

Assessment of Bone Age in Prepubertal Healthy Korean Children: Comparison among the Korean Standard Bone Age Chart, Greulich-Pyle Method, and Tanner-Whitehouse Method

Jeong Rye Kim, MD¹, Young Seok Lee, MD¹, Jeesuk Yu, MD²

Departments of ¹Radiology and ²Pediatrics, Dankook University Hospital, Cheonan 330-715, Korea

Objective: To compare the reliability of the Greulich-Pyle (GP) method, Tanner-Whitehouse 3 (TW3) method and Korean standard bone age chart (KS) in the evaluation of bone age of prepubertal healthy Korean children.

Materials and Methods: Left hand-wrist radiographs of 212 prepubertal healthy Korean children aged 7 to 12 years, obtained for the evaluation of the traumatic injury in emergency department, were analyzed by two observers. Bone age was estimated using the GP method, TW3 method and KS, and was calculated in months. The correlation between bone age measured by each method and chronological age of each child was analyzed using Pearson correlation coefficient, scatterplot. The three methods were compared using one-way analysis of variance.

Results: Significant correlations were found between chronological age and bone age estimated by all three methods in whole group and in each gender (R^2 ranged from 0.87 to 0.9, $p < 0.01$). Although bone age estimated by KS was slightly closer to chronological age than those estimated by the GP and TW3 methods, the difference between three methods was not statistically significant ($p > 0.01$).

Conclusion: The KS, GP, and TW3 methods show good reliability in the evaluation of bone age of prepubertal healthy Korean children without significant difference between them. Any are useful for evaluation of bone age in prepubertal healthy Korean children.

Index terms: Bone age; Greulich-Pyle method; Tanner-Whitehouse method; Korean standard bone age chart

INTRODUCTION

Bone age assessment allows the evaluation of developmental status, sexual maturity and prediction of the ultimate height of children. Bone age assessment plays an important role in diagnosis and monitoring of treatment

Received July 17, 2014; accepted after revision October 28, 2014. The present research was conducted by the research fund of Dankook University in 2014.

Corresponding author: Young Seok Lee, MD, Department of Radiology, Dankook University Hospital, 201 Manghyang-ro, Dongnam-gu, Cheonan 330-715, Korea.

• Tel: (8241) 550-6921 • Fax: (8241) 552-9674
• E-mail: yslee@dkuh.co.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

of growth disorders and endocrinological abnormalities in children by detecting significant discrepancy between bone age and chronological age (1).

For evaluation of bone age, two major methods are generally used globally in clinical practice based on radiograph of left hand-wrist: the Greulich-Pyle (GP) method (2) and the Tanner-Whitehouse 3 (TW3) method (3). Both methods have been used to assess bone age in Korean children (4, 5), although whether these standards can be applied to the current assessment of bone age in children of different ethnicities is questionable. In 1999, Korean standard bone age chart (KS) was published, which was derived from radiographs of 3407 Korean children (6, 7). Since then, the KS has been widely used in Korea.

There have been some comparative studies between the GP method and TW3 method around the world including Korea (8-12). But, no study about KS compared to the GP

or TW3 method has been done. Therefore, the aim of this study was to compare the reliability of these three methods (GP, TW3, and KS) in the evaluation of bone age of Korean children.

MATERIALS AND METHODS

We selected left hand-wrist radiographs of 278 prepubertal Korean children (aged between 7 and 12 years) between January 1, 2002 and December 31, 2012. These radiographs were obtained for the evaluation of the traumatic injury in the emergency department. Radiographs of 66 children were excluded from the study due to inadequate radiographic quality, poor position, bony abnormalities including fracture and those were derived from children who were diagnosed as growth disorders or endocrinological disorders. A total of 212 radiographs were analyzed, which included 135 boys and 77 girls.

Bone age was estimated independently by two observers including one attending radiologist and one resident. Bone age was estimated using the GP method, TW3 method, and KS. The TW3 method was calculated on the average of the Radius-Ulna-Short bone (RUS) score and carpal score. When there were some mismatches between the patient's radiography and image of KS chart or GP atlas, we decided the mean value of two most similar images of KS chart or GP atlas as the patient's bone age. Patient-identifying

information, except sex, was masked during interpretation including subject's chronological age. Bone age and chronological age were calculated in months (Table 1).

The correlation between chronological age and bone age measured by each method was analyzed using Pearson correlation coefficient and scatterplot. The differences between the chronological age and the measured bone ages by each method were calculated for each child, and one-way analysis of variance (ANOVA) was used to test of statistical significance of difference between three methods using those calculated values. Interobserver reliability coefficient was also calculated using intra-class correlation coefficient by each method. All analyses were performed using SPSS software for Windows version 17.0 (SPSS Inc., Chicago, IL, USA). *P* values < 0.01 were considered to be statistically significant.

RESULTS

In the whole group and the group of boys, the average of bone age estimated by KS was the closest to the average of chronological age. But, in girls, the average of bone age estimated by TW3 method was the closest to the average of chronological age (Table 2). Bone age measured by GP method tended to younger than chronological age in the whole group and the group of boys. On the other hand, bone age measured by TW3 method tended to older than

Table 1. Means and Standard Deviations of Chronological Age and Bone Age Measured with GP, TW3, and KS in Different Gender and Age Distribution

Age (Years)	No.		Chronological Age (Months)		Bone Age (Months)					
	Boys	Girls			GP		TW3		KS	
			Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
7	23	18	89.6 (4.1)	89.1 (3.1)	82.6 (10.6)	91.9 (13.7)	92 (9.4)	91.2 (12.0)	87.2 (11.7)	91 (15.6)
8	22	14	102.6 (3.8)	100.9 (4.0)	91.2 (12.7)	104.8 (17.6)	106.6 (12.4)	103.6 (15.6)	97.8 (9.9)	103.3 (13.9)
9	16	10	113.8 (3.9)	112.2 (3.3)	99.8 (13.4)	116.4 (14.8)	113.9 (14.6)	112.9 (17.1)	106.8 (12.8)	116.4 (16.4)
10	24	17	126 (3.1)	127.3 (2.3)	119.8 (14.2)	130.2 (6.8)	133.4 (16.4)	132.1 (9.2)	124.4 (14.2)	135.5 (8.7)
11	21	8	136.2 (3.1)	137.3 (3.7)	136.2 (13.4)	138 (14.3)	146.1 (12.1)	137.9 (12.4)	138.6 (14.1)	143.3 (14.4)
12	28	10	151 (3.5)	150.3 (3.0)	151.5 (10.3)	150.6 (7.2)	159.1 (13.4)	144.5 (3.2)	156.4 (13.4)	157.5 (6.2)

Note.— Figures in parenthesis refer to standard deviations. GP = Greulich-Pyle method, KS = Korean standard bone age chart, TW3 = Tanner-Whitehouse 3 method

Table 2. Average and Standard Deviation of Chronological Age and Bone Age Measured with GP, TW3, and KS

	Chronological Age (Months)	GP Method Bone Age (Months)	TW3 Method Bone Age (Months)	KS Method Bone Age (Months)
All (n = 212)	119.2 ± 21.74	116.4 ± 26.79	123.2 ± 26.26	120.33 ± 27.42
Boys (n = 135)	121.16 ± 21.80	115.34 ± 28.40	126.76 ± 27.38	120.28 ± 27.98
Girls (n = 77)	115.64 ± 21.34	118.29 ± 23.77	117.06 ± 23.06	120.42 ± 26.59

Note.— GP = Greulich-Pyle method, KS = Korean standard bone age chart, TW3 = Tanner-Whitehouse 3 method

Assessment of Bone Age in Korean Children: Comparison among KS, GP, and TW3 Method

chronological age in the whole group and in each gender.

The correlations between chronological age and bone age estimated by each method are shown in Figures 1, 2, and 3. There were significant correlations between chronological

age and bone age estimated by the three methods in the whole group and each gender (R^2 ranged from 0.87 to 0.9, $p < 0.01$). Table 3 summarizes correlation coefficients between chronological age and bone age measured with

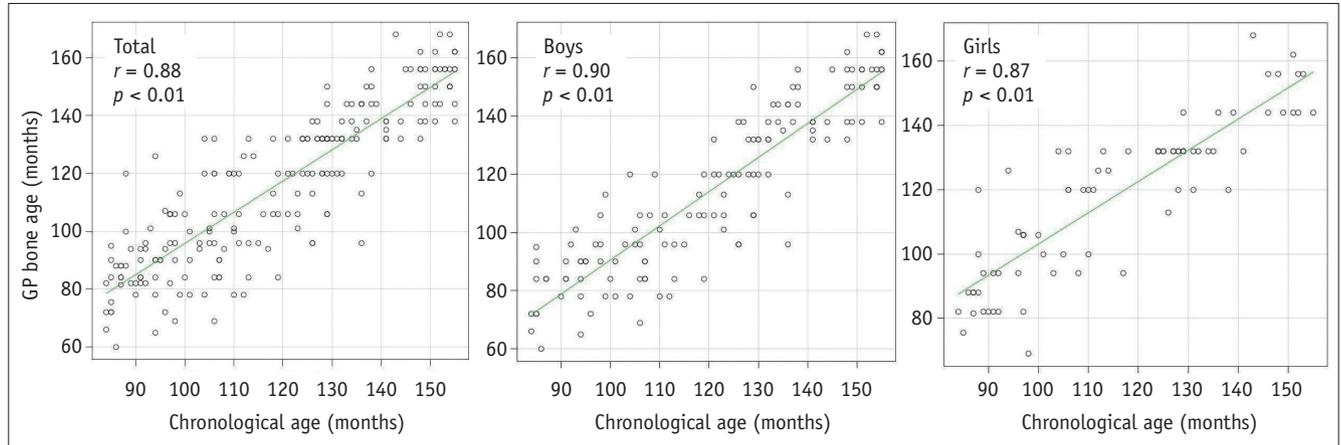


Fig. 1. Correlation between chronological age and bone age measured with Greulich-Pyle (GP) method in whole group and in each gender.

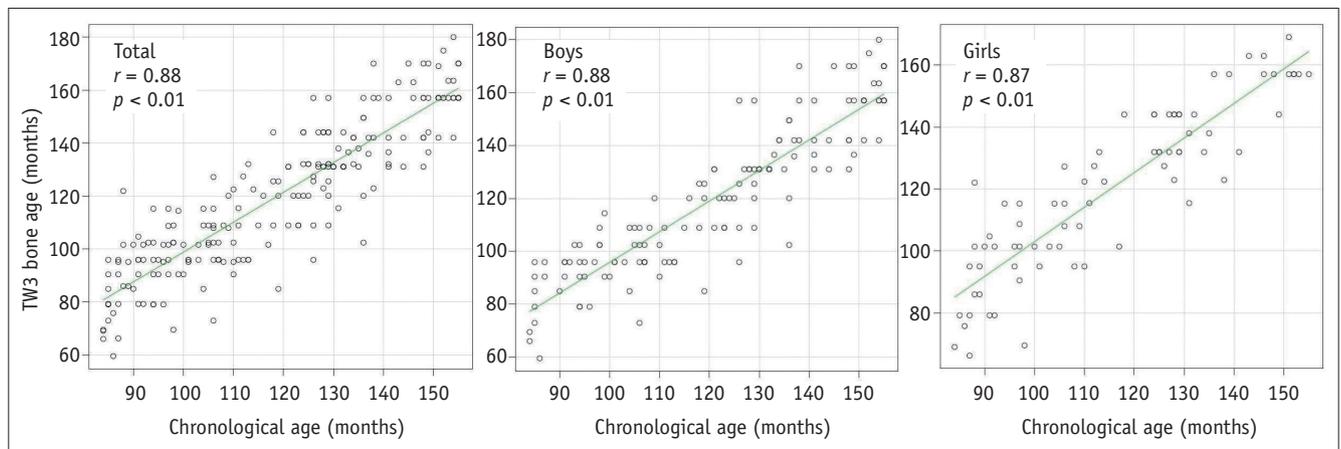


Fig. 2. Correlation between chronological age and bone age measured with Tanner-Whitehouse 3 (TW3) method in whole group and in each gender.

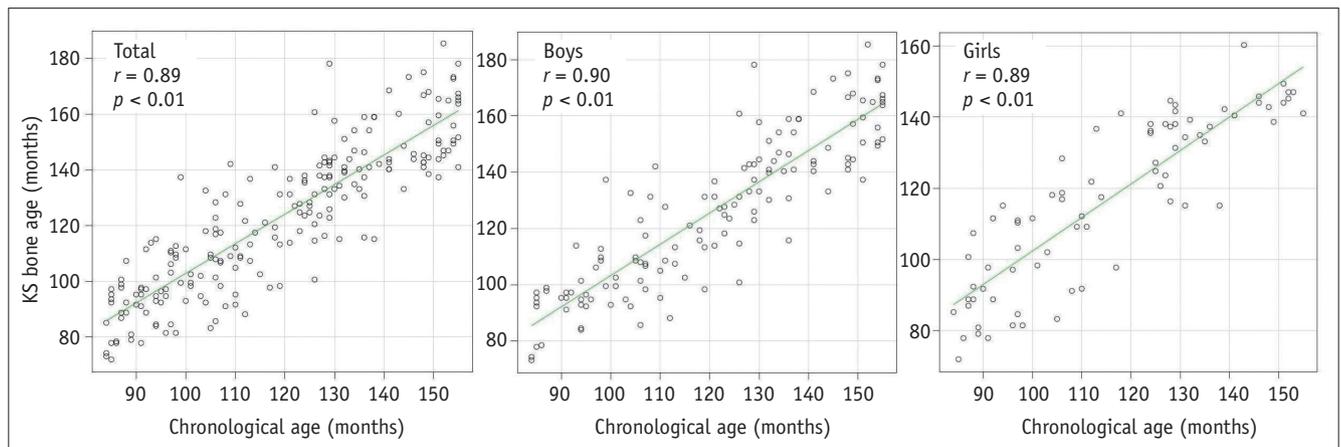


Fig. 3. Correlation between chronological age and bone age measured with Korean standard bone age chart (KS) in whole group and in each gender.

each method. In the whole group and the group of girls, correlation coefficient between chronological age and bone age estimated by KS ($R^2 = 0.89$ in the whole group and the group of girls) was slightly higher than others (GP and TW3, $R^2 = 0.88$ in the whole group and 0.87 in the group of girls). In the group of boys, the correlation coefficient between chronological age and bone age estimated by KS and GP ($R^2 = 0.9$) were slightly higher than TW3 ($R^2 = 0.88$). But, according to the result of one-way ANOVA (Table 4), the difference between three methods was not statistically significant in the whole group ($p = 0.9$) and in each gender ($p = 0.34$ in boys and 0.29 in girls).

Average bone age differences among individual observers are given in Table 5. The average difference in bone age assessments among observers were 0.45 ± 1.79 months by the GP method, 0.45 ± 1.81 months by the TW3 method and

Table 3. Correlation Coefficients between Chronological Age and Bone Age Measured with GP, TW3, and KS

Sex	GP	TW3	KS
All	0.88 ($p < 0.01$)	0.88 ($p < 0.01$)	0.89 ($p < 0.01$)
Boys	0.9 ($p < 0.01$)	0.88 ($p < 0.01$)	0.9 ($p < 0.01$)
Girls	0.87 ($p < 0.01$)	0.87 ($p < 0.01$)	0.89 ($p < 0.01$)

Note.— GP = Greulich-Pyle method, KS = Korean standard bone age chart, TW3 = Tanner-Whitehouse 3 method

Table 4. Comparison of GP, TW3, and KS Using One-Way Analysis of Variance

Group	Method	Mean	SD*	F Ratio	P
All (n = 212)	TW3	10.25	8.36	0.10	0.90
	GP	10.39	8.35		
	KS	10.04	7.75		
Boys (n = 135)	TW3	10.84	9.06	1.08	0.34
	GP	11.01	8.83		
	KS	9.60	8.07		
Girls (n = 77)	TW3	9.19	6.89	1.23	0.29
	GP	9.30	7.37		
	KS	10.81	7.14		

Note.— *Means and standard deviations of differences between chronological age and measured bone ages using each method. GP = Greulich-Pyle method, KS = Korean standard bone age chart, TW3 = Tanner-Whitehouse 3 method

Table 5. Average Interobserver Differences in Bone Age (Months) by GP, TW3, and KS

Group	GP			TW3			KS		
	Mean*	SD*	Range	Mean*	SD*	Range	Mean*	SD*	Range
All	0.45	1.79	0–12	0.45	1.81	0–19.2	0.21	1.19	0–9.8
Boys	0.43	1.70	0–12	0.24	1.09	0–7.8	0.09	0.70	0–6
Girls	0.49	1.95	0–12	0.82	2.6	0–19.2	0.44	1.74	0–9.8

Note.— *Means and standard deviations of differences in measured bone ages by two observers using each method. GP = Greulich-Pyle method, KS = Korean standard bone age chart, SD = standard deviation, TW3 = Tanner-Whitehouse 3 method

0.21 ± 1.19 months by the KS method. The interobserver reliability coefficients were more than 0.95 in all three methods (ranging from 0.996 to 1.000) (Table 6).

DISCUSSION

Bone age reflects current developmental status and sexual maturity of children and allows prediction of the ultimate height. Thus, assessment of bone age is important in the diagnosis and monitoring of the treatment of growth disorders and endocrinologic disorders.

There are various methods for bone age assessment, but two major methods are commonly used globally for bone age assessment in children based on radiography of left hand-wrist: the GP method (2) and the TW3 method (3). The GP method was derived from white children of upper socioeconomic level in United States of America during the 1930s and the second edition was published in 1959. Bone age determination is performed by comparing the left hand-wrist radiograph of a patient with the GP atlas, so it is relatively simple and convenient. The TW method was derived from British children in the 1950s and it was revised as TW3 method in 2001. The TW3 method uses a detailed shape analysis of several bones, and bone age assessment by TW3 method is performed by scoring each ossification centers and plotting the final sum. Thus, some researchers claim that the TW3 method provides high reproducibility and less intra- and inter-observer difference (10–12), but it is time-consuming and difficult to apply.

In Korea, the GP and TW3 methods are generally used for bone age estimation. However, some radiologists and clinicians have certain reservations regarding the applicability of these two methods in the evaluation of skeletal maturity of current Korean children, because subjects of these two methods have different ethnicities from Korean children, and these are long-established. A study by Ontell et al. (4) suggested GP method showed significant discrepancy in black and Hispanic girls, and in Asian and Hispanic boys in late childhood and adolescence.

Table 6. Interclass Correlation Coefficient for Interobserver Reliability by Three Different Methods (GP, TW3, and KS)

Sex	GP	TW3	KS
All	0.999 ($p < 0.01$)	0.999 ($p < 0.01$)	1.000 ($p < 0.01$)
Boys	0.999 ($p < 0.01$)	1.000 ($p < 0.01$)	1.000 ($p < 0.01$)
Girls	0.998 ($p < 0.01$)	0.996 ($p < 0.01$)	0.999 ($p < 0.01$)

Note.— GP = Greulich-Pyle method, KS = Korean standard bone age chart, TW3 = Tanner-Whitehouse 3 method

Also Zhang et al. (5) claimed that there were ethnic and racial differences in growth patterns at certain ages, and bone age measured with GP method was significantly overestimated in Asian and Hispanic children. Asian children tend to attain ultimate height earlier than Western children (13, 14).

Yeon and Kim (6, 7) established KS derived from radiographs of the left hand of 3407 healthy Korean children in 1999. As with the GP method, bone age estimation by KS is performed by comparing the left hand-wrist radiograph of child with KS chart, so it is easy to apply and convenient to use in the out-patient setting.

Despite the passage of more than 10 years since the KS was established, there has been no comparative study between GP, TW3, and KS. However there were several studies about comparison between GP method and TW3 method globally (8-12). Kim et al. (8) investigated the difference of bone age comparing GP and TW3 method, and there was a significant correlation between the GP and TW3 methods without significant difference of intraobserver variation in the bone age assessment of Korean children. So the authors claimed both GP and TW3 methods are useful for estimation of bone age in Korean children.

In this study, we compared the reliability of these three methods (GP, TW3, and KS) in the evaluation of bone age of healthy prepubertal Korean children. The correlations between those three methods and chronological age were strong. Therefore, the GP, TW3, and KS methods are well-suited for the evaluation of the bone age of healthy Korean children. KS showed higher correlation with chronological age than GP or TW3, but the difference between them was not significant.

There are several limitations to the current study. First, the number of subjects was small and age groups including adolescents and preschool age children were not included in the study. Furthermore, while intraobserver variation is important in the study about bone age estimation, this study did not include intraobserver variability.

In conclusion, the KS, GP, and TW3 methods are useful

for evaluation of bone age in Korean children. For the KS to be used as the radiological index in the maturity status of Korean children, a multi-institutional study with a large population may be needed. As there are worldwide efforts to establish reliable computer-assisted bone age estimation, including computer-assisted skeletal age scores based on the TW3 RUS method, we should consider reinforcement and modification of the KS to make it an index of computer-assisted bone age estimation in Korean children.

REFERENCES

- Zerin JM, Hernandez RJ. Approach to skeletal maturation. *Hand Clin* 1991;7:53-62
- Greulich WW, Pyle SI. *Radiographic atlas of skeletal development of hand wrist*. Stanford: Stanford University Press, 1971
- Tanner JM, Healy MJ, Goldstein H, Cameron N. *Assessment of skeletal maturity and prediction of adult height (TW3 method)*. London: W.B Saunders Company, 2001
- Ontell FK, Ivanovic M, Ablin DS, Barlow TW. Bone age in children of diverse ethnicity. *AJR Am J Roentgenol* 1996;167:1395-1398
- 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* 2009;250:228-235
- Yeon KM, Kim IO. *Standard atlas of bone-age in Korean children*. Seoul: Sungmoonbook, 1999
- Yeon KM. Standard bone-age of infants and children in Korea. *J Korean Med Sci* 1997;12:9-16
- Kim SY, Oh YJ, Shin JY, Rhie YJ, Lee KH. Comparison of the Greulich-Pyle and Tanner Whitehouse (TW3) methods in bone age assessment. *J Korean Soc Pediatr Endocrinol* 2008;13:50-55
- Vignolo M, Milani S, DiBattista E, Naselli A, Mostert M, Aicardi G. Modified Greulich-Pyle, Tanner-Whitehouse, and Roche-Wainer-Thissen (knee) methods for skeletal age assessment in a group of Italian children and adolescents. *Eur J Pediatr* 1990;149:314-317
- Cole AJ, Webb L, Cole TJ. Bone age estimation: a comparison of methods. *Br J Radiol* 1988;61:683-686
- Milner GR, Levick RK, Kay R. Assessment of bone age: a comparison of the Greulich and Pyle, and the Tanner and Whitehouse methods. *Clin Radiol* 1986;37:119-121
- Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. Bone age assessment: a large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods. *Arch Dis Child* 1999;81:172-173
- Ashizawa K, Asami T, Anzo M, Matsuo N, Matsuoka H, Murata M, et al. Standard RUS skeletal maturation of Tokyo children. *Ann Hum Biol* 1996;23:457-469
- Zhen OY, Baolin L. Skeletal maturity of the hand and wrist in Chinese school children in Harbin assessed by the TW2 method. *Ann Hum Biol* 1986;13:183-187