Original Article

J Korean Soc Radiol 2022;83(5):1081-1089
https://doi.org/10.3348/jksr.2021.0128
eISSN 2951-0805

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

Comparison of Bone Ages in Early Puberty: Computerized Greulich-Pyle Based Bone Age vs. Sauvegrain Method

초기 사춘기의 골연령 비교: 전산화된 Greulich-Pyle 기반 골연령 대비 Sauvegrain 방법

Sang Young Lee, MD, Soo Ah Im, MD*

Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

ORCID iDs
Sang Young Lee https://orcid.org/0000-0001-8684-034X
Soo Ah Im https://orcid.org/0000-0002-1264-3690

Received July 13, 2021
Revised October 19, 2021
Accepted November 22, 2021

***Corresponding author**
Soo Ah Im, MD
Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea.

Tel 82-2-2258-1444
Fax 82-2-599-6771
E-mail saim@catholic.ac.kr

Purpose To compare the computerized Greulich-Pyle based bone age with elbow bone age.

Materials and Methods A total of 2126 patients (1525 girls; 601 boys) whose elbow bone age was within the evaluable range by the Sauvegrain method, and who simultaneously underwent hand radiography, were enrolled in the study. The 1st-bone age and VUNO score of the hand were evaluated using VUNOMed-BoneAge software. The correlation between the hand and elbow bone age was analyzed according to the child's gender and the probability of 1st-bone age.

Results The correlation between VUNO score and elbow bone age (r = 0.898) was higher than the correlation between 1st-bone age and elbow bone age (r = 0.879). Moreover, the VUNO score showed a better correlation with the elbow bone age in patients with a 1st-bone age probability of less than 70%, or in girls. Elbow bone age was more advanced compared to hand bone age, and this difference increased until the middle of puberty and gradually decreased in the latter half.

Conclusion The computerized Greulich-Pyle based hand bone age showed a significant correlation with the elbow bone age at puberty. However, since the elbow bone age tends to advance faster than the hand bone age, caution is required while judging the bone age during puberty.

Index terms Bone and Bones; Artificial Intelligence; Puberty; Hand; Elbow

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

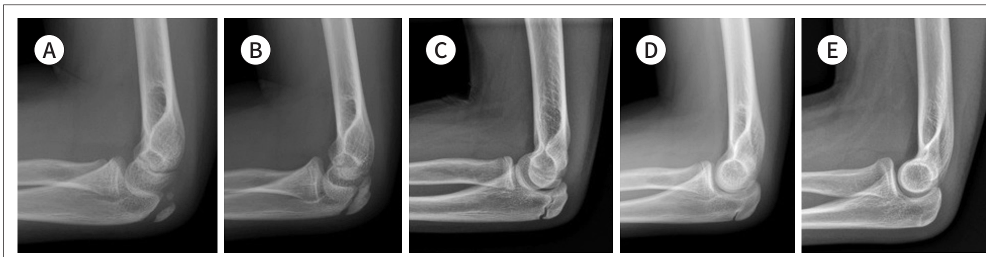
INTRODUCTION

Bone age (BA) evaluation is very important in that reflects the current or future growth status in growth evaluation (1). The Greulich & Pyle (GP) atlas has been the most used BA measurement method (2, 3). Another, Tanner & White (TW) method has less discordance between analysts regardless of the skill level than the GP method (4). However, many doctors use the GP method for its 'simplicity' and 'time efficiency' in clinical field. But, there might be discrepancy of reading BA among analysts because the differences between the reference images are subtle, making it difficult for a detailed BA assessment. In particular, the limitation of GP method in evaluation of adolescent BA appears even more severe in the acceleration phase of puberty (5-7). The acceleration phase shows peak height velocity during the first two years of puberty, i.e., 11–13 years of age in girls and 13–15 years of age in boys. Then, the deceleration phase comes along and the growth rate curve slows down. The problem is that the hand radiograph configuration changes only slightly during the acceleration phase (7).

To reduce reading errors between readers and increase (workflow) efficiency, automated BA assessments have been developed in many countries. There is artificial intelligence (AI) software (VUNOmed-BoneAge, VMBA, Seoul, Korea) by deep learning based on 18940 left-hand radiographs evaluated with GP method of Korean children in 2017 in Korea (8). VMBA is an AI program that provides the three most likely estimated BAs with % in probability and VUNO score (calculated by the summation of all BAs multiplied by each predicted probabilities). However, the significance of VUNO score has not been studied yet.

On the other hand, some hospitals or doctors adopt a method that additionally evaluates the BA of Sauvegrain method in the elbow lateral view to compensate for the inaccuracy of BA measurement by GP method in adolescence (5). This is a simple method of measuring BA using the fact that clear changes occur at 6-month intervals in the olecranon in the adolescence on the lateral view of elbow (Fig. 1). This method showed high accuracy and reproducibility, and the intraobserver difference was lower compared to GP method in determining the BA in puberty (9).

Fig. 1. Elbow bone age according to the Sauvegrain method.



	Girls	Boys	Characteristic morphology of the olecranon apophysis
A	11	13	Two ossification centers
B	11.5	13.5	Half-moon shape
C	12	14	Rectangular shape
D	12.5	14.5	Beginning of fusion
E	13	15	Complete fusion

Although, there have been many studies comparing AI vs. human ratings in reading BA (3, 10-12), it is considered that there may be errors in evaluating the accuracy of the AI of the GP method in a small sized study group, especially in adolescent patients, in whom it is not easy for the human reader to verify and who are in critical zone for appropriate treatment based on accurate BA evaluation.

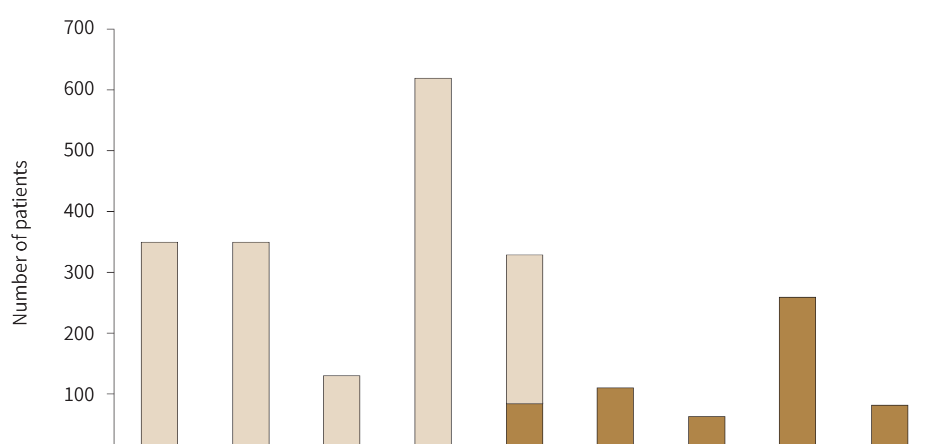
Therefore, the authors tried to compare and correlate each BA using automated BA assessment system (VMBA) and Sauvegrain method, especially in puberty with high discordance of hand BA.

MATERIALS AND METHODS

POPULATION AND STUDY DESIGN

This study was approved by the Institutional Review Board of Seoul St. Mary's hospital. The need for informed consent was waived for its retrospective study design (IRB No. KC20RISI0493). A total of 2274 children (*n* = 1658 in girls, *n* = 616 in boys) who took X-rays of left hand as well as elbow X-ray at the same time from January 2012 to December 2019 was included. These were cases in which an additional elbow lateral radiograph was taken when the BAs of the hand were ambiguous in early puberty patients. Among these patients, 148 patients were excluded based on following criteria: 1) patients whose BA cannot be assessed by the Sauvegrain method on the lateral radiograph of the elbow (*n* = 143) or 2) patients with underlying bone abnormalities such as lymphoproliferative disease (*n* = 3) and achondroplasia (*n* = 2). Finally, 2126 patients (*n* = 1525 in girls, *n* = 601 in boys), who had both X-rays of hand and elbow taken at the same time and whose elbow BA fell within the evaluable category (elbow BA: 11–13 years for girls and 13–15 years for boys), were studied. The distribution of girls and

Fig. 2. Manual analysis of elbow lateral radiographs according to the Sauvegrain method: age and sex distribution of bone ages.



	11	11.5	12	12.5	13	13.5	14	14.5	15
Girls	180	349	132	620	244	0	0	0	0
Boys	0	0	0	0	85	111	64	259	82

Elbow bone age (years)

boys according to elbow BA is detailed in the table (Fig. 2). The chronological age distribution of patients was between 8 years 2 months and 17 years 2 months.

BA ASSESSMENT

To determine BA, the BA reports were extracted after the Digital Imaging and Communications in Medicine (DICOM) images were sent to and analyzed by the software. VMBA displays the three most likely estimated BAs with % in probability and VUNO score (Fig. 3). BAs of elbow lateral radiographs (elbow BA) were manually analyzed by two radiologists (one resident and a 20-year experienced pediatric radiologist) who were blinded to the patient's age in elbow BA reading according to the Sauvegrain method with consensus (Fig. 1). The radiologists assessed 1) the correlation between the 1st-BA and the elbow BA, and 2) the correlation between the VUNO score and the elbow BA by gender, age, and 1st-BA probability in detail.

STATISTICAL ANALYSIS

SSPS Statistics (ver. 21; IBM Corp., Armonk, NY, USA) was used for statistical analysis. Scattered plots and the Pearson correlation coefficient (*r*) were used as the correlation level of measurements for comparison between elbow BA and 1st-BA, and elbow BA and VUNO score from VMBA. In addition, the Bland-Altman analysis was used to show the variability of the measurements and difference between elbow BA and VMBA. *t*-test with *p*-value < 0.05 was

Fig. 3. Screenshot of bone age assessment by VUNOmed-BoneAge software program.
On the right column, three most likely estimated bone ages and each probability of an 11-year-old girl are shown. The 1st-rank bone age of this patient is 12 years with a probability of 61.03%. The 2nd- and 3rd-rank bone ages are 11 years with a probability of 38.88% and 13 years with a probability of 0.09%, respectively. The program also reports the optimal bone age as 11 years 7 months using the VUNO score.

Table 1. Correlation of Elbow Bone Age with Automatically Estimated 1st-Bone Age and VUNO Score

	Pearson Correlation Coefficient (*r*)		
	1st-Bone Age	VUNO Score	*p*-Value
Total	0.879	0.898	< 0.001
Girl	0.626	0.687	< 0.001
Boy	0.682	0.719	< 0.001

used for these analyses. A *p*-value < 0.05 was considered statistically significant.

RESULTS

The correlation coefficient (*r*) between the 1st-BA and the elbow BA was high (*r* = 0.879) regarding the entire group of patients. VUNO score showed higher correlation coefficient with elbow BA (*r* = 0.898) (Table 1). The boys showed relatively high correlation coefficient with 1st-BA (*r* = 0.682) and VUNO score (*r* = 0.719), compared to the girls (*r* = 0.626, 0.687) (Table 1).

Bland-Altman analysis exhibited an average difference between elbow BA and 1st-BA, and between elbow BA and VUNO score (Table 2, Fig. 4). We used elbow BA as a reference to compare 1st-BA and VUNO score. The average difference was 0.421 year (girls 0.499 year, boys 0.224 year) for total group between 1st-BA and elbow BA, and 0.445 year (girls 0.531 year, boys 0.225 year) between VUNO score and elbow BA. We also evaluated the BA differences according to elbow BA by age and gender in detail (Table 3). Elbow BA tended to precede hand BA except for the onset of early puberty. The BA difference was greatest at the mid-puberty (12 years in girls and at 14 years in boys), and the difference gradually increased during the early puberty while it gradually decreased during the late puberty. The BA difference of boys was lower than girls (Table 3).

In addition, we analyzed the correlation between the hand BA and elbow BA in the groups that are above and below the median 70% of the probability of 1st-BA, since the probability of 1st-BA was distributed between 47.26% and 91.29%. When the probability of 1st-BA was below 70%, the VUNO score showed improved results in its association with elbow BA, especially in girls (Table 4).

DISCUSSION

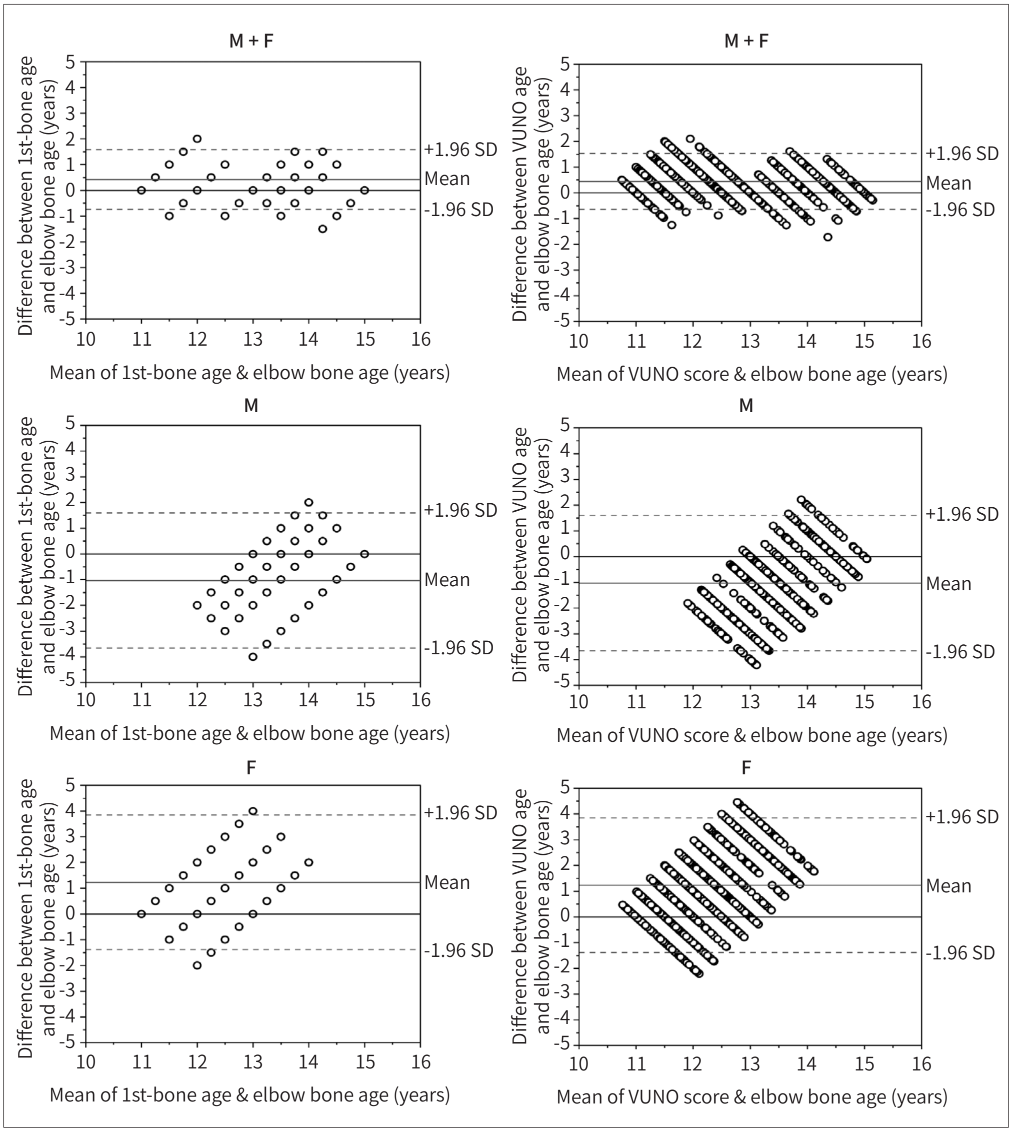
Despite the importance of BA evaluation in early puberty, when rapid growth occurs, there may be large differences between readers in this stage of BA evaluation (5-7). To reduce inter- and intraobserver bias, several automated BA assessments have been developed by machine

Table 2. Bland-Altman Results for Average Difference between Elbow Bone Age with 1st-Bone Age and VUNO Score

	Average Difference (95% CI) (Year)	SD	Slope	*p*-Value
Total				
1st-bone age	0.421 (0.395, 0.447)	0.592	-0.126	< 0.001
VUNO score	0.445 (0.420, 0.469)	0.555	-0.137	< 0.001
Girl				
1st-bone age	0.499 (0.469, 0.528)	0.593	-0.125	< 0.001
VUNO score	0.531 (0.503, 0.559)	0.551	-0.143	< 0.001
Boy				
1st-bone age	0.224 (0.181, 0.267)	0.541	-0.101	< 0.001
VUNO score	0.225 (0.185, 0.265)	0.501	-0.068	< 0.001

CI = confidence interval, SD = standard deviation

Fig. 4. Bland-Altman plots for average difference between elbow bone age with 1st-bone age and VUNO score.
F = female, M = male, SD = standard deviation



learning and deep learning, with high accuracy and efficiency in recent studies (3, 10-14).

Our study shows considerable correlation rate (r = 0.898) of computerized GP based BA with Sauvegrain method in puberty. This is compatible with previous serial studies in 'all age' groups; the correlation rate of 0.91–0.93 was obtained in a study by Pose Lepe et al. (3) with 1500 children using GP method vs. automated BA assessment using BoneXpert software. It is particularly interesting that there was a high correlation 'in puberty' which showed a significant measurement deviation between AI and human ratings in previous reports (8).

In our study, the group of boys exhibited a higher correlation than the group of girls. This might have resulted from the fact that the BA interval by GP method of the group of girls (13, 14, and 15 age) was longer than that of the group of boy (13, 13.5, 14, and 15 age). Also, it demonstrates that using the VUNO score rather than simply using the 1st-BA shows a better cor-

Table 3. Average Difference of Elbow Bone Age with Estimated 1st-Bone Age and VUNO Score in the Group of Age and Sex

Girl	Elbow Bone Age				
	11	11.5	12	12.5	13
Elbow bone age-1st-bone age (year)					
Mean	-0.083	0.337	0.735	0.715	0.484
Standard deviation	0.276	0.370	0.458	0.656	0.521
Elbow bone age–VUNO score (year)					
Mean	0.066	0.408	0.758	0.718	0.451
Standard deviation	0.332	0.362	0.474	0.605	0.521
Boy	**Elbow Bone Age**				
	13	13.5	14	14.5	15
Elbow bone age-1st-bone age (year)					
Mean	-0.200	0.068	0.453	0.369	0.238
Standard deviation	0.318	0.443	0.490	0.581	0.436
Elbow bone age-VUNO score (year)					
Mean	-0.166	0.076	0.437	0.364	0.229
Standard deviation	0.332	0.428	0.500	0.514	0.403

Table 4. Agreement of Elbow Bone Age with Estimated 1st-Bone Age, VUNO Score Categorized by Two Groups: Less Than 70% and 70% or More of 1st-Bone Age Probability

	Correlation Coefficient (*r*)			
	1st-Bone Age < 70%		1st-Bone Age ≥ 70%	
	1st-Bone Age	VUNO Score	1st-Bone Age	VUNO Score
Total	0.859	0.890	0.880	0.891
Girl	0.612	0.721	0.632	0.664
Boy	0.659	0.702	0.411	0.427

$p < 0.05$.

relation with the elbow BA. This is especially evident in the case for girls with a 1st-BA probability of less than 70%. These results suggest that computerized BA assessment can help determine BA when the 1st-BA is ambiguous in girls with wide BA interval in GP atlas.

Another finding from this study was that in puberty, the elbow BA showed a tendency to precede the hand BA, and this pattern was observed with both 1st-BA and the VUNO score with no significant difference. The difference between elbow BA and hand BA was greatest at 12 years in girls and 14 years in boys. Therefore, it can be assumed that elbow demonstrates advanced BA when there is no hand X-ray available for this population. Further research will be needed to determine which of elbow BA or hand BA better reflects the growth of children during puberty.

There are some limitations in this study. First, the lateral elbow X-ray was used as the reference BA. Therefore, there is controversy in comparing the correlation rates between automated BA and reference BA with other studies, which used GP method as the reference. However, short BA category interval of the Sauvegrain method could be thought to affect the correlation rate more specifically than GP method. Second, the study population is from a single center and the male to female ratio and age distribution of study subjects were uneven. Third, since the Sauvegrain method was only applicable for 2 years of BA (11–13 age of girl, 13–15 age of boy), the difference in BA gap beyond 2 years could not be included.

In conclusion, the computerized GP based BA, especially VUNO score, showed considerable correlation with Sauvegrain method at the age of puberty. It can be particularly helpful in adolescent patients with ambiguous 1st-BA or in girls who the GP atlas has wider BA intervals. In addition, since the elbow BA during puberty tends to be faster than the hand BA, caution is needed in judging the BA during this period of age.

Author Contributions

Conceptualization, I.S.A.; data curation, all authors; formal analysis, all authors; investigation, all authors; methodology, all authors; project administration, all authors; resources, all authors; software, all authors; supervision, I.S.A.; validation, I.S.A.; visualization, all authors; writing—original draft, all authors; and writing—review & editing, I.S.A.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Funding

None

REFERENCES

1. Martin DD, Wit JM, Hochberg Z, Sävendahl 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
2. Greulich WW, Pyle SI. *Radiographic atlas of skeletal development of the hand and wrist.* 2nd ed. Paloalto: Stanford University Press 1959
3. Pose Lepe G, Villacrés F, Silva Fuente-Alba C, Guiloff S. Correlation in radiological bone age determination using the Greulich and Pyle method versus automated evaluation using BoneXpert software. *Rev Chil Pediatr* 2018;89:606-611
4. Tanner JM, Healy MJR, Goldstein H, Cameron M. *Assessment of skeletal maturity and prediction of adult height (TW3) method.* 3rd ed. Philadelphia: Saunders 2001
5. Kim DH. Assessment of bone age during pubertal age. *J Korean Soc Pediatr Endocrinol* 2011;16:135-138
6. Eitel KB, Eugster EA. Differences in bone age readings between pediatric endocrinologists and radiologists. *Endocr Pract* 2020;26:328-331
7. Canavese F, Charles YP, Dimeglio A. Skeletal age assessment from elbow radiographs. Review of the literature. *Chir Organi Mov* 2008;92:1-6
8. Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, et al. Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency. *AJR Am J Roentgenol* 2017;209:1374-1380
9. Diméglio A, Charles YP, Daures JP, de Rosa V, Kaboré B. Accuracy of the Sauvegrain method in determining skeletal age during puberty. *J Bone Joint Surg Am* 2005;87:1689-1696
10. Martin DD, Sato K, Sato M, Thodberg HH, Tanaka T. Validation of a new method for automated determination of bone age in Japanese children. *Horm Res Paediatr* 2010;73:398-404
11. Zhang SY, Liu G, Ma CG, Han YS, Shen XZ, Xu RL, et al. Automated determination of bone age in a modern Chinese population. *ISRN Radiol* 2013;2013:874570
12. Martin DD, Meister K, Schweizer R, Ranke MB, Thodberg HH, Binder G. Validation of automatic bone age rating in children with precocious and early puberty. *J Pediatr Endocrinol Metab* 2011;24:1009-1014
13. Pan I, Baird GL, Mutasa S, Merck D, Ruzal-Shapiro C, Swenson DW, et al. Rethinking Greulich and Pyle: a deep learning approach to pediatric bone age assessment using pediatric trauma hand radiographs. *Radiol Artif Intell* 2020;2:e190198

14. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans Med Imaging* 2009;28:52-66

초기 사춘기의 골연령 비교: 전산화된 Greulich-Pyle 기반 골연령 대비 Sauvegrain 방법

이상영 · 임수아*

목적 Greulich-Pyle 기반 전산화된 손 골연령과 팔꿈치 골연령을 비교하고자 하였다.

대상과 방법 팔꿈치 골연령이 Sauvegrain 방법에 의해 평가 가능한 범위 내에 있고, 동시에 손 X선 사진을 촬영한 2126명의 환자(여아 1525명, 남아 601명)를 대상으로 하였다. VUNOMed-BoneAge 소프트웨어를 이용하여 손의 1순위 골연령과 VUNO 점수를 얻었으며, 아동의 성별과 1순위 골연령 확률에 따라 손 골연령과 팔꿈치 골연령의 상관관계를 분석하였다.

결과 VUNO 점수와 팔꿈치 골연령의 상관관계(r = 0.898)가 1순위 골연령과 팔꿈치 골연령의 상관관계(r = 0.879)보다 높았다. 1순위 골연령 확률이 70% 미만이거나 여아인 경우, VUNO 점수를 사용하면 팔꿈치 골연령과 더 좋은 상관관계를 보였다. 팔꿈치 골연령은 손 골연령보다 진행된 경향을 보였으며 그 차이는 사춘기 중반까지 증가하다가 후반에 점차 감소하였다.

결론 사춘기 시기의 Greulich-Pyle 기반 전산화된 손 골연령은 팔꿈치 골연령과 유의한 상관관계를 보였다. 다만 팔꿈치 골연령은 손 골연령보다 빠른 경향이 있어 사춘기의 골연령 판단에 있어 주의가 필요하겠다.

가톨릭대학교 의과대학 서울성모병원 영상의학과