Comparison of classical methods for bone age determination with capitohamate planimetry on wrist x-ray

*Corresponding Author

Hasan YILDIZ

Department Of Anatomy, Faculty of Medicine, Selçuk University, Konya, Turkiye

Orcid : https://orcid.org/0000-0002-5265-807X

E-mail : hsnyldzz0@gmail.com

Phone : +90 545 474 43 81

Author

Hasan YILDIZ[1]*: https://orcid.org/0000-0002-5265-807X

Nadire ÜNVER DOĞAN[1]: https://orcid.org/0000-0001-5696-5547

Mehmet ÖZTÜRK[2]: https://orcid.org/0000-0001-5585-1476

Zeliha FAZLIOĞULLARI[1]: https://orcid.org/0000-0002-5103-090X

Muslu Kazım KÖREZ[3]: https://orcid.org/0000-0001-9524-6115

Ahmet Kağan KARABULUT[1]: https://orcid.org/0000-0002-9635-8829

[1] Department Of Anatomy, Faculty of Medicine, Selçuk University, Konya, Turkiye

[2] Department Of Radiology, Faculty of Medicine, Selçuk University, Konya, Turkiye

[3] Department Of Biostatistics, Faculty of Medicine, Selçuk University, Konya, Turkiye

ORIGINAL ARTICLE

ABSTRACT

Background/aim: The aim of the study is to compare the Greulich-Pyle(GP) and TannerWhitehouse2(TW2) methods, which are used in the left wrist radiography for bone age determination in the pediatric age group, with the Capitohamatum method(CH), due to the importance of determining the bone age in the pediatric period.

Materials and Methods: Direct radiographs of 210 female and 291 male individuals between the ages of 0-15 and without any pathology in the left wrist bones obtained. 501 os capitatum and os hamatum are as on AP direct graphy images were measured by GP, TW2 and CH planimetry methods. The estimated age of each measurement data was calculated by evaluating the relationship between chronological age and gender.

Results: In male individuals, it was determined that the estimates obtained by the GP method were on average 5.1 units lower than the actual ages, and the estimates obtained by the TW2 method were on average 1.7 units higher than the actual ages. In female individuals, age estimations obtained by both GP and TW2 methods were found to be 1.4 and 0.5 units lower, respectively, than the chronological ages. It was determined that the ages estimated by the CH method were almost the same as the chronological ages, and there was no deviation in the estimation method.

Conclusion: According to study findings, it was concluded that the CH method can be used reliably and with high accuracy for chronological age determination of children, and is an alternative estimation method to GP and TW2 atlases in the literature.

Keywords: Age determination, greliuch-pyle, tanner-whitehouse, skeleton, carpal bones

Abbreviations:

- **AIC:** Akaike Information Criteria
- **BIC:** Bayesian Information Criteria
- **CCC:** Concordance correlation coefficient
- **CH:**Capitohamatum method

- **GP:** Greulich-Pyle
- **ICC:** Intra-class correlation coefficient
- $\boldsymbol{R^2}$: Coefficient of determination
- **RMSE:** Root mean square error
- **TW2:** TannerWhitehouse2

1. Introduction

Determination of age by considering bone maturation is a commonly used procedure in our country. Bone age determination is important for various metabolic disorders, nutrition and endocrine disorders and criminal liability and legal capacity in judicial cases [1]. Medically accurate age determination can help diagnose, for example, endocrine disorders as well as being useful in following up the patients receiving hormone therapy. It is also helpful while deciding the right surgical intervention in orthopedics [2].

Bone age is one of biological indicators of maturity used in clinical practice and it is a very important parameter of a child's assessment [20]. Radiologic assessment of bones and its adaptation to present atlases keep its importance as the method in which the most commonly used and real-like values are obtained in age determination in clinical practice. The standards based on the method of detecting ossification points with epiphyseal and diaphyseal lines and detection of the periods of growth plate maturation in bones are used in this method. Hand and wrist are the most preferred site for radiographic assessment in bone age study and the most suitable area with the required conditions to significantly assess radiographies. Skeletal maturity is assessed using the hand and wrist X-rays. The assessment is based on the comparison of ossification and maturity of hand and wrist epiphyses with the radiographs of individuals by using atlases and methods formed according to the current standards [3]. Hand and wrist have recently been used in detection of skeletal maturity periods within the process of growth and is the most suitable site with the conditions required for an effective radiographic assessment. Radiologists analyzes the skeletal maturation by using hand and wrist X-rays. The most commonly preferred methods for bone age determination in hand-wrist radiographs are GP and TW methods [4,5]. Although GP is one of the most widely used bone age

estimation methods, its biggest disadvantages are; high inter- and inter-rater variability, not applicable to some populations, and last version available since 1959 [3,5,20]. The biggest advantage of the TW method is that it assigns a digital score to each stage of maturation of the hand and wrist bones and allows the live expression of the sum of these scores to view the maturity. In addition, the Gök Atlas used in our country; It was created by Şemsi Gök and his colleagues in 1985 by adapting the GP atlas. It is used very frequently by forensic experts [21]. It is reported that in Büken et al. study on Sky Atlas, which included the 11-22 age group, the difference between chronological age and bone age according to age groups in men and women was found to be more than 1 year between the ages of 15-19 in men and between the ages of 11-18 in women [19]. However, since there is the possibility of making different estimates due to some evaluation errors during the examination of radiographs in age determination; Since direct radiography is a 2-dimensional reflection of a 3-dimensional object, any angle error in the position of the area to be radiographed may cause incorrect evaluation of the bones of that region, which is stated as its biggest disadvantage [15,16].

Our study aimed to present the comparative effectiveness of single or multiple assessment methods as well as detect which method was more reliable and applicable by applying GP, TW2 and CH methods to female and male children in the city is Konya. As a result, the aim of the study is to compare the GP and TW2 methods, which are used in the left wrist radiography for bone age determination in the pediatric age group, with the CH, due to the importance of determining the bone age in the pediatric period.

2. Materials And Methods

AP direct radiographs of the hand-wrist that had been stored for the last 5 years in PACS archive of the Department of Radiology (Selçuk University) were retrospectively used in the study. Os capitatum and os hamatum sites of 501 cases aged between 0-15 were measured on the AP direct radiographs. The correlation of these measurements with Greulich and Pyle method, Tanner Whitehouse 2 method, age, gender, and chronological age was retrospectively assessed. Technically improper images or those including pathologies such as a metabolic disorder and fracture were not included in the study. The number of hand-wrist radiographs assessed according to age and gender was showed in Table 1(Table 1).

Chronological age of each case was calculated with the duration between the dates of birth and radiograph. Then, bone ages of the cases were calculated according to the Tanner-Whitehouse 2 and GP methods for all cases. Twenty bones in the hand and wrist were assessed one by one according to their stages of maturation on the radiographs with Tanner-Whitehouse 2 method. The maturation stage of the bones assessed was determined in the method. Each bone has a score according to their stage of maturation and there are separate tables for total scores of female and male individuals. TW2 values that would be used in our study were calculated for all cases. Most commonly accepted TW2 scores were used. However, bone ages are not calculated monthly in this method but expressed in decimals. For example, in a result calculated as 12.5, the place after the decimal point is equal to 5/10 of the year, which means this case is 12 years and 6 months old.

In GP method, the radiograph was matched with the present pictures according to female and male individuals in the atlas for each case. Each picture showed a separate

standard for each gender and the bone age corresponding to the standard was given in months and years. In the study, the previous and next pictures were taken into account for each case and they were matched with the most suitable one.

Os capitatum and os hamatum radiography areas were measured in CH method (Figure 1). The measured os capitatum and os hamatum areas were calculated by summing the CH site in each hand-wrist.

2.1.Statistical Method

All data were digitalized and statistical analysis was performed with *R* version 3.6.0 (The *R* Foundation for Statistical Computing, Vienna, Austria; https://www.r-project.org) software program. Normality of the data was assessed with Anderson-Darling Normality test and Q-Q plots. Linear and polynomial (quadratic) regression equations were set up to predict the chronological ages of the individuals with the areas measured with Capitohamate (CH) planimetry method. Which one of these two regression models was more successful in predicting the chronological age was determined with model fit measures (R^2: Coefficient of determination and RMSE: Root mean square error) and model selection criteria (Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC)). The model with high coefficient of determination and low root means square error value among model fit measures and with low AIC and BIC values among model selection criteria was selected as the correlated model for prediction of chronological age (Quadratic regression model). Chronological age predictions of both male and female individuals were performed with 95% confidence level with the help of selected quadratic regression equation. In addition, chronological ages predicted with GP, TW2 and CH methods and real ages of the individuals were compared with Friedman test and then Bonferroni-corrected Durbin-Conover post-hoc

test. The relationships between the real ages of the individuals and their chronological ages predicted with all methods were assessed with Spearman's *rho* correlation analysis and the agreement between the methods were assessed using the Bland-Altman method, Intra-class correlation coefficient (ICC), Concordance correlation coefficient (CCC), precision, and accuracy values.

3. Results

The regression models that could be used to predict the chronological age with CH areas were given in Table 2 (Table 2). According to the results obtained in Table 2, both linear and polynomial regression models could significantly be used for prediction of chronological age; however, it was observed that quadratic regression model was better at data modelling according to the fit and selection criteria compared with the linear method. The regression coefficients related to the prediction equation obtained with the quadratic regression model and the model that would be set up were significant and coefficient of determination rate was above 90% for both male and female children. According to this information, the equations for chronological age prediction both without separating the genders and for both genders were as follows (Figure 2):

The equation which was set up using all data without separating the genders;

Chronological age (month) = – 6.1429 + 0.4746×(CH) – 0.0002957×$(CH)^2$, R^2=92.68%

For male individuals;

Chronological age (month) = –3.9251 + 0.4623×(CH) – 0.0002950×$(CH)^2$, R^2=92.68%

For female individuals;

Chronological age (month) = $-6.9239 + 0.4622\times(CH) - 0.0002351\times(CH)^2$, R^2=91.51%

The differences between the real ages and chronological ages predicted with different prediction methods for each gender were assessed with Friedman test and then Bonferroni-corrected Durbin-Conover test for multiple comparisons in Table 3 (Table 3).

The agreement and correlations between the real ages of the individuals and age predictions of the methods used for chronological age prediction were assessed in Table 4 (Table 4). All methods were corresponded and correlated in determination of real age. However, for male individuals, the predictions obtained with GP method (mean age: 85.57±54.68) were mean 5.1 units lower than the real ages and the predictions obtained with TW2 method (mean age: 94.08±49.18) were mean 1.7 units higher than the real ages, which were determined with Bland-Altman method (Figure.3). The ages predicted with CH method (mean age: 90.64±47.65) were almost the same with chronological ages (mean age: 90.64±49.50) and there was no deviation in the prediction method (Figure.3).

For female individuals, ages predicted with both GP (mean age: 90.21±54.49) and TW2 (mean age: 91.86±51.24) methods were mean 1.4 and 0.5 units lower than chronological ages (mean age: 91.58±51.89) respectively and there was no deviation in the CH method (mean age: 91.58±49.63) (Figure. 4).

According to the obtained results, there was a statistically significant difference between the real ages and chronological ages predicted with GP and TW2 methods in male children and there was no statistically significant difference between the real ages

and chronological ages predicted with CH method. The ages predicted with GP method were lower than the real ages of children and the ages predicted with TW2 method were higher than the real ages. However, the ages predicted with GP and TW2 methods were significantly different from the ages predicted with CH method.

For female children, there was no significant difference between the real ages and predicted ages obtained with any of the methods and the ages predicted with CH method (mean age: 91.58±49.63) were different from the ages predicted with GP method.

According to these results, it has been concluded that CH method is reliable and have high accuracy for chronological age determination in children and can be a prediction method alternative to GP and TW atlases in literature.

4. Discussion

The most important finding of this study was that CH method is reliable and have high accuracy for chronological age determination in children and can be a prediction method alternative to GP and TW atlases in literature. This methods are accepted as a valid scientific method by courts all over the world [12]. Histological, morphological and radiological methods are used in age determination [13].

In a study conducted on 515 obese children in Brazil, bone age was found to be older than chronological age according to the GP atlas in all groups [17].

He applied the Greulich-Pyle method in Germany and reported that it could be used for children living in West Germany if correction tables were added [18].

Buken et al. investigated whether the Greulich-Pyle (G-P) method was sufficient in estimating forensic age for Turkish children, and as a result, the standard deviation was found to be more than 1 year for girls at the ages of 12, 15, and for boys at the ages of 12, 15, and 18. However, it is not known whether other methods are more useful than this

method. For now, they concluded that this method should be used with caution in cases of possible criminal liability in forensic age diagnosis, unless another method proves to be more useful [19].

In a study including 303 male and 122 female cases between 2009 and 2010 in Iran, both genders were divided into 3 subgroups (6-10, 10-14 and 14-18) and GP atlas was found acceptably accurate and applicable in Iranian female children considering that bone age of female participants was 0.5 months higher [7].

A meta-analysis, which was published in 2019, assessed whether the GP atlas could be applied to all ethnic groups or not and a total of 49 studies between 1950 and 2017 were included in the assessment while 35 of them were found correlated with the meta-analysis. As a result of the meta-analysis, there was no significant difference between the bone age and chronological age in African male cases, Asian female cases, Caucasians, and the Spanish; however, it was emphasized that the GP atlas should be carefully used when applied to Asian male and African female cases [6]. In addition, while the correlation of the studies was being assessed whether the individuals had any diseases or not, the difference found between mean chronological age and mean bone age and only the studies written in English were determined as the selection criteria.

Cases aged between 10 and 22 were included in a study performed in Ethiopia in 2015 and bone age was found 8.7 months lower in male cases and 11.8 months lower in female cases compared with the chronological age in GP atlas, which was not statistically significant and it was stated that new methods must be developed [8]. Only the age range of 0-15 was assessed in our study, which was different from the study above and GP atlas was mean 5.1 months lower in the male group between the ages of 0-15 and mean 1.4

months lower in the female group between the ages of 0-15 compared with the chronological age.

In the study performed by Malina et al. to compare TW2 and TW3 atlases in 1831 young footballers aged between 10 and 17 in 2018, ages of the participants were 0.97 and 1.16 years delayed compared with the chronological age according to TW2 and TW3 atlases. While 42% of the players classified as average according to TW2 atlas were delayed throughout the age range compared with the TW3 atlas, 64% of the cases who grew early according to TW2 atlas were found average compared with the TW3 atlas. They stated that both GP atlas and TW3 atlas would be the first choice for clinical use compared with TW2 [9].

In our study, the GP method measured the age lower and TW2 method measured the age higher in male children and lower in female children. Differently from this study, we think that CH planimetry method is more convenient in our society compared with the other methods. In a study performed on 611 children in Taiwan, while the GP atlas measured the age 1.24 years higher in female children it was measured 0.61 years lower in male children compared with the chronological age [10]. According to a review assessing the GP method applied on 33 female and 37 male participants in the Eastern Uttar Paradesh region of India, age retardation was higher in the male than in the female. It was stated in this study that a larger population was required to apply GP [14].

As it can be seen in the GP atlas in the study performed by Choi et al. on 391 Korean children in 2018 the probability of earlier appearance of os capitatum and os hamatum nuclei was higher in female children than in male children. In addition, they found a strong positive correlation between chronological age and CH planimetry measurement. They stated that there were rising slopes in the planimetry curves of female

and male capitatums and hamatums. The strongest correlations were between os triquetrum, os capitatum and os hamatum bones that were present in all of 20 hands. Mean bone age value predicted with GP method was lower than the one measured with CH planimetry method ($p<0.0001$). The range for 95% confidence interval was between -10.5 and 13.4 months in age prediction with CH planimetry method and between -21.1 and 29.5 months in CP method. While the age was found 1.4 months higher in CH method it was -4.2 months lower in GP method [11].

In our study performed on 501 children, GP method was -5.1 months lower, TW2 method was 1.7 months higher and CH method was 0 months in male children while GP method was -1.7 months lower, TW2 method was 0.5 months higher and CH method was 0 in female children. In addition, our study has the feature of being the study with the highest number of subjects among the studies performed with this method until now.

5. Conclusion

As a result of this study, we performed on 501 children by developing the study by Choi et al. on 391 children, a statistically significant difference was found between male children's chronological ages and chronological ages predicted with GP and TW2 methods, but there was no statistically significant difference between their real ages and chronological ages predicted with CH method. The ages predicted with GP method were lower than the real ages of the children and the ages predicted with TW2 method were higher than the real ages. However, the ages predicted with GP and TW2 were significantly different from the ages predicted with CH method.

For female children, there was no significant difference between their real ages and ages predicted with any of the methods, but the ages predicted with CH method were different from the ages obtained with GP method.

In conclusion, CH planimetry method can be useful for bone age assessment in the city is Konya. Age prediction with a simple application method with 91.83% reliability and within a time shorter than 1 minute has gained an advantage over the methods used until now. In addition, this is the first study performed in our country and as wells being the study performed on CH planimetry method with the highest number of subjects.

The CH planimetry method can be performed automatically after ethnicity and gender parity, can save money in the future if it can be integrated into PACS, and can lead to more precise bone age assessment.

REFERENCES

1. Gökmen E. Radyolojik Yaş Tayini. İstanbul,Türkiye: İstanbul Üniversitesi Fen Fakültesi Prof. Dr. N. Terzioğlu Basım Atölyesi;1990 (in Turkish).
2. Bilgin N, Çekin N, Gülmen M, Alper B. Çukurova Üniversitesi Tıp Fakültesi Adli Tıp Anabilim Dalı'na başvuran yaş tayini olgularının retrospektif değerlendirilmesi. Mersin Üniversitesi Tıp Fakültesi Dergisi. 2003;2: 140-4. (in Turkish).
3. Daş, V. Suriyeli mülteci 18 yaş altı evlilerde kemik yaşının tespitinde gök, greulich-pyle ve tanner-whitehouse atlaslarının kullanılabilirliğinin değerlendirilmesi ve sosyodemografik özellikleri. Uzmanlık tezi, Gaziantep Üniversitesi, Gaziantep, Turkiye. 2019 (in Turkish).
4. 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. Archives of Disease in Childhood 1999; 81:172-173. https://doi.org/10.1136/adc.81.2.172
5. Groell R, Lindbichler F, Riepl T, Gherra L, Roposch A et al. The reliability of bone age determination in central European children using the Greulich and Pyle method. The British Journal of Radiology 1999;72(857): 461-464. https://doi.org/10.1259/bjr.72.857.10505010

6. Alshamrani K, Messina F, Offiah AC. Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. European Radiology 2019;29:2910–23. https://doi.org/10.1007/s00330-018-5792-5
7. Moradi M, Sirous M, Morovatti P. The reliability of skeletal age determination in an Iranian sample using Greulich and Pyle method. Forensic Science International 2012;223(1-3):372. e1-.e4. https://doi.org/10.1016/j.forsciint.2012.08.030
8. Tsehay B, Afework M, Mesifin M. Assessment of reliability of Greulich and Pyle (gp) method for determination of age of children at Debre Markos Referral Hospital, East Gojjam Zone. Ethiopian Journal of Health Sciences. 2017;27(6):631-40. https://doi.org/10.4314/ejhs.v27i6.8
9. Malina RM, Coelho-E-Silva MJ, Figueiredo AJ, Philippaerts RM, Hirose N et al. Tanner–Whitehouse skeletal ages in male youth soccer players: TW2 or TW3? Sports Medicine 2018;48: 991–1008. https://doi.org/10.1007/s40279-017-0799-7
10. Wang YM, Tsai TH, Hsu JS, Chao MF, Wang YT et al. Automatic assessment of bone age in Taiwanese children: A comparison of the Greulich and Pyle method and the Tanner and Whitehouse 3 method. The Kaohsiung Journal of Medical Sciences 2020;36(11): 937-943. https://doi.org/10.1002/kjm2.12268
11. Choi JA, Kim YC, Min SJ, Khil EK. A simple method for bone age assessment: the capitohamate planimetry. European Radiology 2018;28: 2299-307. https://doi.org/10.1007/s00330-017-5255-4
12. Isır A, Buken B, Tokdemir M, Dülger H, Erel Ö et al. Assessing the Age Determination Cases Which Have Been Analyzed at Forensic Medicine Departments of 5 Different Region's Universities in Turkey Between Years 1998-2005 Turkiye Klinikleri Tıp Bilimleri Dergisi 2009;29(2): 304-313 (in Turkish).
13. Baransel Isır A. Adli Tıpta Yaş Tayini. Klinik Gelişim Dergisi Adli Tıp Özel Sayısı 2009;22: 114-121 (in Turkish).
14. Tiwari PK, Gupta M, Verma A, Pandey S, Nayak A. Applicability of the Greulich-Pyle Method in Assessing the Skeletal Maturity of Children in the Eastern Utter Pradesh (UP) Region: A Pilot Study. Cureus. 2020;12(10). doi: 10.7759/cureus.10880. PMID: 33178532; PMCID: PMC7652372.

15. Büken B, Şafak AA, Büken E, Yazici B, Erkol Z et al. Is the Tanner–Whitehouse (TW3) method sufficiently reliable for forensic age determination of Turkish children. Turkish Journal of Medical Sciences 2010;40(5): 797-805. https://doi.org/10.3906/sag-0808-6
16. Yarımoğlu HB, Alper B, Meral B, Çekin N. Yaş tayini uygulamalarında epifiz plağı kapanma derecelerinin incelenmesi. Adli Tıp Bülteni 2005;10(3):84-89 (in Turkish).
17. Artioli TO, Alvares MA, Macedo VSC, Silva TS, Avritchir R et al. Bone age determination in eutrophic, overweight and obese Brazilian children and adolescents: a comparison between computerized BoneXpert and Greulich-Pyle methods. Pediatric Radiology 2019;49(9): 1185-1191. https://doi.org/10.1007/s00247-019-04435-z
18. Kemperdick HF. Determination of skeletal age in children of Western Germany with normal and abnormal growth development. Fortschritte der Medizin 1981 Feb;99(5):152-156. PMID: 7274957.
19. Büken B, Erzengin ÖU, Büken E, Şafak AA, Yazıcı B et al. Comparison of the three age estimation methods: Which is more reliable for Turkish children? Forensic Science International. 2009;183(1-3): 103.e1-.e7. https://doi.org/10.1016/j.forsciint.2008.10.012
20. Prokop-Piotrkowska M, Marszałek-Dziuba K, Moszczyńska E, Szalecki M, Jurkiewicz E. Traditional and new methods of bone age assessment-an overview. Journal of Clinical Research in Pediatric Endocrinology. 2021;13(3): 251. https://doi.org/10.4274/jcrpe.galenos.2020.2020.0091
21. Büken B, Demir F, Büken E. 2001-2003 yılları arasında Abant İzzet Baysal Üniversitesi Düzce Tıp Fakültesi Adli Tıp Anabilim Dalı'na gönderilen yaş tayini olgularının analizi ve adli tıp pratiğinde karşılaşılan güçlükler. Düzce Tıp Fakültesi Dergisi. 2003; 5(2): 18-23 (in Turkish).

Figures

Figure 1 Os capitatum and os hamatum measurement according to the capitohamatum planimetry method

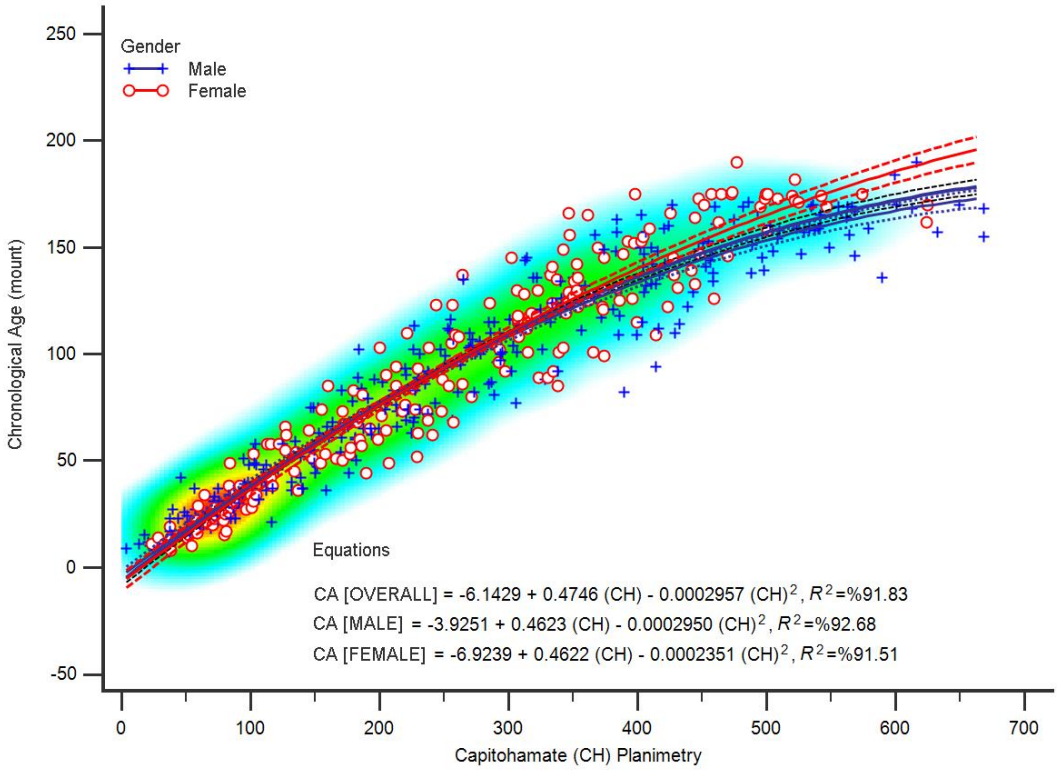
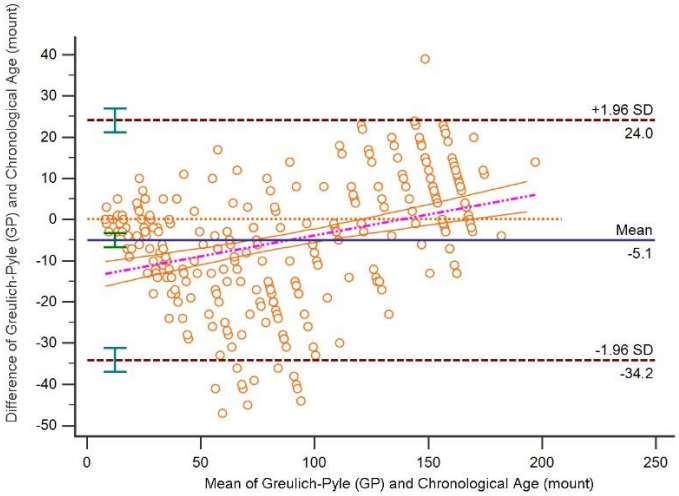
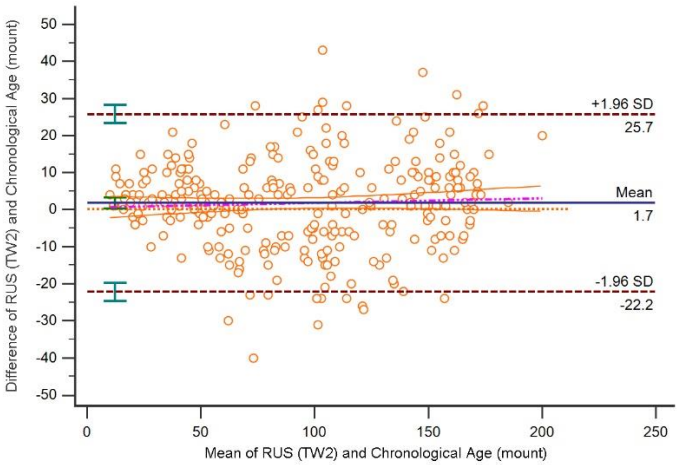


Figure 2 Quadratic regression curves in general and according to genders





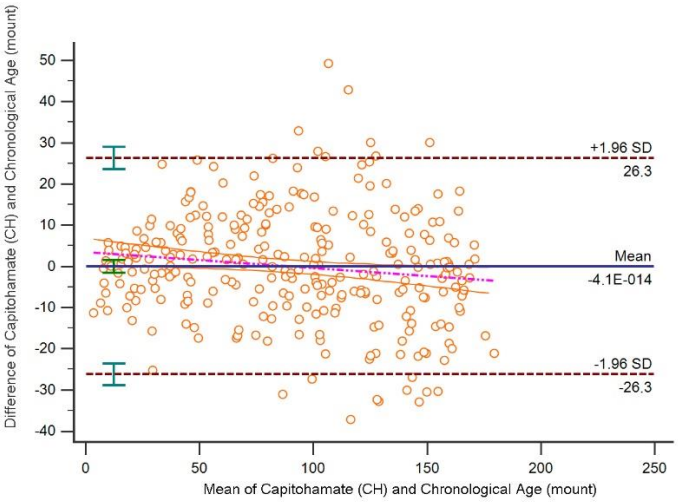


Figure 3 The Bland-Altman Plot revealing the correlation between chronological age (month) and the age predicted methods in male children

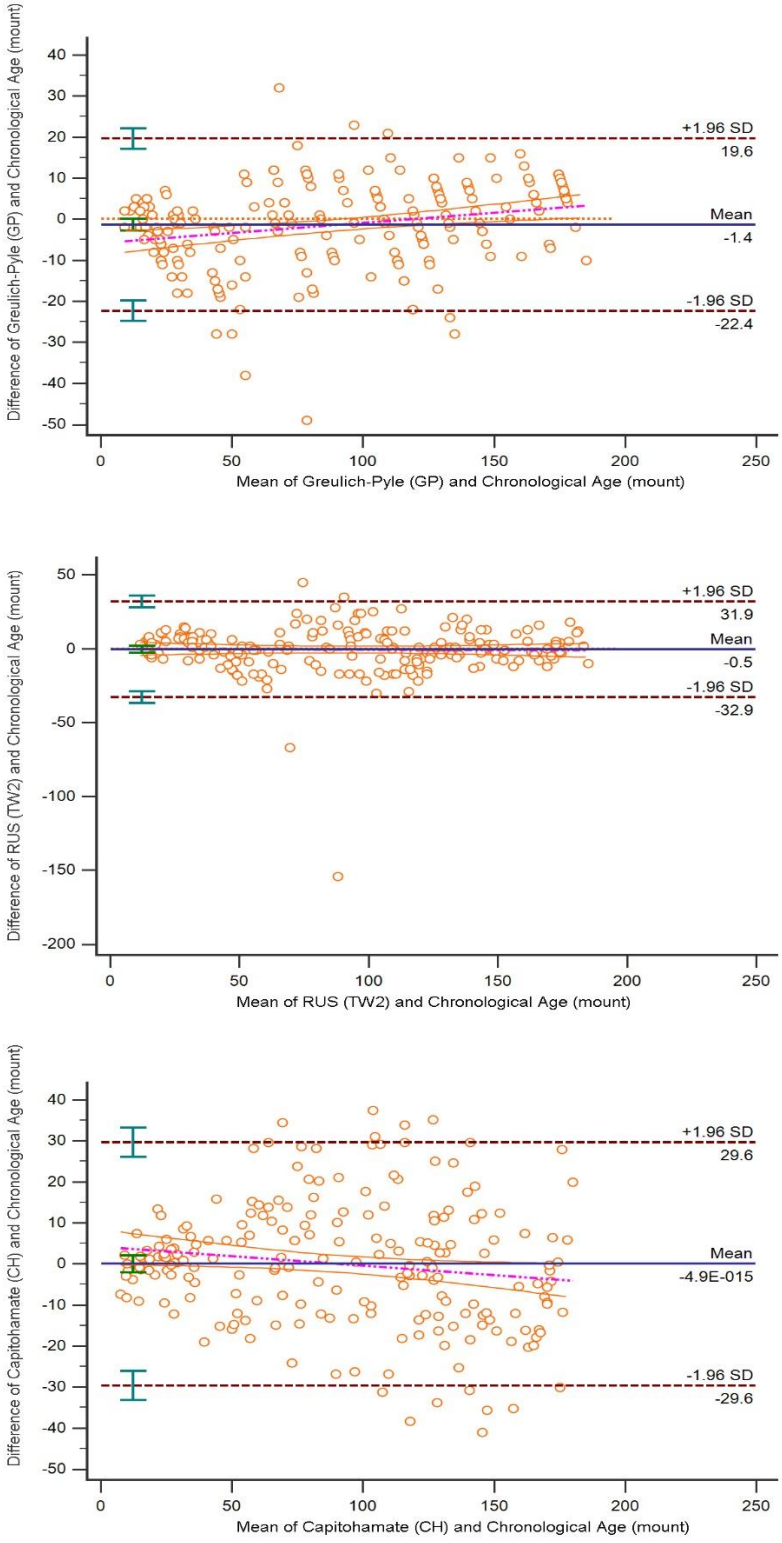


Figure 4 The Bland-Altman Plot revealing the correlation between chronological age (month) and the age predicted methods in female children

Tables

Table 1 Distribution of hand-wrist radiographs by age and gender

Age (months)	Male (*n*=291)	Female (*n*=210)
1-11	10	6
12-23	22	18
24-35	16	20
36-47	27	8
48-59	17	17
60-71	20	15
72-83	23	11
84-95	20	14
96-107	22	12
108-119	22	13
120-131	12	24
132-143	20	20
144-155	21	12
156-167	24	9
168-179	13	19
180	2	2

Table 2 Significance statistics for regression models set up to predict chronological age with CH area

	Regression Coefficients and Model Significances					**Moel Fit Measures**				**Model Selection Criteria**	
Method	t_{β_0}	t_{β_1}	t_{β_2}	***F***		***r***	R^2	**RMSE**		**AIC**	**BIC**
Linear Model											
General	7.99*	65.35*	-	4271.09*		0.946*	0.896	16.29		4215.65	4228.29
Male	8.18*	50.86*	-	2586.88*		0.949*	0.900	15.64		2423.91	2434.92
Female	2.51*	44.55*	-	1985.25*		0.951*	0.905	15.94		1764.92	1774.96
Quadratic Model											
General	–2.32*	31.22*	–11.77*	2793.56*		0.958*	0.918	14.41		4094.88	4111.74
Male	–1.86*	25.81*	–10.27*	1815.61*		0.963*	0.927	13.37		2335.11	2349.78
Female	–2.09*	16.98*	–4.91*	1115.09*		0.957*	0.915	15.09		1743.73	1757.13

$t_{\beta_0,\beta_1,\beta_2}$: Significance values for regression coefficients in linear and quadratic regression models (shows stability and slope coefficients respectively), *F*: Significance value for models, R^2: Coefficient of

determination, RMSE: Root mean square error, AIC: Akaike information criteria, BIC: Bayesian information criteria, * Shows statistical significance (p<0.05).

Table 3 Results on comparison between chronological age and predicted age values obtained with different methods for each gender

	Mean ± SD	**Median (IQR=Q1–Q3)**	***p*-value**	**Multiple Comparison**
Male			<0.001	
Chronological Age (month)	90.64 ± 49.50	90.50 (46 – 136)		GP – TW2
GP Age Prediction (month)	85.57 ± 54.68	72 (36 – 144)		CA – TW2 – CH
TW2 Age Prediction (month)	94.08 ± 49.18	91.50 (51 – 140)		CA – GP – CH
CH Age Prediction (month)	90.64 ± 47.65	92.67 (51.71 – 134.49)		GP – TW2
Female			0.011	
Chronological Age (month)	91.58 ± 51.89	91 (49 – 131)		
GP Age Prediction (month)	90.21 ± 54.49	90 (36 – 132)		TW2 – CH
TW2 Age Prediction (month)	91.86 ± 51.24	96.50 (40 – 134)		GP
CH Age Prediction (month)	91.58 ± 49.63	95.49 (43.80 – 130.69)		GP

CA: Chronological Age, GP: Greulich-Pyle method, TW2: Tanner-Whitehouse 2 method, CH: Capitohamate method, Mean ± SD: mean ± standard deviation, IQR (Q1–Q3): interquartile range (1^{st}quartile – 3^{rd}quartile), *p*-value: calculated with Friedman test, Multiple Comparison: Bonferroni-corrected Durbin-Conover post-hoc tests were used.

Table 4 Results on agreement and correlation statistics between chronological age and predicted age values obtained with different methods for each gender

	CA & GP	**CA & TW2**	**CA & CH**
Male			
B-A Method (95% LoA)	– 5.1 (– 34.2 to 24)	1.7 (– 22.2 to 25.7)	0 (– 26.3 to 26)
ICC (95% CI)	0.959 (0.949 – 0.967)	0.968 (0.960 – 0.975)	0.962 (0.952 – 0.969)
CCC (95% CI)	0.954 (0.944 – 0.963)	0.968 (0.959 – 0.974)	0.962 (0.952 – 0.970)
Precision	0.964	0.969	0.963
Accuracy	0.990	0.999	0.999
Spearman's *rho* (95% CI)	0.971 (0.963 – 0.977)	0.968 (0.960 – 0.975)	0.962 (0.952 – 0.970)
Female			
B-A Method (95% LoA)	– 1.4 (– 22.4 to 19.6)	– 0.5 (– 32.9 to 31.9)	0 (– 29.6 to 29.6)
ICC (95% CI)	0.979 (0.973 – 0.984)	0.948 (0.932 – 0.960)	0.955 (0.942– 0.966)
CCC (95% CI)	0.979 (0.973 – 0.984)	0.948 (0.932 – 0.960)	0.956 (0.942– 0.965)
Precision	0.981	0.948	0.957
Accuracy	0.998	0.999	0.999
Spearman's *rho* (95% CI)	0.981 (0.975 – 0.985)	0.949 (0.933 – 0.961)	0.956 (0.943 – 0.966)

B-A Method (95% LoA): Bland-Altman Method (95% Limits of Agreement), ICC: Intra-class correlation coefficient, CCC: Concordance correlation coefficient, 95% CI: 95% Confidence Intervals, CA: Chronological Age, GP: Greulich-Pyle method, TW2: Tanner-Whitehouse 2 method, CH: Capitohamate method