Article:

Alshamrani, K., Hewitt, A. and Offiah, A.C. orcid.org/0000-0001-8991-5036 (2019) Applicability of two bone age assessment methods to children from Saudi Arabia. Clinical Radiology. ISSN 0009-9260

https://doi.org/10.1016/j.crad.2019.08.029

Applicability of Two Bone Age Assessment Methods to Children from Saudi Arabia

Khalaf Alshamrani [1], Amanda Hewitt [2], Amaka C. Offiah [1,2]

[1] Department of Oncology & Metabolism, University of Sheffield, UK

[2] Sheffield Children's NHS Foundation Trust, Western Bank, Sheffield, UK

Correspondence: Khalaf Alshamrani, Academic Unit of Child Health, Damer Street, Sheffield, S10 2TH. Kalshamrani1@sheffield.ac.uk

Declarations of interest: none

Funding: This work was supported by Najran University (grant number 787) as part of Khalaf Alshamrani PhD study.

Author Contributions:

1- Khalaf Alshamrani
- study concepts and design
- literature research
- data collection
- data analysis
- statistical analysis
- manuscript preparation

2- Amanda Hewitt
- data analysis

3- Amaka C. Offiah
- data analysis
- manuscript editing
- guarantor of integrity of the entire study

Abstract

Background:

The Greulich & Pyle (G&P) and Tanner & Whitehouse (TW) methods are frequently used to determine bone age. The question to be raised is, “Are these standards applicable to children of different ethnicity to those on which they are based?”

Methods:

Bone age was assessed using the G&P and TW3 methods, firstly by independent manual rating of 2 observers, followed by a single observer using the BoneXpert software programme. In total, 420 hand trauma radiographs for [redacted] (220 males, 329 left, age range 1 to 18 years) performed in the period January 2012– September 2016 were assessed. Paired sample *t* test was used to compare the difference between mean bone age (BA) and mean chronological age (CA) and to compare the difference between manual and BoneXpert ratings. Statistical analysis was undertaken using SPSS v.25.

Findings:

We found a statistically significant difference between BA and CA in males when using G&P (mean difference -0.36 ± 1 years, p <0.01) and TW3 (mean difference -0.22 ±0.9 years, p=0.03) methods but not in females for either G&P (mean difference 0.13 ± 1.2 years) or TW3 (mean difference 0.08 ± 1.1 years). In males, BoneXpert results conformed to the manual ratings for TW3 but not for G&P, for which the mean difference between manual and BoneXpert ratings was -0.27 ±0.5 years (p<0.01).

Interpretation:

Our results indicate that manual and BoneXpert-derived G&P and TW3 bone age assessment can be applied with no modification to [redacted] females. However, only TW3 BoneXpert-derived BA can be applied without caution to [redacted] males.

Keywords: Greulich & Pyle Atlas, Tanner & Whitehouse method, bone age, ethnicity

Introduction

The determination of bone age is a routine diagnostic procedure usually required to identify growth disorders in children and plan for therapeutic procedures. It is important to assess bone age using a reliable method, one of which is the assessment of bone age from a left hand radiograph (1). Two approaches are widely used to assess bone age from a left-hand radiograph, namely the Greulich and Pyle (G&P) and the Tanner and Whitehouse (TW3) methods (2,3). The data that were used to establish the G&P atlas and the TW3 standard came from healthy children of north American and western European origin and was collected around 4 and 9 decades ago. In addition to potential secular change, ethnicity and socioeconomic status are factors that have an impact on children's bone age. Therefore, one question to be raised when using these standards is, "Are they relevant to a current population of different ethnicity and/or socioeconomic status to the children used to develop the standards?"

The G&P and TW3 methods were initially (and still most commonly) based on a subjective approach that is likely to suffer from variations in rating between assessors due to different levels of competence, with their reliability partially dependent on the skill of the assessor. To eliminate observer variation and reduce rating time, BoneXpert software was introduced in 2009. This is an automated software programme that calculates bone age according to the G&P and TW3 methods (4). However, although the software has been validated in Caucasian (5,6), African-American (6), Hispanic and Asian-Chinese (6,7), studies on other populations are limited. Therefore, this study will assess the applicability of the G&P and TW3 to children from [redacted] using both subjective (manual) rating and BoneXpert software.

Methods

Hand radiographs performed on children aged between 1 and 18 years old presenting to the Emergency Department of xxx between January 1st, 2012 and September 30th 2016 following trauma were retrospectively identified from the Picture Archiving and Communication System.

All radiographs were acquired via a computerised radiography system, and were in DICOM format. Studies with a specific request for BA estimation were excluded. Emergency Department notes were scrutinised and any child with an underlying disorder was excluded. Demographic data including sex and age at the time of the radiograph were recorded. Only radiographs of [redacted] were included and were confirmed using the national ID included within the health ID (8,9). All the radiographs were assessed first manually and then using the BoneXpert software. Ethical approval was obtained from [redacted]

Manual rating

Observers 1 and 2 independently assessed bone age from all radiographs without knowledge of chronological age using the G&P method. When the patient's bone age was assessed to lie between two adjacent standards, the intermediate value was assigned as the bone age. Observers 1 and 3 assessed the radiographs using the RUS (radius, ulna and short bone) method. The time interval between Observer 1's G&P and TW3 reads was at least three months. To determine intra-observer reliability, a random sample of 43 radiographs (22 males) were assessed by each observer 1 month following their initial reads.

The maximum potential TW3 bone age score is 1000, which corresponds to an adult standard, while the minimum potential score is 42, which corresponds to 2 years of age. In this study, radiographs that were assigned as adult or did not achieve the minimum score were excluded. Additionally, for both G&P and TW3 reads, radiographs were excluded when bone age could not be assigned as a result of poor positioning or artefact.

BoneXpert rating

All radiographs were exported into an external hard drive and a standalone version of BoneXpert (Visiana, Holte, Denmark, v2.5.1.1) was used to determine bone age (G&P and TW3). Age was limited to 15 years in females and 17 years in males because the software does not provide a precise G&P reading above these ages. The default ethnicity for analysing the radiographs was Caucasian, as the software does not include ethnicity-specific standard deviation scores (SDS).

Statistical analysis

Statistical analysis was undertaken using SPSS version 24 for PC (IBM, Armonk, New York). Inter-observer reliability was assessed using interclass correlation coefficient. The mean variation for BA and CA was determined for each child by subtracting BA from CA (BA-CA). Therefore, a positive value indicates advanced BA, whereas a negative value indicates delayed BA, compared to CA. Paired sample *t* test was used to test the significance of the differences between BA and CA for each method and to test the significance of the differences between manual and BoneXpert ratings for each method. This analysis was undertaken separately for both males and females.

Results

G&P atlas:

Concerning manual G&P ratings, 420 radiographs (220 males) were assessed by each observer. Tables 1 to 3 summarise the number of radiographs assessed by age and sex. The inter-class correlation coefficient (ICC) showed a high correlation between the two observers with coefficients of 0.984 for females and 0.991 for males. No significant intra-observer difference was identified ($p=0.772$). In this regards, readings from the first observer were used when comparing the BA to CA using the G&P atlas.

BA was lower than CA in 48% of females and 61% of males, while being equal in 1% of males. The mean difference between BA and CA ranged from 37 months underestimation to 36 months overestimation in both females and males. On average, G&P underestimated males by 0.31 years/4 months ($p < 0.01$) and overestimated females by 0.1 years/1 month ($p = 0.089$) (Table 1).

With the cohort divided into yearly intervals, G&P overestimated females aged from 1 to 5 years by between 0.5 and 6 months, apart from at 3 years of age. After 5 years of age, G&P consistently underestimated females by between 3 and 8 months until 9 years of age, with underestimation being statistically significant ($p <0.05$) at 6 years of age (Table 2). The G&P

atlas then overestimated females by between 1 and 13 months with overestimation being statistically significant ($p < 0.05$) at 12 and 13 years of age.

G&P underestimated males from 1 to 13 years by between 2 and 13 months, apart from at 4 years. This underestimation was statistically significant ($p<0.05$) at the ages of 7,8,9 and 10 years (Table 3). After the age of 13 years, G&P overestimated males, but this did not reach statistical significance.

BoneXpert, was not able to analyse 208 (50%) of the radiographs, thus only 212 radiographs (114 males) were included in the final analysis. BoneXpert overestimated G&P BA in females by 2 months ($p = 0.06$) and underestimated G&P BA in males by 2.5 months ($p < 0.05$). Mean difference between BA and CA ranged from 32 months underestimation to 30 months overestimation in both females and males.

With the cohort divided into yearly intervals, G&P BA derived by BoneXpert followed a similar pattern of under/overestimation as the manual rating in females, however, no statistical significance was found, apart from at the age of 13 where the software significantly overestimated females ($p<0.05$) (Table2). In males, in contrast to manual rating BoneXpert overestimated males aged between 2 and 6 years by between 1 and 4 months. BoneXpert underestimated G&P BA in males aged between 7 and 12 years, with underestimation being statistically significant ($p<0.01$) at ages 8 and 9 years (Table3).

The G&P manual rating was lower than BoneXpert derived G&P by an average of 0.27 years/3 months in males ($p < 0.01$) and 0.1 years/1 month ($p = 0.184$) in females. Bland Altman plots comparing manual and BoneXpert ratings in females and males using G&P are illustrated in Figures 1a and 1b

TW3 Method:

Concerning manual TW3 ratings, 67 radiographs were excluded from analysis for the following reasons; (a) 43 radiographs achieved the maximum score (26 females), (b) 14 radiographs did not reach the minimum score (6 females), (c) 11 radiographs were poorly positioned, such that bone age could not be determined. In total, 353 radiographs were included in the final analysis (Tables 1, 4 and 5). The intra-class correlation coefficient indicated a high correlation between the two observers (0.972 for females and 0.963 for males). As there is no significant intra-observer difference was identified (p=0.351), readings from the first observer was used when comparing BA to CA.

BA was lower than CA in 44% of females and 56% of males, while being equal in 1% of females. The mean difference between BA and CA ranged from 30 months underestimation to 28 months overestimation in both females and males. On average, TW3 underestimated males by 0.22 years/2.5 months ($p < 0.01$) and overestimated females by 0.1 years/1 month ($p = 0.413$) (Table 1).

With the cohort divided into yearly intervals, TW3 overestimated females aged from 1 to 13 years by between 0.5 and 7 months, apart from at 6,7 and 8 years, with overestimation being statistically significant ($p < 0.05$) at 11 and 12 years of age (Table 4). In contrast, TW3 underestimated males aged 5 to 11 years, with underestimation being statistically significant ($p < 0.05$) at 8 and 9 years. After the age of 11 years, TW3 overestimated males by between 1 to 6 months, with overestimation being statistically significant ($p < 0.05$) at 13 years.

Concerning BoneXpert, additional 5 radiographs (2 females) were excluded as the radiographs achieved the maximum score according to the BoneXpert-derived TW3 BA. BoneXpert overestimated TW3 BA in females by an average of 1 month, while underestimating males by 2 months (Table 1). Mean difference between BA and CA ranged from 28 months underestimation to 30 months overestimation in both males and females. Breaking the cohort into yearly intervals showed that similar to manual ratings, the software overestimated TW3 BA in females aged between 10 and 13 years, being statistically significant at age of 8 years (Table 4). In males, BoneXpert underestimated TW3 BA in males aged between 7 and 12 years, being statistically significant at the age of 9 years (Table 5). Mean BA using the manual TW3 method was lower than TW3 derived by BoneXpert by 1 month, with no significant difference between the two methods in both males and females.

BoneXpert and manually-derived TW3 are compared as Bland Altman plots in Figures 2a and 2b.

Discussion

Using a reliable method to determine bone age is crucial for clinical and legal purposes. Hence we sought to analyse the applicability of G&P and TW3 bone age standards to [redacted] children, who are of different ethnicity to the population used to generate these two standards. We also sought to compare manual rating to BoneXpert, which software programme has not previously been used in the [redacted] ethnic group.

In relation to G&P, underestimation by an average of 4 months and 2.5 months was observed in males using manual rating and BoneXpert, respectively. In females, both manual rating and BoneXpert, overestimated their age by 1 month and 2 months respectively. These findings are in line with the study by Alhadlaq et al. who found that the bone age of children from [redacted] aged 9 to 15 tended to be lower than chronological age by 8 months (10). In other Asian populations, a large number of studies have shown that the G&P atlas is not applicable due to the large differences between bone age and chronological age (7,10,19–27,11–18). Generally, the G&P atlas seems to underestimate Asian boys during early and mid-childhood and overestimate boys during adolescence. The findings of these studies are summarised beside our findings in Table 6.

Similar to the G&P atlas, the TW3 method underestimated females and males in younger age groups, and overestimated females and males after the age of 9 and 12 years, respectively. Although, there was no significant different between BA and CA when using the TW3 method in females, the TW3 underestimated BA in males by an average of 2.5 months. These finding were also recently observed in the Thai population (27). Other studies on Asians have shown that young adults are reaching the end of maturity prior to the age observed through the TW3 method (7,16). The mean difference between BA and CA observed in similar research that focused on Asian populations is summarised in Table 6.

One of the main factors that has an impact on skeletal maturation rate is ethnicity (18,23,28–30). This impact has been shown by studies that sought to test the applicability of the methods on two different ethnic groups residing in the same region (13,17,31). One of these studies showed that the G&P atlas was only applicable to Asian children between 7 and 13.5 years (13). Additionally, it seems that Asian children mature sooner than Caucasian children, especially between 10 and 13 years of age in girls and between 11 and 15 years of age in boys (17).

Socioeconomic status is another factor that may affect skeletal maturation. Bone age is usually delayed in children of low and advanced in those of high socioeconomic status (32). Some authors suggest that the inapplicability of the bone age standards is more likely to be due to differences in socioeconomic status than ethnicity. For example, Asians-Japanese children living in Japan were skeletally delayed between the age of 5 and 18 years in comparison to the Caucasian children who lived in Cleveland (US) at all age groups (33). However, Greulich argued that this was not due to ethnicity, as Japanese children living in California were skeletally delayed only between 5 to 7 years, which was attributed to less favourable environmental conditions, which can be interpreted as low socioeconomic status (33).

This delay was attributed to unfavourable environments of which socioeconomic status is part. Although BoneXpert agreed with the manual rating in the overall over/underestimation pattern, there was a statistically significant difference between the two methods in males but not in females. This may be due to the method by which BoneXpert calculates G&P bone age; the software does not include the carpal bones in its assessment. In our study, male radiographs in the younger age groups appeared to show less maturity in the carpal compared to the other bones of the hand (Figure 3). This has also been highlighted in other populations, in which carpal maturation pattern has influenced bone age assessment results (10,14,34). However, the value of the carpal bones in bone age assessment has been questioned due to the poor correlation between carpal bone development and chronological age. Johnston and Jahina concluded that the accuracy of bone age assessment increased when the carpal bones were illuminated (35). If this is the case, then the BoneXpert-dervied BA results in the current study are more reliable than the manual results for which all hand and carpal bones were assessed. BoneXpert could not assess approximately half of all radiographs, mainly because the images were post-processed using a sharpening algorithm, which gave them excessively sharp borders, rendering them unreadable by the software.

The relatively small number of radiographs included in each age group for Bonexpert analysis compared to manual rating, may have contributed to the differences between BoneXpert and manually-derived BA.

The limitations of this study include 1) socioeconomic status was not reported due to insufficient information; 2) hospital notes were not reviewed to ascertain full health in the children (although radiology and ED notes were scrutinised) 3) both left and right hand radiographs were used; traditionally BA has been assessed from left hand radiographs, however, it has been shown that there is no significant difference in G&P or TW3 BA between left and right hands ((36)) and so this should not have affected our results and 4) only certain age groups were included in BoneXpert analysis, namely between 2 and 15 years old in females and between 2.5 and 17 years in males. This was unavoidable because the software tool is unable to read images from younger age groups due to limited ossification or non-ossification of epiphyses, while its dependability is questionable when used in older age groups. Having said that, due to recruitment method (children attending an Emergency Department with hand trauma) and the high rejection rate of the software within as a result of insufficient image quality, some of the age groups included in the BoneXpert analysis had fewer than 5 radiographs (Tables 2 to 5), and the results of this study in these age groups should be treated with caution.

Conclusion

Our results indicate that the G&P and TW3 manual and BoneXpert methods can be applied to [redacted] females. However, significant differences between BA and CA were apparent in [redacted] males for manual and BoneXpert-derived G&P and TW3 BA but not for BoneXpert-derived TW3 BA.

Tables legend

Table 1: Mean difference (±SD) in years, between BA and CA in females and males

Table 2: Mean difference (±SD) in years, between GP BA (manual and BoneXpert) and CA in females

Table 3: Mean difference (±SD) in years, between GP BA (manual and BoneXpert) and CA in males

Table 4: Mean difference (±SD) in years, between TW3 BA (manual and BoneXpert) and CA in females

Table 5: Mean difference (±SD) in years, between TW3 BA (manual and BoneXpert) and CA in males

Table 6: Mean difference between BA and CA in studies that assessed the reliability of the G&P and TW3 methods in Asian Children

Figure legends:

Figure 1: Bland Altman plot comparing manual and BoneXpert ratings using the G&P method.

a) Females b) males

Figure 2: Bland Altman plot comparing manual and BoneXpert ratings using the TW3 method.

a) Females b) males

Figure 3: DP L hand radiograph of a male, chronological age 5 years and 7 months, showing less maturity in the carpal area compared to the other bones of the hand

References

1. Ritz-Timme S, Cattaneo C, Collins MJ, Waite ER, Schütz HW, Kaatsch HJ, et al. Age estimation: The state of the art in relation to the specific demands of forensic practise [Internet]. Vol. 113, International Journal of Legal Medicine. 2000 [cited 2017 Jan 30]. p. 129–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10876982

2. Greulich W, Pyle I. Radiographic Atlas of Skeletal Development of the Hand and Wrist. Stanford University Press. London: Stanford University Press; 1959.

3. Tanner J, Healy M, Goldstein H, Cameron N. Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method). London: WB Saunders; 2001. 110 p.

4. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert Method for Automated Determination of Skeletal Maturity. IEEE Trans Med Imaging [Internet]. 2009 Jan [cited 2018 Jun 18];28(1):52–66. Available from: http://ieeexplore.ieee.org/document/4530646/

5. Van Rijn RR, Lequin MH, Thodberg HH. Automatic determination of Greulich and Pyle bone age in healthy Dutch children. Pediatr Radiol. 2009;39(6):591–7.

6. Thodberg HH, Sävendahl L. Validation and reference values of automated bone age determination for four ethnicities. Acad Radiol. 2010;17(11):1425–32.

7. 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(1):201–5.

8. Blinded

9. Blinded

10. Blinded

11. So LL, Yen PK. Secular trend in skeletal maturation in southern Chinese girls in Hong Kong. Z Morphol Anthropol [Internet]. 1990 [cited 2017 Jan 30];78(2):145–53. Available from: http://www.jstor.org/stable/25757279

12. Y So LL, Lisa So BL. Correlation of skeletal maturation with stature and body weight of Southern Chinese girls Correlation of skeletal maturation with stature and body weight of Southern Chinese girls in Hong Kong With 3 tables in the text. E Schweizerbart'sche Verlagsbuchhandlung [Internet]. 1991;78(3):307–12. Available from: http://www.jstor.org/stable/25757319

13. Ontell FK, Ivanovic M, Ablin DS, Barlow TW. Bone age in children of diverse ethnicity. Am J Roentgenol. 1996;167(6):1395–8.

14. Krailassiri S, Anuwongnukroh N, Dechkunakorn S. Relationships between Dental Calcification Stages and Skeletal Maturity Indicators in Thai Individuals. Angle Orthod. 2002;72(2):155–66.

15. Chiang K, Chou A, Yen P, Ling C. The reliability of using Greulich-Pyle method to determine children's bone age in Taiwan. Tzu Chi Med J [Internet]. 2005;(6):15–8. Available from: http://www.tzuchi.com.tw/tzuchi/tcmj/94-6/1-4.pdf

16. Griffith J, Cheng J, Wong E. Are Western skeletal age standards applicable to Hong Kong Chinese? – a comparison of Greulich and Pyle and TW3 methods. Hong Kong Med J. 2007;13(3):S28-32.

17. Zhang A, Sayre JW, Vachon L, Liu BJ, Huang HK. Racial Differences in Growth Patterns of Children Assessed on the Basis of Bone Age. Radiology [Internet]. 2009 [cited 2017 Jan 30];250(1):228–35. Available from: http://pubs.rsna.org/doi/abs/10.1148/radiol.2493080468

18. Zafar AM, Nadeem N, Husen Y, Ahmad MN. An appraisal of greulich-pyle atlas for skeletal age assessment in Pakistan. J Pak Med Assoc [Internet]. 2010 [cited 2017 Jan 30];60(7):552–5. Available from: http://www.jpma.org.pk/PdfDownload/2168.pdf

19. Moradi M, Sirous M, Morovatti P. The reliability of skeletal age determination in an Iranian sample using Greulich and Pyle method. Forensic Sci Int. 2012;223(1–3).

20. Soudack M, Ben-Shlush A, Jacobson J, Raviv-Zilka L, Eshed I, Hamiel O. Bone age in the 21st century: UIs Greulich and Pyle's atlas accurate for Israeli children? Vol. 42, Pediatric Radiology. 2012. p. 343–8.

21. Patil ST, Parchand MP, Meshram MM, Kamdi NY. Applicability of Greulich and Pyle skeletal age standards to Indian children. Forensic Sci Int. 2012;216(1–3):200.e1-200.e4.

22. Awais M, Nadeem N, Husen Y, Abdul-Rehman Y, Beg M, Khattak S. Comparison between Greulich-Pyle and Girdany-Golden Methods for Estimating Skeletal Age of Children in Pakistan. J Coll Physicians Surg Pakistan [Internet]. 2014 [cited 2017 May 18];24(12):889–93. Available from: https://www.jcpsp.pk/archive/2014/Dec2014/03.pdf

23. 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.

24. Mughal AM, Hassan N, Ahmed A. The applicability of the Greulich & Pyle Atlas for bone age assessment in primary school-going children of Karachi, Pakistan. Pakistan J Med Sci. 2014;30(2):409–12.

25. Rai V, Saha S, Yadav G, Tripathi AM, Grover K. Dental and skeletal maturity- A biological indicator of chronologic age. J Clin Diagnostic Res. 2014;8(9):ZC60-ZC64.

26. Mohammed RB, Rao DS, Goud AS, Sailaja S, Thetay AAR, Gopalakrishnan M. Is

Greulich and Pyle standards of skeletal maturation applicable for age estimation in South Indian Andhra children? J Pharm Bioallied Sci [Internet]. 2015;7(3):218–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26229357%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4517325

27. Benjavongkulchai S, Pittayapat P. Age estimation methods using hand and wrist radiographs in a group of contemporary Thais. Forensic Sci Int [Internet]. Elsevier Ireland Ltd; 2018;287:218.e1-218.e8. Available from: https://doi.org/10.1016/j.forsciint.2018.03.045

28. Mora S, Boechat MI, Pietka E, Huang HK, Gilsanz V. Skeletal Age Determinations in Children of European and African Descent: Applicability of the Greulich and Pyle Standards. Pediatr Res [Internet]. 2001 [cited 2017 Jan 30];50(5):624–8. Available from: http://www.nature.com/doifinder/10.1203/00006450-200111000-00015

29. 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. Vol. 154, Forensic Science International. 2005. p. 3–12.

30. Cole TJ, Rousham EK, Hawley NL, Cameron N, Norris SA, Pettifor JM. Ethnic and sex differences in skeletal maturation among the Birth to Twenty cohort in South Africa. Arch Dis Child [Internet]. BMJ Publishing Group; 2015 Feb [cited 2018 Mar 13];100(2):138–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25409981

31. Loder RT. Applicability of the Greulich and Pyle Skeletal Age Standards to Black and White Children of Today. Arch Pediatr Adolesc Med [Internet]. 1993 [cited 2017 Jan 30];147(12):1329. Available from: http://archpedi.jamanetwork.com/article.aspx?doi=10.1001/archpedi.1993.02160360071022

32. Schmeling A, Schulz R, Danner B, Rösing FW. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Legal Med [Internet]. Springer-Verlag; 2006 Mar 13 [cited 2018 Mar 13];120(2):121–6. Available from: http://link.springer.com/10.1007/s00414-005-0007-4

33. Greulich-Pyle. A comparison of the physical growth and development of American born and native Japanese children. Am J Phys Antrhopol. 1957;15(4):489–515.

34. Acheson R, Vicinus J, Fowler G. Studies in the reliabilityof assessing skeltal maturity from x-ray: Part III . Greulich-Pyle atlas and Tanner-Whitehouse method contrasted. Hum Biol. 1966;38(3):204–18.

35. Johnston FE, Jahina SB. The Contribution of the Carpal Bones to the Assessment of Skeletal Age. Am J Phys Anthropol [Internet]. 1965;23:349–54. Available from: https://onlinelibrary.wiley.com/doi/epdf/10.1002/ajpa.1330230408

36. Martin DD, Neuhof J, Jenni OG, Ranke MB, Thodberg HH. Automatic determination of left- and right-hand bone age in the first Zurich longitudinal study. Horm Res Paediatr [Internet]. 2010 [cited 2018 Jan 22];74(1):50–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20395658

Figure 1a

Click here to access/download;Figure;Fig. 1a.tif

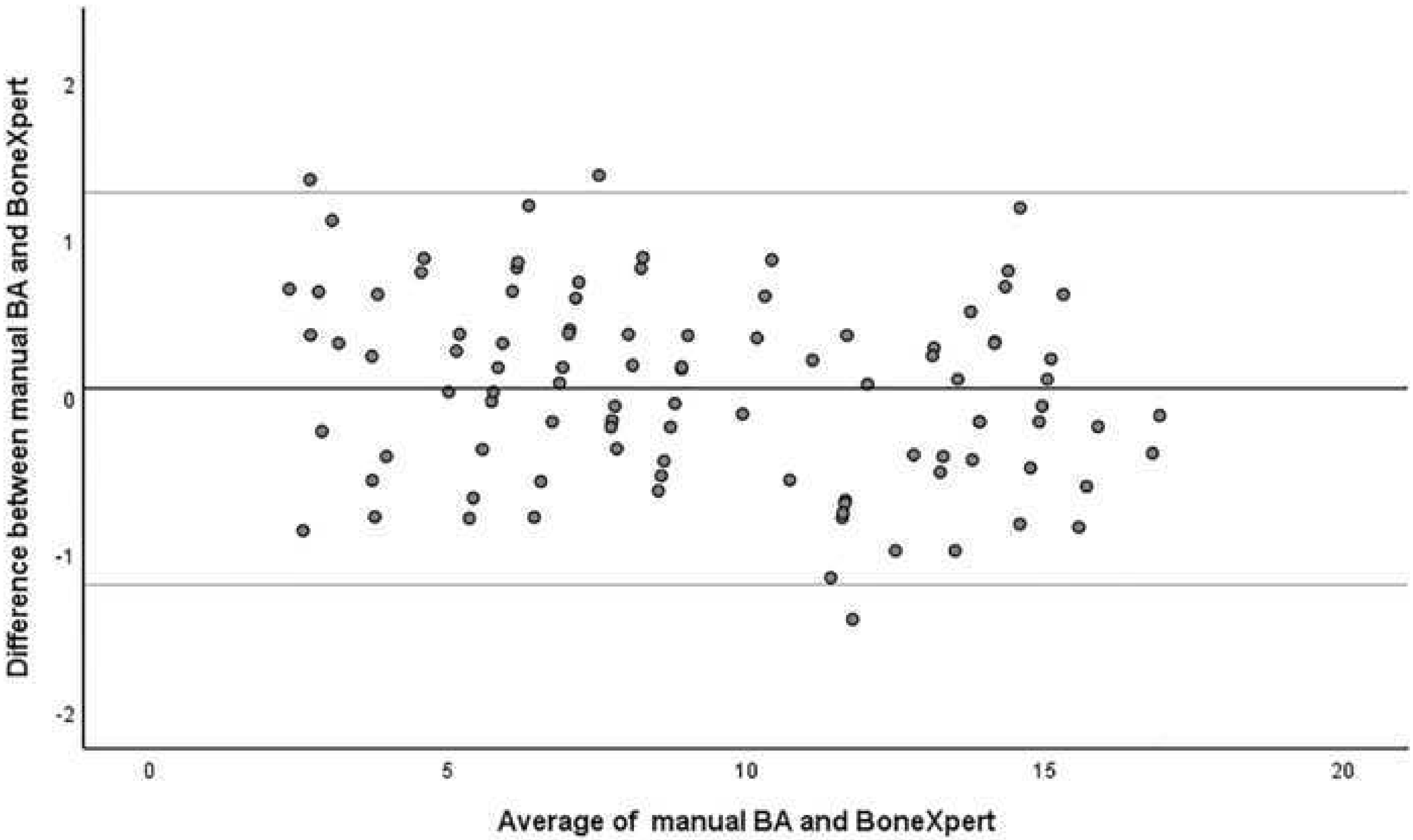

Figure 1b

Click here to access/download;Figure;Fig. 1b.tif

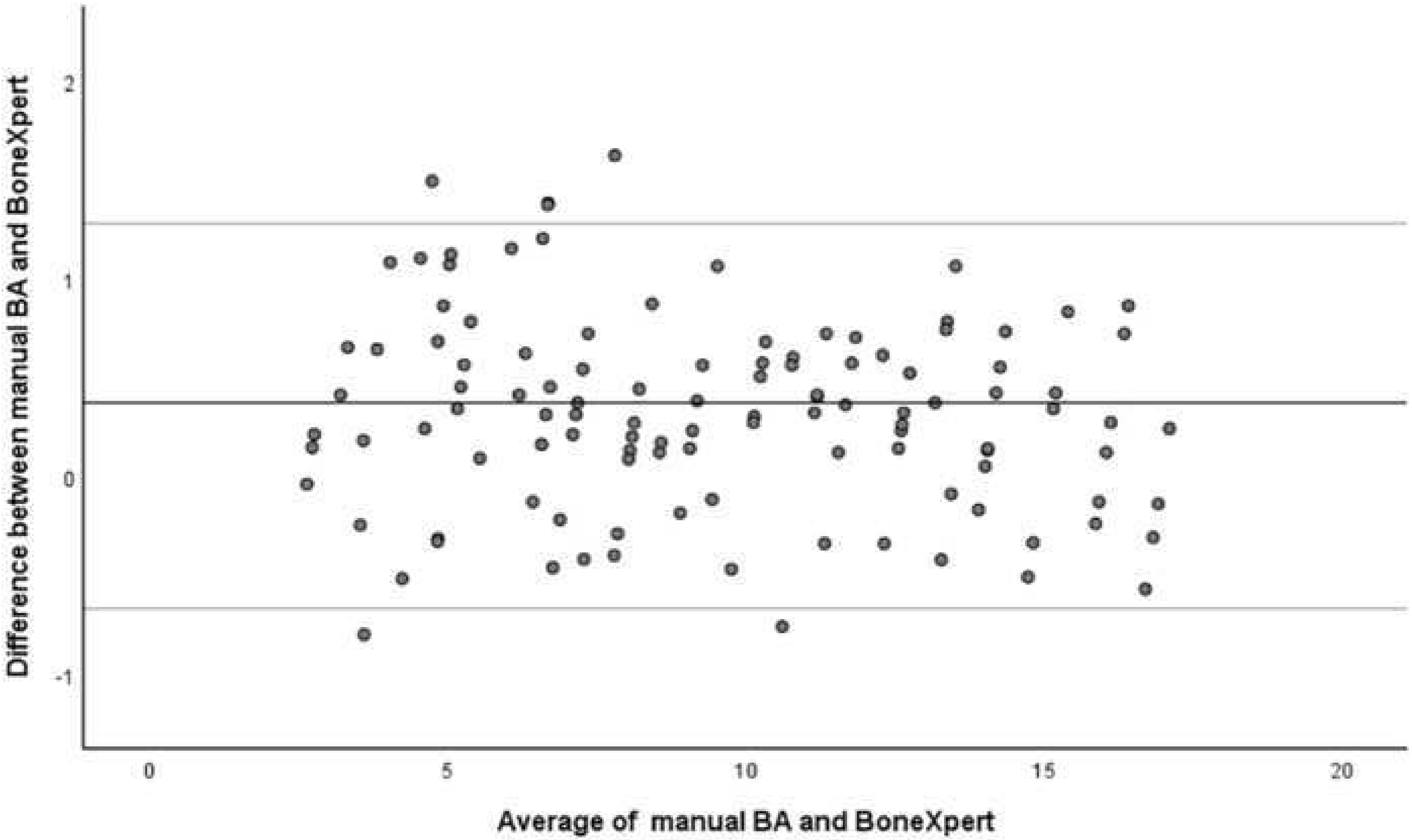

Figure 2a

Click here to access/download;Figure;Fig. 2a.tif

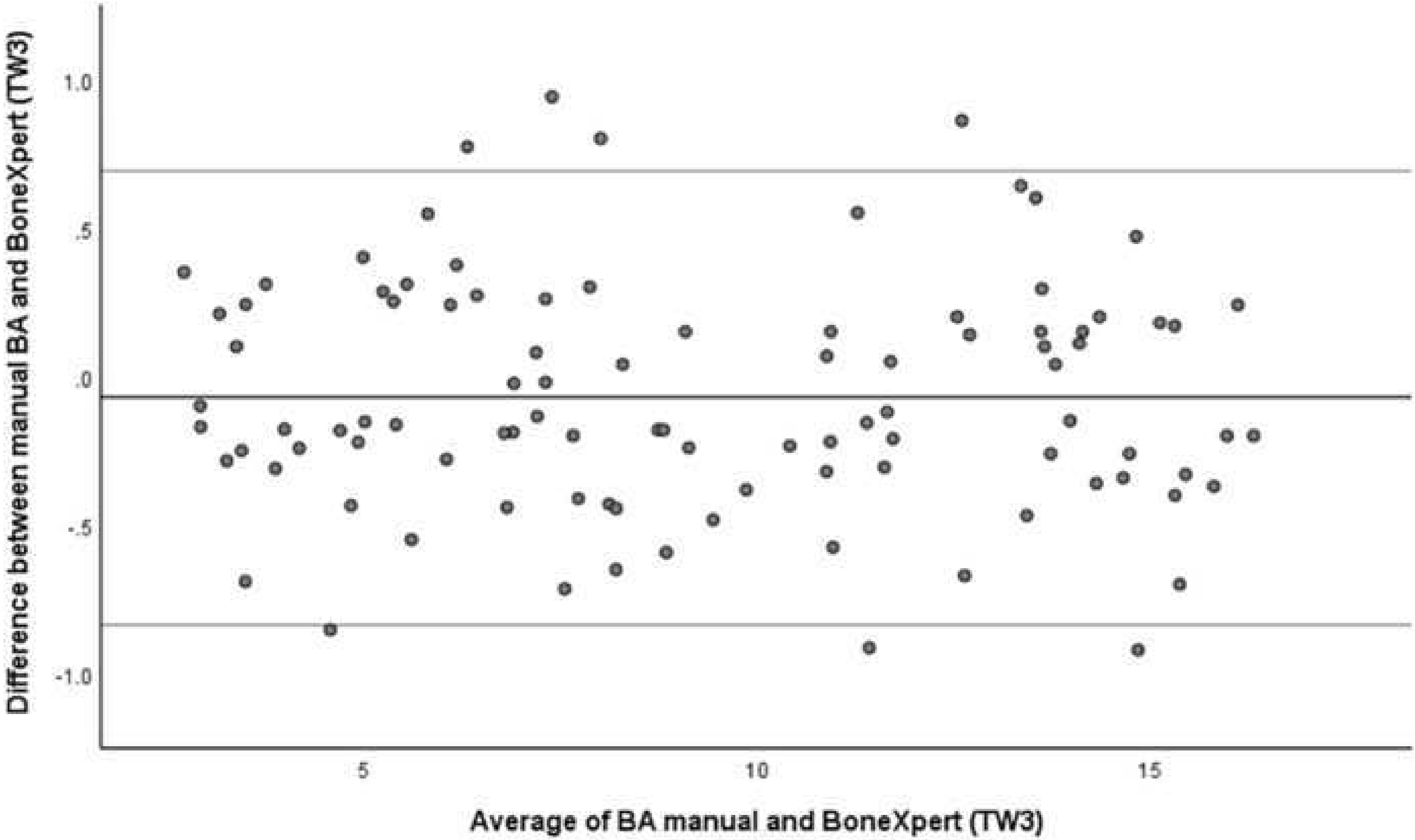

Figure 2b

Click here to access/download;Figure;Fig. 2b.tif

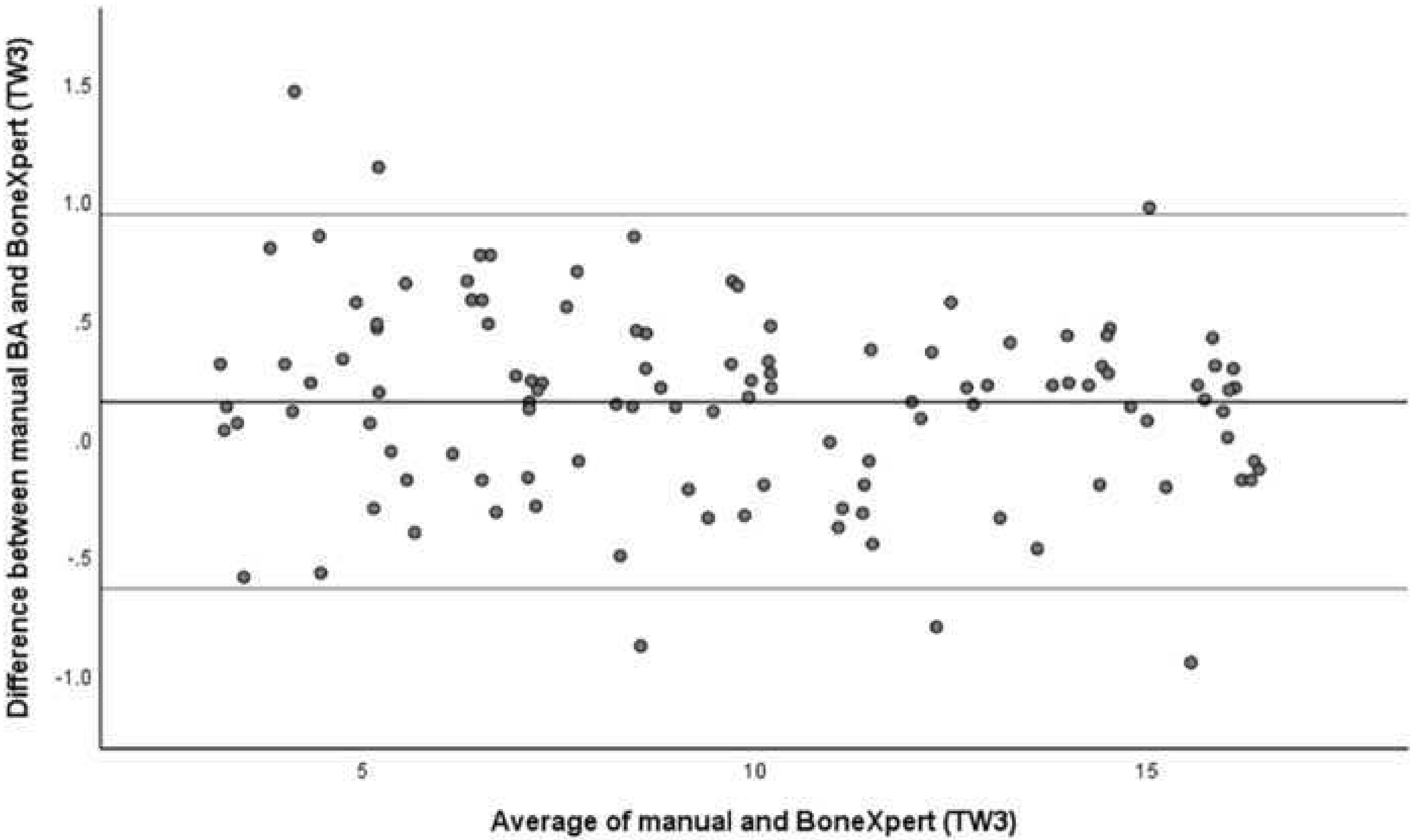

Figure 3 Click here to access/download;Figure;Fig. 3.tif

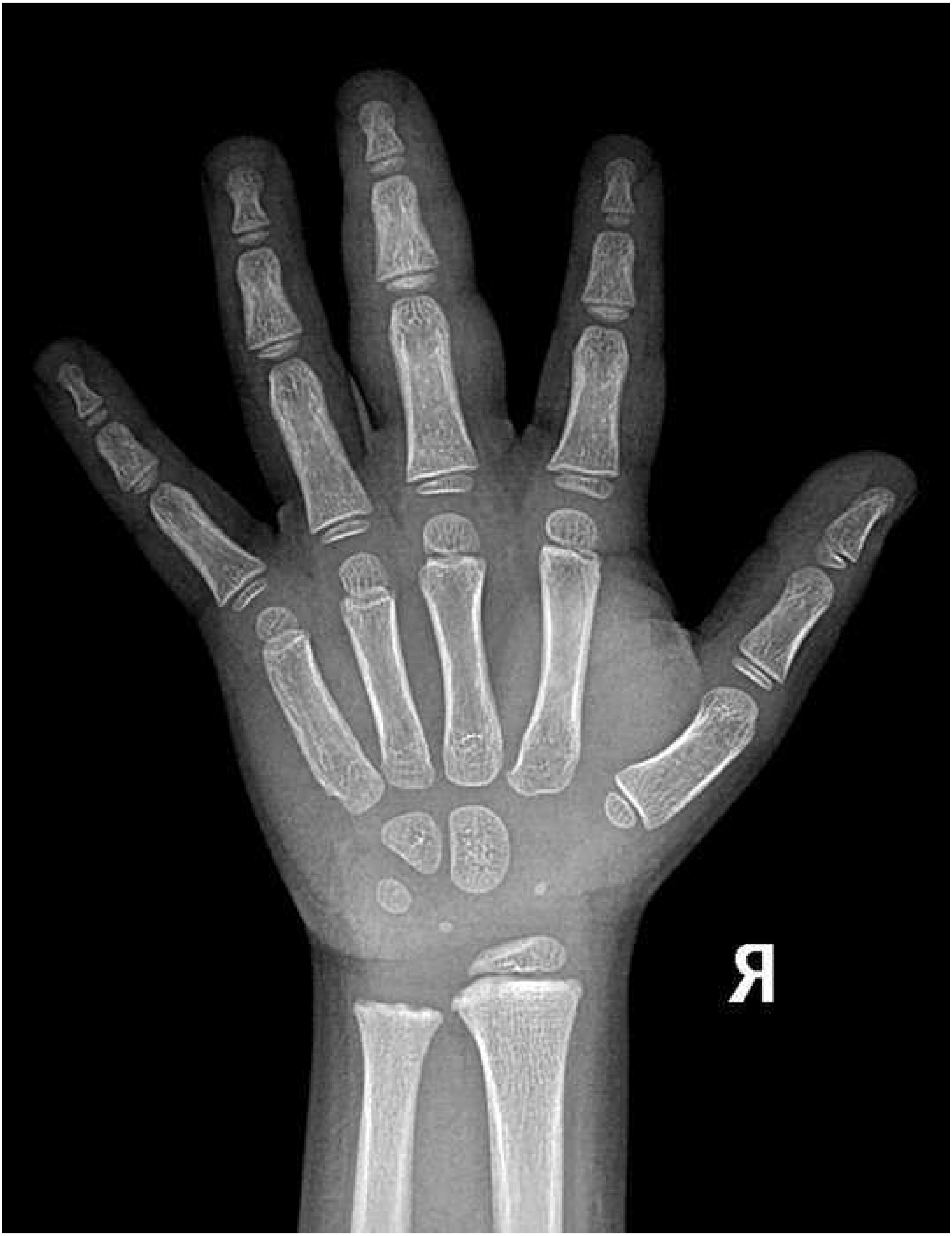

Table 1

Table 1: Mean difference (±SD) in years, between BA and CA in females and males

Manual rating	**Sex**	**No**	**Mean CA (±SD)**	**Mean BA (±SD) (Observer 1)**	**Mean difference BA-CA**	**p value**
G&P BA vs CA	Female	200	10.21 (± 4.4)	10.34 (±4.8)	0.13 (± 1.2)	0.089
	Male	220	10.48 (± 4.8)	10.12 (±5.2)	-0.36 (± 1.0)	<0.01
TW3 BA vs CA	Female	164	8.80 (±3.6)	8.88 (± 3.8)	0.08 (± 1.1)	0.413
	Male	189	9.59 (± 4.4)	9.37 (±4.7)	-0.22 (± 0.9)	0.03
BoneXpert rating	**Sex**	**No**	**Mean CA (±SD)**	**Mean BA (±SD) (Observer 1)**	**Mean difference BA-CA**	**p value**
G&P BA vs CA	Female	98	9.02 (± 3.7)	9.18 (± 4.0)	0.16 (± 1.0)	0.06
	Male	114	9.89 (± 3.9)	9.68 (± 4.0)	-0.21 (± 0.8)	0.03
TW3 BA vs CA	Female	96	8.45 (±3.38)	8.58 (±3.6)	0.13 (± .9)	0.22
	Male	111	9.85 (± 3.9)	9.73 (± 3.9)	-0.12 (± 0.9)	0.09

Table 2

Table 2: Mean difference (±SD) in years, between GP BA (manual and BoneXpert) and CA in females

	Manual Rating				BoneXpert Rating			
Age (years)	**No**	**Mean difference**	**(±SD)**	**p value**	**No**	**Mean difference**	**(±SD)**	**p value**
1	4	0.04	0.43	0.86		-	-	-
2	4	0.48	0.65	0.24	3	0.68	0.46	0.16
3	9	-0.41	0.87	0.20	5	0.12	0.44	0.25
4	11	0.03	0.75	0.89	6	0.38	0.62	0.43
5	12	0.21	0.71	0.11	8	0.42	0.79	0.20
6	13	-0.68	1.02	0.03	7	0.32	1.19	0.21
7	14	-0.25	1.10	0.47	6	-0.02	0.96	0.91
8	14	-0.36	0.95	0.18	9	-0.38	0.82	0.08
9	17	-0.41	1.40	0.23	11	-0.29	1.47	0.52
10	13	0.22	1.63	0.65	5	0.47	0.92	0.38
11	15	0.71	1.48	0.08	10	0.35	1.02	0.36
12	14	1.10	1.20	0.00	9	0.89	1.26	0.08
13	16	0.83	1.47	0.04	6	0.98	1.16	0.03
14	11	0.46	1.37	0.29	6	0.41	1.24	0.37
15	12	0.56	1.50	0.22	7	0.02	1.01	0.96
16	8	0.18	1.32	0.72	-	-	-	-
17	8	0.01	0.73	0.97	-	-	-	-
18	5	-0.12	0.34	0.13	-	-	-	-

Table 3

Table 3: Mean difference (±SD) in years, between G&P BA (manual and BoneXpert) and CA in males

	Manual Rating				BoneXpert			
Age (years)	**No**	**Mean**	**(±SD)**	**p value**	**No**	**Mean**	**(±SD)**	**p value**
1	5	-0.30	0.66	0.37		-	-	-
2	7	-0.20	0.63	0.40	3	0.29*	0.59	0.61
3	14	-0.26	0.85	0.28	7	0.04	0.62	0.83
4	11	0.33	0.53	0.07	6	0.41	0.58	0.14
5	13	-0.35	0.59	0.06	8	0.25	0.58	0.24
6	10	-0.21	0.65	0.39	6	0.11	0.63	0.69
7	15	-0.72	1.00	0.01	10	-0.31	0.88	0.18
8	12	-1.12	1.20	0.01	8	-0.97	1.06	0.01
9	14	-1.03	1.09	<0.00	9	-0.97	1.13	<0.01
10	12	-0.84	1.16	0.02	6	-0.72	1.07	0.09
11	15	-0.43	0.92	0.08	7	-0.17	1.03	0.48
12	14	-0.57	1.05	0.11	8	-0.36	0.91	0.30
13	13	-0.38	0.98	0.13	8	0.07	1.11	0.72
14	12	0.33	1.28	0.44	6	0.26	1.05	0.48
15	16	0.51	1.08	0.11	12	0.17	1.16	0.53
16	15	0.56	1.13	0.10	7	0.40	0.71	0.04
17	13	0.22	0.85	0.35	3	-0.24	0.64	0.34
18	9	0.07	0.77	0.78	-	-	-	-

Table 4

Table 4: Mean difference (±SD) in years, between TW3 BA (manual and BoneXpert) and CA in females

	Manual Rating				BoneXpert			
Age (years)	**No**	**Mean**	**(±SD)**	**p value**	**No**	**Mean**	**(±SD)**	**p value**
2	4	0.66*	0.32	0.03	2	0.21*	0.21	0.04
3	9	0.28	0.48	0.12	5	0.19	0.34	0.20
4	11	0.35	0.66	0.11	6	0.30	0.78	0.44
5	12	0.08	0.51	0.59	8	-0.19	0.64	0.53
6	13	-0.35	0.73	0.08	7	-0.12	0.88	0.70
7	12	-0.21	0.75	0.37	6	-0.15	0.98	0.73
8	14	-0.26	0.90	0.31	9	-0.63	0.76	0.04
9	15	0.14	1.11	0.60	11	-0.27	1.16	0.45
10	13	0.22	1.27	0.56	5	0.82*	1.02	0.06
11	15	0.59	0.87	0.02	10	0.53	1.18	0.24
12	14	0.68	0.97	0.00	9	0.81	0.96	0.05
13	14	0.16	1.16	0.09	6	0.80	0.91	0.03
14	11	-0.07	0.38	0.08	6	0.28	0.71	0.12
15	7	-0.53*	0.34	0.02	6	-0.1*	0.35	0.15

*less than 5 radiographs

Table 5

Table 5: Mean difference (±SD) in years, between TW3 BA (manual and BoneXpert) and CA in males

	Manual Rating				BoneXpert			
Age (years)	**No**	**Mean**	**(±SD)**	**p value**	**No**	**Mean**	**(±SD)**	**p value**
2	9	0.44	0.68	0.14	3	0.82*	0.31	0.17
3	11	0.05	0.47	0.72	7	0.48	0.48	0.04
4	12	0.02	0.56	0.89	6	0.62	0.80	0.12
5	10	-0.11	0.50	0.43	8	0.16	0.46	0.32
6	13	-0.33	0.46	0.09	6	0.04	0.50	0.87
7	12	-0.23	0.72	0.22	10	-0.26	0.63	0.23
8	14	-0.84	1.00	0.01	8	-0.44	0.88	0.20
9	12	-0.58	0.92	0.03	9	-0.68	0.78	0.03
10	15	-0.43	0.96	0.13	6	-0.59	0.73	0.08
11	14	-0.17	1.13	0.58	7	-0.21	0.90	0.46
12	13	0.06	1.04	0.84	8	-0.27	1.36	0.59
13	12	0.58	1.09	0.05	8	0.47	1.30	0.25
14	16	0.46	1.11	0.23	6	0.73	1.08	0.16
15	14	0.22	0.68	0.22	12	0.12	0.65	0.59
16	9	-0.16	0.36	0.03	7	-0.21*	0.19	0.07
17	3	-0.85	0.25	0.00	-	-	-	-

* less than 5 radiographs

Table 6

Table 6: Mean difference between BA and CA in studies that assessed the reliability of the G&P atlas and TW3 method in Asian children

Study	Origin/ ethnicity	Age (years)	N	Mean BA-CA (years)	p value
So & Yen 1990	Chinese	11.9-12.3	F=117	F= 0.6	< 0.01
So & Yen 1991	Chinese	11.9-12.3	F=117	F= 0.6	NR
Ontell et al, 1996	Asian	1-18	M=63 F=30	M= -0.03 F= 0.27	M= < 0.05 (after age of 3 years) F= > 0.05
Krailassiri et al, 2002	Thai	7-19	M=139 F=222	M= -0.8 F= 0.8	NR
Chiang et al, 2005	Taiwan	7-19	M=230 F=140	M= 0.82 F= -0.3	M= < 0.05 (at age of 3,4,6,7,8,10,12-17 years) F= <0.05 (at age of 2, 13 -15 years)
Al-Hadlaq et al, 2007	Saudi Arabian	7-15	M=115	M= -0.71	M= <0.05
Griffith et al, 2007	Chinese	0-18	M=650 F=366	M= 0.25 F= 0.15	M= < 0.05 (at age of 3-7,10 ,13,14,17 years) F= < 0.05 (at age of 3,4,9,10,12,13 years)
Zhang et al, 2009	Asian	0-18	M=165 F=166	M= 0.41 F= 0.24	M= < 0.05 F= < 0.05
Zafar et al, 2010	Pakistan	0-18	M=535 F=354	M= 0.1 F=- 0.19	M= < 0.05 (at all ages except after age of 13 years) F= < 0.05
Moradi et al, 2012	Iran	6-18	M=303 F=122	M= 0.37 F=- 0.04	M= 0.63 F= 0.59
Soudack et al, 2012	Iseral	0-18	M=375 F=304	M= 0.16 F=-0.04	M= <0.05 F= 0.188
Patil et al, 2012	India	1-19	M=194 F=181	M= 0.69 F= 0.64	M= < 0.05 (at age of 4,5,9,10,13,15 years) F= <0.5 (at age of 2,5,6,15 years)
Awais et al, 2014	Pakistani	0-18	M=136 F=147	M= -1.3 F= 0.06	M=<0.001 F= 0.695
Mansourvar et al, 2014	Asian American	1-8	M=48	M= 0.87	M= <0.05

Mughal et al, 2014	Pakistan	4.5-9.5	M=139 F=81	M= -1.3 F= 0.55	M=<0.001 F=<0.001
Rai et al, 2014	India	5-15	M=75 F=75	M= -0.07 F= -0.33	F=<0.01 M=<0.01
Kim et al, 2015	Korean	7-12	M=135 F=77	M= -0.48 F= -0.02	NR
Mohammed et al, 2015	South India	9-20	M=330 F=330	M= -0.23 F= 0.02	M= < 0.05 (at age of 11,13,16-19 years) F= <0.5 (at age of 10,11,15-19 years)
Benjvongkulchia et al 2018	Thai	8-20	M=172 F=193	M= 0.42 F= 0.90	M=<0.001 F=<0.001
		TW3 method			
Ashizawa et al, 2005	Beijing	6-16	M=631 F=642	M= 0.07 F= 0.11	NR
Griffith et al, 2007	Hong Kong	0-18	M=645 F=329	M= 0.22 F= 0.3	M= < 0.05 (at age of 2,4,7,10 ,13,14,17 years) F= < 0.05 (at age of 3,11-18 years)
Kim et al, 2015	Korean	7-12	M=135 F=77	M= 0.41 F= 0.12	NR
Benjvongkulchia et al 2018	Thai	8-20	M=172 F=193	M= -0.12 F= 0.40	M=<0.001 F=<0.001

A positive value for the mean difference between BA and CA indicates advanced while a negative value indicates delayed bone age compared to chronological age, M = males, F = females, NR = not reported