

Standardization of Weightage Assigned to Different Segments of the Hand X-ray for Assessment of Bone Age by the Greulich–Pyle Method

Chirantap Oza¹, Anuradha V. Khadilkar^{1,2}, Pranay Goel³, Tim Aeppli⁴, Shruti Mondkar¹, Nikhil Shah^{1,5}, Nikhil Lohiya⁶, Hemchand Krishna Prasad⁷, Prashant Patil⁸, Chidvilas More¹, Neha Kajale^{1,2}, Vaman Khadilkar^{1,2,9}, Lars Säwendahl⁴

¹Department of Growth and Bone, Hirabai Cowasji Jehangir Medical Research Institute, Pune, Maharashtra, India, ²Department of Health Sciences, Savitribai Phule Pune University, Pune, Maharashtra, India, ³Department of Biology, Indian Institute of Science Education and Research Pune, Maharashtra, India, ⁴Department of Women's and Children's Health, Karolinska Institutet, Astrid Lindgren Children's Hospital, Stockholm, Sweden, ⁵Department of Paediatrics, Cloudnine Hospital, Malad, Mumbai, Maharashtra, India, ⁶Department of Pediatric Endocrinology, Consultant Paediatric Endocrinologist, Division of Growth and Endocrinology, Silver Lining Paediatric Super Speciality Centre for Growth Development and Endocrine Care, Lokmat Square, Nagpur, Maharashtra, India, ⁷Department of Paediatric Endocrinology, Mehta Multispeciality Hospitals, Chennai, Tamil Nadu, India, ⁸Department of Pediatric Endocrinology, Consultant Paediatric Endocrinologist, SRCC NH CHILDREN'S Hospital, Mumbai and Apollo Hospital, Navi Mumbai, Maharashtra, India, ⁹Department of Pediatric Endocrinology, Senior Paediatric Endocrinologist, Jehangir Hospital, Pune, Maharashtra, India

Abstract

Introduction: Bone age (BA) assessment is important in evaluating disorders of growth and puberty; the Greulich and Pyle atlas method (GP) is most used. We aimed to determine the weightage to be attributed by raters to various segments of the hand x-ray, namely, distal end of radius-ulna (RU), carpals, and short bones for rating bone age using the GP atlas method. **Methods:** 692 deidentified x-rays from a previous study (PUNE-dataset) and 400 from the Radiological Society of North America (RSNA-dataset) were included in the study. Mean of BA assessed by experienced raters was termed reference rating. Linear regression was used to model reference age as function of age ratings of the three segments. The root-mean-square-error (RMSE) of segmental arithmetic mean and weighted mean with respect to reference rating were computed for both datasets. **Results:** Short bones were assigned the highest weightage. Carpals were assigned higher weightage in pre-pubertal PUNE participants as compared to RSNA, vice-versa in RU segment of post-pubertal participants. The RMSE of weighted mean ratings was significantly lower than for the arithmetic mean in the PUNE dataset. **Conclusion:** We thus determined weightage to be attributed by raters to segments of the hand x-ray for assessment of bone age by the GP method.

Keywords: Bone age, Greulich and Pyle, segment, weightage

INTRODUCTION

Bone age (BA) assessment is an important tool used in medicine, particularly for disorders of growth and puberty as it correlates better with height velocity, muscle mass, menarche and bone mineral mass than chronological age (CA).^[1] For the estimation of BA, a radiograph of the hand and wrist of the non-dominant hand is assessed using a traditional manual method or an automated method.^[2] The manual method involves a comparison of obtained radiograph with radiographs in reference atlases; the two most important publications in this field are by Greulich and Pyle (GP) and by Tanner, Whitehouse and Healy (TW).^[3,4]

The GP atlas was created based on radiographs of the hands of Caucasian upper-middle-class children living in Cleveland, Ohio, United States of America, between the years 1931 and 1942.^[5,6] The GP method is the most popular method among

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

Submitted: 06-Jun-2023

Revised: 30-Aug-2023

Accepted: 27-Nov-2023

Published: 29-Apr-2024

Supplementary material available online

Access this article online

Quick Response Code:



Website:
<https://journals.lww.com/indjem/>

DOI:
10.4103/ijem.ijem_237_23

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: Oza C, Khadilkar AV, Goel P, Aeppli T, Mondkar S, Shah N, *et al.* Standardization of weightage assigned to different segments of the hand X-ray for assessment of bone age by the Greulich–Pyle method. *Indian J Endocr Metab* 2024;28:160-6.

clinicians and radiologists for the assessment of BA; the assessment is relatively quick and easy to learn but reports suggest that there is considerable inter- and intra-observer variability in the ratings obtained.^[2] The GP atlas apart from the standard method (i.e., radiograph of the hand and wrist is compared with a standard radiograph from the GP atlas) also suggests a procedure to assess BA of a hand X-ray film by the rater by adapting the assessment as per his/her experience/preference. As per the atlas, after finding a standard X-ray plate that resembles closely the film to be assessed, the rater should make a detailed comparison of individual bones in the following order: distal end of radius and ulna, carpals, metacarpals and phalanges. Each bone may be either less or more advanced than its counterpart in the standard plate from the atlas. Thus, it is not merely sufficient to assign a single BA which corresponds most closely to the majority of bony centres visible in it.^[3] There may, in some radiographs, be a considerable difference in the BA of various hand bones; this is possibly because the tubular bones (metacarpals, phalanges, radius and ulna) mature by processes in the complex growth plates, while the carpals mature by a simpler process of bone apposition. These two mechanisms of growth of the skeleton are governed by different sets of hormones. The tubular bone maturation is more sensitive to sex hormones than the maturation of the carpals.^[7] In a recent study, we observed the mean advancement of carpal BA to be -1.1 ± 0.4 years which may be due to differences in the ethnicity/environment in Pune (in 2015) vs Ohio (1940).^[7]

Further, there is no standardization for assigning weights to bones of the wrist and hand; thus, a rater may assign different weights to different bones as per his/her experience/preference. This variation is particularly due to the carpal bones as some raters may ignore the carpals and others may assign 50% weight to the carpals during the assessment. Raters using the carpals reduce their weightage at higher maturity but again, there is no standardized method for assigning these weightages.^[8] Variation by ethnicity and earlier development of secondary sexual characteristics due to secular trends further complicate the problem.^[9-11] That BA estimation using the GP method may be imprecise in children of Indian ethnicity has been documented earlier by our group.^[12]

Rating each bone individually and assigning separate weightage to each of the 28 bones of the hand and wrist are cumbersome and time-consuming. We thus, similar to the study by Carpenter *et al.*,^[13] for the purpose of BA assessment by the GP method, divided the hand and wrist bones according to the anatomical differences and maturational disparity, into three segments. We divided the whole hand X-ray film into three segments: 1) the distal ends of radius and ulna, 2) the carpals and 3) the tubular bones, that is metacarpals and phalanges. Given that there is no standard weightage attributed to each bone of the wrist and hand film in the GP atlas method and no particular adjustment for racial/ethnic variations and secular trends exists till date, our study aimed to standardize weightage assigned to segments of the hand while using the

GP method of BA assessment. Our specific objective was to determine the weightage to be attributed by raters to various segments of the hand X-ray namely, distal end of long bones of the forearm (radius and ulna), carpals and short bones of hand (metacarpals and phalanges) for rating BA using the GP atlas method.

MATERIALS AND METHODS

Subjects

We extracted (from a previous study, children who had height and weight within reference range for age and gender)^[14,15] de-identified left-hand radiographs of 692 healthy Indian subjects aged 1–19 years as per their BA assessment (referred henceforth as the PUNE dataset)^[12] and 400×-rays from the Radiological Society of North America (RSNA) dataset (available as free download from <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pediatric-bone-age-challenge-2017>; referred to henceforth as the RSNA dataset, accessed on May 2022).^[16]

BA measurement

The mean of the BA assessed by the GP method (using whole hand) by raters was considered as reference rating. The true rating was defined as the average rating of a large number of expert raters; the root mean square error (RMSE) of the reference rating relative to the true rating was obtained as the standard deviation of the manual ratings divided by the square root of the number of raters.^[7] The mean of the rating of the three segments of the X-ray is referred to as the arithmetic mean rating.

For the GP atlas assessment of the PUNE dataset, three paediatric endocrinologists (HP, PP and NL) who were skilled in the assessment of BA (with >10 years of experience) rated the overall (whole hand) BA. Further, an average of rating by two paediatric endocrinologists (VK and CO—experience 25 years and 5 years, respectively) who rated the same X-rays in a segmental manner, that is assigned a GP age to wrist (distal end of radius and ulna), carpals and hand bones (metacarpals and phalanges), was computed.^[13]

Similarly, for the RSNA dataset, the average of rating by multiple raters was considered reference rating (as available with the RSNA dataset) and two expert paediatric endocrinologists (LS and TA—experience around 25 years and 5 years, respectively) rated the X-rays segmentally.

Pubertal status

As the details on pubertal status were not available, we considered average age at onset of puberty in girls and boys for the PUNE dataset as 10.5 and 11.5 years, respectively.^[17,18] The age at onset of puberty for the RSNA dataset was considered 9.5 years and 11 years for girls and boys, respectively.^[19,20] The PUNE boys and girls under the age of 11.5 years and 10.5 years, respectively, and RSNA boys and girls under the age of 11 years and 9.5 years, respectively, are referred to as group A, while others as group B.

Statistical analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS) software for Windows (version 26.0, IBM Statistics Data Editor, IBM Corp., Armonk, NY). We assessed the correlation between arithmetic mean rating by the segmental method and reference rating (the average of our multiple manual ratings as the reference rating) using the Pearson correlation coefficient and the Bland–Altman analysis. Linear regression was used to model reference age as a function of the age ratings of the three segments. The regression analysis was run separately for the PUNE and RSNA datasets, girls and boys, and for pubertal status. All three segments were included as covariates for the below-puberty sub-groups in both girls and boys, while carpals were excluded in the over-puberty sub-groups. Intercept terms were excluded if not significant. Robust regression was used throughout to improve fits. The analysis was conducted in MATLAB R2022, using the function ‘fitlm’. We used stepwise regression to find the best-fit combination of variables and weightage in MATLAB R2022. The accuracy was defined as ‘the degree to which the information correctly described the phenomena it was designed to measure’. The accuracy was computed using the RMSE to assess how close the observed data (reference rating) were to values predicted by the model of weighted segmental rating. To compute RMSE, calculate the residual (difference between prediction and truth) for each data point, compute the norm of residual for each data point, compute the mean of residuals and take the square root of that mean. We applied an independent-sample *t*-test to compare the accuracy of the weighted segmental rating with arithmetic mean rating. The mean absolute deviation (i.e., the average distance between each data point and the mean) of segmental arithmetic mean and weighted mean with respect to reference rating was computed. A *P*-value of <0.05 was considered significant.

Ethical Aspect

The original study (PUNE) was approved by the Institutional Ethics Committee, and all participants gave written informed consent (original approval was dated 7 August 2019). This study was approved by the Ethics Committee of the Indian Institute of Science Education and Research, Pune (IHEC/Admin/2021/014).

RESULTS

PUNE dataset

A total of 692 de-identified X-rays of Indian children aged 1 to 19 years were included in the study. The sample composition is shown in Table 1a. The regression model results are illustrated in Table 2a to obtain the reference rating using the weighted segmental method of rating. Short bones had the highest weightage irrespective of gender and pubertal status (>50%). Carpals were given 27% and 36% weightage in group A boys and girls, respectively. The radius and ulna segment received 21% and 17% weightage in group B boys and girls, respectively. We observed high correlation between arithmetic mean and weighted mean of 0.995. The correlation of arithmetic mean and weighted mean with

Table 1a: Study population by BA and gender for the PUNE dataset

Age-group	Boys	Girls	Total
1-2	12	3	15
2-3	19	27	46
3-4	41	30	71
4-5	35	19	54
5-6	18	27	45
6-7	26	14	40
7-8	18	20	38
8-9	24	28	52
9-10	14	10	34
10-11	17	16	33
11-12	21	27	48
12-13	25	9	34
13-14	24	18	52
14-15	14	9	23
15-16	9	28	37
16-17	21	15	36
17-18	20	26	46
18-19	8		8
Total	366	326	692

Table 1b: Study population by BA and gender for the RSNA dataset

Age-group	Boys	Girls	Total
0-1	6	7	13
1-2	14	10	24
2-3	10	13	23
3-4	10	10	20
4-5	10	10	20
5-6	10	10	20
6-7	10	10	20
7-8	10	10	20
8-9	10	10	20
9-10	9	10	19
10-11	17	24	41
11-12	10	17	27
12-13	14	12	26
13-14	10	13	23
14-15	10	10	20
15-16	10	10	20
16-17	10	8	18
17-18	7	6	13
18-19	7	6	13
Grand total	194	206	400

the reference rating is illustrated in Figure 1. It was similar for both methods, that is 0.989 for arithmetic mean and 0.990 for weighted mean. The mean absolute deviation of segmental arithmetic mean and weighted mean with respect to reference rating was 0.64 and 0.51 years ($P < 0.05$). The Bland–Altman plots for arithmetic mean rating and the weighted mean by the segmental method and the reference rating are illustrated in Figure 2. The SD among the manual raters was 0.49 years (data not shown); thus, the SD

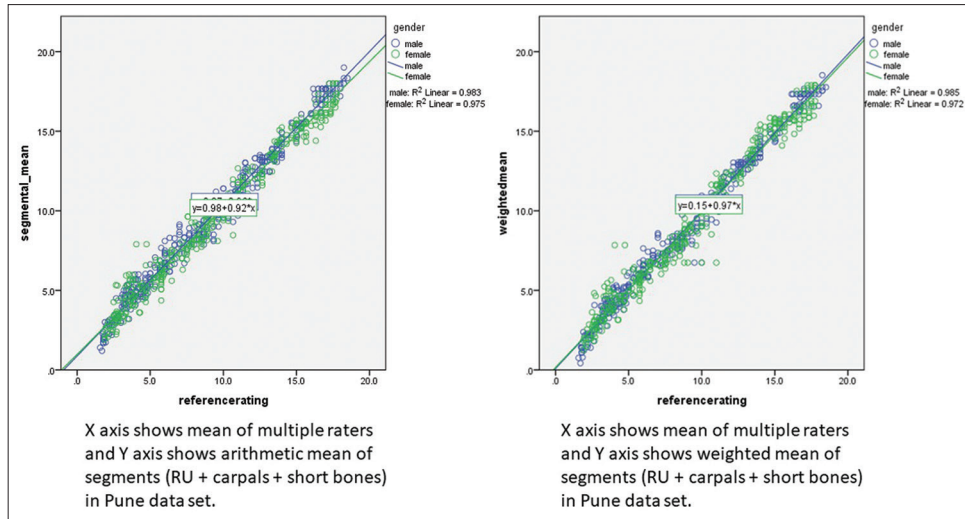


Figure 1: The correlation of arithmetic mean and weighted mean with the reference rating in Pune dataset

Table 2a: Weightage attributed to segments of wrist and hand films to obtain closest result to reference rating – segregated by gender and pubertal status for the PUNE dataset

Age and puberty	Intercept	Radius–ulna	Carpals	Short bones
Boys <11.5 years	-0.62	0.08	0.27	0.64
Boys >11.5 years		0.21		0.75
Girls <10.5 years	-0.45	0.11	0.36	0.51
Girls >10.5 years		0.17		0.81

Table 2b: Weightage attributed to segments of wrist and hand films to obtain closest result to reference rating – segregated by gender and pubertal status for the RSNA dataset

Age and puberty	Intercept	Radius–ulna	Carpals	Short bones
Boys <11 years	0.16	0.23	0.15	0.56
Boys >11 years		0.32		0.67
Girls <9.5 years		0.30	0.15	0.54
Girls >9.5 years		0.30		0.70

between reference and true ratings was 0.28 years. For the PUNE dataset, the RMSE of weighted mean ratings was significantly lower ($P < 0.05$) than for the arithmetic mean [Table 3a].

RSNA dataset

A total of 400 de-identified X-rays of Caucasian children selected randomly (aged 0 to 19 years) were used for the study. The sample composition is shown in Table 1b. The results of the regression model are illustrated in Table 2b to obtain reference rating using the weighted segmental rating. Carpals were given 15% weightage in group A children irrespective of gender. Short bones received highest weightage in subjects of both genders irrespective of age (>50%). The radius and ulna segment received 32% and 30%

weightage in group B boys and girls, respectively. We observed very high correlation between arithmetic mean and weighted mean of 0.996. The correlation of arithmetic mean and weighted mean with reference rating is illustrated in Figure 3. It was similar for both methods, that is 0.990 for arithmetic mean and 0.986 for weighted mean. The mean absolute deviation of arithmetic mean and weighted mean with respect to reference rating was 0.40 and 0.42 years ($P = 0.526$). The Bland–Altman curves are illustrated in Figure 4. The SD among six manual raters was 0.68 years; thus, the SD between reference and true ratings was 0.31 years.^[21] The RMSE is shown in Table 3b. For the RSNA dataset, there was a trend towards RMSE of weighted mean ratings being higher than that for the arithmetic mean [Table 3b] ($P = 0.241$). An algorithmic flowchart of assessing bone age x-ray by weighted segmental method is shown in Figure 5.

DISCUSSION

We found that on regression analysis, carpals were assigned significant weightage in group A for both genders in the PUNE and RSNA datasets; however, more weightage to carpals was assigned in the PUNE dataset (27% in boys and 36% in girls) as compared to the RSNA dataset (15% for both genders). The metacarpal and phalangeal segments were assigned the highest weightage in group B subjects of both genders in both datasets. In the RSNA dataset, a significantly higher weightage was assigned to the radius–ulna segment (30% for both genders) as compared to the PUNE dataset (21% in boys and 17% in girls). We found a high correlation between weighted mean rating using three segments of the hand and wrist X-ray with reference rating for both the PUNE and RSNA datasets. The RMSE of weighted mean ratings with respect to the true rating was significantly lower than for the arithmetic mean for the PUNE dataset, while it was higher for the RSNA dataset.

Similar to our study, a study on younger children (less than 10 years old) also reported that in the study group as a whole, hand BA more closely approximated CA than did wrist or carpal BA.

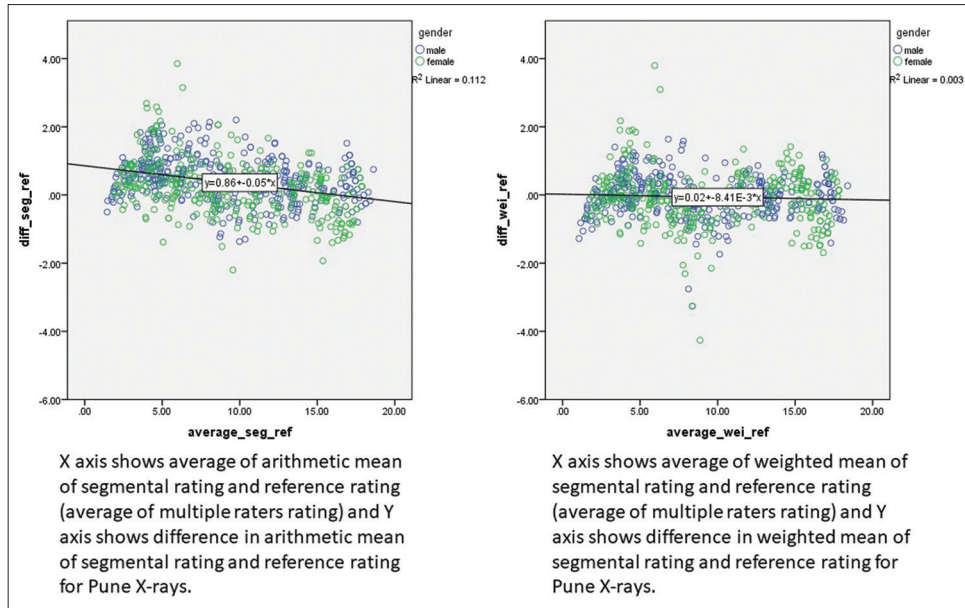


Figure 2: Bland Altman plots for arithmetic mean rating and the weighted mean by the segmental method and the reference rating in Pune dataset

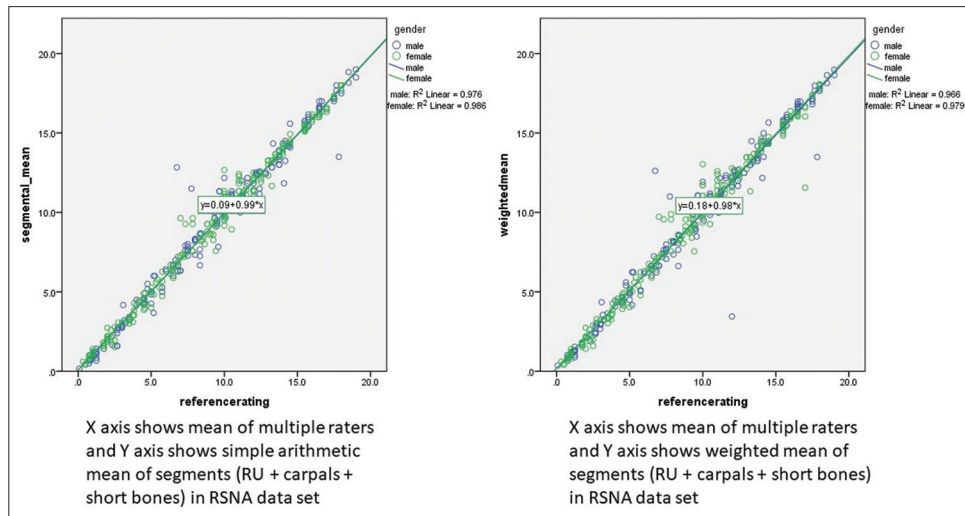


Figure 3: The correlation of arithmetic mean and weighted mean with reference rating in RSNA dataset

Table 3a: Root mean square errors of the arithmetic mean and weighted mean methods of BA estimation with respect to reference rating and true rating for the PUNE dataset

Rating/method	Arithmetic mean	Weighted mean
Reference	0.82	0.70
True	0.78	0.64

Table 3b: Root mean square error of the arithmetic mean and weighted mean methods of BA estimation with respect to reference rating and true rating for the RSNA dataset

Rating/method	Arithmetic mean	Weighted mean
Reference	0.70	0.84
True	0.63	0.78

The authors of the study also observed that rapid assessment of BA using only the carpals or giving higher weightage to carpals leads to underestimation of the BA.^[13] Moreover, the maturity of carpal bones varies greatly and hence is a major reason for variation in BA estimation among raters. We also, in a previous study, observed retardation of carpal BA as compared to tubular BA, an effect that was more pronounced in boys.^[7] A study to understand maturational disparity between hand–wrist bones in Hong Kong Chinese children by Lee *et al.*^[22] divided the 28 bones of the hand and wrist into six groups so that each group consisted of bones from the same row. The groups were as follows: (1) distal radius and distal ulna, (2) seven carpals, (3) five metacarpals, (4) five proximal phalanges, (5) four middle phalanges and (6) five distal phalanges. The authors observed that when the mean skeletal age (SA) of all the six groups was considered together, a tendency

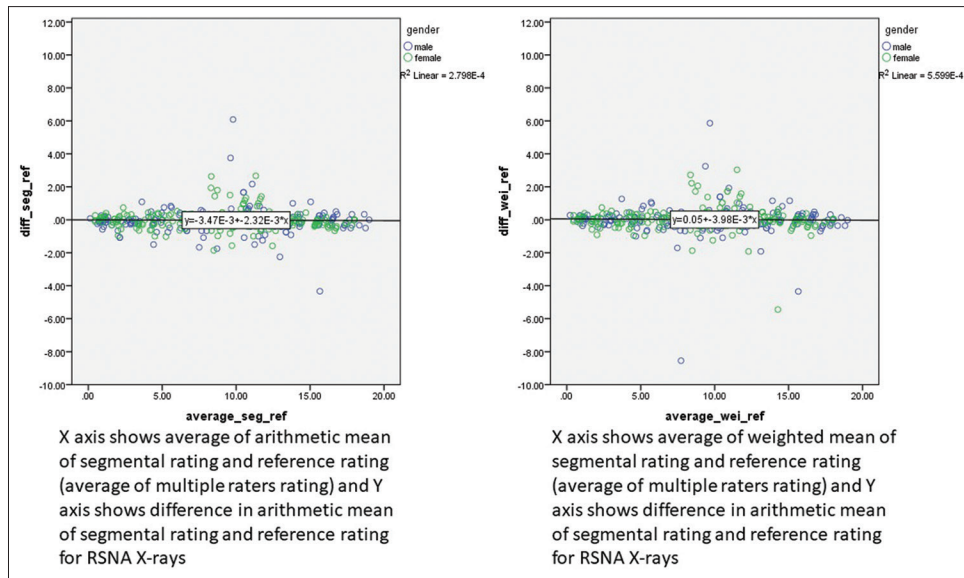


Figure 4: Bland Altman plots for arithmetic mean rating and the weighted mean by the segmental method and the reference rating in RSNA dataset

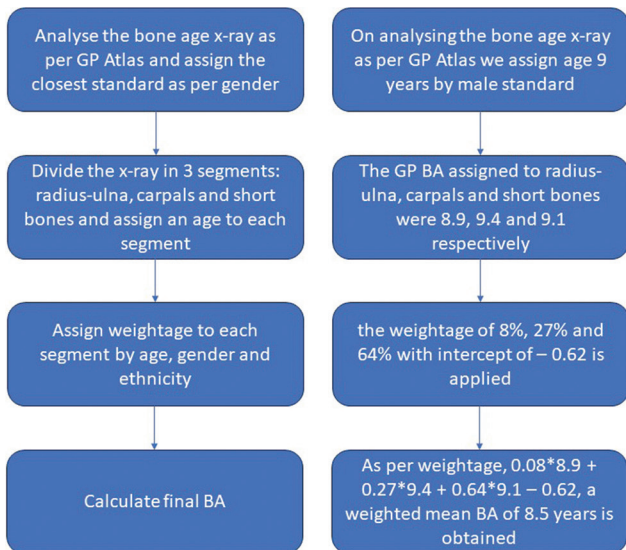


Figure 5: An algorithmic flowchart of assessing bone age x-ray by weighted segmental method

was observed for the two proximal rows of bones (distal radio-ulna and carpals) to be more retarded, the two distal rows (middle and distal phalanges) were relatively more advanced, and the two groups in-between (metacarpals and proximal phalanges) were intermediate in their maturation. Their SA values of metacarpals and phalanges were very similar to that of the overall SA.

On performing multiple linear regression analysis, we obtained the highest weightage for short bones of the hand, that is the metacarpals and phalanges. Lee *et al.*^[22] also noted several significant differences between the overall SA and the mean SA of the various bone groups, particularly the carpals. They observed that if the SA of the present group were rated by exclusion of the carpals, a significantly higher estimate of

overall SA was obtained, while, if the influence of the carpals was increased by weighing them to as much as 50% of the total score, a considerably lower estimate of the overall SA was obtained. In an attempt to establish backward compatibility to the reference rating using automated methods in children of Indian ethnicity, our group has previously published data on various composite ages with variable weightage on carpal BA (0%, 25%, 50%, 75% and 100%). We found that the best results were obtained with 50% weightage to carpals. However, these ratings were conducted using artificial intelligence software and not manually and arbitrary weightage was assigned to carpals and tubular bones.^[7] Another study has also observed that a better fit between individual BA and total BA was obtained for distal radius-ulna and metacarpals as opposed to carpals.^[23]

A study, which aimed to simplify the Tanner-Whitehouse III (TW3) method, formed a grouped-TW algorithm (GTA) wherein group 1 was composed of the maturity pattern of the radius and the middle phalange of the third and fifth digits, and three weights were obtained by data mining, yielding a result similar to that of the TW3 method. Subsequently, new BA assessment tables were constructed for boys and girls by linear regression while maintaining accuracy.^[24] We used a similar concept but the rating was made using the GP method and weightages as provided by the TW3 method were not used. Also, the previously referred study was conducted on children of Taipei as opposed to our study being performed on children of Indian (PUNE dataset) and Caucasian (RSNA dataset) ethnicity. Another Chinese study attributed maturational disparity mainly to middle phalanges and metacarpals. However, they observed that the BA of the proximal and distal phalanges was closer to the CA in both genders at most ages.^[25]

To the best of our knowledge, this is the first study that has established backward compatibility to estimate the weightage

attributed to bones of wrist and hands classified into three groups based on anatomy in children of Indian ethnicity. The limitation of this study is lack of details of pubertal status of individual subjects whose X-rays were used in the study, unknown ethnicity of subjects whose data were used from the RSNA dataset, lack of data on infants (0–1 years old) in PUNE dataset and data from a single centre in the PUNE dataset that may not be nationally representative. Moreover, validation studies with larger cohorts and multicentric analysis are required to assess the utility of this method of BA assessment. However, we present a representative case wherein the segmental weightage method of BA assessment was used in Appendix A.

CONCLUSION

We have determined the weightage to be attributed by raters to various segments of the hand X-ray for assessment of BA by the GP method. The weightages differed with ethnicity, gender and pubertal status. Paediatric endocrinologists and radiologists in India while rating BA by the GP method should consider rating the X-rays segmentally and then computing the BA based on weightage assigned.

Authors' contribution

All the authors were involved in designing and planning the project, data collection, manuscript writing and final approval of the manuscript, data analysis and interpretation of the data.

Acknowledgement

We thank the Radiological Society of North America for the access to use the imaging datasets and annotations for the purposes of academic research and education, and other non-commercial purposes.

Financial support and sponsorship

The project was funded by grants from the Department of Biotechnology (No BT/PR41073/Swdn/135/4/2020) and Vinnova (No 2020-03609).

Conflicts of interest

There are no conflicts of interest.

REFERENCES

- 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.
- Prokop-Piotrkowska M, Marszałek-Dziuba K, Moszczyńska E, Szalecki M, Jurkiewicz E. Traditional and new methods of bone age assessment-an overview. *J Clin Res Paediatr Endocrinol* 2021;13:251-62.
- Greulich WW, Pyle S. Radiographic atlas of skeletal development of the hand and wrist. Stanford: Stanford University Press; 1959.
- Tanner JM, Whitehouse RH, Healy M. A new system for estimating skeletal maturity from the hand and wrist with standards derived from a study of 2600 healthy British children. Paris: Centre International de L'enfance; 1962.
- Satoh M. Bone age: Assessment methods and clinical applications. *Clin Paediatr Endocrinol* 2015;24:143-52.
- Mughal AM, Hassan N, Ahmed A. Bone age assessment methods: A critical review. *Pak J Med Sci* 2014;30:211-5.
- Oza C, Khadilkar AV, Mondkar S, Gondhalekar K, Ladkat A, Shah N, *et al.* A comparison of bone age assessments using automated and manual methods in children of Indian ethnicity. *Pediatr Radiol* 2022;52:2188-96.
- Roche AF, Rohmann G, French NY, Davila H. Effect of training on replicability of assessments of skeletal maturity (Greulich-Pyle). *Am J Roentgenol Radium Ther Nucl Med* 1970;108:511-5.
- Euling SY, Herman-Giddens ME, Lee PA, Selevan SG, Juul A, Sorensen TI, *et al.* Examination of US puberty-timing data from 1940 to 1994 for secular trends: Panel findings. *Pediatrics* 2008;121(Suppl 3):S172-91.
- Herman-Giddens ME, Steffes J, Harris D, Slora E, Hussey M, Dowshen SA, *et al.* Secondary sexual characteristics in boys: Data from the Pediatric Research in Office Settings Network. *Pediatrics* 2012;130:e1058-68.
- 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.
- Shah N, Khadilkar V, Lohiya N, Prasad HK, Patil P, Gondhalekar K, *et al.* Comparison of bone age assessments by Greulich-Pyle, Gilsanz-Ratib, and Tanner Whitehouse methods in healthy Indian children. *Indian J Endocrinol Metab* 2021;25:240-6.
- Carpenter CT, Lester EL. Skeletal age determination in young children: Analysis of three regions of the hand/wrist film. *J Pediatr Orthop* 1993;13:76-9.
- Khadilkar VV, Khadilkar AV. Revised Indian academy of pediatrics 2015 growth charts for height, weight and body mass index for 5–18-year-old Indian children. *Indian J Endocrinol Metab* 2015;19:470–6.
- WHO AnthroPlus for Personal Computers Manual: Software for Assessing Growth of the World's Children and Adolescents. Geneva: WHO; 2009. Available from: <https://www.who.int/toolkits/growth-reference-data-for-5to19-years/application-tools>.
- Halabi SS, Prevedello LM, Kalpathy-Cramer J, Mamonov AB, Bilbily A, Cicero M, *et al.* The RSNA pediatric bone age machine learning challenge. *Radiology* 2018;290:498-503.
- Khadgawat R, Marwaha RK, Mehan N, Surana V, Dabas A, Sreenivas V, *et al.* Age of onset of puberty in apparently healthy school girls from Northern India. *Indian Paediatr* 2016;53:383-7.
- Agarwal DK, Agarwal KN, Upadhyay SK, Mittal R, Prakash R, Rai S. Physical and sexual growth pattern of affluent Indian children from 5 to 18 years of age. *Indian Paediatr* 1992;29:1203-82.
- Eckert-Lind C, Busch AS, Petersen JH, Biro FM, Butler G, Bräuner EV, *et al.* Worldwide secular trends in age at pubertal onset assessed by breast development among girls: A systematic review and meta-analysis. *JAMA Paediatr* 2020;174:e195881. doi: 10.1001/jamapediatrics.2019.5881.
- Styne DM. Puberty, obesity and ethnicity. *Trends Endocrinol Metab*. 2004 Dec; 15(10):472-8.
- Martin DD, Calder AD, Ranke MB, Binder G, Thodberg HH. Accuracy and self-validation of automated bone age determination. *Sci Rep* 2022;12:6388.
- Lee MM. Maturation disparity between hand-wrist bones in Hong Kong Chinese children. *Am J Phys Anthropol* 1971;34:385-95.
- Jiménez-Castellanos J, Carmona A, Catalina-Herrera CJ, Viñuales M. Skeletal maturation of wrist and hand ossification centers in normal Spanish boys and girls: A study using the Greulich-Pyle method. *Cells Tissues Organs* 1996;155:206-11.
- Hsieh CW, Liu TC, Wang JK, Jong TL, Tiu CM. Simplified radius, ulna, and short bone-age assessment procedure using grouped-Tanner-Whitehouse method. *Pediatr Int* 2011;53:567-75.
- Zhang J, Lin F, Ding X. Maturation disparity between hand-wrist bones in a Chinese sample of normal children: An analysis based on automatic BoneXpert and manual Greulich and Pyle atlas assessment. *Korean J Radiol* 2016;17:435-42.

APPENDIX A

Example of calculations of BA by the segmental weightage method

Case 1: For the X-ray image of boy, the reference rating (average rating of multiple expert paediatric endocrinologists rating the BA hand X-ray) by the GP method was 8.3 years. On individually rating the segments based on images and description in the GP atlas, the GP BA assigned to radius–ulna, carpals and short bones was 8.9, 9.4 and 9.1, respectively. The simple arithmetic mean of these segments is 9.1 years. However, using the weightage as mentioned in Table 2b, $0.08*8.9 + 0.27*9.4 + 0.64*9.1 = 0.62$, a weighted mean BA of 8.5 years was obtained which is closer to the reference rating.

