skeletal age, Tanner Whitehouse, TW-3

INTRODUCTION

Bone age (BA) is a measure of the skeletal maturity of an individual and tells us about the growth potential of a child.^[1] The other applications of BA include height prediction and estimation of age in children seeking asylum in other countries and in competitive sports where chronological age (CA) may be unknown.^[1-3]

The two major methods of BA assessment used commonly are i) Atlas method [Gruelich-Pyle (GP) and Gilsanz-Ratib (GR)] and ii) Scoring of epiphyses method [Tanner Whitehouse-3 (TW-3) method]. GP Atlas was prepared from 1,000 Caucasian children of North European descent born between 1917 and 1942 in the USA.^[4] GR atlas, a relatively new method developed in 2005,

was produced by the creation of artificial, idealized, averaged sex, and age-specific images of skeletal development.^[5] The TW method was derived from British children in 1950 and was revised as the TW-3 method in 2001 (inclusion of children from different ethnic populations and abolition of 20 bone score).

Address for correspondence: Dr. Anuradha Khadilkar, Hirabai Cowasji Jehangir Medical Research Institute, Jehangir Hospital, 32 Sassoon Road, Pune - 411 001, Maharashtra, India. E-mail: anuradhavkhadilkar@gmail.com

Submitted: 31-Dec-2020

Revised: 25-Feb-2021

Accepted: 05-Jul-2021

Published: 26-Oct-2021

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Shah N, Khadilkar V, Lohiya N, Prasad HK, Patil P, Gondhalekar K, *et al.* Comparison of bone age assessments by Gruelich-Pyle, Gilsanz-Ratib, and Tanner Whitehouse methods in healthy Indian children. *Indian J Endocr Metab* 2021;25:240-6.

Access this article online

Quick Response Code:



Website:
www.ijem.in

DOI:
10.4103/ijem.IJEM_826_20

TW method, in contrast to other methods, is not based on pattern matching; rather, it is based on the stage of maturity of epiphyses of long and short bones of the hand or carpal bones depending on the method used respectively (13 bone method or 20 bone method, respectively).^[6]

Although GP atlas is the most commonly used method of BA estimation worldwide, it is based only on Caucasian children and the data are over 7 decades old. Hence, there are concerns regarding its accuracy and reliability in different ethnic populations. This has been discussed in a systematic review, which showed that the GP atlas could be imprecise and should be used with caution in Asian/African children.^[7] This makes the use of the TW-3 method of BA estimation an attractive option in children of different ethnicities (Indian children), as, during its development, Asian children's bones were also included. However, only a handful of Indian studies have compared GP atlas and TW methods of estimation, and these studies were based on the TW-2 method and were performed several decades ago.^[8] In addition, most of the studies done worldwide comparing different methods of BA estimation have used correlation analysis.^[5,9-11] Correlation coefficients measure the strength of association between two variables and not the agreement between them. Also, the wider the range of values analyzed (in our study 2–16.5 years), the greater is the correlation obtained.^[12] The objectives of this study were thus 1) to assess the relationship of bone age with chronological age as assessed by different methods (GP, GR, and TW3) in healthy Indian children and to study differences in genders and 2) to assess which of the three methods of BA assessment is most suitable for Indian children.

METHODOLOGY

Subjects

This was a cross-sectional, observational study (June–September 2018). Five schools and preschools were randomly selected from Pune district (western India) and approached for permission to carry out anthropometric measurements and BA X-rays. After approval from the institutional ethics committee and consent from the management of each school, written consent was obtained from the parent or caregiver of each child before study commencement.

All children in the schools approached were offered the study. The selection criteria included that the child should be physically and mentally healthy and that the parents gave consent and children (greater than 7 years) gave assent. Standing height using a portable stadiometer (Leicester Height Meter, Child Growth Foundation, UK) was measured to the nearest millimeter and weight was measured using an electronic scale (nearest 100 g). Body mass index (BMI) was computed by dividing weight in kilograms by height in meter square; height, weight, and BMI were converted to Z-scores.^[13] Children with a past history of chronic illness, previous fracture to the left hand, and whose height, weight, and BMI were >2 or <-2 SD for that age and gender

were excluded. Of the 1,016 children, 908 children (477 boys and 431 girls) met the inclusion and exclusion criteria. As the TW-3 method generates Z-scores in boys ≤16.5 years and girls ≤16 years,^[6] 57 children (who were older) were excluded. The study was carried out on 851 children (438 boys, 413 girls), who then underwent X-rays for BA assessment.

CA was computed by subtracting the difference between the date of birth and the date of X-ray taken. The 851 children were then divided into prepuberty and puberty groups for analysis; girls between 2 to 9 as prepubertal and 9.1 to 18 as pubertal, in case of boys between 2 to 11 as prepubertal and 11.1 to 18 as pubertal. We performed Preece-Baines growth modeling (PBG) of anthropometric data to evaluate the timing of pubertal height spurt.^[14] The PBGM is a family of curves that conform to the shape of the human growth curve and is one of the most widely accepted models of human growth. The model is represented by the formula:

$$\text{Height} = H1 - (2[H1 - H0]) / (-eS0[t-\text{theta}] + eS1[t-\text{theta}])$$

Where t = age. The equation parameter, $H1$ is the estimated adult height. The parameters $S1$ and $S0$ are rate constants, and $H0$ and theta are related to the height and age at the take-off of the adolescent growth spurt. The first and second derivatives of the equation are used to calculate the biological parameters viz. the age at peak height velocity (PHV), age at take-off ($T0$), the velocity at PHV, and velocity at $T0$. By using this model, it is possible to confirm the timing of pubertal height spurt in any given population, which is proxy to Tanner stage 2 in girls and Tanner stage 3 in boys. In our study, the age of pubertal spurt in boys was 13.9 and in girls, it was 11.7 years, respectively, which corresponds to the average timing of onset of puberty in Indian children.^[15]

Bone age X-ray

Roentgenograms of the left hand and wrist were performed by exposing them in posteroanterior (P-A) position with the elbow and wrist on the same axis, thumb at an angle of 30°, and fingers not too close or widely spaced. X-ray was centered overhead of the third metacarpal bone from a distance of 76 cm. The radiographic examinations were performed using a portable X-ray generator machine (Philips, Eindhoven, Netherlands) operating at 50 mA and the digital X-rays were generated by the standard computed radiography system (Fuji, Tokyo, Japan) by the same technician.

Bone age assessment

As this was a blinded study, subject identifying information, except gender of children, was masked during BA estimation of X-rays including subject's CA. Initially, 30 randomly selected X-rays, different from the ones used in the study, were analyzed by three methods of BA assessment (GP atlas, GR atlas, and TW-3 method) to check for interobserver variation. TW-3 method was based on age calculated from the Radius-Ulna-Short bone (RUS) score of 13 bones. These assessments were carried out independently by four pediatric endocrinologists (VK, NL, HP, and PP) trained in

methods of BA estimation. To assess intrarater variation, the 30 radiographs were analyzed initially and then were reevaluated by the same four observers independently by the three methods after a 4-week interval.^[11] If there were statistically significant inter-rater (disagreement among the four observers) and intra-rater differences, they were checked for clinical significance. If the differences were within the range of SDs given in each of the methods of BA assessment, it was considered as clinically not significant and if they were out of range, it was considered as clinically significant.^[4-6] As there were no significant clinical interobserver and intraobserver variations, all the 851 X-rays (blinded except for gender) were analyzed by the four observers independently by all three methods

Statistical analysis

For statistical analysis, the final BA for each of the three methods for each child was the mean of readings calculated by all four observers for that particular child for that method. These mean bone ages were then converted into Z-scores. For GP and GR atlas, Z-scores were calculated using mean ages and standard deviations given in the atlas.^[4,5] For the TW-3 method, Z-scores were calculated by the software provided with the atlas from the CA, gender, and RUS scores.^[6]

Statistical analyses were carried out using SPSS (v. 26, Chicago, IL, USA). Outcome variables were tested for normality before performing analyses. Interobserver variation was assessed using one-way ANOVA and intraobserver variation was tested using paired sample *t*-test. Correlations between CA and BA measured by each method were analyzed using Pearson correlation coefficients. The correlation coefficient of 0.7 to 1 was defined as strong, 0.4 to 0.7 as moderate, and <0.4 as weak.^[16] For purpose of deciding which method of BA was most suitable in our cohort, tests of proportions, and root mean square (RMS) deviations were computed. Z-scores of differences between BA and CA in the range of +1 and -1 were defined as within normal limits and anything greater than or less than 1 SD was defined as outside normal limits (4, 5). This was then calculated as the proportion of children within normal limits by each of the three methods and the differences between these proportions were computed using related sample Cochran's Q test. The method with the highest proportion of correct classification within normal limits was considered the best fit. Similarly, RMS deviations were also calculated for the differences between BA and CA to confirm the results of the test of proportions. The method with the least RMS deviation represented the most suitable method of BA estimation among the three methods. *P* values <0.05 were considered statistically significant.

RESULTS

Initially, bone ages on the 30 children's X-rays (15 boys and 15 girls) as assessed by all four observers using three different methods of BA assessment were analyzed for interobserver and intraobserver variation. Although there were statistically

significant intraobserver differences for GP, GR, and TW-3 methods, none of these were clinically significant (as judged by SD's given in the methods, e.g., for GP method, VK and PP showed a difference of 3.4 and 4.1 months, respectively, which though statistically significant did not reach the clinical significance of 9.0 and 10.7 months, respectively, given by the GP atlas). There were no statistically or clinically significant interobserver differences for the GP, GR atlas, and TW-3 method ($P > 0.1$). Based on these observations, there were no clinically significant interobserver and intraobserver differences, the mean SD of BA assessment by the observers for each method was considered for further analysis.

Data on 851 children (438 boys, 413 girls), aged 2–16.5 including mean CA, anthropometric parameters, and BA assessed by different methods (GP atlas, GR atlas, and TW-3 method) has been illustrated in Tables 1 and 2.

The correlations between CA and BA for each method of assessment for different age groups and genders have been described in Table 3. There were statistically significant correlations between CA and BA for all methods (GP atlas, GR atlas, and TW3 method) ($P < 0.05$) as a whole and as per gender and age group [Table 3].

When we compared, proportions of successful classification of differences between CA and BA within normal limits (BA should lie within ± 1 SD of CA) by different methods, the TW-3 method overall was the most suitable method ($P < 0.05$) [Figure 1]. Similarly, for girls and boys, the TW-3 method was the most suitable method of BA assessment (boys, $P = 0.08$; and girls, $P < 0.05$) [Figure 1]. On dividing groups based on whether puberty was achieved or not, the TW-3 method was again the most applicable in prepubertal boys ($P < 0.05$), prepubertal girls ($P > 0.1$), and pubertal girls ($P < 0.05$) [Figure 2]. However, in pubertal boys, the GR atlas method was the most suitable ($P < 0.05$). When we computed RMS deviation on differences between CA and BA for each method of BA assessment, results showed a similar trend to the proportions test. The TW-3 method had

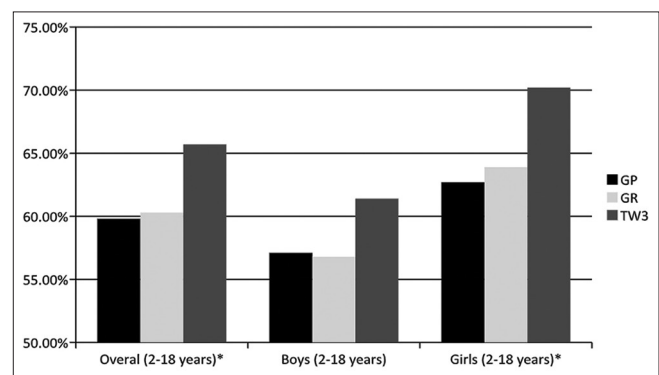


Figure 1: Proportion of successful classification of bone ages within ± 1 SD (i.e., normal limits) of the chronological age using different methods of bone age assessment across all children overall and also according to genders (* $P < 0.05$). GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

Table 1: Demographic characteristics and mean bone age by different methods in boys

Age groups	n	Chronological age	Height Z-scores	Weight Z-scores	BMI Z-scores	Mean GP age	Mean GR age	Mean TW3 age
2+	22	2.4±0.3	-0.4±1.1	-0.9±0.7	0.2±0.9	2.3±0.5	2.2±0.5	2.5±0.6
3+	29	3.5±0.3	-0.5±0.9	-0.6±1.0	-0.2±0.7	3.1±0.8	2.8±0.8	3.3±0.8
4+	28	4.5±0.2	-0.7±0.8	-0.4±0.8	-0.6±0.6	3.9±0.5	3.7±0.6	4.4±0.5
5+	24	5.8±0.3	-1.0±0.7	-0.5±0.9	-1.0±0.8	4.9±0.7	4.9±0.9	5.2±0.8
6+	35	6.5±0.3	-1.1±0.6	-0.8±0.7	-0.9±0.7	5.4±1.1	5.5±1.2	6.0±1.2
7+	30	7.5±0.3	-0.8±0.7	-0.5±0.7	-0.8±0.9	6.6±1.1	6.8±1.3	7.0±1.0
8+	30	8.4±0.3	-1.0±0.5	-0.5±0.7	-1.1±0.6	7.9±1.6	8.2±1.9	8.2±1.3
9+	30	9.4±0.3	-0.9±0.8	-0.9±0.7	-0.6±0.8	8.8±1.3	9.6±1.1	8.9±0.8
10+	30	10.4±0.3	-1.3±0.6	-1.1±0.6	-1.0±0.7	9.6±1.6	10.1±1.6	9.5±1.4
11+	31	11.4±0.3	-1.1±0.8	-0.8±0.9	-0.9±0.7	11.0±1.4	11.7±1.5	10.8±1.4
12+	30	12.4±0.3	-1.0±0.8	-0.9±0.7	-0.8±0.9	12.1±1.0	12.7±1.0	11.5±1.3
13+	29	13.5±0.2	-0.8±0.8	-0.8±1.0	-0.7±0.8	13.0±0.9	13.3±0.6	12.6±1.2
14+	29	14.4±0.3	-0.7±0.7	-0.9±1.0	-0.5±0.7	13.6±1.1	13.8±0.7	13.4±1.2
15+	38	15.4±0.3	-0.6±0.8	-0.6±0.9	-0.5±0.9	15.0±1.4	14.9±1.2	14.9±1.2
16+	23	16.3±0.2	-0.7±0.8	-0.9±1.1	-0.4±0.8	16.2±1.3	15.9±1.3	15.7±1.2
Total	438	9.6±4.2	-0.9±0.8	-0.7±0.8	-0.7±0.8	9.0±4.4	9.2±4.5	9.1±4.0

GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

Table 2: Demographic characteristics and mean bone age by different methods in girls

Age groups	n	Chronological Age	Height Z-scores	Weight Z-scores	BMI Z-Scores	Mean GP age	Mean GR age	Mean TW3 age
2+	21	2.5±0.3	-0.7±1.0	-0.4±0.8	-0.3±0.8	2.6±0.6	2.3±0.6	2.8±0.6
3+	28	3.5±0.2	-0.5±1.0	-0.2±1.0	-0.3±0.7	3.5±0.8	3.2±0.9	3.6±0.5
4+	28	4.5±0.3	-0.6±1.0	-0.5±0.9	-0.4±0.9	3.9±0.7	3.5±0.8	4.0±0.7
5+	28	5.8±0.2	-0.6±0.7	-0.3±0.7	-0.6±0.7	5.4±0.7	5.4±0.8	5.3±0.7
6+	32	6.4±0.3	-0.6±0.7	-0.4±0.7	-0.6±0.7	6.1±0.9	6.3±1.2	6.2±1.1
7+	28	7.4±0.3	-0.6±1.0	-0.4±0.8	-0.6±1.1	6.6±0.8	7.3±1.0	7.0±1.1
8+	32	8.5±0.3	-1.0±0.6	-0.9±0.7	-0.8±0.7	8.1±1.1	8.9±0.8	8.5±1.0
9+	29	9.5±0.3	-1.3±0.6	-1.0±0.8	-1.0±0.6	8.7±1.5	9.2±1.1	9.1±1.4
10+	30	10.4±0.3	-0.9±0.7	-0.7±0.7	-0.8±0.8	10.1±1.5	10.3±1.0	10.5±1.3
11+	28	11.4±0.3	-1.0±0.7	-0.8±0.8	-0.8±0.6	11.1±1.1	11.0±0.8	11.2±0.9
12+	32	12.4±0.3	-0.8±0.8	-0.5±0.8	-0.7±0.8	12.7±1.4	12.7±1.5	12.5±1.2
13+	30	13.5±0.3	-0.8±0.9	-0.5±0.9	-0.7±0.7	13.8±1.9	14.0±2.0	13.4±1.4
14+	29	14.4±0.3	-0.8±0.6	-0.6±0.6	-0.7±0.6	14.8±0.7	15.1±0.8	14.3±0.6
15+	38	15.4±0.3	-0.8±0.9	-1.0±0.9	-0.5±0.8	15.7±1.1	16.0±1.2	14.7±0.6
Total	413	9.3±4.0	-0.8±0.8	-0.6±0.8	-0.6±0.8	9.1±4.3	9.3±4.5	9.1±4.1

GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

Table 3: Correlation of chronological age with mean bone age calculated by different methods across various groups

Groups	n		Mean GP	Mean GR	Mean TW3
Overall (2-18 years)	851		0.965	0.967	0.967
Boys (2-18 years)	438	Pearson's	0.967	0.967	0.967
Girls (2-18 years)	413	Correlation	0.969	0.972	0.967
Boys (2-10.9 years)	258	Coefficient	0.910	0.917	0.923
Girls (2-8.9 years)	197		0.906	0.922	0.909
Boys (11-18.9 years)	180		0.834	0.797	0.835
Girls (9-18.9 years)	216		0.883	0.892	0.870

All correlations above were significant, P<0.05. GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

the least RMS deviation [Table 4], thus suggesting that the TW-3 method was possibly the most suitable method for BA assessment of children in our study.

Figures 3 and 4 summarize the relationship of BA with CA using three different methods of assessment. In boys, till the age of 9 years, BA estimated by TW-3 method was closest to the CA (maximum delay = 6 months). However, in pubertal years, GR atlas estimated bone ages were closest to the CA, whereas GP atlas and TW-3 method underestimated the BA. As compared with this, TW-3 bone ages were close to the CA in girls irrespective of their pubertal status. Interestingly, GP and GR atlases underestimate BA till 12 years, after which they overestimate the BA.

Table 4: Root mean square (RMS) error values of the bone ages calculated by different methods of assessment across various groups

	n	GP	GR	TW3
Overall (2-18 years)	851	1.21	1.18	1.11
Boys (2-18 years)	438	1.26	1.21	1.14
Girls (2-18 years)	413	1.16	1.14	1.00
Prepubertal boys (2-10.9 years)	258	1.28	1.29	1.06
Prepubertal girls (2-8.9 years)	197	0.92	0.97	0.87
Pubertal boys (11-18 years)	180	1.21	1.09	1.37
Pubertal girls (9-18 years)	216	1.35	1.28	1.12

GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

DISCUSSION

Our cross-sectional, observational study in healthy Indian children shows that in methods used for BA assessment, the TW-3 method was overall the most suitable. This statement holds true for girls irrespective of pubertal status and prepubertal boys. However, for boys in puberty, the GR method was more suitable and BA assessment using the TW-3 method was less precise and needs to be used with caution. Also, irrespective of the method used, BA is underestimated at all ages in boys and girls till the pubertal growth spurt. In girls, after a pubertal growth spurt, BA rapidly advanced and was overestimated by all methods.

GP method, the most commonly used method worldwide, is based on Caucasian children studied in the early 20th century. Thus, when the GP method is used irrespective of the ethnicity of the child, it may lead to errors and imprecision in its estimation. Similar findings of BA estimation using the GP method being imprecise in Indian children and children belonging to different ethnicities have been documented earlier.^[7,17-24] In our study, GP atlas underestimated BA in prepubertal children (irrespective of pubertal status) and overestimated BA in girls after 12 years. Similar findings of BA being underestimated in Korean, French, and Indian children have been documented in studies by Kim *et al.*, Zabet *et al.*, and Patel *et al.*, respectively.^[10,20,23]

There have been many studies worldwide that have compared different methods of BA estimation.^[9-11,24-26] Studies using correlation coefficients (which only measure the strength of association and not an agreement between methods) done by Kim *et al.*,^[9] Lin *et al.*,^[11] and Kim *et al.*^[10] showed no differences between various BA methods. Similarly, in our study, the correlation between all three methods of BA estimation was very strong. However, when we used the test of proportions (number of X-rays where estimated BA was within the normal limits, i.e., ± 1 SD of CA) and the RMS deviations, the TW-3 method was the most suitable method (except for pubertal boys). Similar findings of the TW method being more suitable than GP were reported by Bull *et al.*,^[12] Pinchi *et al.*, and Buken *et al.*^[24,25] The main reasons for TW-3 being more accurate as compared with other methods, are that the TW-3

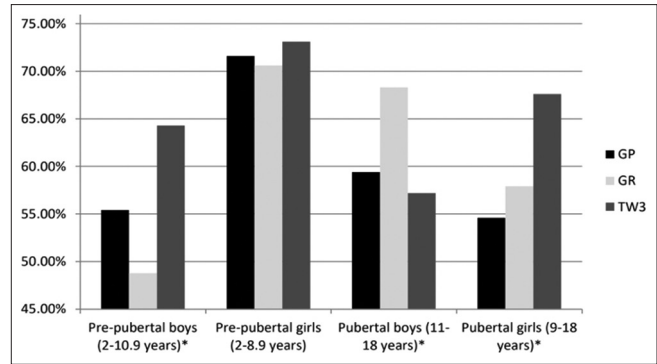


Figure 2: Proportion of successful classification of bone ages within ± 1 SD (i.e., normal limits) of the chronological age using different methods of bone age assessment across both prepubertal and pubertal groups separately (* $P < 0.05$). GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

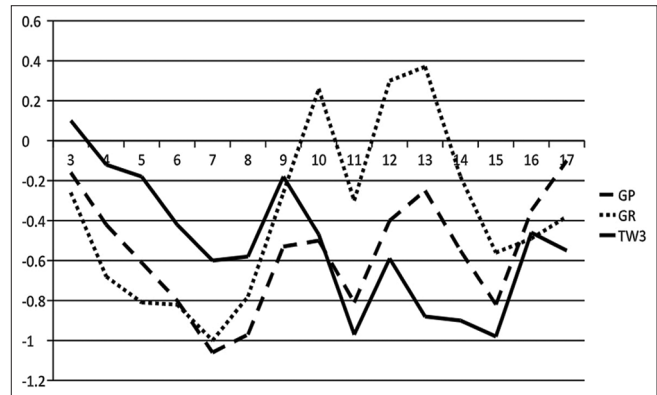


Figure 3: Graphical representation of the difference between bone age and chronological age (in years) in boys using the three different methods of bone age assessment (x-axis: chronological age in years and y-axis: the difference between bone age and chronological age in years). GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

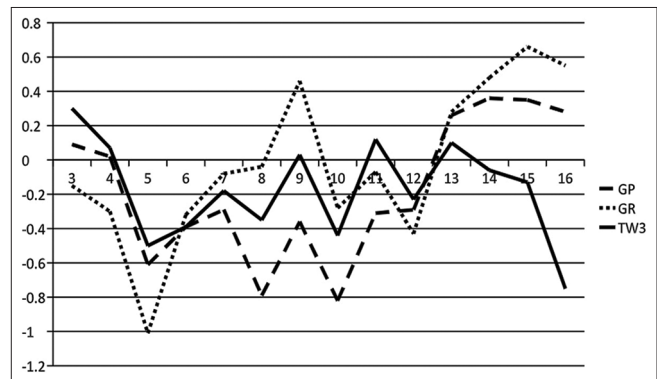


Figure 4: Graphical representation of the difference between bone age and chronological age (in years) in girls using the three different methods of bone age assessment (x-axis: chronological age in years and y-axis: the difference between bone age and chronological age in years). GP: Greulich and Pyle atlas, GR: Gilsanz and Ratib atlas, TW: Tanner Whitehouse method

method is based on a strong mathematical base rather than on pattern matching (GP and GR method) and multiethnic

populations (Europeans, Americans, and Asians) have been used in its development. Furthermore, the TW method has been updated over time (currently TW-3 is in use) and reflects secular trends making it more applicable.^[6] Also, the TW-3 method gives BA at 1 monthly interval as compared with GP and GR methods where the differences can range from 3 months to 1 year depending on the age of the child^[4-6]

As GR atlas is a relatively new method of BA estimation, only a handful of studies have assessed its applicability and there are no studies on Indian children.^[11,27] In the study by Schmidt *et al.*, GR atlas overestimated BA in 14- to 18-year-old girls by up to 0.6 years and underestimated BA in 14–18 boys by up to 0.5 years.^[27] In our study, when we assessed children older than 13 years, in girls, BA was overestimated by up to 0.7 years and in boys, was underestimated by up to 0.6 years. Interestingly, we found that the GR method was less accurate as compared with GP and TW-3 methods in the prepubertal years, and the accuracy improved in puberty.

There is a paucity of studies on Indian children using the TW method. These studies were performed a few decades ago and used the TW-2 method of estimation.^[8] Prakash *et al.* showed that when Indian children were assessed using the TW-2 method, RUS scores for 6 to 7 year old, and 13- to 14-year-old boys were advanced, and in 8- to 12-year-old boys, the RUS scores were on the 50th centile. In the same study, girls' RUS maturity scores were slightly advanced except at 6 years. To our knowledge, ours is the first study using the TW-3 method for estimation of BA in Indian children.

Interestingly, in our study BA estimated in boys using any of the methods was delayed as compared with the CA (GP much more than TW-3) except for a few age groups where the GR method was mildly advanced. On the other hand, in Indian girls, BA was underestimated by all methods till about 12 years (although less than in boys) and after 12 years it rapidly advanced. This pattern of skeletal development is peculiar and very different from other Asian populations. In Chinese populations, GP and GR methods underestimated BA in 5- to 10-year-old boys and overestimated after 12 years, whereas, after 5 years in girls the BA was overestimated (less than in Chinese boys).^[11] In Thai children (8–16 years), where BA was assessed by GP and TW-3 methods, BA was underestimated in boys till 12 years of age after which it was overestimated, and in girls the BA was overestimated at all ages.^[26] Given these differences noted in BA estimation of different Asian populations, country-specific standards may be more useful in determining BA.

The strengths of our study include a large number of children analyzed, we also studied children from both extremes of age (preschool and adolescent). To our knowledge, this is the first study assessing the applicability and suitability of GR and TW-3 methods of BA assessment in Indian children. Also, this is the first Indian study, comparing BA in Indian children using three different methods (GP, GR, and TW-3 methods). Finally, more appropriate statistical methods (test of proportions and

RMS deviation) to compare BA assessment by different methods (as compared with the correlation coefficients used by most studies) have been demonstrated in our study.

Our study is limited by the fact that children involved in the study were from a single-center and these findings may be variable in different parts of the country. Furthermore, the absence of sexual maturity data in our cohort is another limitation of our study; we have tried to overcome this limitation by using the Preece-Baines modeling to evaluate the timing of pubertal height spurt (which correlates with sexual maturity). This age-based presumption of pubertal status has its limitations; as in any given population, there may be subjects of early and late puberty, which may affect the bone ages. Hence, the results of this study based on age-based classification of the pubertal status should be interpreted with caution and need to be validated with future studies.

In conclusion, BA is underestimated in Indian boys irrespective of the method used. In Indian girls, bone age is underestimated till pubertal growth spurt, after which there is a rapid advancement of BA. Among the three methods of bone age assessment (Grulich Pyle, Gilsanz Ratib, and Tanner Whitehouse-3), BA's estimated by the TW-3 method were closest to the chronological age in our cohort (except for pubertal boys). Hence, it seems reasonable to recommend the use of the TW-3 method for BA estimation in Indian girls of all ages. Although the TW-3 method may be used in younger Indian boys, it should be used with caution in older boys, especially till an Indian standard bone age atlas is developed.

Declaration of patient consent

The authors certify that they have obtained all appropriate patient consent forms. In the form the patient(s) has/have given his/her/their consent for his/her/their images and other clinical information to be reported in the journal. The patients understand that their names and initials will not be published and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Satoh M. Bone age: Assessment methods and clinical applications. *Clin Pediatr Endocrinol* 2015;24:143-52.
2. Topor LS, Feldman HA, Bauchner H, Cohen LE. Variation in methods of predicting adult height for children with idiopathic short stature. *Pediatrics* 2010;126:938-44.
3. Martin DD, Wit JM, Hochberg Z, Säwendahl L, Van Rijn RR, Fricke O, *et al.* The use of bone age in clinical practice-Part 1. *Horm Res Paediatr* 2011;76:1-9.
4. Greulich WW, Pyle SI. *Radiographic Atlas of Skeletal Development of the Hand and Wrist*. 2nd ed. Stanford: Stanford University Press; 1959.
5. Gilsanz V, Ratib O. *Hand Bone Age*. Los Angeles, CA: Springer; 2005.
6. Tanner JM, Healy MJR, Goldstein H, Cameron N. Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 method).

- London, UK: Saunders; 2001.
7. Alshamrani K, Messina F, Offiah AC. Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. *Eur Radiol* 2019;29:2910-23.
 8. Prakash S, Cameron N. Skeletal maturity of well-off children in Chandigarh, North India. *Ann Hum Biol* 1981;8:175-80.
 9. Kim SY, Oh YJ, Shin JY, Rhie YJ, Lee KH. Comparison of the Greulich-Pyle and Tanner Whitehouse (TW3) methods in bone age assessment. *J Korean Soc Pediatr Endocrinol* 2008;13:50-5.
 10. Kim JR, Lee YS, Yu J. Assessment of bone age in prepubertal healthy Korean children: Comparison among the Korean standard bone age chart, Greulich-Pyle method, and Tanner-Whitehouse method. *Korean J Radiol* 2015;16:201-5.
 11. Lin FQ, Zhang J, Zhu Z, Wu YM. Comparative study of Gilsanz-Ratib digital atlas and Greulich-Pyle atlas for bone age estimation in a Chinese sample. *Ann Hum Biol* 2015;42:523-7.
 12. Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. Bone age assessment: A large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods. *Arch Dis Child* 1999;81:172-3.
 13. Indian Academy of Pediatrics Growth Charts Committee, Khadilkar V, Yadav S, Agrawal KK, Tamboli S, Banerjee M, *et al.* Revised IAP growth charts for height, weight and body mass index for 5- to 18-year-old Indian children. *Indian Pediatr* 2015;52:47-55.
 14. Preece MA, Baines MJ. A new family of mathematical models describing the human growth curve [published correction appears in *Ann Hum Biol* 2013;40:298-9]. *Ann Hum Biol* 1978;5:1-24.
 15. Khadilkar V, Khadilkar A, Arya A, Ekbote V, Kajale N, Parthasarathy L, *et al.* Height velocity percentiles in Indian children aged 5-17 years. *Indian Pediatr* 2019;56:23-8.
 16. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med* 2018;18:91-3.
 17. Patil ST, Parchand MP, Meshram MM, Kamdi NY. Applicability of Greulich and Pyle skeletal age standards to Indian children. *Forensic Sci Int* 2012;216:200.e1-4. doi: 10.1016/j.forsciint. 2011.09.022.
 18. Shilpa PH, Sunil RS, Sapna K, Kumar NC. Estimation and comparison of dental, skeletal and chronologic age in Bangalore south school going children. *J Indian Soc Pedod Prev Dent* 2013;31:63-8.
 19. Manzoor Mughal A, Hassan N, Ahmed A. The applicability of the Greulich and Pyle Atlas for bone age assessment in primary school-going children of Karachi, Pakistan. *Pak J Med Sci* 2014;30:409-11.
 20. Zabet D, Rérolle C, Pucheux J, Telmon N, Saint-Martin P. Can the Greulich and Pyle method be used on French contemporary individuals? *Int J Legal Med* 2015;129:171-7.
 21. Garamendi PM, Landa MI, Ballesteros J, Solano MA. Reliability of the methods applied to assess age minority in living subjects around 18 years old. A survey on a Moroccan origin population. *Forensic Sci Int* 2005;154:3-12.
 22. Mansourvar M, Ismail MA, Raj RG, Kareem SA, Aik S, Gunalan R, *et al.* The applicability of Greulich and Pyle atlas to assess skeletal age for four ethnic groups. *J Forensic Leg Med* 2014;22:26-9.
 23. Patel PS, Chaudhary AR, Dudhia BB, Bhatia PV, Soni NC, Jani YV. Accuracy of two dental and one skeletal age estimation methods in 6-16 year old Gujarati children. *J Forensic Dent Sci* 2015;7:18-27.
 24. Pinchi V, De Luca F, Ricciardi F, Focardi M, Piredda V, Mazzeo E, *et al.* Skeletal age estimation for forensic purposes: A comparison of GP, TW2 and TW3 methods on an Italian sample. *Forensic Sci Int* 2014;238:83-90.
 25. Büken B, Erzenin OU, Büken E, Safak AA, Yazici B, Erkol Z. Comparison of the three age estimation methods: Which is more reliable for Turkish children? *Forensic Sci Int* 2009;183:103.e1-7. doi: 10.1016/j.forsciint. 2008.10.012.
 26. Benjavongkulchai S, Pittayapat P. Age estimation methods using hand and wrist radiographs in a group of contemporary Thais. *Forensic Sci Int* 2018;287:218.e1-8. doi: 10.1016/j.forsciint. 2018.03.045.
 27. Schmidt S, Nitz I, Schulz R, Tsokos M, Schmeling A. The digital atlas of skeletal maturity by Gilsanz and Ratib: A suitable alternative for age estimation of living individuals in criminal proceedings? *Int J Legal Med* 2009;123:489-94.