การประเมินอายุจากภาพรังสีมือและข้อมือในกลุ่มคนไทย

นายสรรพัชญ เบญจวงศ์กุลชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชารังสีวิทยาช่องปากและแม็กซิลโลเฟเชียล ภาควิชารังสีวิทยา
คณะทันตแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559

AGE ESTIMATION METHODS USING HAND AND WRIST RADIOGRAPHS IN A GROUP OF THAI

Mr. Sunpatch Benjavongkulchai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Oral and Maxillofacial Radiology
Department of Radiology
Faculty of Dentistry
Chulalongkorn University
Academic Year 2016

Thesis Title	AGE ESTIMATION METHODS USING HAND AND WRIST RADIOGRAPHS IN A GROUP OF THAI
By	Mr. Sunpatch Benjavongkulchai
Field of Study	Oral and Maxillofacial Radiology
Thesis Advisor	Pisha Pittayapat, D.D.S., Ph.D.

Accepted by the Faculty of Dentistry, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

....Dean of the Faculty of Dentistry

(Assistant Professor Suchit Poolthong, D.D.S., Ph.D.)

THESIS COMMITTEE

....Chairman

(Associate Professor Soontra Panmekiate, D.D.S., Ph.D.)

....Thesis Advisor

(Pisha Pittayapat, D.D.S., Ph.D.)

....External Examiner

(Associate Professor Sangsom Prapayasatok, D.D.S.)

สรรพัชญ เบญจวงศ์กุลชัย : การประเมินอายุจากภาพรังสีมือและข้อมือในกลุ่มคนไทย (AGE ESTIMATION METHODS USING HAND AND WRIST RADIOGRAPHS IN A GROUP OF THAI) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อ. ทญ. ดร. พิสชา พิทยพัฒน์, 87 หน้า.

บทนำ: การประเมินอายุถูกนำมาใช้ในหลายกรณี เช่น การติดตามการเจริญเติบโต การลงทะเบียนผู้อพยพ การตัดสินคดีทางกฎหมาย และการระบุเอกลักษณ์บุคคล โดยอายุตามปฏิทินจะถูกนำมาเทียบเคียงกับระดับการเจริญเติบโตของกระดูก ซึ่งการประเมินอายุจากภาพรังสีมือและข้อมือเป็นวิธีที่ใช้กันมาก เนื่องจากเป็นวิธีที่ไม่ซับซ้อน ไม่แพง และไม่ล่วงล้ำ ทั้งนี้ ความถูกต้องของผลการประเมินอายุยังคงเป็นข้อสงสัย เนื่องมาจากประเด็นหลักเกี่ยวกับความแตกต่างของชาติพันธุ์

วัตถุประสงค์: เป้าหมายหลักของการศึกษานี้เพื่อประเมินความถูกต้องและความน่าเชื่อถือของวิธีการประเมินอายุที่ใช้บ่อยในกลุ่มคนไทย และมีเป้าหมายรองในการเปรียบเทียบระหว่างผลที่ได้จากแต่ละวิธีเพื่อหาความสัมพันธ์กับกลุ่มคนไทย

วิธีดำเนินการ: รวบรวมภาพรังสีมือและข้อมือจากคลินิกรังสีวิทยา คณะทันตแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย คัดเลือกเฉพาะภาพรังสีที่ถ่ายระหว่างปี พ.ศ. 2554 – 2559 และผู้ป่วยไม่มีประวัติโรคประจำตัวที่มีผลต่อการพัฒนากระดูก จากนั้นทำการประเมินอายุจากภาพรังสีด้วยวิธีของ ‘Greulich and Pyle’ ‘Tanner and Whitehouse (TW3-RUS)’ และ ‘Fishman’ ผลที่ได้จากการประเมินแต่ละวิธีและอายุตามปฏิทินของผู้ป่วยจะถูกนำมาเปรียบเทียบโดยใช้ Friedman’s rank test ($p<0.05$) ตามด้วย Wilcoxon signed ranks test ร่วมกับ Bonferroni correction

ผลการศึกษาและบทอภิปราย: อายุที่ได้จากการประเมินทั้ง 3 วิธี มีความแตกต่างจากอายุตามปฏิทินอย่างมีนัยสำคัญ ($p<0.002$) ยกเว้นวิธีของ ‘Tanner and Whitehouse’ ในผู้ชาย ซึ่งไม่มีความแตกต่างอย่างมีนัยสำคัญ โดยภาพรวมผลการประเมินแต่ละวิธีมีแนวโน้มที่จะมากกว่าอายุจริง แม้ว่าจะพบแนวโน้มที่จะต่ำกว่าอายุจริงบ้างในบางกลุ่มที่มีอายุน้อย ซึ่งความคลาดเคลื่อนที่พบนี้เป็นข้อสนับสนุนถึงปัจจัยทางชาติพันธุ์ และความผันเปลี่ยนที่เกิดขึ้นตามกาลเวลา

บทสรุป: วิธีการของ ‘Tanner and Whitehouse’ TW3-RUS สามารถประเมินอายุได้ถูกต้องกว่าวิธีอื่นเมื่อใช้กับกลุ่มคนไทยร่วมสมัย อย่างไรก็ดีการปรับปรุงวิธีการประเมินเพิ่มเติมเพื่อลดความคลาดเคลื่อนในกลุ่มวัยรุ่นตอนปลายก็เป็นที่แนะนำ

ภาควิชา	รังสีวิทยา	ลายมือชื่อนิสิต ....
สาขาวิชา	รังสีวิทยาช่องปากและแม็กซิลโลเฟเชียล	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
ปีการศึกษา	2559	

5875839832 : MAJOR ORAL AND MAXILLOFACIAL RADIOLOGY
KEYWORDS: AGE ESTIMATION / HAND / WRIST / RADIOGRAPH / THAI

SUNPATCH BENJAVONGKULCHAI: AGE ESTIMATION METHODS USING HAND AND WRIST RADIOGRAPHS IN A GROUP OF THAI. ADVISOR: PISHA PITTAYAPAT, D.D.S, Ph.D., 87 pp.

Introduction: Age estimation takes part in a variety of situations, such as growth observation, immigrant registration, legal penalty judgment and body identification. Chronological age usually corresponds with stages of skeletal and dental development. Age estimation by using hand and wrist radiography is the first method of choice for many cases because it is uncomplicated, inexpensive and non-invasive. However, the validity of the result is still questionable mainly because of the influence of ethnic differences.

Objectives: The primary aim of this study was to evaluate the accuracy and reliability of the commonly used age estimation methods on Thai population. The secondary aims were to compare the results between different methods and to find any possible relationship on Thai population.

Materials and methods: Hand and wrist radiographs from the Department of Radiology, Faculty of Dentistry, Chulalongkorn University were collected. The radiographs were taken between 2011 – 2016 and the patients must not have history of diseases that affect skeletal development. 'Greulich and Pyle', 'Tanner and Whitehouse (TW3-RUS)', and 'Fishman' method were applied for each radiograph. The results from each method and the chronological age were compared using Friedman's rank test ($p<0.05$) followed by a Wilcoxon signed ranks test with Bonferroni correction.

Results and Discussion: The estimated ages from the three methods were significantly different from the chronological age ($p<0.002$), except for 'Tanner and Whitehouse' in male which showed no significant difference. Overall, the results from each method tended to be overestimated although the tendency of underestimation was found particularly in younger groups. The presence of discordance supported the influence of ethnical factor and secular change on skeletal maturation rate.

Conclusions: Tanner and Whitehouse TW3-RUS method was more accurate in predicting the age of contemporary Thais; however, a further adaptation of the method to reduce the discrepancies in late teenagers is recommended.

Department:	Radiology	Student's Signature	....
Field of Study:	Oral and Maxillofacial Radiology	Advisor's Signature	....
Academic Year:	2016		

ACKNOWLEDGEMENTS

The Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72nd anniversary of his Majesty King Bhumibala Aduladeja is gratefully acknowledged.

CONTENTS

Page

Page

Chapter 1 Introduction

Background

Age estimation is an essential method taking part in a variety of situations, such as growth observation, immigrant registration, legal penalties judgment and body identification. In medicine, pediatricians collected data to find norms and, in consequence, proposed age estimation methods. The methods are used to represent anatomical developmental status of a child and compared with his or her chronological (calendar) age to assess whether any developmental problem is present or not (Berk, 2012, Black and Ferguson, 2016, Kliegman et al, 2015). In case of immigrants and legal judgment, it is essential to define the age of each individual since cares and penalties are differently sentenced based on levels of maturity (Aggrawal, 2009, Franklin et al, 2015, Smith and Brownlees, 2011). In Thailand, criminals who are between 15 and 18 years old will be sentenced for only half of the adult's penalty, while ones who are between 18 and 20 years old will have 1/3 reduced penalty (Thailand, criminal code, article 75-76). In human identification and missing persons, predicting the age of the victims helps narrow down the search for possible identity of the victims (Black and Ferguson, 2016, Sharma, 2007).

Age estimation can be achieved by evaluating several aspects of an individual, thus it may be classified into 4 major groups: medical, physical, psycho-social, and documentation (Smith and Brownlees, 2011). In the medical aspect, chronological age corresponds with stages of skeletal and dental development. Many bones in the whole body can be observed as an indicator such as, hand, wrist, clavicle, scapular, and hip. Eruption and morphological development of teeth are also considered as indicators. Physical age estimation focuses on pubertal growth and anthropometric measurements without radiographs, such as, weight and height. Psycho-social aspect relies on family history, social activities, and reaction to particular situations. Documentation helps to predict the age by giving records or information about the subject to the date of birth. Using every aspect is being suggested (Smith and Brownlees, 2011), but in many cases, this would rather be impossible, especially for deceased persons.

Hand and wrist radiography, one of the methods in medical aspect, is useful for age estimation. Since it is uncomplicated, inexpensive and non-invasive, this technique is the first choice for many cases (Hackman, 2012, Smith and Brownlees, 2011). However, the validity of the result is still questionable (Buken et al, 2007, Buken et al, 2009, Buken et al, 2010, de Sousa Dantas et al, 2015, Dharman.M.K et al, 2014, Haiter-Neto et al, 2006, Mohammed et al, 2014a, Zhang et al, 2009). The reason is that the common age estimation methods are based on the United States and European population (Buken et al, 2009, Fishman, 1982, Fishman, 1987, Franklin et al, 2015, Greulich and Pyle, 1959, Haiter-Neto et al, 2006, Tanner et al, 1975) which the trait and proportion of ethnicities coincide with just some parts of the whole world population, also different from Thai. Therefore, the deviation between the result and the real age may be greater and thus, make it significant in some country such as in Eastern world (Black and Ferguson, 2016, Manzoor Mughal et al, 2014). Moreover, these methods were invented many decades ago or based on old data which belonged to the generations different from the present days. Some novel studies showed a trend of changes in the representative chronological age for each stages of growth, generation to generation (Berk, 2012, Franklin et al, 2015, Smith and Brownlees, 2011). Therefore, the first aim of this present study is to evaluate the suitability of the commonly used age estimation methods on Thai population. The second aim is to compare the age prediction results between different methods on Thai population. Finally, the results of this study will be compared with other studies to find any possible relationship.

Chapter 2 Review of literature

Hand and wrist

Human hand and wrist consists of 27 bones for each side of the body (Figure 1). The 19 bones of one hand can be counted into 5 metacarpuses, 5 proximal phalanges, 4 middle phalanges (absent in thumb finger), and 5 distal phalanges. The rest 8 bones, called carpal bones, belong to the wrist and are defined as capitate, hamate, pisiform, triquetrum, lunate, scaphoid, trapezium, and trapezoid (Figure 2). There is an exceptionally calcified mass found on the thumb called a sesamoid bone. Radius bone and ulnar bone are adjacent to the wrist and found on hand and wrist radiographs thus, are also used as developmental indicators. Metacarpuses and phalanges are developed in the twelfth week of pregnancy, thus, are present at birth. The first carpal bone present in this region is capitate which can be found in 2-month newborns. The last bone appear is the sesamoid bone which can be found in 10-12 year-old children (Kliegman et al, 2015, Schoenwolf et al, 2014). However, changes in bone morphology still continue even after every bone is already present (Figure 3).

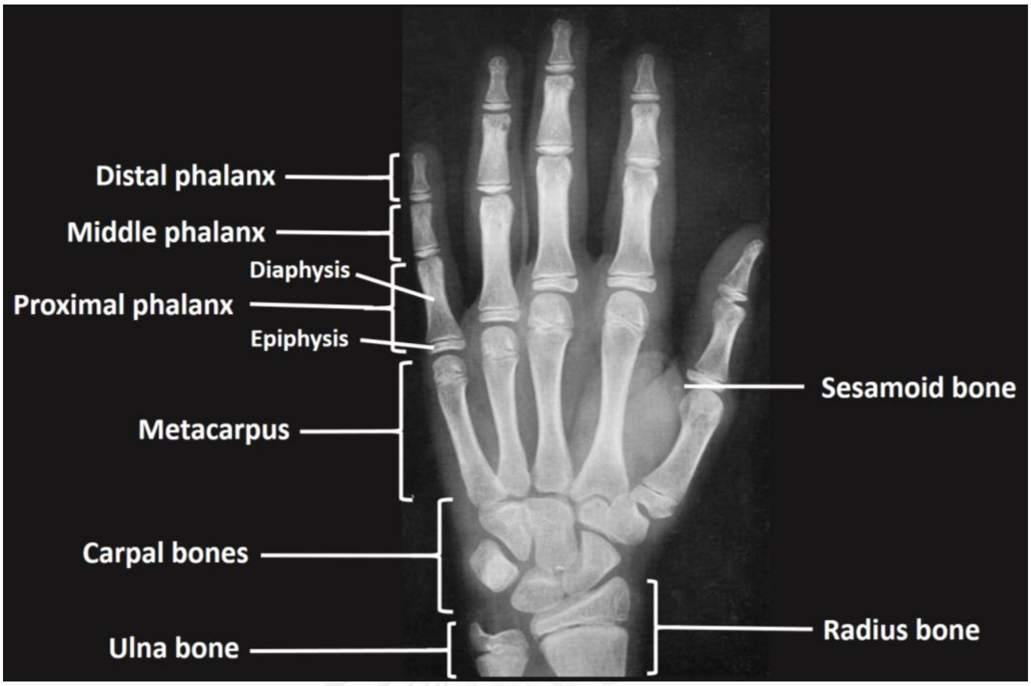


Figure 1 Hand and wrist bones of a 13-year-old boy labeled with name of the bones (Greulich and Pyle, 1959).

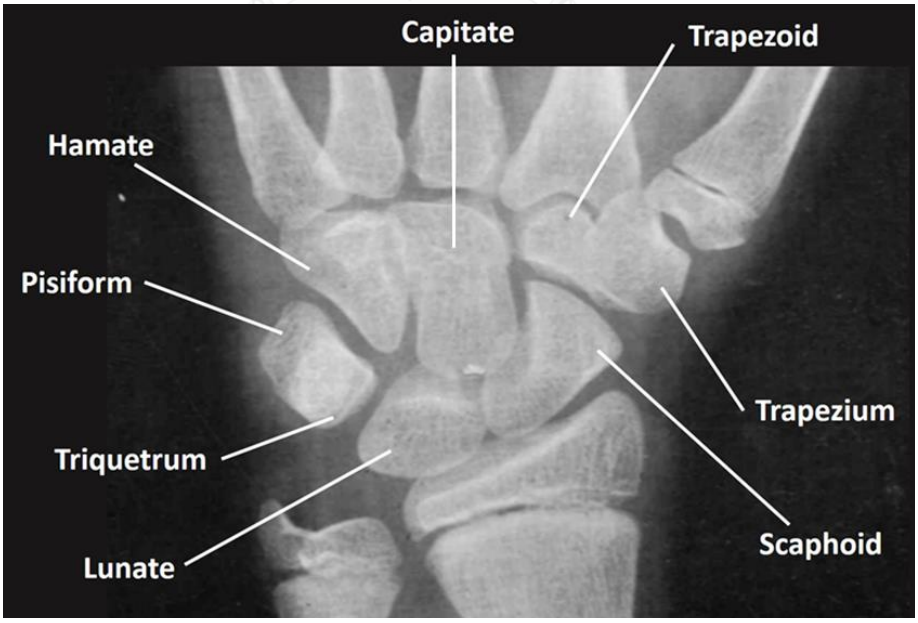


Figure 2 Carpal bones of a 13-year-old boy labeled with name of the bones (Greulich and Pyle, 1959).

A long bone (such as phalange, metacarpus and radius bone) is composed of two substructures called diaphysis and epiphysis. Diaphysis is the long thin central compartment which presents before birth, while epiphysis is the terminal head compartment which presents later after birth. During long bone development, the diaphysis gains greater length while the epiphysis changes its appearance over the course of development. The stages of epiphyseal development are presenting, widening, capping (cover) at the end of diaphysis, and fusing of the epiphysis with the diaphysis (Schoenwolf et al, 2014). Therefore, obvious staging can be found on long bone development. Unlike long bones, other bones like carpal bones change their shape during development too, but the staging is not so empirical.

The last change found in hand and wrist region is a complete fusion of distal epiphysis with diaphysis of radius bone at the age of 17 years in female and 19 years in male (Greulich and Pyle, 1959). Therefore, hand and wrist cover almost 20 years of human development, from the time of newborn to the end of teenager.

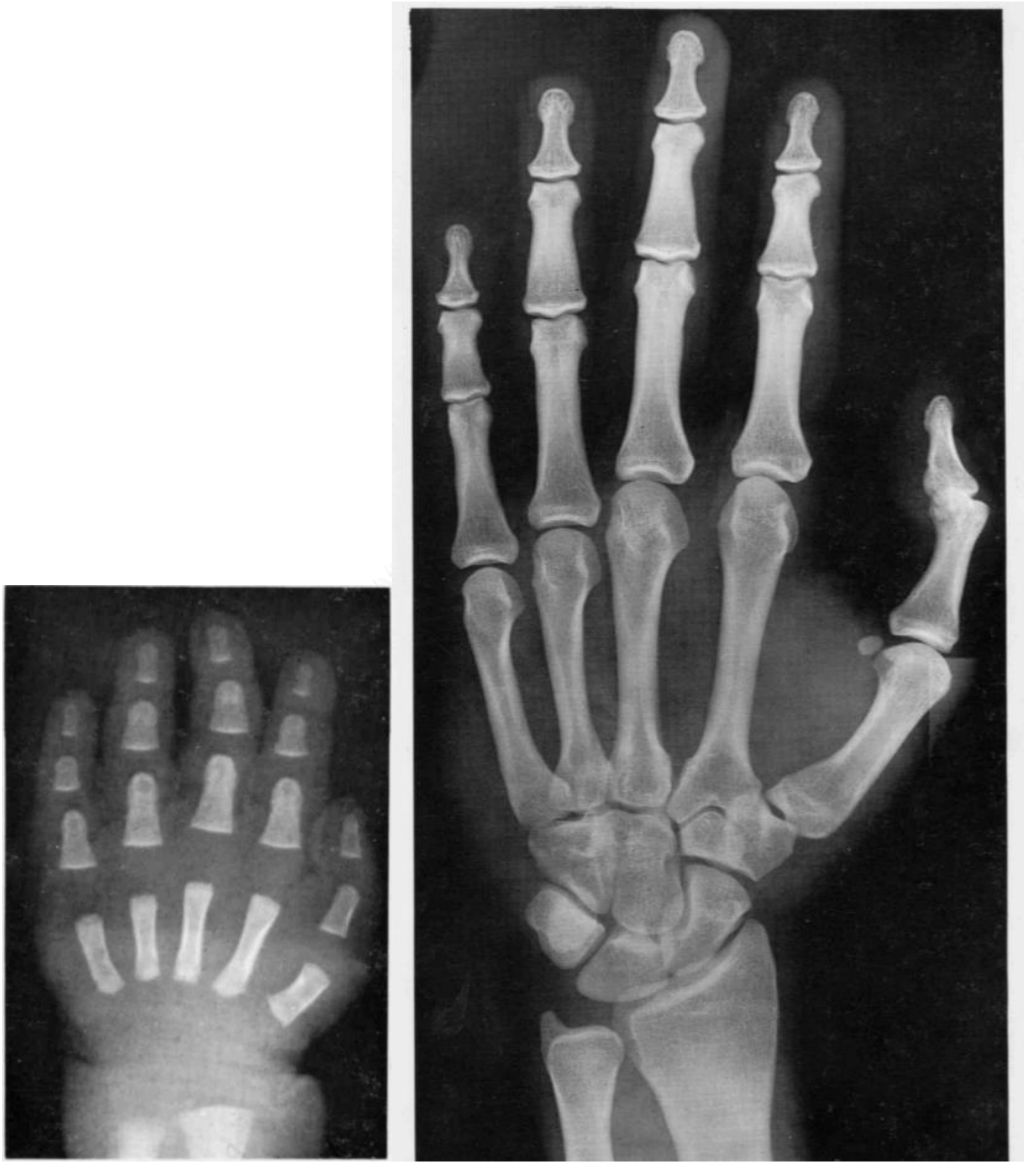


Figure 3 Comparison of hand and wrist bones of a newborn boy with a 19-year-old boy, showing changes in number and shape of the bones (Greulich and Pyle, 1959).

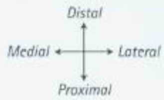
Age estimation using hand and wrist radiographs

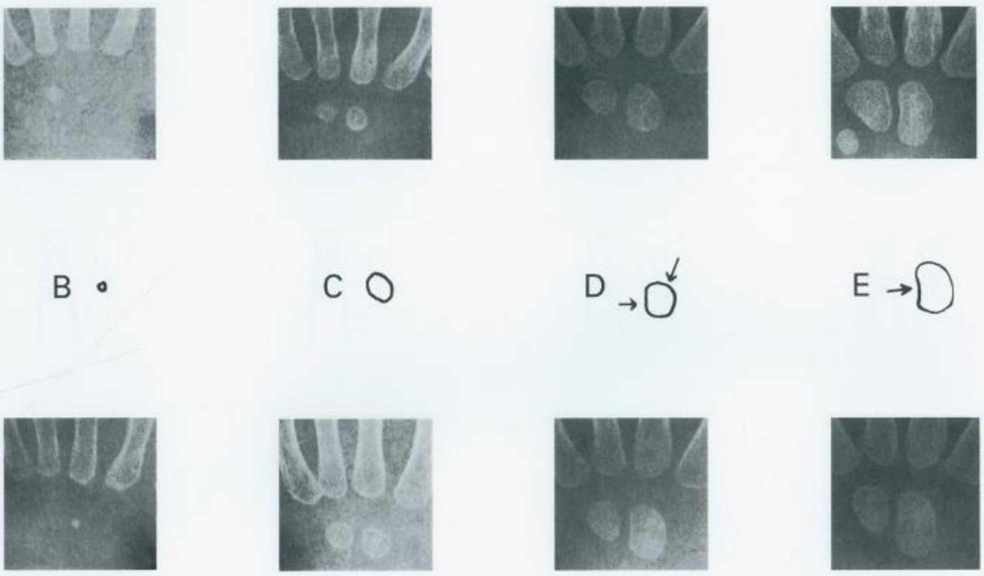
Many age estimation methods using hand and wrist radiographs have been proposed. Each method has its own pros and cons relying on which of the main concept it belongs. In general, there are 3 main concepts that are commonly used: visual comparison, numerical system and dichotomous tracing.

The most well-known visual comparison method was introduced by Greulich and Pyle in 1959 (Greulich and Pyle, 1959). It is the most popular age estimation method for hand and wrist radiographs using its own atlas as a key (Franklin et al, 2015, Manzoor Mughal et al, 2014, Satoh, 2015, Zhang et al, 2009). In this method, hand and wrist radiographs were taken from children with known definite chronological age. The best representative radiographs for each age were selected and collected into the "Radiographic Atlas of Skeletal Development of the Hand and Wrist". Therefore, if there is an adolescent or child who has unknown real age, his or her hand and wrist radiograph can be taken and then a comparison with the atlas to find the matching image can be done. The age of that individual will be concluded to be similar to the age described on that atlas picture. Moreover, the atlas also described the possible deviation for each age (Greulich and Pyle, 1959). The advantage of this method is that it does not require much time to achieve the result because the outlying images can be ruled out very early in the process, leaving only fewer images for making decision. However, the atlas does not cover cases with variations, some bones may develop faster or slower, and that makes the result sometimes questionable (Acheson, 1954, Black and Ferguson, 2016, Manzoor Mughal et al, 2014).

Since the concept of image comparison had been criticized, the concept of numerical system using sum score table was raised (Acheson, 1954). Several methods were proposed, but the most famous one is Tanner and Whitehouse method (Franklin et al, 2015, Satoh, 2015, Smith and Brownlees, 2011). This method refers to stages of skeletal growth focusing on 20 regions on hand and wrist bones. Each stage of each region is represented by a number (Figure 4). The numbers corresponding with the present bone stage from all regions are then summed together and compared with the sum score table correlated with the chronological age (Figure 5) (Tanner et al, 2001, Tanner et al, 1975). This method has been introduced in 3 editions called “TW1”, “TW2” and “TW3”. The most popular edition is the second one (TW2) which is based on the same data as the first edition (TW1) and uses the same sum score table, but has some changes in the staging description. Both TW1 and TW2 composes of 3 scoring systems; “radius, ulna, and selected metacarpal and phalanges (RUS) score” relying on 13 bones (radius, ulna, and 1^{st}, 3^{rd}, 5^{th} metacarpuses and phalanges), “carpal (CAR) score” relying on 7 carpal bones (capitate, hamate, triquetrum, lunate, scaphoid, trapezium and trapezoid, excepting pisiform bone), "20-bone score" (composition of RUS and CAR score) (Hackman, 2012, Tanner et al, 1975). The latest edition (TW3) was published in 2001. TW3 uses the new data which covers more varieties of ethnicity resulting in the new sum score table and excludes "20-bone score" (Hackman, 2012, Tanner et al, 2001). As Tanner and Whitehouse methods rely on each region of the bones separately, a definite result can be made even in case of variation. The drawback is that they consume so much time to rate, calculate and find the correlated age (Black and Ferguson, 2016, Manzoor Mughal et al, 2014, Satoh, 2015).

Capitate





BOYS' SCORES		GIRLS' SCORES
Carp 100	**Stage B** (i) The centre is just visible as a single deposit of calcium, or more rarely as multiple deposits. The border is ill-defined.	Carp 84
Carp 104	**Stage C** (i) The centre is distinct in appearance and oval in shape, with a smooth, continuous border.	Carp 88
Carp 106	**Stage D** (i) The maximum diameter is half or more the width of the radial metaphysis. (ii) The border adjacent to the hamate is now flat or only slightly convex. (iii) The border adjacent to the second metacarpal is also beginning to become distinct from the hamate border so that the centre now appears somewhat D-shaped.	Carp 91
Carp 113	**Stage E** (i) The hamate border is now concave and slightly thickened. (ii) The bone has lengthened, so that the longitudinal diameter is distinctly greater than the transverse, but less than the distance from its proximal border to the radial metaphysis.	Carp 99

Figure 4 Developmental stages of the capitate bone by pictures and radiographs. Each stage is matched to a score by sex (Tanner et al, 2001).

Maturity score	Bone Age ('yr')	Maturity score	Bone Age ('yr')	Maturity score	Bone Age ('yr')
42	2.0	214	7.0	427	12.0
46	2.1	216	7.1	434	12.1
50	2.2	219	7.2	441	12.2
55	2.3	222	7.3	448	12.3
60	2.4	225	7.4	455	12.4
66	2.5	228	7.5	462	12.5
70	2.6	231	7.6	470	12.6
75	2.7	234	7.7	478	12.7
80	2.8	237	7.8	485	12.8
86	2.9	240	7.9	493	12.9
91	3.0	243	8.0	501	13.0
94	3.1	246	8.1	511	13.1
98	3.2	250	8.2	520	13.2
101	3.3	253	8.3	530	13.3
105	3.4	256	8.4	540	13.4
108	3.5	259	8.5	550	13.5
112	3.6	262	8.6	560	13.6
116	3.7	265	8.7	570	13.7
120	3.8	268	8.8	581	13.8
124	3.9	272	8.9	592	13.9
129	4.0	275	9.0	603	14.0
132	4.1	279	9.1	615	14.1
134	4.2	283	9.2	628	14.2
137	4.3	287	9.3	641	14.3
140	4.4	291	9.4	655	14.4
143	4.5	295	9.5	668	14.5
146	4.6	299	9.6	682	14.6
149	4.7	303	9.7	697	14.7
152	4.8	308	9.8	711	14.8
155	4.9	312	9.9	726	14.9
158	5.0	316	10.0	741	15.0
161	5.1	321	10.1	755	15.1
164	5.2	325	10.2	769	15.2
166	5.3	330	10.3	783	15.3
169	5.4	334	10.4	798	15.4
172	5.5	339	10.5	813	15.5
175	5.6	344	10.6	828	15.6
177	5.7	348	10.7	843	15.7
180	5.8	353	10.8	859	15.8
183	5.9	358	10.9	875	15.9
186	6.0	363	11.0	891	16.0
189	6.1	369	11.1	912	16.1
191	6.2	375	11.2	933	16.2
194	6.3	381	11.3	955	16.3
197	6.4	387	11.4	977	16.4
200	6.5	394	11.5	1000	16.5
202	6.6	400	11.6		
205	6.7	406	11.7		
208	6.8	413	11.8		
211	6.9	420	11.9		

Figure 5 RUS sum score table for boys comparing sum score and its corresponding age (Tanner et al, 2001).

King and colleagues analyzed the reproducibility of Tanner and Whitehouse method (TW2) and Greulich and Pyle method using 50 radiographs and 3 observers (King et al, 1994). The results showed inter-observer error 0.74 year and 0.96 year for TW2 and 'Greulich and Pyle' respectively, but were not statistically significant. They also reported that the average time required for TW2 20-bone score was 7.9 minutes, while Greulich and Pyle method consumed just 1.4 minutes in average (King et al, 1994).

The third concept of hand and wrist age estimation methods, using a dichotomous system, was developed by Fishman in 1982 (Fishman, 1982). This method combines staging of the bones and expresses them as 11 signs of skeletal maturity indicators (SMIs), in 6 regions of interest (Figure 6). The 11 SMIs are matched with their respective age in chronological order. For applicable methodology, the SMIs are systemized into a dichotomous observation system. The SMIs will be observed by priority in the system which means that a positive result of the first SMI will lead to the next one while a negative result will lead to the other one. The final positive SMI will indicate the resulted age (Figure 7, 8) (Fishman, 1982, Fishman, 1987). Fishman first published the idea in 1982, at that time the number of subjects was at 1168 (Fishman, 1982). Later in 1987, Fishman published a re-evaluated study using more than 4000 subjects and had some changes in the results (Fishman, 1987). Because the points of interest are limited to 11 signs and cut off by dichotomous division, this method seems to be a by-pass to the result and should be a time saving method. The downside of this method is that the results corresponding with the signs are also limited to 11.

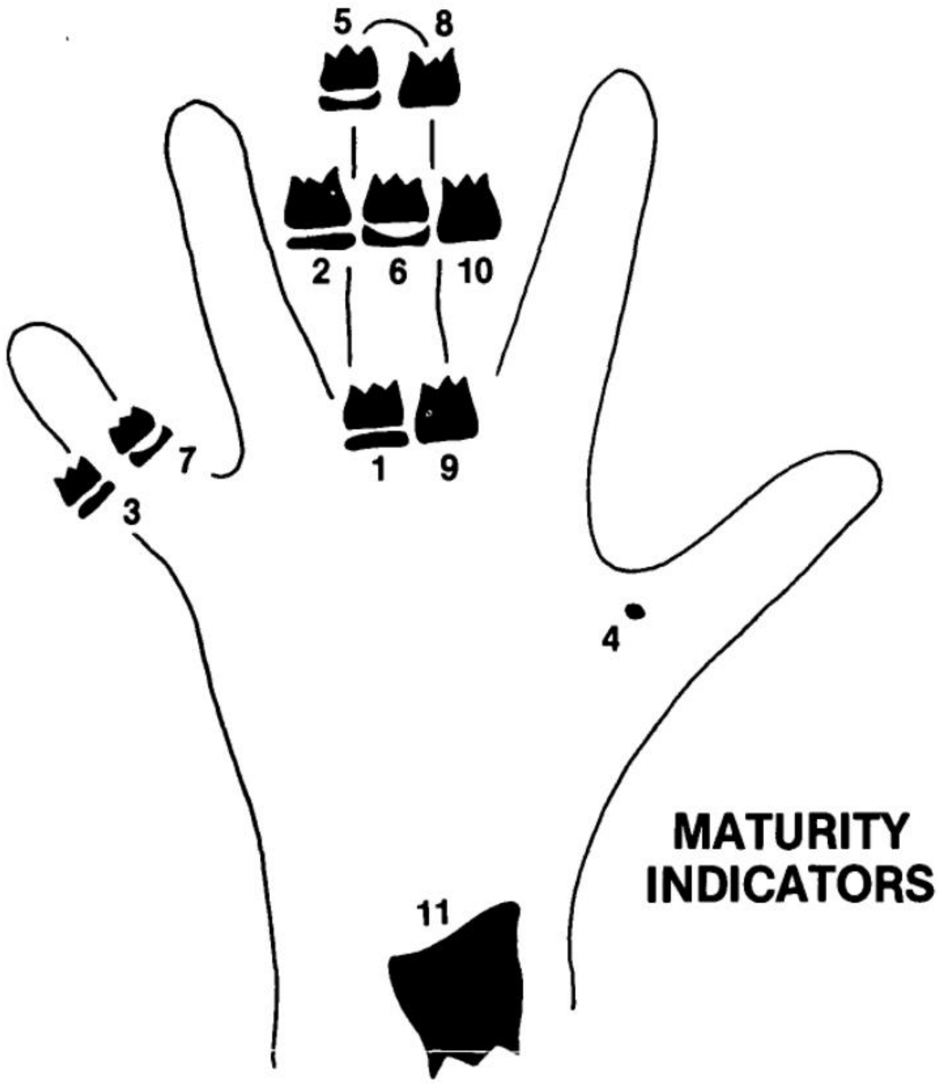


Figure 6 The 6 regions of interest for the 11 observational signs as skeletal maturity indicators (Fishman, 1982).

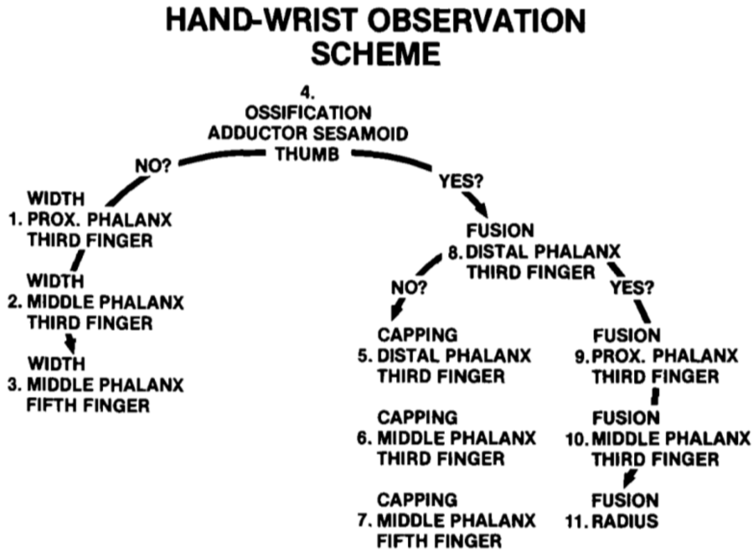


Figure 7 The "hand-wrist observation scheme" was designed for more applicable method based on dichotomous observation system. Ossification of the sesamoid bone (SMI 4) is the first SMI to be considered - if absent, width of the proximal phalanx of the third finger (SMI 1) will be the next SMI. However, if present, fusion of the distal phalanx of the third finger (SMI 8) will be the next SMI (Fishman, 1982).

Average Age at each SMI for Early, Average and Late Maturers FEMALE

SMI	Early	Average	Late
1	9.157	10.230	11.302
2	9.800	10.722	11.643
3	9.914	10.872	11.831
4	10.072	11.041	12.011
5	10.549	11.708	12.867
6	10.802	11.955	13.108
7	11.369	12.512	13.655
8	12.260	13.263	14.265
9	12.636	13.967	15.298
10	13.626	14.890	16.154
11	14.791	16.508	18.225

Average Age at each SMI for Early, Average and Late Maturers MALE

SMI	Early	Average	Late
1	10.248	11.368	12.488
2	10.955	11.921	12.887
3	11.109	12.211	13.312
4	11.351	12.574	13.797
5	11.835	13.029	14.224
6	12.571	13.779	14.987
7	13.230	14.429	15.629
8	14.185	15.194	16.203
9	14.289	15.485	16.682
10	14.958	16.174	17.390
11	15.815	17.242	18.669

Figure 8 The tables comparing SMI numbers and average ages which are defined as the results for female and male (Fishman, 1987).

No matter which method or which concept is used for age estimation from hand and wrist radiographs, there are still some factors influencing the predicted age. The factors that should be taken into account are genetic variations, underlying disorders, generation differences and observer dependence (Hackman, 2012, Smith and Brownlees, 2011).

Genetic variations

Genetic variations affect the progress on physiological development, including skeletal development, correlating with ages (Berk, 2012, Hackman, 2012). At least 2 aspects must be taken into account: sex and ethnicity. By sex, many previous studies found that females usually grow faster than males (Berk, 2012, Beunen et al, 2006, Black and Ferguson, 2016, Bogin and MacVean, 1983, Cole et al, 2015, Fitzpatrick, 2004, Flory, 1935, Flory, 1936, Molinari et al, 2004). By ethnicity, there are many studies finding differences in timing of growth spurt and rate of skeletal growth between ethnic groups (Black and Ferguson, 2016, Buken et al, 2007, Buken et al, 2010, Cole et al, 2015, Hackman, 2012, Manzoor Mughal et al, 2014, Smith and Brownlees, 2011, Zhang et al, 2009).

Flory mentioned sex differences in skeletal growth based on hand and wrist radiographs from over 6000 children in Chicago (Flory, 1935). He stated that females had earlier skeletal development than males by evaluating 5 aspects. The first aspect was the difference at birth. He found that the incidence of presence of carpal bone ossification center at birth was 8 % in girls while only 2 % in boys, thus girls were more likely to have early development. Second, by qualitative difference, he found that female standards were younger than male standards with similar radiographic appearance. As epiphyseal closure was used as the maturity indicator, he found that 50% of female subjects were 17 years old while 50% of male subjects were 18.5 years old. Third, by difference in bone growth, there was a crossing point among growth charts of scaphoid, lunate and triquetrum bones which could be used as a reference point and he found that this point correlated with 8.5 years in girls while 10 years in boys. Fourth, by ossification ratios, he found that the ratio between area of ossification and area of carpal quadrilateral reached 1.0 before 12 years old in girls, and before 14 years old in boys. Fifth, the author evaluated the variability in skeletal development. The variability in the population was greater by increasing age until the maturity stage began. The variability started to decrease after the maturity stage began in the population. The peak of the variability was marked. He found that females were at most variable at 12 years old while males were at 14 years old. The female population reached zero variability at 19 years old while male still did not reach the zero point at 20 years old (Flory, 1935). Flory also published a review of the sex difference in skeletal growth (Flory, 1936).

Bogin and MacVean studied on factors influencing skeletal growth, stature, and cognitive status in Guatemalan children (Bogin and MacVean, 1983). The data included socio-economic status, sex, general intelligence and reading ability. They found that socio-economic status effects mainly on intellectual ability with minor effects on physical development, which higher socio-economic status results in greater developments. However, sex was shown to be strongly related to difference in skeletal maturity, which girls were more advance than boys in the same age group (Bogin and MacVean, 1983).

Fitzpatrick stated that sex influences skeletal growth by roles of sex hormones: estrogen and androgen (Fitzpatrick, 2004). Estrogen tends to play a major role controlling bone growth in both male and female through functions of osteoblasts which take action in bone formation. Estrogen induces osteoblast cell division and proliferation, and also increases expression of alkaline phosphatase and type I collagen which are contribute to bone development. Androgen hormones, such as dihydrotestosterone and dehydroepiandrosterone, influence skeletal growth by increase osteoblast proliferation and expression of bone regulating cytokines and growth factors although the effect seems weaker than estrogen. Due to difference in sex hormones between male and female, different skeletal developmental rate is evident (Fitzpatrick, 2004).

Cole and colleagues compared bone age between Caucasian and African residence in Johannesburg, South Africa, using Tanner and Whitehouse (TW3) RUS score (Cole et al, 2015). Relying on the advantage of scoring system, a growth acceleration stage can be recognized by a great increase of RUS score between adjacent chronological age groups. They found that Caucasian male reached growth acceleration stage earlier than African male with the same acceleration rate, while Caucasian female reached the acceleration stage a little earlier but with lower acceleration rate than African female. Moreover, the study showed that girls reached skeletal maturation 1.9 years faster than boys (Cole et al, 2015).

Buken and colleagues have made many studies to prove the availability of age estimation methods in Turkish population (Buken et al, 2007, Buken et al, 2009, Buken et al, 2010). First they studied with Greulich and Pyle method and found significant overestimation 0.17 - 1.1 years in 11, 12, 14, and 16-year-old girl with underestimation 0.43 years in 18-year-old girl (Buken et al, 2007). Using the same method with Turkish boys, they found significant underestimation 0.58 years in 13-year-old group, overestimation 0.88 – 0.98 years in 15 – 17-year-old group, and underestimation 0.48 years in 19-year-old group (Buken et al, 2007). Later they studied with Tanner and Whitehouse method (TW3), and found significant underestimation 0.44 and 0.43 years in 13 and 14 years old Turkish boy respectively (Buken et al, 2010). Another study was done to compare applicability of these prior methods and Gök method (referred as the most preference method by Turkish pediatricians) and found either overestimation or underestimation by each method for several age groups (Buken et al, 2009).

Zhang and colleagues studied the reliability of Greulich and Pyle method in Los Angelis residence with an intention to create a new digital hand atlas with racial variation (Zhang et al, 2009). From the study, they found that residence with either Asian or

Hispanic origin matured sooner than whom with African or Caucasian origin. The results showed 0.24 years and 0.41 years overestimation in girl and boy residence with Asian origin respectively, and 0.24 years and 0.30 years overestimation in girl and boy residence with Hispanic origin, respectively (Zhang et al, 2009).

Underlying diseases

Diseases and anomalies can disturb physiological development. Some diseases affect absorption and utilization of nutrients essential for bone formation (Bacchetta et al, 2012, Hackman, 2012, Pludowski et al, 2009, Zemel et al, 2007).

Zemel and colleagues conducted a longitudinal study on children aged 0-18 years old who affected with sickle cell anemia (Zemel et al, 2007). The subjects were evaluated annually for 4 years. Measurements were made for stature, weight, skinfold thickness, Tanner's pubertal stage, TW3 bone age, and hematologic statuses. In the end of the study, they reported lack of BMI and weight in patients with sickle cell anemia, also having 1-2 years hindered growth and pubertal on set. They also found that in older age group, boys had greater delay in skeletal development. Moreover, comparison between patients with and without blood transfusion was made and the results showed that subjects with blood transfusion showed less severity in bone development retardation, although delays in pubertal onset were not significantly different (Zemel et al, 2007).

Pludowski and colleagues found that primary hypertension, hypertension originated in childhood, associated with accelerated maturation (Pludowski et al, 2009). They estimated the bone age of 108 Polish patients, 54 with primary hypertension and 54 without hypertension, based on Polish hand and wrist radiographic atlas. The differences between bone age and chronological age from the hypertension patients were compared with their normal counterpart patients. The mean difference between bone age and chronological age (bone age – chronological age) in normal group was 0.45 years, in prehypertension was 1.59 years, in stage 1 hypertension was 1.77 years, and in stage 2 hypertension was 2.38 years. This could be inferred that the degree of acceleration was higher due to the severity of primary hypertension (Pludowski et al, 2009).

Bacchetta and colleagues had written a review on consequences of chronic kidney disease on children growth and development (Bacchetta et al, 2012). Effects of chronic kidney disease on bone and mineralization consist of 3 main features: abnormal calcium, phosphorus, hormones and vitamin D metabolism; abnormal bone growth and quality; and soft tissue calcification. The mechanisms including altered growth hormone metabolism, vitamin D (essential for calcium absorption) deficiency, hypoparathyroidism, hypogonadism, and malnutrition lead to growth retardation (Bacchetta et al, 2012).

Genetic disorders related to DNA transcription and translation provide biochemical substances which control whole body development including bones. For example, Turner's syndrome results in delayed and abnormal pubertal stage related with growth spurt onset. Because a pubertal onset relates to growth acceleration stage, hormonal deficiency such as gonadotrophin deficiency and growth hormone deficiency/surplus also affect pubertal maturation and bone development (Satoh, 2015, Stanhope and Fry, 2000).

Therefore, patients affected by these disorders would have delayed or advanced growth rate, also with organ malformation or dysfunction. In addition, chronic diseases can hinder the development since the body utilizes nutrition rather for regeneration than further growth (Hackman, 2012).

Hewitt, Westropp and Acheson wrote a report about effects of illness on skeleton development from Oxford Child Health Survey (Hewitt et al, 1955). They collected data about yearly height gains, illnesses, radiological striation of long bones (radiographic lines appearing after a long bone stops development, also called 'Harris lines'), and skeletal maturity progressions. The recorded illnesses included exanthema, respiratory tract infection, skin diseases and other infection. They found that subjects affected by illnesses had slightly reduced skeletal maturity. However, it was not statistically significant. They stated that illnesses influenced mainly on the growth of bones, by the meaning of size, while the time of maturation was still the same thus, resulted in short stature in the end of development (Hewitt et al, 1955).

Generation differences

Generation differences also affect the accuracy of age estimation. Studies on recent generation showed age overestimation using the long time-practiced age estimation methods. A hypothesis for this issue is that the reference populations are from different generation (Franklin et al, 2015, Halvorsen and Kristensen, 2003, Satoh, 2015, Smith and Brownlees, 2011). Many authors have mentioned changes in children developmental rate between generations, which tend to be faster in younger generations (Black and Ferguson, 2016, Hawley et al, 2009, Hsieh et al, 2013, Lai et al, 2008).

Higher developmental rate in younger generations was found in every part of the world (Berk, 2012). Berk stated in the study that children in recent generation were greater in their stature than their elder relatives, when at the same age, due to faster physical development. In addition, they also reached puberty earlier, explicitly in females that earlier age of menstruation was found (Figure 9) (Berk, 2012).

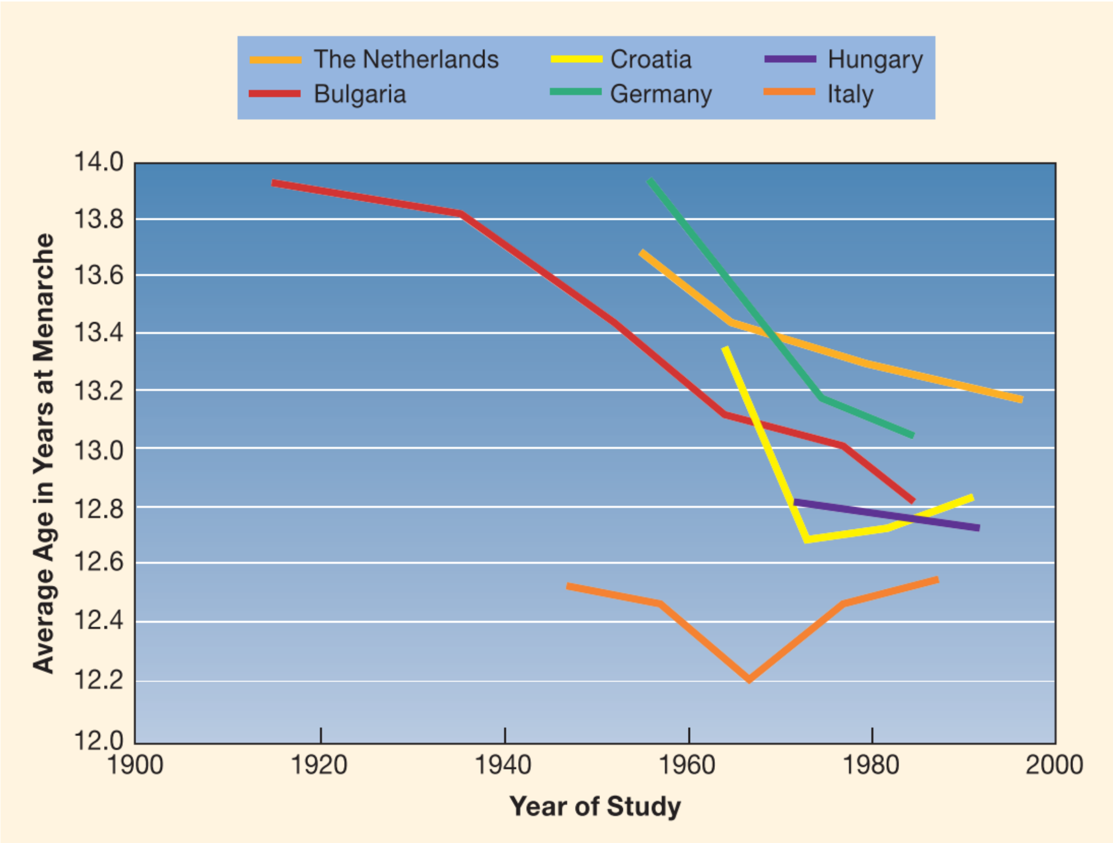


Figure 9 The diagram showing average ages at menarche in population from various countries which tend to be younger in later generation (Berk, 2012).

Lai and colleagues studied the relationship between age of menarche and skeletal development based on 8 – 18.9 years old orthodontic patients in Taiwan (Lai et al, 2008). They compared the age of menarche with skeletal age from NTUH-SMI (National Taiwan University Hospital skeletal maturation index) and CVMI (cervical vertebral maturity index). As the development progresses, the NTUH-SMI rates from stage 1 to 9 and the CVMI rates from stage I to VI. From the study, the average age of menarche was 11.97 years old and all the patients had menstruation at 14 years old. They found that menarche appeared to be a critical point between NTUH-SMI stage 4 and 5, and CVMI stage III and IV. They inferred that the NTUH-SMI stage 4 correlated with peak height velocity, which one would gain height so fast, and the average age in this stage was 11.12 years old. Menarche tended to appear 1 year after reaching peak height velocity. Moreover, they compared the average age of menarche with previous studies and found 0.9 and 0.14 years earlier than ones from 1987 and 2005, respectively (Lai et al, 2008).

Hawley and colleagues compared skeletal age of South African children between 1962 and 2001 (Hawley et al, 2009). The skeletal age was achieved by hand and wrist radiography using Greulich and Pyle method. The study was limited to 9, 10 and 11-year-old children and comparisons were made in each chronological age group. They found that both African and Caucasian samples in later generation were more advance in skeletal maturation when compared within their respective ethnic. Caucasian males were advance 3.4 months in average, but were significant only in 9-year-old group which was about 5.16 months. Caucasian females were advance 2.0 months in average, but, like the males, were significant only in 9-year-old group which was about 5.16 months too. Both African males and females had significant advance in all age groups: African males were 9.7 months and African females were 15.8 months in average (Figure 10). In addition, they mentioned changes in nutrition, healthcare, and socio-economic statuses as the factors, especially in African population (Hawley et al, 2009).

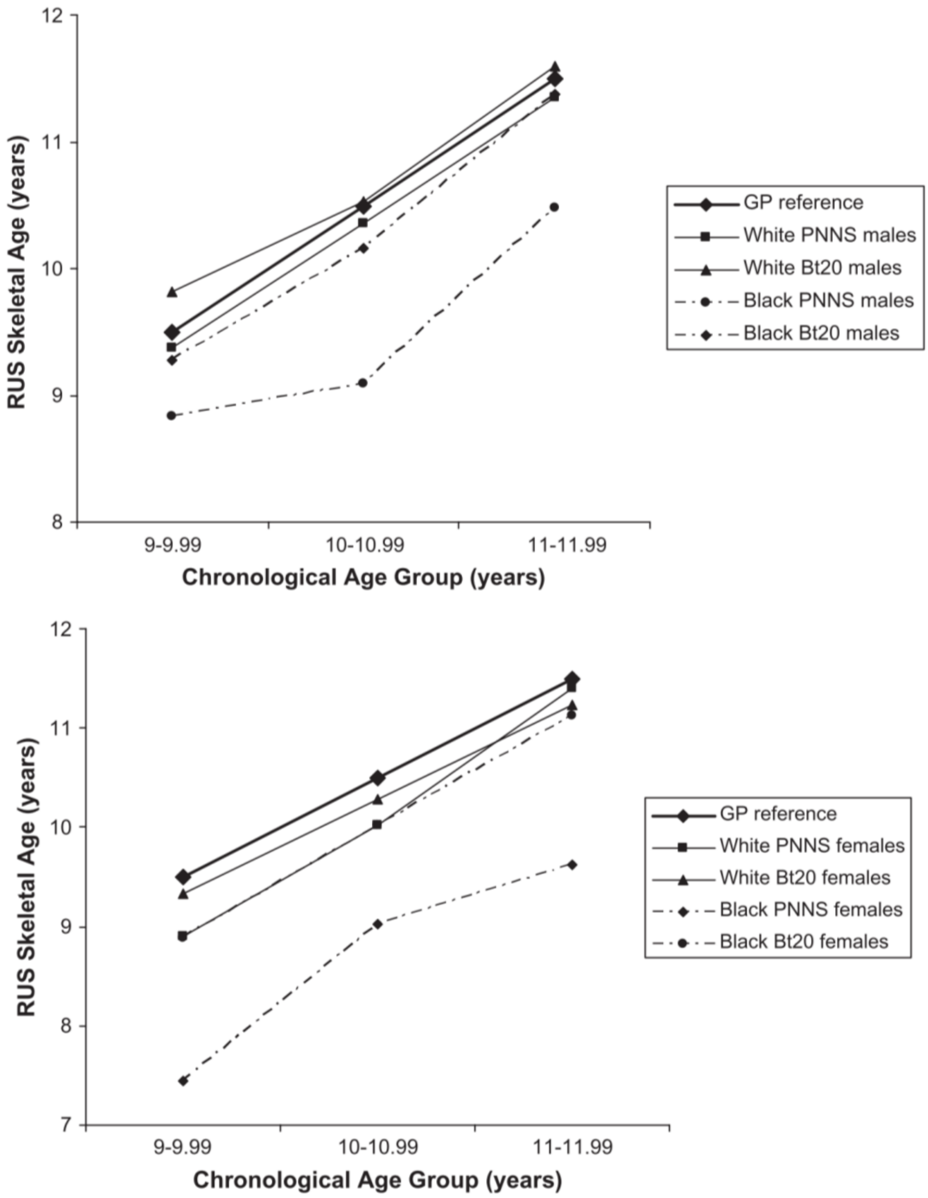


Figure 10 A diagram showed changes in skeletal age in South African males and females sample, PNNS as Praetoria National Nutrition Survey (1962-1965), Bt20 as The Birth to Twenty sample (2001) (Hawley et al, 2009).

Hsieh and colleagues compared skeletal age between subjects in mid-1960s and mid-2000s in Taiwan using Tanner and Whitehouse method (TW3, CAR-score) (Hsieh et al, 2013). They found that at the same chronological age, subjects from mid-

2000s had significantly higher CAR-score than from mid-1960s, meaning that people in mid-2000s demonstrated more advance skeletal development than in the past (Hsieh et al, 2013).

There are still some controversies as shown in a study done by Hackman and Black in 2013 (Hackman and Black, 2013). They studied on the reliability of Greulich and Pyle method on modern Scottish population. The results showed that the differences between bone age and chronological age, in average, were 1.95 and 1.63 months underestimation for females and males respectively. However, the discrepancy may be worse as 13.38 months underestimation for 16 - 20 years old female group and 11.05 months overestimation in 16-year-old male group by means. In conclusion, Greulich and Pyle atlas still showed a strong correlation with modern Scottish population by the growth pattern, but the concern about age differences was also mentioned (Hackman and Black, 2013).

The variation may be caused by changes in food consumption, development in medical sciences and technology which improve illness prevention and accelerates healing process. The socio-economic and nutritional factors must be taken into consideration (Berk, 2012, Bogin and MacVean, 1983, Hackman, 2012, Hawley et al, 2009, Lai et al, 2008, Smith and Brownlees, 2011).

Observer dependence

Furthermore, using the same method, but different observers may generate different results (Hackman, 2012) because some methods rely heavily on visual recognition and personal decisions. Studies on inter-observer reliability of the age estimation methods showed some disagreement between observers (Alcina et al, 2015, Kim et al, 2008, Paxton et al, 2013).

Kim and colleagues compared the reliability of 'Greulich and Pyle method' and 'Tanner and Whitehouse method' (Kim et al, 2008). They used 102 radiographs and 3 observers, 100 radiographs for each one. For the results, the inter-observer errors were 0.51 years for Greulich and Pyle method, and 0.54 years for Tanner and Whitehouse method. The intra-observer errors were 0.48 and 0.45 years for 'Greulich and Pyle method' and 'Tanner and Whitehouse method', respectively (Kim et al, 2008).

Paxton, Lamont and Stillwell studied on the reliability of Greulich and Pyle method based on Australian population (Paxton et al, 2013). The results showed underestimation 1.5 months in males and 3.7 months in females. From this study, the inter-observer difference was 2.7 months in average, not statistically significant (Paxton et al, 2013).

Alcina and colleague studied the application of Greulich and Pyle method in Spanish population and also analyzed inter-observer and intra-observer error (Alcina et al, 2015). It was found that, although not clinically significant, inter-observer mean difference was less than a month but the deviation might be up to 6 months. The intra-observer result appeared in the same direction. Comparing with chronological age, this study found differences approximately half a year for girls and 4 months for boys (Alcina et al, 2015).

These studies directly proved the reliability of the age estimation methods. However, in each individual case, any chance of errors must be treated with caution. For example, using Greulich and Pyle method, just a little discrepancy may result as much as 1-year deviation since adjacent radiographs in some part of the Atlas show 1-year interval (Figure 11).

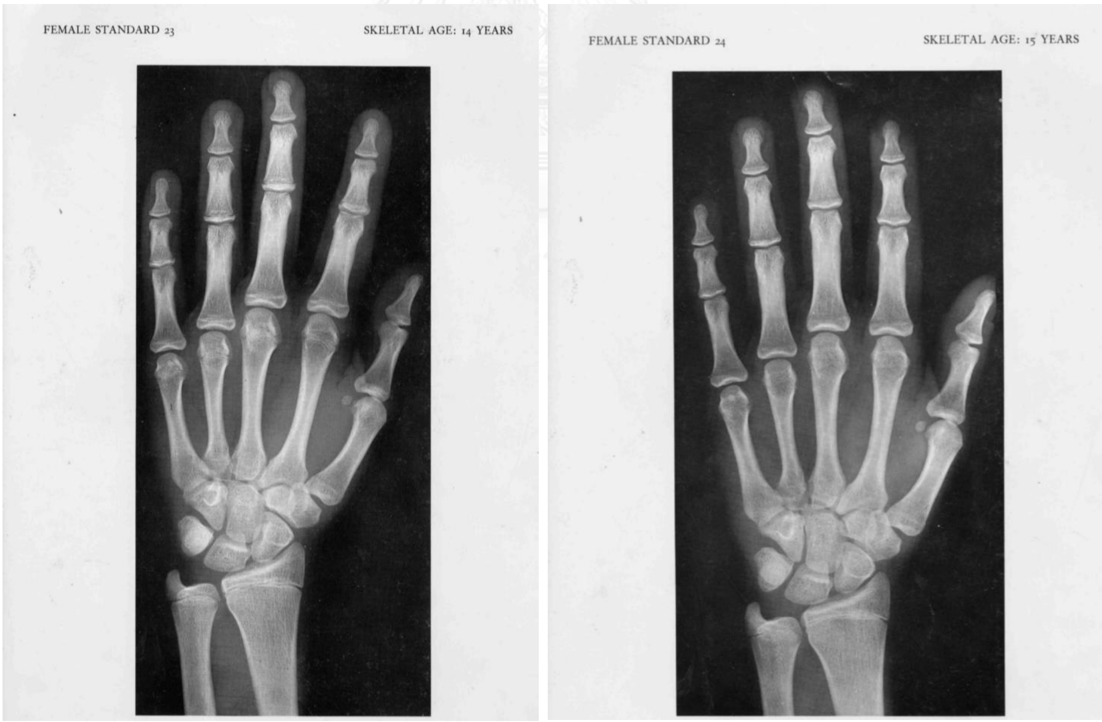
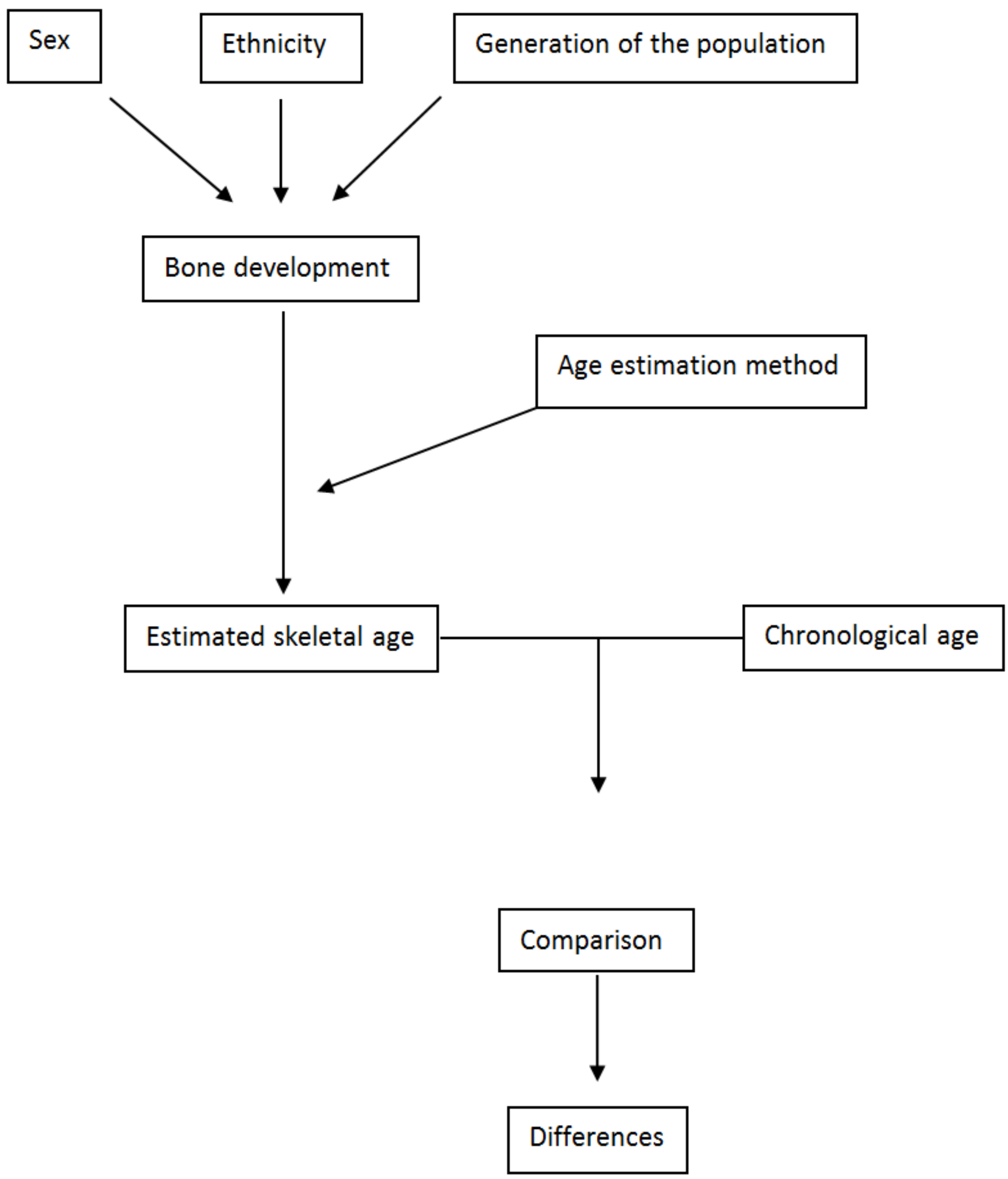


Figure 11 An example of consecutive radiographs in Greulich and Pyle atlas, 15 years old next to 14 years old, which have 1-year interval (Greulich and Pyle, 1959).

Conceptual framework



Research question

Are there any differences between chronological age and estimated age using three commonly used hand and wrist age estimation methods in a group of contemporary Thai?

Hypotheses

Null hypothesis: none of the tested methods of which the estimated age is different from the chronological age in a group of contemporary Thai

Alternative hypothesis: there is at least one of the tested methods of which the estimated age is different from the chronological age in a group of contemporary Thai

Objectives

1. To evaluate the difference between the chronological age and the estimated age by using three commonly used hand and wrist age estimation methods in a group of contemporary Thai
2. To evaluate the accuracy and reliability of each age estimation method for age prediction in a group of contemporary Thai

Chapter 3 Materials and methods

The study protocol was approved by the ethical committee, Faculty of Dentistry, Chulalongkorn University No.022/2017.

Study design

Analytical study, cross-sectional study, observational study

Study population

Source of population:

Thai patients aged 8 - 20 years old, from the dental hospital, Faculty of Dentistry, Chulalongkorn University

Method of recruitment:

Hand and wrist radiographs from patients who needed orthodontic treatment or some other treatments at the Department of Radiology, Faculty of Dentistry, Chulalongkorn University. Radiographs of left or right hand and wrist are both acceptable since no significant different was found when using them for age estimation (Baer and Djrkatz, 1957, Greulich and Pyle, 1959, Hackman and Black, 2012).

Selection criteria:

To define "Contemporary Thai"

- The patients must have declared Thai nationality.
- The hand and wrist radiographs were taken between 1 January 2011 – 31 December 2016.
- Date of birth of all subjects must be between 1991 – 2008.

Patients with history of systemic diseases which affect skeletal development were excluded.

Sample size:

At least 100 female and 100 male subjects were collected.

The sample size was calculated based on the results from prior study (Alcina et al, 2015) by the *formula:*

$$n = \frac{(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2 \sigma^2}{\Delta^2}$$

Eq.1 Sample size estimation formula for testing dependent means (Chow et al, 2007)

SD. (σ) = 1.11

Delta (Δ) = 0.32

α = 0.05

β = 0.20

The minimal sample size was 95 per sex, but to assure the statistical power, the total sample size was set at minimum 100 subjects per sex.

Study place:

Radiology clinic, Faculty of Dentistry, Chulalongkorn University

Study procedure

The hand and wrist radiographs were taken by Carestream™ CS 8000c and CS 9000c x-ray machine (Carestream Health, Inc, Rochester, NY, USA). After sending to the picture archiving and communication system, the images were visualized on a high-definition monitor screen using Infinitt® PACS software (Infinitt Healthcare Co., Ltd., Seoul, South Korea).

The cases were listed based on the acquisition date. Age estimations were done with the hand and wrist radiographs by all 3 methods; Greulich and Pyle (2nd edition) (Greulich and Pyle, 1959), Tanner and Whitehouse (TW3 RUS score) (Tanner et al, 2001), and Fishman (1987) (Fishman, 1987). During the estimation process the observers were blinded from the true (chronological) age leaving only the sex of the patients to be known.

Two pre-calibrated observers participated in the age estimation process. The first observer used all 3 age estimation methods with the whole samples. Twenty-percent of the samples were randomly selected for intra-observer analysis. The second observer only performed age estimation with this group of samples for inter-observer analysis.

After all samples were analyzed, the chronological ages were revealed and compared with the estimated age. The subjects were categorized by the chronological age, 1-year-old-ranged for each group. In each group, the mean ages for each estimation method and the chronological age were calculated. Descriptive analyses were analyzed and presented. The accuracy of each age estimation method was presented as mean differences between the chronological age and the estimated age.

A test for normality was done using Kolmogorov-Smirnov test. Friedman's rank test was done to find any significant differences between estimated ages by each method. The significance was set at $p < 0.05$. Wilcoxon signed ranks test with Bonferroni correction was later used to compare the estimated ages by each pair of methods. Weighted kappa statistics was used for intra-observer and inter-observer reliability analysis.

Concerning that the difference between the estimated age and chronological age might vary among age groups, the analysis was also done separately for each age group. Moreover, the categorical data will be the reference for further adaptation on the age estimation methods.

Chapter 4 Results

Descriptive statistics

A total of 610 patients had hand and wrist radiographs taken during 2011 – 2016. Due to a lack of information, 123 patients were excluded. According to the selection criteria, 33 patients were excluded as their hand and wrist radiographs were not taken at 8 – 20 years old, and other 89 patients were excluded because of their conflicting nationalities, race and influencing diseases. Therefore, only 365 subjects, consisting 193 females and 172 males, were included for age estimation (Table 1).

Table 1 shows the final number of female and male subjects for each age group.

Age (years)	Female	Male	Total
8 – 8.99	11	11	22
9 – 9.99	22	20	42
10 – 10.99	27	21	48
11 – 11.99	29	33	62
12 – 12.99	36	23	59
13 – 13.99	31	34	65
14 – 14.99	12	10	22
15 – 15.99	12	9	21
16 – 16.99	7	5	12
17 – 17.99	3	3	6
18 – 18.99	1	1	2
19 – 19.99	1	2	3
20 – 20.99	1	0	1
Total	193	172	365

The mean chronological age of the subjects was 12.26 ± 2.27 years. After the application of age estimation process, 53 subjects had less skeletal maturity than the least SMI (SMI 1) in Fishman method, thus were counted as missing data for Fishman method. As the results of the age estimation process, the overall mean age from Greulich and Pyle method was 12.93 ± 2.97 years, and from Tanner and Whitehouse method was 12.42 ± 2.62 years. The mean age for each age group are shown in Table 2, Figure 12. The missing data from Fishman method were found in 8 – 8.99 to 13 - 13.99 years age groups, hence the mean ages in these groups were not calculated. In order to compare with the chronological age, the mean differences are shown in Table 3.

Table 2 shows the mean chronological age and mean estimated age with standard deviation (SD) for each age group.

Age group (year)	Mean Greulich and Pyle ± SD (year)	Mean Tanner and Whitehouse ± SD (year)	Mean Fishman ± SD (year)	Mean chronological age ± SD (year)
8 – 8.99	8.65 ± 2.15	8.22 ± 2.00	N/A	8.63 ± 0.30
9 – 9.99	9.65 ± 1.95	9.61 ± 1.56	N/A	9.59 ± 0.24
10 – 10.99	10.81 ± 1.79	10.49 ± 1.58	N/A	10.48 ± 0.31
11 – 11.99	12.33 ± 1.63	11.88 ± 1.53	N/A	11.46 ± 0.32
12 – 12.99	13.16 ± 1.54	12.85 ± 1.58	N/A	12.38 ± 0.27
13 – 13.99	14.55 ± 1.45	14.13 ± 1.45	N/A	13.49 ± 0.28
14 – 14.99	15.68 ± 1.36	15.11 ± 1.10	14.93 ± 1.04	14.51 ± 0.36
15 – 15.99	16.62 ± 1.16	15.47 ± 0.84	15.74 ± 0.75	15.42 ± 0.33
16 – 16.99	17.90 ± 1.28	15.46 ± 0.78	15.91 ± 0.89	16.34 ± 0.29
17 – 17.99	18.33 ± 0.52	15.75 ± 0.82	16.70 ± 0.44	17.37 ± 0.20
18 – 18.99	18.50 ± 0.71	15.75 ± 1.06	16.88 ± 0.52	18.58 ± 0.37
19 – 19.99	19.67 ± 0.58	16.00 ± 0.87	17.00 ± 0.42	19.53 ± 0.23
20 – 20.99	18.00*	15.00*	16.50*	20.15*
overall	12.93 ± 2.97	12.42 ± 2.62	N/A	12.26 ± 2.27

* only one subject in the age group

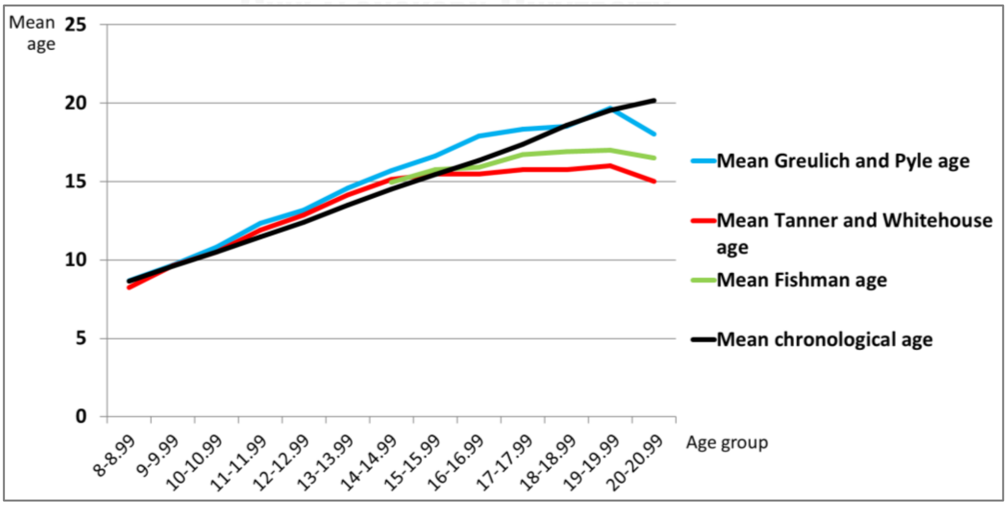


Figure 12 shows means of the estimated ages from each method and mean chronological age.

Table 3 shows the mean difference with standard deviation (SD) between the chronological age and the estimated age for each method and each age group.

Age group (year)	[Greulich and Pyle] – [chronological age] ± SD (year)	[Tanner and Whitehouse] – [chronological age] ± SD (year)	[Fishman] – [chronological age] ± SD (year)
8 – 8.99	0.02 ± 2.04	- 0.41 ± 1.91	N/A
9 – 9.99	0.03 ± 1.93	- 0.00 ± 1.58	N/A
10 – 10.99	0.33 ± 1.72	0.01 ± 1.53	N/A
11 – 11.99	0.87 ± 1.56	0.42 ± 1.48	N/A
12 – 12.99	0.78 ± 1.46	0.47 ± 1.51	N/A
13 – 13.99	1.06 ± 1.42	0.64 ± 1.42	N/A
14 – 14.99	1.17 ± 1.26	0.60 ± 1.05	0.42 ± 1.00
15 – 15.99	1.20 ± 1.16	0.05 ± 0.81	0.32 ± 0.82
16 – 16.99	0.67 ± 1.27	- 0.88 ± 0.69	- 0.43 ± 0.93
17 – 17.99	0.96 ± 0.38	- 1.62 ± 0.80	- 0.67 ± 0.28
18 – 18.99	- 0.08 ± 0.33	- 2.83 ± 0.69	- 1.70 ± 0.14
19 – 19.99	- 0.87 ± 0.58	- 3.53 ± 0.86	- 2.54 ± 0.45
20 – 20.99	- 2.15*	- 5.15*	- 3.64*
overall	0.67 ± 1.60	0.16 ± 1.56	N/A

* only one subject in the age group

In female subjects, the mean chronological age was 12.27 ± 2.26 years. The mean estimated ages were 13.18 ± 2.59 and 12.69 ± 2.18 years using ‘Greulich and Pyle’ method and ‘Tanner and Whitehouse’ method, respectively (Table 4, figure 13). These methods showed an overestimating trend with mean difference 0.90 ± 1.29 years for Greulich and Pyle method, and 0.41 ± 1.36 years for Tanner and Whitehouse method (Table 5).

Table 4 shows the mean chronological age and mean estimated age with standard deviation (SD) for each age group in female subjects.

Age group (year)	Mean Greulich and Pyle age ± SD (year)	Mean Tanner and Whitehouse age ± SD (year)	Mean Fishman age ± SD (year)	Mean chronological age ± SD (year)
8 – 8.99	9.26 ± 2.00	8.68 ± 2.03	N/A	8.63 ± 0.30
9 – 9.99	10.02 ± 1.57	9.90 ± 1.15	N/A	9.64 ± 0.30
10 – 10.99	11.39 ± 1.45	11.22 ± 1.31	11.93 ± 0.80	10.49 ± 0.31
11 – 11.99	12.62 ± 1.11	12.48 ± 0.91	12.44 ± 0.67	11.44 ± 0.30
12 – 12.99	13.32 ± 1.13	13.04 ± 1.29	13.22 ± 1.16	12.39 ± 0.26
13 – 13.99	14.90 ± 1.13	14.53 ± 0.65	14.42 ± 0.90	13.51 ± 0.29
14 – 14.99	15.17 ± 1.11	14.58 ± 0.89	14.42 ± 0.93	14.45 ± 0.37
15 – 15.99	16.50 ± 1.17	15.00 ± 0.00	15.56 ± 0.83	15.41 ± 0.32
16 – 16.99	16.86 ± 1.22	15.00 ± 0.00	15.82 ± 0.86	16.21 ± 0.21
17 – 17.99	18.00 ± 0.00	15.00 ± 0.00	16.50 ± 0.00	17.32 ± 0.15
18 – 18.99	18.00*	15.00*	16.50*	18.31*
19 – 19.99	18.00*	15.00*	16.50*	19.49*
20 – 20.99	18.00*	15.00*	16.50*	20.15*
overall	13.18 ± 2.59	12.69 ± 2.18	N/A	12.27 ± 2.26

* only one subject in the age group

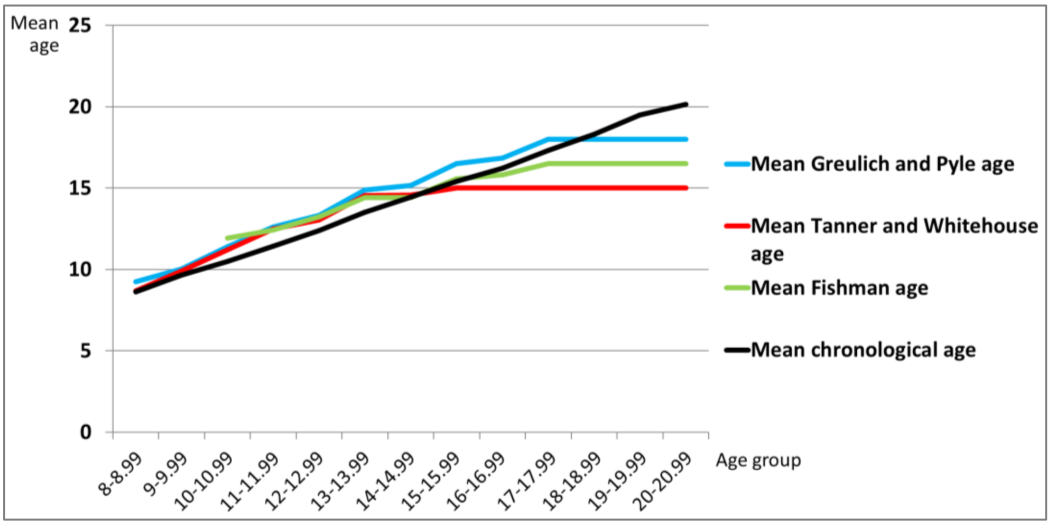


Figure 13 shows means of the estimated ages from each method and mean chronological age in female.

Table 5 shows the mean difference with standard deviation (SD) between the chronological age and the estimated age for each age group in female subjects.

Age group (year)	[Greulich and Pyle] – [chronological age] ± SD (year)	[Tanner and Whitehouse] – [chronological age] ± SD (year)	[Fishman] – [chronological age] ± SD (year)
8 – 8.99	0.63 ± 1.89	0.05 ± 1.96	N/A
9 – 9.99	0.37 ± 1.51	0.26 ± 1.13	N/A
10 – 10.99	0.90 ± 1.33	0.73 ± 1.24	1.43 ± 0.71
11 – 11.99	1.18 ± 1.10	1.04 ± 0.91	1.00 ± 0.68
12 – 12.99	0.93 ± 1.06	0.66 ± 1.21	0.84 ± 1.09
13 – 13.99	1.39 ± 1.13	1.22 ± 0.66	0.90 ± 0.92
14 – 14.99	0.72 ± 1.16	0.13 ± 0.95	-0.03 ± 0.99
15 – 15.99	1.09 ± 1.24	-0.41 ± 0.32	0.16 ± 0.97
16 – 16.99	0.65 ± 1.34	-1.21 ± 0.21	- 0.40 ± 1.00
17 – 17.99	0.68 ± 0.15	-2.32 ± 0.15	- 0.82 ± 0.15
18 – 18.99	-0.31*	-3.31*	1.81*
19 – 19.99	-1.49*	-4.49*	-2.98*
20 – 20.99	-2.15*	-5.15*	-3.64*
overall	0.90 ± 1.29	0.41 ± 1.36	N/A

* only one subject in the age group

For male subjects, the mean chronological age was 12.24 ± 2.30 years. The mean age estimated by Greulich and Pyle method was 12.65 ± 3.32 years, showing 0.42 ± 1.86 years overestimation (Table 6, 7). In contrast, Tanner and Whitehouse method resulted in 12.12 ± 3.01 years mean age that was 0.12 ± 1.73 years underestimated (Table 6, 7).

Table 6 shows the mean chronological age and mean estimated age with standard deviation (SD) for each age group in male subjects.

Age group (year)	Mean Greulich and Pyle age ± SD (year)	Mean Tanner and Whitehouse age ± SD (year)	Mean Fishman age ± SD (year)	Mean chronological age ± SD (year)
8 – 8.99	8.05 ± 2.22	7.76 ± 1.96	N/A	8.63 ± 0.30
9 – 9.99	9.25 ± 2.27	9.30 ± 1.90	N/A	9.59 ± 0.24
10 – 10.99	10.05 ± 1.94	9.54 ± 1.40	N/A	10.46 ± 0.31
11 – 11.99	12.08 ± 1.96	11.35 ± 1.76	N/A	11.47 ± 0.33
12 – 12.99	12.91 ± 2.02	12.55 ± 1.95	N/A	12.37 ± 0.30
13 – 13.99	14.22 ± 1.63	13.76 ± 1.85	N/A	13.46 ± 0.26
14 – 14.99	16.30 ± 1.42	15.74 ± 1.01	15.55 ± 0.83	14.58 ± 0.36
15 – 15.99	16.78 ± 1.20	16.10 ± 0.98	15.98 ± 0.58	15.44 ± 0.37
16 – 16.99	17.20 ± 1.48	16.10 ± 0.89	16.04 ± 1.01	16.51 ± 0.35
17 – 17.99	18.67 ± 0.58	16.50 ± 0.00	16.89 ± 0.62	17.42 ± 0.27
18 – 18.99	19.00*	16.50*	17.20*	18.84*
19 – 19.99	19.00 ± 0.00	16.50 ± 0.00	17.20 ± 0.00	19.56 ± 0.32
overall	12.65 ± 3.32	12.12 ± 3.01	N/A	12.24 ± 2.30

* only one subject in the age group

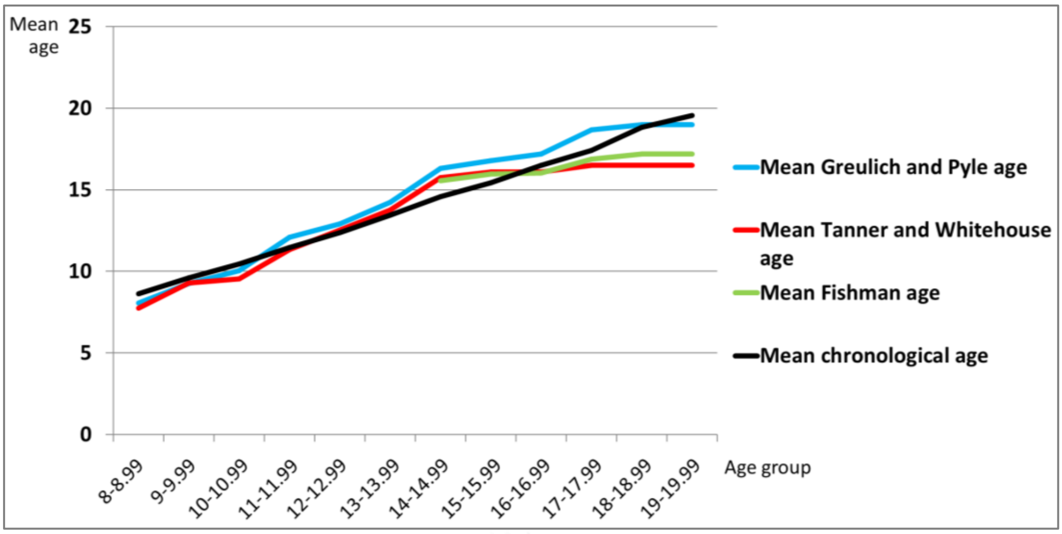


Figure 14 shows means of the estimated ages from each method and the mean chronological age in male.

Table 7 shows the mean difference with standard deviation (SD) between the chronological age and the estimated age for each age group in male subjects.

Age group (year)	[Greulich and Pyle] – [chronological age] ± SD (year)	[Tanner and Whitehouse] – [chronological age] ± SD (year)	[Fishman] – [chronological age] ± SD (year)
8 – 8.99	-0.58 ± 2.09	-0.88 ± 1.83	N/A
9 – 9.99	-0.34 ± 2.29	-0.29 ± 1.95	N/A
10 – 10.99	-0.41 ± 1.91	-0.92 ± 1.39	N/A
11 – 11.99	0.60 ± 1.85	-0.12 ± 1.67	N/A
12 – 12.99	0.54 ± 1.94	0.18 ± 1.88	N/A
13 – 13.99	0.76 ± 1.59	0.30 ± 1.80	N/A
14 – 14.99	1.72 ± 1.21	1.16 ± 0.92	0.97 ± 0.73
15 – 15.99	1.34 ± 1.12	0.66 ± 0.89	0.54 ± 0.55
16 – 16.99	0.69 ± 1.33	-0.41 ± 0.88	-0.47 ± 0.94
17 – 17.99	1.25 ± 0.30	-0.92 ± 0.27	-0.53 ± 0.34
18 – 18.99	0.16*	-2.34*	-1.64*
19 – 19.99	-0.56 ± 0.32	-3.06 ± 0.32	-2.31 ± 0.32
overall	0.42 ± 1.86	-0.12 ± 1.73	N/A

* only one subject in the age group

Comparison between estimated age and chronological age

Overall comparison

The Kolmogorov-Smirnov test rejected normal distribution ($p < 0.001$ for GP, TW and Fi, and $p = 0.021$ for chronological age); therefore, the estimated ages and the chronological age were compared using non-parametric statistics.

The results from each method and the chronological age were divided into 3 parts due to conditions.

Age 8 – 13.99 years (mean Fishman age was absent)

For this group, only Greulich and Pyle, Tanner and Whitehouse, and chronological age were statistically analyzed by Friedman's rank test. The result showed significant difference ($p < 0.001$) among the two methods and the chronological age. Wilcoxon signed ranks test with Bonferroni correction was done for multiple comparisons. The level of significance was set at $p < 0.017$. The results showed significant difference between each pair ($p < 0.001$). The results from Wilcoxon signed ranks test when done on each age group showed that the differences between the estimated age and the chronological age were not statistically significant in the groups under 11 years old (Table 8).

Age 14 – 16.99 years (mean Fishman age was present)

In these groups, the results from Fishman method were included. Friedman's rank test showed significant difference ($p < 0.001$) among the three age estimation methods and the chronological age. As a result of Bonferroni correction, the significance level was set at $p < 0.008$. Only Greulich and Pyle age showed a significant difference ($p < 0.001$) when compared with the chronological age and the other methods. The results from Wilcoxon signed ranks test for each age group are shown in Table 9.

Age 17 – 20.99 years

The data was not appropriate for statistical analysis due to lack of subjects.

Table 8 shows the results from Wilcoxon signed ranks test for 8 – 13.99 years old subjects for each age group between 8-13.99 years old ($p < 0.017$).

Age group (year)	Chronological age vs Greulich and Pyle (p value)	Chronological age vs Tanner and Whitehouse (p value)	Tanner and Whitehouse age vs Greulich and Pyle (p value)
8 – 8.99	0.910	0.445	0.123
9 – 9.99	0.930	0.750	0.653
10 – 10.99	0.112	0.821	0.041
11 – 11.99	< 0.001*	0.007*	< 0.001*
12 – 12.99	< 0.001*	0.004*	< 0.001*
13 – 13.99	< 0.001*	< 0.001*	< 0.001*

* $p < 0.017$

Table 9 shows the results from Wilcoxon signed ranks test for 14 – 16.99 years old subjects for each age group between 14-16.99 years old (p < 0.008).

Age group (year)	Chronological age vs Greulich and Pyle (p value)	Chronological age vs Tanner and Whitehouse (p value)	Chronological age vs Fishman (p value)	Tanner and Whitehouse vs Greulich and Pyle (p value)	Fishman vs Tanner and Whitehouse (p value)	Fishman vs Greulich and Pyle (p value)
14 – 14.99	0.001*	0.011	0.026	0.001*	0.014	< 0.001*
15 – 15.99	0.001*	0.903	0.079	< 0.001*	0.875	0.001*
16 – 16.99	0.117	0.006*	0.272	0.003*	0.385	0.002*

* p < 0.008

Females

In female subjects, the results from Fishman method were absent only in 8 - 8.99 and 9 - 9.99 years groups, and there were only 7 subjects in 16 - 16.99 years group that was considered to be insufficient. Therefore, the comparisons were made in 2 parts as 8 – 9.99 years which Fishman age was excluded, and 10 – 15.99 years which included Fishman age.

For 8 – 9.99 years old female, there was no statistically significant difference among Greulich and Pyle, Tanner and Whitehouse, and the chronological age using Friedman's rank test (p = 0.499).

For 10 – 15.99 years females, the results from Friedman's rank test showed significant difference (p < 0.001) among the three methods and the chronological age. The level of significance was set at p < 0.008 after Bonferroni correction. Wilcoxon signed ranks test showed significant difference (p = 0.001) for each pair of them, except for the difference between Fishman and Tanner and Whitehouse methods which was not significant (p = 0.034). The results from Wilcoxon signed ranks test for each age group are shown in Table 10.

Table 10 shows the results from Wilcoxon signed ranks test for 10 – 15.99 years old female subjects separated by age groups (p < 0.008).

Age group (year)	Chronological age vs Greulich and Pyle (p value)	Chronological age vs Tanner and Whitehouse (p value)	Chronological age vs Fishman (p value)	Tanner and Whitehouse vs Greulich and Pyle (p value)	Fishman vs Tanner and Whitehouse (p value)	Fishman vs Greulich and Pyle (p value)
10 – 10.99	0.005*	0.008*	< 0.001*	0.313	< 0.001*	0.003*
11 – 11.99	< 0.001*	< 0.001*	< 0.001*	0.057	0.803	0.288
12 – 12.99	< 0.001*	0.003*	< 0.001*	0.003*	0.078	0.186
13 – 13.99	< 0.001*	< 0.001*	< 0.001*	0.010	0.034	< 0.001*
14 – 14.99	0.099	0.209	0.433	0.006*	0.029	0.002*
15 – 15.99	0.028	0.002*	0.530	0.004*	0.374	0.012

* p < 0.008

Males

In male subjects, the results were categorized into 3 parts as 8 – 13.99 years, 14 – 15.99 years, and over 16 years due to presence of Fishman age and lack of subjects.

For 8 – 13.99 years old male, only Greulich and Pyle, Tanner and Whitehouse, and the chronological age were included. The result from Friedman's rank test showed statistically significant difference (p < 0.001). Since the level of significance was set at p < 0.017 by Bonferroni correction, Wilcoxon signed ranks test showed only that 'Greulich and Pyle' age was significantly different from 'Tanner and Whitehouse' age (p < 0.001).

Divided by age groups, the results showed significant differences only in 11 – 11.99 years and 13 – 13.99 years groups (Table 11).

Table 11 shows the results from Wilcoxon signed ranks test for 8 – 13.99 years old male subjects divided by age groups (p < 0.017).

Age group (year)	Chronological age vs Greulich and Pyle (p value)	Chronological age vs Tanner and Whitehouse (p value)	Tanner and Whitehouse vs Greulich and Pyle (p value)
8 – 8.99	0.424	0.213	0.533
9 – 9.99	0.411	0.601	0.776
10 – 10.99	0.502	0.014*	0.064
11 – 11.99	0.012*	0.943	< 0.001*
12 – 12.99	0.070	0.323	0.028
13 – 13.99	0.007*	0.266	0.001*

* p < 0.017

For 14 – 15.99 years old male, the results from all the three methods and the chronological age were compared by Friedman's rank test, and significant difference (p < 0.001) was found. Bonferroni correction showed that the level of significance was at p < 0.008 and Wilcoxon signed ranks test showed that the difference in every pair was statistically significant (p = 0.003), except for the difference between Fishman and Tanner and Whitehouse methods which was not significant (p = 0.016). The results from Wilcoxon signed ranks test for 14 – 14.99 years and 15 – 15.99 years groups are shown in Table 12.

Table 12 shows the results from Wilcoxon signed ranks test for 14 – 15.99 years old male subjects separated by age groups ($p < 0.008$).

Age group (year)	Chronological age vs Greulich and Pyle (*p* value)	Chronological age vs Tanner and Whitehouse (*p* value)	Chronological age vs Fishman (*p* value)	Tanner and Whitehouse vs Greulich and Pyle (*p* value)	Fishman age vs Tanner and Whitehouse (*p* value)	Fishman vs Greulich and Pyle (*p* value)
14 – 14.99	0.007	0.022	0.013	0.059	0.092	0.012
15 – 15.99	0.021	0.110	0.028	0.009	0.098	0.019

* $p < 0.008$

Comparison by methods

The estimated ages from each particular method were compared only with the chronological ages. Therefore, the level of significance for Wilcoxon signed ranks test was set at $p < 0.05$ without Bonferroni correction as each comparison included only a single pair (chronological age vs estimated age).

Greulich and Pyle method

The overall mean Greulich and Pyle age was 12.93 ± 2.97 years which was 0.67 ± 1.60 years higher than the mean chronological age (Table 2 and 3). The result from Wilcoxon signed ranks test showed significant difference ($p < 0.001$) from the chronological age.

Wilcoxon signed ranks test was done for each age group and the results showed statistically significant differences ($p < 0.001$) between Greulich and Pyle age and the chronological age from 11 – 11.99 years to 15 – 15.99 years groups (Table 13). The age groups over 17 years were excluded due to lack of subjects.

In female subjects, the mean Greulich and Pyle age was 13.18 ± 2.59 years which was 0.90 ± 1.29 years overestimation (Table 4 and 5) and Wilcoxon signed ranks test showed statistically significance (*p* < 0.001). The results separated by the age groups showed significant differences (*p* < 0.028) in 10 – 10.99 years to 15 – 15.99 years groups, except for 14 – 14.99 years group which was not statistically significant (*p* = 0.099) (Table 13). The age groups over 16 years were excluded due to the lack of subjects.

In male subjects, the mean Greulich and Pyle age was 12.65 ± 3.32 years showing 0.42 ± 1.86 years overestimation (Table 6 and 7). The result from Wilcoxon signed ranks test showed statistically significant difference (*p* = 0.001) between Greulich and Pyle age and the chronological age. Separated by the age groups, the results showed significant differences in 11 – 11.99 years to 15 – 15.99 years groups, except for 12 – 12.99 years group which was not statistically significant (*p* = 0.070) (Table 13).

Table 13 shows the results from Wilcoxon signed ranks test between Greulich and Pyle age and the chronological age for each age group ($p < 0.05$).

Age group	Chronological age – Greulich and Pyle age (*p* value)		
(year)	overall	females	Males
8 – 8.99	0.910	0.374	0.424
9 – 9.99	0.930	0.200	0.411
10 – 10.99	0.112	0.005*	0.502
11 – 11.99	< 0.001*	< 0.001*	0.012*
12 – 12.99	< 0.001*	< 0.001*	0.070
13 – 13.99	< 0.001*	< 0.001*	0.007*
14 – 14.99	0.001*	0.099	0.007*
15 – 15.99	0.001*	0.028*	0.021*
16 – 16.99	0.117	N/A	N/A

* $p < 0.05$

Tanner and Whitehouse method

The overall mean Tanner and Whitehouse age was 12.42 ± 2.62 years (Table 2). The difference between the estimated age and the mean chronological age was 0.16 ± 1.56 years (Table 3), and Wilcoxon signed ranks test showed statistically significance ($p = 0.002$).

Separated by age groups, the results from Wilcoxon signed ranks test were significant ($p \leq 0.011$) from 11 – 11.99 years to 16 – 16.99 years, except for 15 – 15.99 years group which showed non-significance ($p = 0.903$) (Table 14).

For female subjects, the mean Tanner and Whitehouse age was 12.69 ± 2.18 years showing 0.41 ± 1.36 years overestimation (Table 4 and 5) which was statistically significant ($p < 0.001$) by Wilcoxon signed ranks test. The analysis for each age group showed significant differences ($p \leq 0.008$) in 10 – 10.99 years to 15 – 15.99 years groups, except for 14 – 14.99 years group ($p = 0.209$) (Table 14).

For male subjects, the mean Tanner and Whitehouse age was 12.12 ± 3.01 years that was 0.12 ± 1.73 years underestimated (Table 6 and 7). The difference between the estimated age and the chronological age was not statistically significant ($p = 0.725$). The result for each age group showed significant differences only in 10 – 10.99 years ($p = 0.014$) and 14 – 14.99 years ($p = 0.022$) groups (Table 14).

Remarkable deviations were found related to particular age groups, hence the whole result can be categorized into 3 parts as following (Figure 15):

Part I, 8 – 11.99 years old, overestimating trend was found in female while underestimation was found in male and the overall result appeared to be consistent with the chronological age.

Part II, 12 – 15.99 years old, overestimation was found in both sexes thus, the overall result also appeared to be overestimated. Female subjects were found more overestimation in earlier age groups while male subjects were found more in later age groups.

Part III, over 16 years old, the results turned to underestimation since the highest predictable ages were 15 years old in female and 16.5 years old in male.

Table 14 shows the results from Wilcoxon signed ranks test between Tanner and Whitehouse age and the chronological age for each age group ($p < 0.05$).

Age group (year)	Chronological age – Tanner and Whitehouse age (*p* value)		
	overall	females	males
8 – 8.99	0.445	0.929	0.213
9 – 9.99	0.750	0.277	0.601
10 – 10.99	0.821	0.008*	0.014*
11 – 11.99	0.007*	< 0.001*	0.943
12 – 12.99	0.004*	0.003*	0.323
13 – 13.99	< 0.001*	< 0.001*	0.266
14 – 14.99	0.011*	0.209	0.022*
15 – 15.99	0.903	0.002*	0.110
16 – 16.99	0.006*	N/A	N/A

* $p < 0.05$

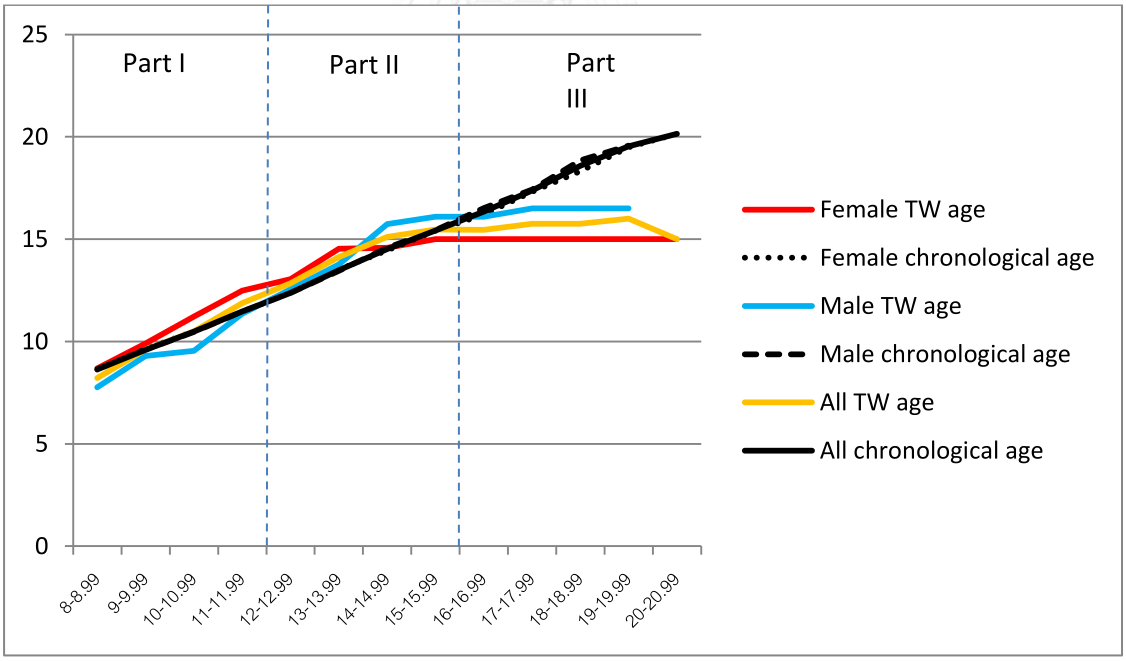


Figure 15 shows the mean estimated Tanner and Whitehouse ages and the mean chronological. The trends of over/underestimation can be categorized into 3 parts: part I 8-11.99, part II 12-15.99 and part III over 16.

Fishman method

The estimated age from Fishman method was unachievable in 53 subjects in 8 – 13.99 years groups because the skeletal development was less than the skeletal condition for the lowest predictable age in Fishman method. Therefore, the mean Fishman ages in some groups and the overall mean Fishman age were missing. Comparisons with the chronological age were done only in 14 – 16.99 years groups.

In 14 – 16.99 years subjects, the results from Wilcoxon signed ranks test showed marginally significant difference (p = 0.057) between Fishman age and the chronological age. The comparison in each age group showed statistically significant difference (p = 0.026) only in 14 – 14.99 years group.

For female subjects, the missing Fishman ages were found only in 8 – 9.99 years groups, so Wilcoxon signed ranks test could be done in 10 – 15.99 years groups while 16 years and over were excluded due to the limitation in subject quantity. The mean Fishman age ± SD from these groups was 13.37 ± 1.44 years that was 0.87 years overestimated. The result from Wilcoxon signed ranks test showed significant difference (p < 0.001) between Fishman age and the chronological age. The comparison in each age group showed significant differences (p < 0.001) in 10 – 10.99 years to 13 – 13.99 years groups (Table 15).

For male subjects, the statistical analysis was done only in 14 – 15.99 years groups. The mean estimated age ± SD from these groups was 15.76 ± 0.74 years which was 0.77 years higher than the chronological age, and was statistically significant (p = 0.001). Comparing in each age group, the results showed statistically significant difference in both groups (p = 0.013 and 0.028) (Table 15).

Table 15 shows the results from Wilcoxon signed ranks test between Fishman age and the chronological age for each age group in female and male subjects ($p < 0.05$).

Age group (year)	Chronological age – Fishman age (p value)	
	females	males
10 – 10.99	< 0.001*	N/A
11 – 11.99	< 0.001*	N/A
12 – 12.99	< 0.001*	N/A
13 – 13.99	< 0.001*	N/A
14 – 14.99	0.433	0.013*
15 – 15.99	0.530	0.028*

* $p < 0.05$

The result from Fishman method was also categorized by SMIs to make a comparison with the original data from Fishman's publication (Fishman, 1987). The results for each SMI are shown in Table 16. Female subjects rapidly changed from SMI 1 to 6 at around 10 years old since the subjects were separated into several groups while the range of mean age was only 1.34 year, including subjects around 11 years old with suspected delayed skeletal maturation. Male subjects were found with the same trend but at older mean ages. When compared SMI 1- 10 with Fishman's data (female 10.23 – 14.89 years, male 11.37 – 16.17 years) (Fishman, 1987), contemporary Thais were approximately 1 year more advanced in skeletal maturation (Table 17). The SMI 11 was inconsistent since this indicator, fusion of distal radius epiphysis, is lifelong persistent.

Table 16 shows mean chronological ages and standard deviations (SD) for each Fishman's SMI.

SMI	overall		female		male	
	n	Chronological age ± SD (year)	n	Chronological age ± SD (year)	n	Chronological age ± SD (year)
missing	53	10.02 ± 1.24	9	8.93 ± 0.66	44	10.24 ± 1.22
1	13	10.76 ± 1.45	6	9.68 ± 0.55	7	11.68 ± 1.35
2	19	10.48 ± 1.12	5	9.69 ± 0.65	14	10.76 ± 1.14
3	17	10.53 ± 1.17	5	11.02 ± 0.90	12	10.33 ± 1.25
4	21	11.57 ± 1.07	5	10.30 ± 0.52	16	11.97 ± 0.87
5	5	10.97 ± 1.57	3	10.35 ± 1.86	2	11.90 ± 0.32
6	25	10.90 ± 1.39	19	10.59 ± 1.21	6	11.89 ± 1.55
7	92	12.14 ± 1.40	54	11.40 ± 1.15	38	13.19 ± 1.01
8	22	12.86 ± 0.90	16	12.61 ± 0.88	6	13.51 ± 0.63
9	8	13.17 ± 0.80	7	13.12 ± 0.85	1	13.51*
10	68	14.34 ± 1.27	48	13.94 ± 1.17	20	15.29 ± 0.97
11	22	17.03 ± 1.72	16	16.55 ± 1.67	6	18.31 ± 1.18

* only one subject in the SMI group

Table 17 shows mean differences with standard deviations (SD) comparing between Fishman age and the chronological age for each Fishman's SMI

SMI	overall		female		male	
	n	[Fishman] – [chronological age] ± SD (year)	n	[Fishman] – [chronological age] ± SD (year)	n	[Fishman] – [chronological age] ± SD (year)
1	13	0.08 ± 1.11	6	0.55 ± 0.55	7	-0.31 ± 1.35
2	19	1.07 ± 1.04	5	1.03 ± 0.65	14	1.08 ± 1.17
3	17	1.29 ± 1.48	5	-0.15 ± 0.90	12	1.89 ± 1.25
4	21	0.64 ± 0.79	5	0.74 ± 0.52	16	0.61 ± 0.87
5	5	1.27 ± 1.33	3	1.35 ± 1.86	2	1.13 ± 0.32
6	25	1.49 ± 1.29	19	1.36 ± 1.21	6	1.89 ± 1.55
7	92	1.17 ± 1.09	54	1.11 ± 1.15	38	1.24 ± 1.01
8	22	0.93 ± 0.93	16	0.65 ± 0.88	6	1.68 ± 0.63
9	8	0.99 ± 0.88	7	0.85 ± 0.85	1	1.97*
10	68	0.93 ± 1.11	48	0.95 ± 1.17	20	0.88 ± 0.97
11	22	-0.32 ± 1.59	16	0.05 ± 1.67	6	-1.07 ± 1.18

* only one subject in the SMI group

Intra-observer and inter-observer analysis

Weighted kappa analysis was used to analyze the reliability of age estimation methods. For Greulich and Pyle method, the estimated ages were related with separated images in the radiographic atlas and the numbers of months between consecutive images were not constant, thus the result should be considered as ordinal information. For Tanner and Whitehouse method, although the possible results were continuous with 0.1 year interval, the prior scores were based on the classification on skeletal development which was ordinal. Therefore, the reliability test for Tanner and Whitehouse method was done with each individual bone. For Fishman method, the estimated ages were matched with the 'skeletal maturity indicators' that were also ordinal.

The results from weighted kappa analysis ranged between 0.66 - 0.92 showing good inter and intra-observer reliability. Most of the intra-observer results were higher than the inter-observer. The highest result was intra-observer fifth proximal phalange of Tanner and Whitehouse method which was 0.92 while the lowest result was inter-observer ulna bone of the same method which was 0.66 (Table 18).

Table 18 shows results from weighted kappa analysis for each age estimation method

Methods		Intra-observer	Inter-observer
Greulich & Pyle		0.89	0.82
Tanner and Whitehouse 3 RUS	Radius	0.86	0.73
	Ulna	0.83	0.66
	1st metacarpus	0.85	0.88
	3rd metacarpus	0.84	0.72
	5th metacarpus	0.87	0.77
	1st proximal phalange	0.90	0.88
	3rd proximal phalange	0.91	0.85
	5th proximal phalange	0.92	0.84
	3rd middle phalange	0.88	0.84
	5th middle phalange	0.91	0.87
	1st distal phalange	0.87	0.83
	3rd distal phalange	0.84	0.84
	5th distal phalange	0.90	0.84
Fishman		0.91	0.85

Chapter 5 Discussion and conclusion

Discussion

The present study compared the accuracy and reliability of age estimated by three commonly used age estimation methods: Greulich & Pyle, Tanner and Whitehouse 3 RUS and Fishman method, in a group of Thai children and adolescents.

The estimated ages from the three methods were significantly different from the chronological age ($p < 0.002$), except for 'Tanner and Whitehouse' in male. Overall, 'Greulich and Pyle' and 'Tanner and Whitehouse' age showed less difference from the chronological age around 10 years old. Higher degrees of overestimation were found in older age groups from 11 years old for both methods (Table 3). However, 'Tanner and Whitehouse' and Fishman methods showed only little derangement at around 15 years old since the highest estimation limits were reached in female subjects while Greulich and Pyle method resulted in strong overestimation until 18 years old. The overall result showed possible higher skeletal maturity rate in contemporary Thais over 10 years old.

In females, all the three methods resulted in overestimation for younger subjects and the discrepancies were greater around 11 – 13 years old. For older groups, the overestimations were less and turned to underestimations at over 15 years old due to the highest predictable limit of each method thus, the differences between the estimated age and the chronological age at 14 – 14.99 years group were not statistically significant. Greulich and Pyle age turned underestimation at the age groups older than other methods since the highest predictable age was 18 years old while other methods were 15 (Tanner and Whitehouse 3 RUS) and 16.5 (Fishman) years old. Therefore, contemporary Thai females over 8 years old obviously showed advanced skeletal maturation with higher maturation rate.

In males, ‘Greulich and Pyle’ and ‘Tanner and Whitehouse’ ages showed underestimation at younger age groups and turned to overestimation at around 11 years old. All the three methods showed underestimation at older age groups due to the predictable limits; ‘Tanner and Whitehouse’ and Fishman age turned to underestimation at 16 – 16.99 years group as the highest estimated ages were around 17 years old, while Greulich and Pyle age turned at 19 – 19.99 years group as the highest limit was 19 years old. The results showed that contemporary Thai males had retarded skeletal maturation in childhood but with higher maturation rate that the teenagers turned advanced in skeletal maturation.

Comparisons were performed between three age estimation methods. Tanner and Whitehouse method appeared to be the most accurate method as it had lower mean difference from the chronological age than Greulich and Pyle method and wider age range coverage than Fishman method (Figure 12). The results appeared in the same way both for females and males (Figure 13 and 14). The overall difference was only 0.16 years and not much deviation was found either for males or females. However, remarkable overestimations and underestimations were found depending on the age groups. The difference between the estimated age and the chronological age should be considered separately in 3 parts (Figure 15). The first part was 8 – 11.99 years old which females were overestimated while males were underestimated and the overall Tanner and Whitehouse age was rather similar to the chronological age. In the second part, both females and males were overestimated and the discrepancy can be over a year. The third part was over 16 years old that the results were turning underestimated due to the limitation of the method.

Greulich and Pyle method resulted in the highest estimated age totally that more overestimation found in both females and males, except for 8 – 10 years old males and late teenagers that this method appeared to be the most accurate as other methods tended to underestimate.

Fishman method showed relatively more accurate results than Tanner and Whitehouse 3 RUS method in late teenagers. However, the results covered only a short age range making the application very limited.

In this study, 53 subjects were recorded as missing data for Fishman's method since their radiographs show less maturity than the SMI 1. This seems to be the limitation of Fishman's method that the first SMI starts at around 10 years old, thus the subjects with less maturity were only estimated as less than 10 years old. However, some of these subjects can be matched with the SMI 2 and some of them were predicted as around 11 years old using other methods. This finding showed a remarkable effect of difference between methods. Furthermore, many subjects were found with a narrowing epiphysis at the proximal phalange of the third finger, which the epiphysis did not reach the width of the diaphysis even they were in the capping stage (Figure 16). Therefore, the SMI 1 may not be a suitable indicator in Thai population.

Figure 16 A hand and wrist radiograph showing narrowing epiphysis of proximal phalange of the third finger that is in capping stage while the width is less than the diaphysis.

Contributing factors

Genetic variations

The results showed statistically significant differences between the estimated age and the chronological age. This may prove the effect of ethnicity on skeletal maturation that Thais as a part of Asian population are different from European and contemporary American population. The studies in other Asian countries showed coincident findings, despite of some varying degree of difference that express ethnical diversity even in the same continent (Chiang et al, 2005, Griffith et al, 2007, Kim et al, 2015, Mohammed et al, 2014b, Saade et al, 2017, Zafar et al, 2010).

Chiang and colleagues (Chiang et al, 2005) studied the reliability of Greulich and Pyle method in Taiwan. They found 0.18 – 1.48 years overestimation in 9 – 16 years old females, although only 13 – 15 years groups were found statistically significance. This present study showed 0.37 – 1.39 years overestimation in 8 – 17 years old females which statistically significance found in 10 – 15 years groups, except for the 14 years group. For males, it was found that there was 0.84 – 1.61 years underestimation in 8 – 11 years groups and 0.13 – 1.28 years overestimation in 12 – 17 years groups. Statistically significant difference was found in almost every group (Chiang et al, 2005). This present study showed 0.34 – 0.58 years underestimation in 8 – 10 years groups and 0.54 – 1.72 years overestimation in 11 – 17 years groups, although significance was found only in cases of overestimation (Table 4 - 7). However, the subjects were limited to less than 10 subjects for each age group under 12 years old (Chiang et al, 2005).

Griffith and colleagues (Griffith et al, 2007) studied the application of ‘Greulich and Pyle’ method and ‘Tanner and Whitehouse 3’ methods in Hong Kong. For Greulich and Pyle, they found 0.3 – 1.12 years overestimation in 9 – 16 years old females that the highest discordance found around 12.5 years old. The present study showed similar results 0.37 – 1.39 years overestimation in 8 – 17 years old females with highest discordance at 13 years old group. In males, they found 0.12 - 0.4 years underestimation in 8 – 9 years old boys and 0.13 – 0.82 years overestimation in 10 – 18 years old boys. In this present study, 0.34 – 0.58 years underestimation was found in 8 – 10 years old males while 0.54 – 1.72 years overestimation was found in 11 – 17 years old males, showing somewhat more overestimation. For Tanner and Whitehouse 3 RUS method, both studies showed overestimation under 14.5 years females; Griffith and colleagues found significant overestimation (0.37 – 1.04 years) in 9 – 12 years females while this present study showed significant overestimation (0.66 – 1.22 years) in 10 – 13 years females. In males, they found significant overestimation (0.46 – 0.69 years) in 10, 13 and 14 years old boys (Griffith et al, 2007) while, in this present study, significant overestimation (1.16 years) was only found in 14 years old (Table 4 – 7).

Zafar and colleagues (Zafar et al, 2010) studied the accuracy of Greulich and Pyle method in Karachi, Pakistan. They found 0.67 and 0.83 years overestimation in late childhood (8.42 – 13.25 years) and adolescent (13.25 – 18 years) females respectively, while this present study showed mostly around 1 year overestimation in 8 – 17 years old females. For males, they found 0.55 years underestimation in late childhood (7.58 – 13.25 years) subjects. However, their study showed only 0.28 years overestimation in adolescent (13.25 – 18 years) males and was not statistically significant (Zafar et al, 2010). This present study found about half a year underestimation in 8 – 10 years groups, but was not statistically significant, while significant overestimation was found over 11 years age groups (Table 4 – 7). Nonetheless, both studies showed trends of underestimation in late childhood and overestimation in teenage males. It should be remarked that Pakistan is located in Western part of South Asia where the population trait may rather be similar to European population.

Mohammed and colleagues (Mohammed et al, 2014b) studied the correlation between skeletal maturity stages and chronological ages in South India. Their results were categorized by each SMI related to Fishman's method. From their study, the mean chronological ages were 9.98 – 12.13 years in females at SMI 1 – 7 while Fishman's reference population was 10.23 – 12.51 years. However, at SMI 8 – 10, their results were 13.97 – 15.45 years that were higher than from Fishman (13.26 – 14.89 years). For males, only SMI 1 and 4 that their results showed 1.2 and 1.35 years lower mean age. In others hand, for the rest SMIs, their results showed higher mean chronological ages (Mohammed et al, 2014b). In general, their subjects showed relatively delayed skeletal maturity while this present study showed around 1 year advanced skeletal maturity (Table 17).

Kim, Lee and Yu (Kim et al, 2015) studied the reliability of 'Greulich and Pyle', 'Tanner and Whitehouse 3 RUS' and Korean standard chart in 7 – 12 years old Korean children. They found good correlation between the estimated age from these three methods and the chronological age, and there were no significant difference between the three methods, although the accuracy on age prediction was not literally analyzed. For Greulich and Pyle method, their result showed only 0.03 – 0.35 years overestimation in girls. There was 0.52 – 1.17 years underestimation for 7 – 10 years old boys and 0.04 years overestimation in 12 years old boys that the turning point was at 11 years old. Both males and females results from this present study appeared in the trend similar to their study but with higher skeletal maturity, around half a year in average. For Tanner and Whitehouse method, their result presented in opposite way to this present study. For girls, they found 0.05 – 0.4 years overestimation in 7 – 11 years old girls and 0.48 years underestimation in 12 years old girls. For boys, they found 0.01 – 0.83 years overestimation in 7 – 12 years old children (Kim et al, 2015). This present study showed increasing overestimation tendency till 13 years old in females and found underestimation in 8 – 11 years old males (Table 4 – 7).

Saade and colleagues (Saade et al, 2017) evaluated the applicability of Greulich and Pyle method and Fishman method in Lebanese children. For Greulich and Pyle, they found 0.58 – 0.95 years overestimation in 8 – 14.9 years old females which more derangement found in older age groups. In males, they found 0.13 and 0.21 years underestimation in 8 – 10 and 12 – 14 years old children, respectively, while 0.26 and 0.20 years overestimation were respectively found in 10 – 12 and 14 – 16.5 years old children thus showing no statistically significant. For Fishman, 0.37 – 1.84 years overestimation was found in 8 – 14.9 years old females which less derangement found around 12 – 14 years group. In males, 0.28 – 2.25 years overestimation was found in 8 – 16.5 years old subjects showing less derangement in older age groups (Saade et al, 2017). This present study showed similar trends for both methods with higher degree of overestimation, except for Greulich and Pyle method in males which clearly showed more overestimating trend in older age groups (Table 4 – 7). This also should be remarked as Lebanon is located in Western Asia that the ethnicity might be different from Eastern Asia.

Compared with studies from other regions of the world, the results were shown in various ways (Alcina et al, 2015, Arciniega Ramos et al, 2013, Buken et al, 2010, de Sousa Dantas et al, 2015, Hackman and Black, 2013, Pinchi et al, 2014).

Buken and colleagues (Buken et al, 2010) studied the reliability of Tanner and Whitehouse 3 RUS method in Turkey. For females, 11 – 15 years old subjects were recruited. The result showed 0.07 – 0.33 years underestimation. For males, 11 – 16 years old subjects were recruited. They found 0.07 – 0.44 years underestimation, except for 12 and 15 years old groups which were 0.14 and 0.02 years overestimation respectively (Buken et al, 2010). Their study showed trends of underestimation which was in contrast with this present study which mainly showed trends of overestimation (Table 4 – 7).

Hackman and Black (Hackman and Black, 2013) studied the reliability of Greulich and Pyle method in Scotland. They found trends of overestimation in younger females and underestimation in younger males, which were in the same way as this present study. However, their results showed up to 0.47 years overestimation in 9 – 16 years old females while this present study showed 0.37 – 1.39 years overestimation in female subjects of the same age groups. For males, they found less than 1 month underestimation in 10 – 12 years old males and up to 0.92 years overestimation during 13 – 17 years old. This present study found continuous change from 0.41 years underestimation in 10 years group to 1.25 years overestimation in 17 years group that showed a little higher maturation rate. They found no significant difference between the estimated age and the chronological age (Hackman and Black, 2013) while this present study found significant difference between Greulich and Pyle age and the chronological age in both sexes (Table 4 – 7).

Arciniega-Ramos and colleagues (Arciniega Ramos et al, 2013) compared Fishman skeletal age and the chronological age in 8 - 14 years old patients in Mexico. Their result showed significant difference between them as around 1 year overestimation. However, no more detail on the difference was shown (Arciniega Ramos et al, 2013). This present study also showed around 1 year overestimation for each SMI (Table 17).

Pinchi and colleagues (Pinchi et al, 2014) studied the accuracy of age estimation methods in Italy. Although mean estimated ages were not presented, they found significant differences between 'Greulich and Pyle' age and the chronological age in 9 – 12 years old females. This present study found significant differences in 10 – 13 and 15 years old females that showed coincidence with their result. For males, they found significant differences in 8 – 9 and 12 - 13 years groups while this present study showed significant differences in 11 and 13 – 15 years old males. For Tanner and Whitehouse 3 RUS method, they found significant differences only in 11, 13 and 15 years old females, and 10 years old males. This present study found significant difference in 10 – 13 and 15 years old females, and 10 and 14 years old males thus, was considered as similar to theirs. However, in total, they found significant difference only between Greulich and Pyle age and the chronological age in females (Pinchi et al, 2014), while this present study found significant differences in both methods, except for Tanner and Whitehouse 3 RUS in males (Table 4 – 7).

Alcina and colleagues (Alcina et al, 2015) studied the application of Greulich and Pyle in 10 – 18 years old Spanish. They found 0.02 – 1.04 years overestimation in 10 – 16 years old females which the most derangement found around 13 years old, and

0.24 – 0.57 years underestimation in 17 – 18 years groups. For males, they found less than 1 month derangement in 12 - 13 years groups while other groups were found 0.21 – 0.64 years overestimation, except 18 years group which showed 0.98 years underestimation (Alcina et al, 2015). This present study found similar trend in female subjects with somewhat more overestimation, but male subjects were found underestimation at 10 years group and higher degree of overestimation in older groups (Table 4 - 7).

De Sousa Dantas and colleagues (de Sousa Dantas et al, 2015) studied the reliability of Greulich and Pyle method in 5 – 18 years old Brazilian. They found 8 months (0.67 year) overestimation in average for females which was statistically significant. However, there was only less than a month difference in males which, in turn, was not significant (de Sousa Dantas et al, 2015). This present study found significant overestimation in both females and males which were 0.90 and 0.42 years, respectively (Table 4 – 7). Although these two studies based on different age ranges, an eleven-year intersection should be enough to infer that contemporary Thais were relatively more advanced in skeletal maturation.

A meta-analysis on the accuracy of age estimation methods was done by Serinelli and colleagues (Serinelli et al, 2011). They divided the studied population into Caucasian and Mongoloid groups. For Caucasian, they found non-significant overestimation in males for both 'Greulich and Pyle' and 'Tanner and Whitehouse 3 RUS' methods. In Caucasian females, non-significant overestimation was found for 'Greulich and Pyle' method and non-significant underestimation was found for 'Tanner and Whitehouse 3 RUS'. For Mongoloid, they found non-significant underestimation in males for both methods while significant overestimation was found in females for 'Greulich and Pyle' method and non-significant overestimation for 'Tanner and Whitehouse 3 RUS' (Serinelli et al, 2011). Although the meta-analysis showed mostly non-significant deviations, each analysis included contradicting studies that discordances between nations were present. In addition, the studies included for Mongoloid analyses were mostly from Brazil with only one Asian study for each method thus the result may not absolutely represent the whole Mongoloid ethnics since the majorities of Brazilian were Caucasian (branca) and mixed-ancestry, according to a report from The Brazilian Institute of Geography and Statistics (Petruccelli and Saboia, 2013).

Generation difference

Another factor which may contribute to the discrepancy was the generation difference. This study refers 'contemporary' as the people born between 1991 and 2008 while the reference population of the age estimation methods applied in this study were mostly born before 1990. Comparisons with previous studies in Thai population (Krailassiri et al, 2002, Mathurasai and Viteporn, 1983, Panmethis et al, 2010) illustrated secularly advancing skeletal maturation. However, no previous study on Tanner and Whitehouse 3 method was found in Thai population.

Mathurasai and Viteporn (Mathurasai and Viteporn, 1983) analyzed skeletal ages from hand and wrist radiographs of subjects in Bangkok, in 1983. The skeletal ages were related to the radiographic atlas of 'Greulich, Waterhouse and Pyle' (Greulich et al, 1971), a revised version of Greulich and Pyle atlas, which each representative radiograph can be matched to both male and female skeletal ages. They referred median as the result for each age groups. Although the mean chronological age and difference were not reported, their results rigidly showed underestimation in 8 – 16 years old females and 8 – 11 years old males, while 15 – 16 years old males was overestimated. Their representative radiographs for each 1-year age group were shown in the report (Mathurasai and Viteporn, 1983). Therefore, in order to compare, the radiographs were re-estimated using Greulich and Pyle method by this present author. The Greulich and Pyle ages were distinctly underestimated in 8 – 9 years old females while not clearly underestimated in 10 – 16 years old subjects. In males, only 8 – 10 years old subjects showed strong (around 2 years) underestimation while 11 – 16 years old subjects appeared to be accurate. Contrary, contemporary Thai females showed overestimation from 8 to 16 years old, and males showed less underestimation at 8 – 10 years old and around 1 year overestimation at 11 - 16 years old (Table 4 – 7). Therefore, contemporary Thais were more advanced in skeletal maturation.

Krailassiri, Anuwongnukroh and Dechkunakorn (Krailassiri et al, 2002) studied the relationships between dental calcification stages and skeletal maturity indicators in

Thai subjects. Their skeletal maturity indicators were related to Fishman method and matched with skeletal age from Greulich and Pyle method. However, only few details on the skeletal age were shown. From their result, the mean chronological ages for females were 9.7 years old at MP_3 stage (Fishman's SMI 2), 10.2 years old at S stage (Fishman's SMI 4), 11.4 years old at MP_{3cap} stage (Fishman's SMI 6), 12.6 years old at DP_{3u} stage (Fishman's SMI 8) and 14.1 years old at MP_{3u} stage (Fishman's SMI 10). For male subjects, the mean chronological ages were 11.2, 11.6, 13.2, 14.3 and 15.4 years old respectively. Compared with their results, this present study found a little lower mean chronological age at SMI 2 through SMI 11 in females and distinctly lower mean chronological age in males (Table 16). However, for Greulich and Pyle age, their results showed 0.6 years underestimation to 1.6 years overestimation in females and 0.4 – 2.1 years overestimation in males (Krailassiri et al, 2002). Comparing with this present study, no obvious advanced or delayed development was found in females while male subjects in this present study tended to be less advanced for Greulich and Pyle method (Table 4 - 7).

Panmethis, Sirapallop and Tepprasong (Panmethis et al, 2010) studied the application of Greulich and Pyle method by two radiologists in Lamphun Hospital, Thailand, in 2010. In females, they found 0.24 and 0.26 years underestimation in 6 – 12 years old subjects, and 0.94 and 0.11 years overestimation in 12 – 19 years subjects. In males, they found 0.90 and 0.85 years underestimation in 6 – 12 years old subjects, and

0.50 and 0.44 years overestimation in 12 – 19 years old subjects (Panmethis et al, 2010). This present study found overestimation (0.83 years in average) in 8 -11.99 years old females and more overestimation (1.03 years in average) in 12 – 18.99 years old females. For males, this present study found less underestimation (0.02 years in average) in 8 – 11.99 years old subjects and more overestimation (0.88 years in average) in 12 – 18.99 years old subjects (Table 4 – 7). Although geographic difference may be a contributing factor as Lumphun Hospital is in northern region while Chulalongkorn University locates in central part of Thailand, this present study showed advancing change in skeletal maturity.

Variations of bone development

Skeletal developmental variation could be a cause of decisional dilemma especially in visual comparison based on the Greulich and Pyle method. In this study, developmental variations were remarkably found on ulna and carpal bones. The development of ulna was distinctly delayed in some cases which the epiphysis of ulna was absent while other bones were progressively developed. The variation of ulna was also found in the morphological pattern since some cases showed a wide ulna epiphysis but poorly recognized styloid process while some cases showed smaller epiphysis with prominent styloid process (Figure 17). Variations of ulnar epiphysis in Thai subjects were also mentioned in the report of Mathurasai and colleagues (Mathurasai and Viteporn, 1983). Thus, the development of ulna bone may be related to ethnic groups.

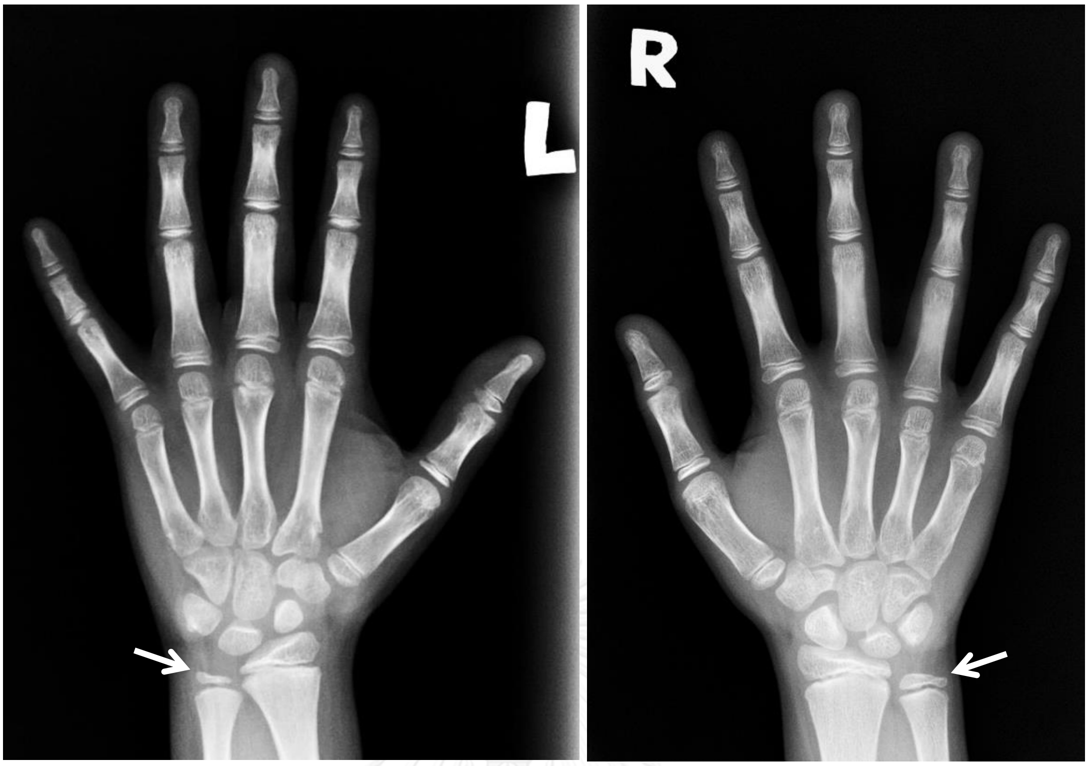


Figure 17 A hand and wrist radiograph showing ulna epiphysis with prominent styloid process (left) and a hand and wrist radiograph showing wider ulna epiphysis with poorly defined styloid process (right).

Observer dependence

Since the three age estimation methods in this study rely on visual recognition, the derangements found in the results were probably contributed by subjective decisional variation. Intra-observer and inter-observer reliability tests were applied in response to the concern. Weighted kappa analysis was selected, as ordinal data was collected for the age estimation, and the results showed good agreement for both intra-observer and inter-observer with slightly lower reliability for inter-observer agreement.

This was expected. The results might be resulted from experiences of the observers; therefore, good calibration prior to the observation is highly recommended.

Time consumption

Time consumption should be another issue of concern. Although Tanner and Whitehouse method was shown as the most accurate method, it tended to be the most time consuming one. During the study, an hour of application was taken to find the results of approximately 10 subjects using Tanner and Whitehouse 3 RUS method, while over 40 subjects were done in an hour using Fishman method. In case of a comprehensive examination on an individual subject, the most accurate method would be preferred with only little concern about the time. However, in some situations where time limitation is an important factor, a faster method may be preferred. Therefore, selecting a proper method for each situation is still important. This is also a challenge for further development to find a more accurate method with reduced time consumption.

Study limitations

Positioning errors

The hand and wrist radiographs in this study were primarily taken for a simple evaluation on skeletal growth before starting an orthodontic treatment; therefore, the position of the hand was sometimes not properly adjusted. Some bones may not be aligned parallel to the x-ray detector, causing overlapping and superimposition. The overlapping of structures made it difficult to observe the stages of bone development, especially the capping stage of an epiphysis. The thumb finger was the most problematic part as the finger torsion was quite different from which was illustrated for the age estimation methods. However, the overall quality of the radiographs was still acceptable.

Number of samples

Limited number of radiographs in some age groups was included in this study due the lack of radiographs in the database. Orthodontists only refer patients to have hand and wrist radiographs taken during certain age when the growth of patients must be determined. Patients who are younger than 8 years old and older than 17 years old are usually not referred for hand and wrist radiographs. More samples are needed to improve the strength of the study especially in some age groups.

Clinical applications

The present study has shown that Tanner and Whitehouse method was the most accurate age estimation method for contemporary Thai children and adolescents. Therefore, Tanner and Whitehouse 3 RUS method is suggested for age estimation using hand and wrist radiographs in contemporary Thais aged 8 - 14 years old. However, the results in 13 – 14 years old subjects should be carefully considered as more than half a year overestimation might be evident. Nonetheless, the concept of Fishman method may also be considered for further development because of its simple procedure. More studies are needed to further evaluate the applications of the methods and possibly find a proper adaptation of the methods which is specific for Thai population.

The hand and wrist age estimation method can be applied in various situations as a complement with other information. In immigrant registration, the estimated age is made in conclusion from physical examination and written documents while skeletal age and psychological evaluation may be included in particular subjects. In case of an unknown body, hand and wrist skeletal age is considered with estimated ages from other parts of the body including dental age before making the final result. In growth observation and legal judgment, only specific cases, such as contradicting physical signs or absence of age-defying document, will have hand and wrist skeletal age estimation. In addition, this study was based on subjects without diseases affecting skeletal development thus, the findings from this study may only represent contemporary Thai subjects with normal skeletal development.

Conclusion

From this Tanner and Whitehouse 3 RUS method was found to be the most accurate among the three age estimation methods. Therefore, it was recommended for age estimation in contemporary Thais aged over 8 years old, but with careful considerations because of the varying degree of differences among age groups. Further study on the adaptation of the method is necessary for more accurate results.

REFERENCES

1. Acheson RM. A method of assessing skeletal maturity from radiographs; a report from the Oxford child health survey. J Anat. 1954;88:498-508.
2. Aggrawal A. Estimation of age in the living: in matters civil and criminal. J Anat. 2009:1-17.
3. Alcina M, Lucea A, Salicrú M, Turbón D. Reliability of the Greulich & Pyle method for bone age estimation in a Spanish sample. J Forensic Leg Investig Sci. 2015;1:1-6.
4. Arciniega Ramos NA, Ballesteros Lozano M, Meléndez Ocampo A. Comparative analysis between dental, skeletal and chronological age. Rev Mex Ortodon. 2013;1:e33-e7.
5. Bacchetta J, Harambat J, Cochat P, Salusky IB, Wesseling-Perry K. The consequences of chronic kidney disease on bone metabolism and growth in children. Nephrol Dial Transplant. 2012;27:3063-71.
6. Baer MJ, Djrkatz J. Bilateral asymmetry in skeletal maturation of the hand and wrist: a roentgenographic analysis. Am J Phys Anthropol. 1957;15:181-96.
7. Berk LE. Child development. 9 ed. London, UK: Pearson Education; 2012.
8. Beunen GP, Rogol AD, Malina RM. Indicators of biological maturation and secular changes in biological maturation. Food Nutr Bull. 2006;27:S244-56.
9. Black S, Ferguson E. Forensic anthropology: 2000 to 2010. Boca Raton, FL: CRC Press; 2016.
10. Bogin B, MacVean RB. The relationship of socioeconomic status and sex to body size, skeletal maturation, and cognitive status of Guatemala City school children. Child Dev. 1983;54:115-28.
11. Buken B, Safak AA, Yazici B, Buken E, Mayda AS. Is the assessment of bone age by the Greulich-Pyle method reliable at forensic age estimation for Turkish children? Forensic Sci Int. 2007;173:146-53.

12. Buken B, Erzengin OU, Buken E, Safak AA, Yazici B, Erkol Z. Comparison of the three age estimation methods: which is more reliable for Turkish children? Forensic Sci Int. 2009;183:103 e1-7.

13. Buken B, Safak AA, Buken E, Yazici B, Erkol Z, Erzengin ÖU. Is the Tanner–Whitehouse (TW3) method sufficiently reliable for forensic age determination of Turkish children?*. Turk J Med Sci. 2010;40:797-805.

14. Chiang K-H, Chou AS-B, Yen P-S, Ling C-M, Lin C-C, Lee C-C, et al. The reliability of using Greulich-Pyle method to determine children's bone age in Taiwan. Tzu Chi Med J. 2005;17:417-20.

15. Chow SC, Wang H, Shao J. Sample Size Calculations in Clinical Research. 2 ed. Boca Raton, FL: CRC Press; 2007.

16. 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. 2015;100:138-43.

17. de Sousa Dantas I, dos Anjos Pontual A, Almeida MSC, de Lucena MIHM, Beltrão RT, de Moraes Ramos-Perez FM, et al. Evaluation of the Greulich and Pyle method in the determination of bone age and chronological age in a Brazilian population. Derecho y Cambio Social 2015;12:1-14.

18. Dharman.M.K N, Moses JC, Olivete CJ, Dias MA. Survey on different bone age estimation methods. Int j adv res comput sci softw eng 2014;4:1128-31.

19. Fishman LS. Radiographic evaluation of skeletal maturation. A clinically oriented method based on hand-wrist films. Angle Orthod. 1982;52:88-112.

20. Fishman LS. Maturational patterns and prediction during adolescence. Angle Orthod. 1987;57:178-93.

21. Fitzpatrick LA. Sex differences in skeletal development. Advances in Molecular and Cell Biology. Volume 34. Amsterdam, Netherland: Elsevier; 2004. p. 229-45.

22. Flory CD. Sex differences in skeletal development. Child Dev. 1935;6:205-12.

23. Flory CD. Osseous Development in the Hand as an Index of Skeletal Development. Monogr Soc Res Child Dev. 1936;1:i+iii-ix+1-141.

24. Franklin D, Flavel A, Noble J, Swift L, Karkhanis S. Forensic age estimation in living individuals: methodological considerations in the context of medico-legal practice. Res Rep Forensic Med Sci. 2015;5:53-66.

25. Greulich W, Pyle S, Waterhouse A. A radiographic standard of reference for the growing hand and wrist. Chicago, IL: Press of Case Western Reserve University; 1971.

26. Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. Stanford, CA: Stanford University Press; 1959.

27. Griffith J, Cheng J, Wong E. Are western skeletal age standards applicable to the Hong Kong Chinese population? A comparison of the Greulich and Pyle method and the Tanner and Whitehouse method. Hong Kong Med J. 2007;13:S28-32.

28. Hackman L. Age estimation in the living: a test of 6 radiographic methods: University of Dundee; 2012.

29. Hackman L, Black S. Does mirror imaging a radiograph affect reliability of age assessment using the Greulich and Pyle atlas? J Forensic Sci. 2012;57:1276-80.

30. Hackman L, Black S. The reliability of the Greulich and Pyle atlas when applied to a modern Scottish population. J Forensic Sci. 2013;58:114-9.

31. Haiter-Neto F, Kurita LM, Menezes AV, Casanova MS. Skeletal age assessment: a comparison of 3 methods. Am J Orthod Dentofacial Orthop. 2006;130:435 e15-20.

32. Halvorsen K, Kristensen S. Separated children in Europe programme workshop on age assessment and identification. Bucharest; 2003.

33. Hawley NL, Rousham EK, Norris SA, Pettifor JM, Cameron N. Secular trends in skeletal maturity in South Africa: 1962-2001. Ann Hum Biol. 2009;36:584-94.

34. Hewitt D, Westropp CK, Acheson RM. Oxford child health survey; effect of childish ailments on skeletal development. Br J Prev Soc Med. 1955;9:179-86.

35. Hsieh CW, Liu TC, Jong TL, Tiu CM. Long-term secular trend of skeletal maturation of Taiwanese children between agricultural (1960s) and contemporary (after 2000s) generations using the Tanner-Whitehouse 3 (TW3) method. J Pediatr Endocr Met. 2013;26:231-7.

36. 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;16:201-5.

37. 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;13:50-5.

38. King DG, Steventon DM, O'Sullivan MP, Cook AM, Hornsby VP, Jefferson IG, et al. Reproducibility of bone ages when performed by radiology registrars: an audit of Tanner and Whitehouse II versus Greulich and Pyle methods. Br J Radiol. 1994;67:848-51.

39. Kliegman RM, Stanton BF, St Geme JW, Schor NF. Nelson textbook of pediatrics. 20 ed. Philadelphia, PA: Elsevier; 2015.

40. Krailassiri S, Anuwongnukroh N, Dechkunakorn S. Relationships between dental calcification stages and skeletal maturity indicators in Thai individuals. Angle Orthod. 2002;72:155-66.

41. Lai EHH, Chang JZC, Yao CCJ, Tsai SJ, Liu JP, Chen YJ, et al. Relationship between age at menarche and skeletal maturation stages in Taiwanese female orthodontic patients. J Formos Med Assoc. 2008;107:527-32.

42. Manzoor Mughal A, Hassan N, Ahmed A. Bone age assessment methods: a critical review. Pak J Med Sci. 2014;30:211-5.

43. Mathurasai W, Viteporn S. The relationship between the chronological age and growth of the hand and wrist bones in Thai, at the age of 8 - 16. Chulalongkorn University, Orthodontics; 1983 November. Report No.

44. Mohammed RB, Kalyan VS, Tircouveluri S, Vegesna GC, Chirla A, Varma DM. The reliability of Fishman method of skeletal maturation for age estimation in children of South Indian population. J Nat Sci Biol Med. 2014a;5:297-302.

45. Mohammed RB, Reddy MA, Jain M, Singh JR, Sanghvi P, Thetay AA. Digital radiographic evaluation of hand-wrist bone maturation and prediction of age in South Indian adolescents. HAND. 2014b;9:375-83.

46. Molinari L, Gasser T, Largo RH. TW3 bone age: RUS/CB and gender differences of percentiles for score and score increments. Ann Hum Biol. 2004;31:421-35.

47. Panmethis M, Sirapallop O, Tepprasong N. Bone age Assessment by Greulich and Pyle Method in Thai Children at Lamphun Hospital. Journal of Health Science-วารสาร วิชาการ สาธารณสุข. 2010;19:190-7.

48. Paxton ML, Lamont AC, Stillwell AP. The reliability of the Greulich–Pyle method in bone age determination among Australian children. J Med Imaging Radiat Oncol. 2013;57:21-4.

49. Petruccelli JL, Saboia AL. Características étnico-raciais da população: classificações e identidades. RJ, Brazil: IBGE; 2013.

50. Pinchi V, De Luca F, Ricciardi F, Focardi M, Piredda V, Mazzeo E, et al. Skeletal age estimation for forensic purposes: A comparison of GP, TW2 and TW3 methods on an Italian sample. Forensic Sci Int. 2014;238:83-90.

51. Pludowski P, Litwin M, Niemirska A, Jaworski M, Sladowska J, Kryskiewicz E, et al. Accelarated skeletal maturation in children with primary hypertension. Hypertension. 2009;54:1234-9.

52. Saade A, Baron P, Noujeim Z, Azar D. Dental and Skeletal Age Estimations in Lebanese Children: A Retrospective Cross-sectional Study. J Int Soc Prev Community Dent. 2017;7:90-7.

53. Satoh M. Bone age: assessment methods and clinical applications. Clin Pediatr Endocrinol. 2015;24:143-52.

54. Schoenwolf GC, Bleyl SB, Brauer PR, Francis-West PH. Larsen's human embryology. Philadelphia, PA: Elsevier Health Sciences; 2014.

55. Serinelli S, Panetta V, Pasqualetti P, Marchetti D. Accuracy of three age determination X-ray methods on the left hand-wrist: a systematic review and meta-analysis. Leg Med (Tokyo). 2011;13:120-33.

56. Sharma RK. Concise textbook of forensic medicine & toxicology. New Delhi, India: Elsevier India Pvt. Limited; 2007.

57. Smith T, Brownlees L. Age assessment practices: a literature review & annotated bibliography. New York, NY: Child Protection Section, United Nations Children’s Fund; 2011. p. 1-78.

58. Stanhope R, Fry V. Growth and growth disorders. 3rd ed. London, UK: The Child Growth Foundation; 2000.

59. Tanner J, Healy M, Goldstein H, Cameron N. Assessment of skeletal maturity and prediction of adult height (TW3). London, UK: WB Saunders; 2001.

60. Tanner JM, Whitehouse R, Marshall W, Healty M, Goldstein H. Assessment of skeletal maturity and prediction of adult height (TW2 method). Cambridge, MA: Academic Press; 1975.

61. Zafar AM, Nadeem N, Husen Y, Ahmad MN. An appraisal of Greulich-Pyle Atlas for skeletal age assessment in Pakistan. J Pak Med Assoc. 2010;60:552-5.

62. Zemel BS, Kawchak DA, Ohene-Frempong K, Schall JI, Stallings VA. Effects of delayed pubertal development, nutritional status, and disease severity on longitudinal patterns of growth failure in children with sickle cell disease. Pediatr Res. 2007;61:607-13.

63. 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;250:228-35.

APPENDIX

VITA

Mr. Sunpatch Benjavongkulchai was born on 26 April 1990 at Chulalongkorn Hospital, Bangkok, Thailand. He graduated from the Faculty of Dentistry, Chulalongkorn University, with second honor on Doctor of Dental Surgery in 2014, and worked as an instructor in the Department of Radiology, Faculty of Dentistry, Chulalongkorn University for 1 year.