

Original Article

Utilization of an Artificial Intelligence Program Using the Greulich-Pyle Method to Evaluate Bone Age in the Skeletal Maturation Stage

Jihoon Kim¹, Hyejun Seo^{1,2}, Soyoung Park³, Eungyung Lee^{1,3}, Taesung Jeong^{1,3}, Ok Hyung Nam^{4,5}, Sungchul Choi^{4,5}, Jonghyun Shin^{1,3}

¹Department of Pediatric Dentistry, School of Dentistry, Dental and Life Science Institute, Pusan National University, Yangsan, Republic of Korea

²Department of Dentistry, Ulsan University Hospital, Ulsan, Republic of Korea

³Department of Pediatric Dentistry, Dental Research Institute, Pusan National University Dental Hospital, Yangsan, Republic of Korea

⁴Department of Pediatric Dentistry, Kyung Hee University Dental Hospital, Seoul, Republic of Korea

⁵Department of Pediatric Dentistry, School of Dentistry, Kyung Hee University, Seoul, Republic of Korea

Abstract

The purpose of this study was to measure bone age using an artificial intelligence program based on the Greulich-Pyle (GP) method to find out the bone age corresponding to each stage of cervical vertebral maturation (CVM) and the middle phalanx of the third finger (MP3). This study was conducted on 3,118 patients who visited pediatric dentistry at Kyung Hee University Dental Hospital and Pusan National University Dental Hospital from 2013 to 2021. The CVM stage was divided into 5 stages according to the classification by Baccetti, and the MP3 stage was divided into 5 stages according to the methods of Hägg and Taranger. Based on the GP method, bone age was evaluated using an artificial intelligence program. The pubertal growth spurt in the CVM stage was CVM II and III. The mean bone age in CVM II was 11.00 ± 1.81 years for males and 10.00 ± 1.49 years for females, and in CVM III, 13.00 ± 1.46 years for males and 12.00 ± 1.44 years for females ($p < 0.0001$). The pubertal growth spurt in the MP3 stage was MP3 - G stage. The bone age at the MP3 - G stage was 13.14 ± 1.07 years for males and 11.40 ± 1.09 years for females ($p < 0.0001$). Bone age evaluation using artificial intelligence is worth using in clinical practice, and it is expected that a faster and more accurate diagnosis will be possible. [J Korean Acad Pediatr Dent 2023;50(1):89-103]

Keywords

Bone age, Cervical vertebral maturation, Middle phalanx of the third finger, Greulich-Pyle method, Artificial intelligence, Pubertal growth spurt

ORCID

Jihoon Kim

<https://orcid.org/0000-0001-5817-0617>

Hyejun Seo

<https://orcid.org/0000-0003-2514-8688>

Soyoung Park

<https://orcid.org/0000-0003-3264-4014>

Eungyung Lee

<https://orcid.org/0000-0001-8313-8571>

Taesung Jeong

<https://orcid.org/0000-0002-0431-5574>

Ok Hyung Nam

<https://orcid.org/0000-0002-6386-803X>

Sungchul Choi

<https://orcid.org/0000-0001-7221-2000>

Jonghyun Shin

<https://orcid.org/0000-0002-9777-0196>

Article history

Received October 7, 2022

Revised November 29, 2022

Accepted November 30, 2022

© 2023 Korean Academy of Pediatric Dentistry

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Corresponding author: Jonghyun Shin

Department of Pediatric Dentistry, School of Dentistry, Pusan National University, 20 Geumo-ro, Mulgeum-eup, Yangsan, 50612, Republic of Korea

Tel: +82-55-360-5183 / Fax: +82-55-360-5174 / E-mail: jonghyuns@pusan.ac.kr

Funding information

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2020R1G1A1011629).

서론

소아 청소년기의 성장 상태를 평가하는 데는 키, 체중, 머리 둘레, 체질량지수, 피부 주름 두께, 치아의 맹출 시기와 같은 다양한 방법이 존재한다. 개인의 성장은 다양한 요인이 복합적으로 작용하므로 소아 청소년의 성장 평가는 유전적, 환경적 배경, 출생, 체중, 성장 속도의 정상 혹은 비정상, 골령 측정 등의 여러 요소가 평가되어야 한다[1,2]. 이 중에서 골령은 소아 청소년의 현재와 미래의 성장 양상을 예측할 수 있어 성장 평가에 매우 중요하다[3].

소아 청소년의 골 성숙도를 평가하는 것은 개인의 다양한 발달 및 성장 상태를 이해할 수 있으며, 이를 바탕으로 부정교합의 적절한 치료 시기를 결정하는 데 있어 중요한 역할을 한다. 또한 개인의 성장을 평가하고 예측하는 것은 치료 결과나 장기적인 예후에도 영향을 미칠 수 있으므로 치료 전 정확한 평가가 중요하다[4,5].

개인의 성장 발육을 평가하기 위해 사용되는 발육 연령과 연대 연령은 발달 상태에 따른 개인 차이가 크고, 인종과 세대에 따른 차이가 있어 정확한 평가가 어렵다[6,7]. 따라서 보다 정확한 평가를 위해 이용되는 골 성숙도 평가는 성장에 따라 예측할 수 있는 골화 중심의 변화를 이용하는데, 골화 중심의 변화 정도가 모든 사람에게서 비슷하게 진행되고 재현성이 우수하므로 성장기 어린이의 성장 발달의 평가 방법으로 가장 많이 이용된다[8,9].

임상에서 골 성숙도를 평가하기 위해 일반적으로 사용하는 방법은 경추골과 수완부골의 방사선사진을 이용하는 것이다. 경추골 성숙도 평가는 교정 치료 전 진단 자료로 촬영하는 측방 두부규격 방사선사진을 이용한다[10,11]. 수완부 방사선사진은 측방두부규격 방사선사진과 함께 골 성숙 단계 평가를 위한 수단으로 사용되며, 판독이 용이하고 재현성이 우수하며 방사선 노출이 비교적 적은 장점이 있지만 일반적으로 촬영하는 측방두부규격 방사선사진과 별도로 추가 촬영해야 하는 단점이 존재한다[12,13].

수완부골은 다양한 부위의 골들이 성숙함에 따라 각각의 구성 골들의 화골 현상의 개시 및 정도와 골의 변화 양상이 다르므로 개개인의 골 성숙도를 유추할 수 있다[14,15]. 수완부골을 이용해 골령을 평가하는 기존의 방법 중에서 가장 대표적인 것으로 Greulich-Pyle (GP)와 Tanner-Whitehouse (TW)방법이 있다. GP 방법은 표본 방사선사진이 나열된 GP 도감과 수완

부 방사선사진을 대조하여 골령을 평가하고, TW 방법은 수완부의 부위별 골 성숙도에 따른 점수를 합산하여 골령을 측정한다[16,17]. TW 방법은 GP 방법에 비해 정확하다고 알려져 있지만 숙련되기까지 다소 시간이 걸리며, 평가 시간이 오래 걸려 임상에서 바로 적용하기에 어려운 단점이 있다. 이에 비해 GP 방법은 골령을 빠르고 간편하게 측정할 수 있어 임상에서 더 쉽게 이용이 가능한 장점이 있다[18-20]. 이러한 GP 방법의 장점을 이용한 골령 평가의 임상적 적용에 관한 기존의 여러 연구를 통해 다양한 인종 및 지역에서 GP 방법의 적용 가능성과 신뢰도를 확인할 수 있었다[21-24].

최근 인공지능의 발전으로 치의학의 여러 분야에 인공지능을 활용한 진단 기술이 적용되고 있고 술자의 진단에 보조 역할을 하고 있다[25,26]. 인공지능이란 기계가 사람처럼 합리적인 생각과 행동을 할 수 있는 시스템을 갖도록 만드는 기술을 말한다[28]. 인공지능의 범주 내에 속하는 하위 개념으로 기계학습(machine learning)과 심층학습(deep learning)이 있다. 기계학습은 주어진 데이터와 알고리즘을 이용해 규칙을 학습하고 판단하는 시스템이다. 심층학습은 기계학습의 하위 개념으로 인간의 뇌에 있는 뉴런의 정보 처리 방식에서 영감을 받은 인공 신경망(artificial neural network)을 기반으로 하여 기계가 스스로 학습하고 데이터를 분석 및 판단하는 시스템이다[27,28]. 이러한 인공지능 기술은 골령을 평가하는 데도 이용되고 있으며, 이를 활용함에 따라 임상에서 보다 빠르고 쉽게 골령을 평가할 수 있게 되었고, 정확도 역시 신뢰할 만한 수준이었다[29,30].

소아 청소년의 골격 발달에서 경추골과 수완부골의 성숙단계 변화를 사춘기 성장과 비교 시 상호 연관성이 있었으며, 이를 연구한 기존의 연구는 대부분이 소아 청소년의 경추골 성숙단계와 수완부골의 단계를 통해 골 성숙단계를 비교한 논문이었다[31-33]. 하지만 임상에서 흔히 사용되는 경추골과 수완부골의 성숙단계의 평가만으로는 잔여 성장의 양이나 각 단계에 해당하는 상세한 골령을 파악하기 어려워 정확한 성장을 평가하는 것은 한계가 있었다[34].

그러나 골령 평가 시 임상에서 쉽게 이용 가능한 GP 방법을 통해 골 성숙도와 연관 지은 국내 아동 대상의 연구는 드물었고, 특히 인공지능의 심층학습 기술을 이용하여 골령을 평가하고 이를 활용한 연구는 드물었다. 따라서 이번 연구에서는 GP 방법을 기반으로 한 인공지능 프로그램을 이용해 골령을 자동으로 평가하고, 경추골 성숙도(Cervical vertebral

maturation, CVM)단계 및 중지 중절골(Middle phalanx of the third finger, MP3)의 골 성숙도를 평가하여 각 단계에 해당하는 골령을 파악하고자 하였다. 이를 통해 국내 소아 청소년의 사춘기 성장기에 해당하는 부정교합 환자의 더욱 정확한 성장 평가와 치료 계획 수립에 도움이 되는 기초 자료를 제공하고 국내 아동을 대상으로 인공지능을 이용한 골령 평가의 임상적 적용 가능성을 제시하고자 하였다.

연구 재료 및 방법

이 연구는 경희대학교 치과병원 임상 연구 윤리 위원회 (Institutional Review Board, IRB)의 승인(IRB No.: KH-DT21011)과 부산대학교 치과병원 연구 윤리 심의 위원회(IRB)의 승인을 받아 시행되었다(IRB No.: PNUDH-2021-032).

1. 연구 대상

이 연구는 2013년 6월부터 2021년 9월까지 경희대학교 치과병원 소아치과와 부산대학교 치과병원 소아치과에 내원한 만 7

세부터 18세 사이의 남자 1,529명, 여자 1,589명, 총 3,118명의 소아, 청소년을 대상으로 하였고(Table 1), 동일한 날에 촬영한 측방두부규격 방사선사진과 수완부 방사선사진을 이용하였다. 경추골 및 수완부 골의 이상을 나타내는 전신 질환이 있는 환자, 판독이 불가능한 방사선사진은 연구 대상에서 제외하였다.

2. 연구 방법

이번 연구의 전체적인 과정은 다음과 같다(Fig. 1).

1) 경추골 성숙도 단계 평가

측방두부규격 방사선 사진을 이용한 CVM 단계의 평가는 Baccetti 등[35]이 제시한 방법으로 경추골 2, 3, 4번의 형태에 따라 총 5단계로 분류하였다(Fig. 2).

2) 중지 중절골 성숙 단계 평가

수완부 방사선 사진은 Hägg와 Taranger[36]가 제안한 MP3 방법을 이용해 중지 중절골의 골단과 골간단의 관계에 따라 총 5단계(MP3 - F, FG, G, H, I)로 분류하였다(Fig. 3).

Table 1. Sample distribution of chronological age and bone age for gender

CA (year)	n		BA (year)	n	
	Male	Female		Male	Female
4.0 - 4.9	-	-	4.0 - 4.9	7	-
5.0 - 5.9	-	-	5.0 - 5.9	36	29
6.0 - 6.9	-	-	6.0 - 6.9	139	167
7.0 - 7.9	386	378	7.0 - 7.9	238	304
8.0 - 8.9	371	356	8.0 - 8.9	281	276
9.0 - 9.9	246	233	9.0 - 9.9	193	56
10.0 - 10.9	200	196	10.0 - 10.9	186	232
11.0 - 11.9	130	173	11.0 - 11.9	182	192
12.0 - 12.9	70	74	12.0 - 12.9	100	132
13.0 - 13.9	38	61	13.0 - 13.9	88	93
14.0 - 14.9	28	42	14.0 - 14.9	53	51
15.0 - 15.9	37	44	15.0 - 15.9	16	32
16.0 - 16.9	12	27	16.0 - 16.9	3	22
17.0 - 17.9	10	4	17.0 - 17.9	5	3
18.0 - 18.9	1	1	18.0 - 18.9	2	-
Total	1529	1589		1529	1589

CA: Chronological age; BA: Bone age.

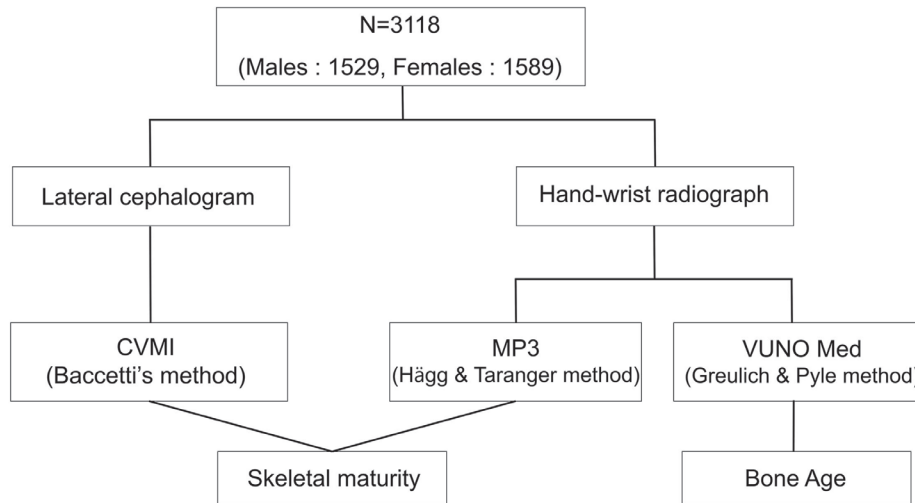


Fig. 1. Schematic summary of the study protocol. CVMI: Cervical vertebral maturation index; MP3: Middle phalanx of the third finger.

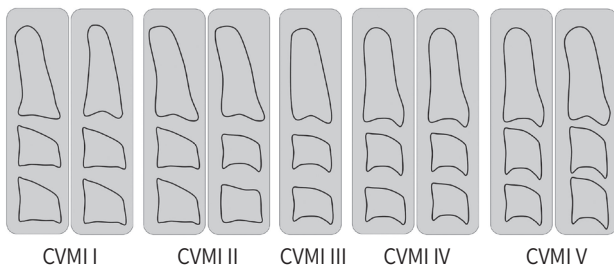


Fig. 2. 5 stages of cervical vertebral maturation. CVMI: Cervical vertebral maturation index.

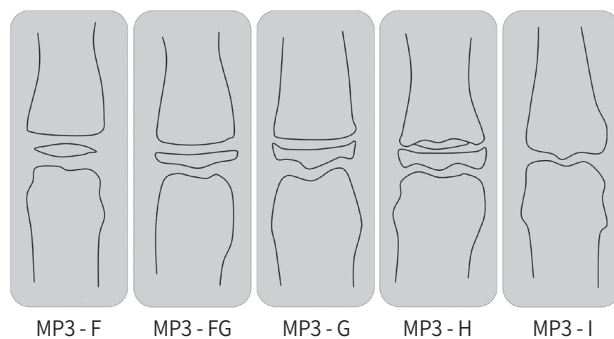


Fig. 3. 5 stages of the middle phalanx of the third finger. MP3: Middle phalanx of the third finger.

3) 골령 평가

수완부 방사선 사진을 이용한 골령 평가를 위해 심층학습 기반의 인공지능 프로그램인 VUNO Med-BoneAge (Version 1.0.3, VUNO Inc., Seoul, South Korea)를 이용하였다(Fig. 4). 이 프로그램은 GP 방법을 기반으로 하며, 입력된 수완부 방사선 사진을 자동으로 평가하여 확률적으로 가장 가능성이 높은 3개의 추정되는 골령 값을 제시한다[30]. 골령의 결정은 프로그램 통해 평가된 골령 중 가장 높은 확률로 제시된 값을 선택하였다.

4) 통계 분석

통계 분석에는 SPSS 26.0 (SPSS Inc., IBM, Chicago, IL, USA)을 사용하였다. 성별에 따른 연대 연령, 골령, CVM 및 MP3 단계의 통계학적 유의성 분석을 위해 Mann whitney U test를 시행하였고, 이들 간의 상관분석을 위해 Spearman's rank-order test를 시행하였다. CVM 단계 및 MP3 단계의 조사자 내 일치도 검사를 위해 소아치과 전공의 1명이 2주 간격을 두고 50명을 무작위로 선정하여 급내 상관 계수(Intra-class correlation coefficient, ICC)를 평가하였다. ICC 값의 평가는 구간별로 0.5 미만일 경우 poor, 0.5 - 0.75일 경우 moderate, 0.75 - 0.9일 경우 good, 0.9 이상일 경우 excellent로 나누었다. ICC 값은 CVM 단계 0.98, MP3 단계 0.97로 모두 0.9 이상으

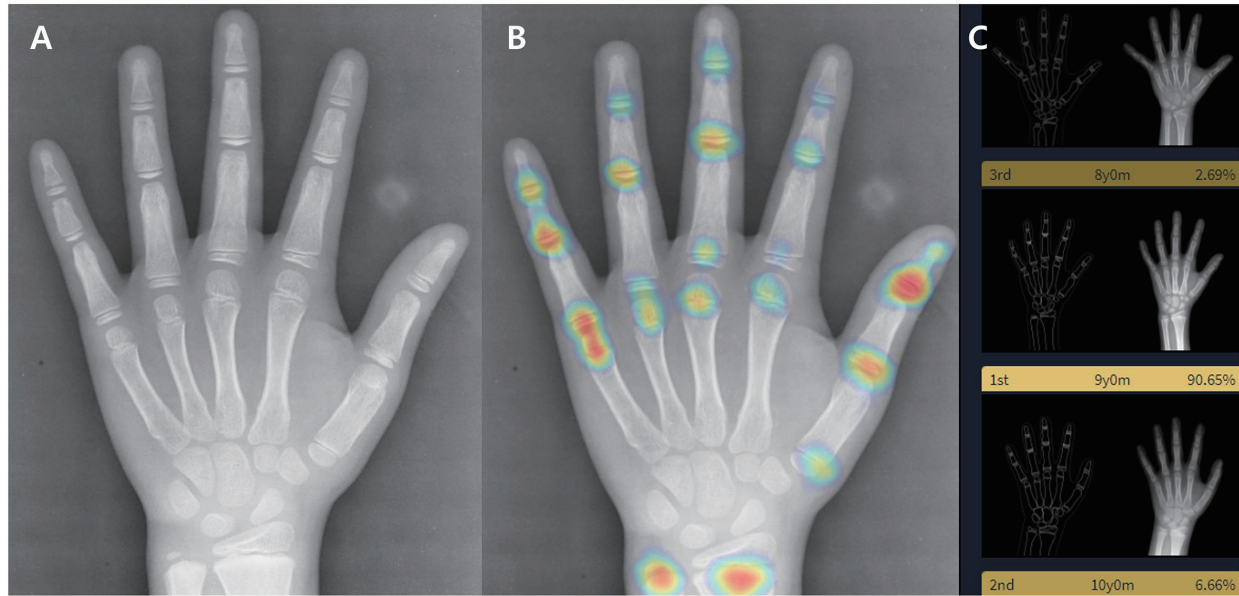


Fig. 4. Automatic bone age assessment using an artificial intelligence program based on the Greulich-Pyle method. (A) Analyze the hand radiographs entered into the software and (B) the shape and density of each bone part were used to form a heat map, then (C) the 3 most probable estimated bone age values are presented as percentages.

로 높은 신뢰도를 나타내었다. 또한 인공지능 프로그램으로 평가한 골령과 조사자가 평가한 골령의 일치도를 평가하기 위해 100명을 무작위로 선정하여 수동으로 골령을 평가하였고 ICC 값은 0.95로 높은 신뢰도를 나타내었다.

연구 성적

1. 성별에 따른 평균 연대 연령과 골령 비교

성별에 따른 평균 연대 연령과 골령은 남자에서 9.53세, 9.32세, 여자에서 9.76세, 9.79세로 남자에게서는 골령이 연대 연령보다 평균 0.21세 낮았고, 여자에게서는 골령이 연대 연령보다 평균 0.03세 높았다. 남자보다 여자에서 연대 연령은 평균 0.23세, 골령은 0.47세 높았다(Table 2). 연대 연령과 골령 비교 시 여자에게서는 유의한 차이는 없었고(Table 2, $p = 0.7642$), 남자에게서는 유의한 차이가 있었다(Table 2, $p = 0.002$). 골령과 연대 연령의 상관성은 남자 0.85, 여자 0.88로 남녀 모두 높은 상관관계를 나타냈다(Table 3, $p < 0.0001$).

Table 2. The mean and standard deviation of chronological age and bone age for gender

Gender	Male	Female	p
CA (year)			
Mean \pm SD	9.53 \pm 2.13	9.76 \pm 2.21	0.002
(range)	(7.00 - 18.00)	(7.00 - 18.17)	
BA (year)			
Mean \pm SD	9.32 \pm 2.51	9.79 \pm 2.33	0.0002
(range)	(4.50 - 18.00)	(5.00 - 17.00)	
p	0.002	0.7642	

p value from Mann-Whitney U test.

CA: Chronological age; BA: Bone age; SD: Standard deviation.

Table 3. Spearman rank correlation coefficient between chronological age and bone age measured with an artificial intelligence program based on the Greulich-Pyle method

	Male		Female		Total	
	r	p	r	p	r	p
CA and BA	0.85	<0.0001	0.88	<0.0001	0.87	<0.0001

p value from Spearman's rank-order test.

CA: Chronological age; BA: Bone age.

2. CVM 단계에 따른 골령 분포

CVM 각 단계에 해당하는 성별에 따른 골령 및 표준편차는 다음과 같다(Table 4, Fig. 5). 남자, 여자 각각 CVM I 단계의 평균 골령은 8.00세, 8.17세, CVM II 단계에서 11.00세, 10.00세, CVM III 단계에서 13.00세, 12.00세, CVM IV 단계에서 14.38세, 14.17세, CVM V 단계에서 17.00세, 16.00세로 CVM I 단계를 제외한 나머지 단계에서 여자가 남자보다 낮은 평균 골령을 보였다.

3. MP3 단계에 따른 골령 분포

MP3 각 단계에 해당하는 성별에 따른 골령 및 표준편차는 다음과 같다(Table 5, Fig. 6). 남자, 여자 각각 MP3 - F 단계의 평균 골령은 8.33세, 8.06세, MP3 - FG 단계에서 11.74세, 9.90세, MP3 - G 단계에서 13.14세, 11.40세, MP3 - H 단계에서 13.92세, 12.50세, MP3 - I 단계에서 15.50세, 14.58세로 MP3의 모든 단계에서 여자가 남자보다 낮은 평균 골령을 보였다.

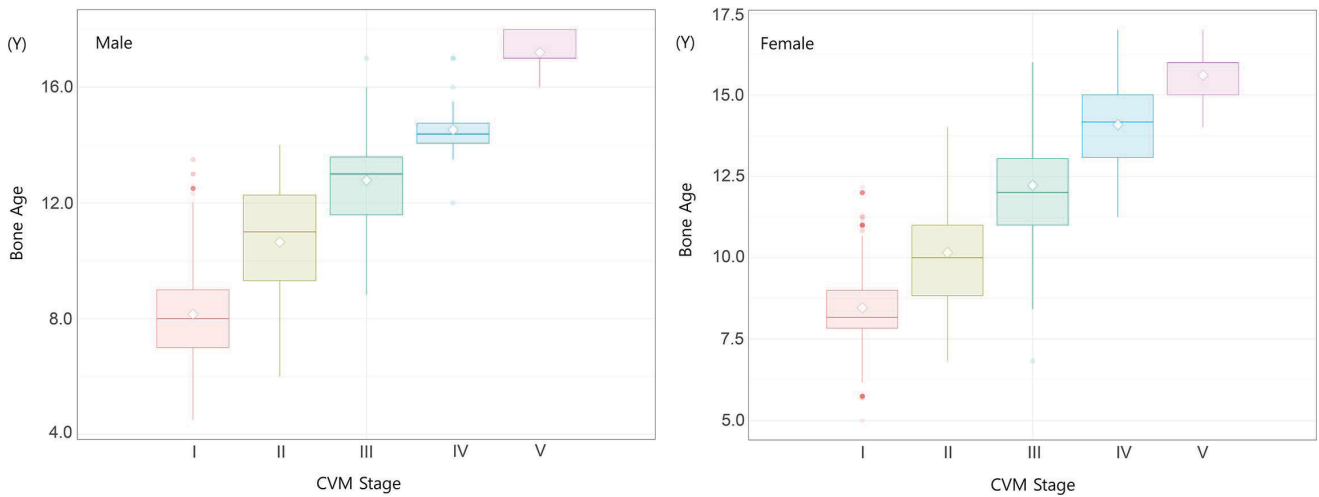


Fig. 5. Bone age distribution according to cervical vertebral maturation stage in males and females. CVM: Cervical vertebral maturation.

Table 4. The mean bone age and standard deviation according to cervical vertebral maturation stage

Stage	Gender	n	BA	p
			Mean ± SD (year)	
CVM I	Male	1031	8.00 ± 1.72	0.093
	Female	930	8.17 ± 1.37	
CVM II	Male	276	11.00 ± 1.81	< 0.0001
	Female	304	10.00 ± 1.49	
CVM III	Male	189	13.00 ± 1.46	< 0.0001
	Female	231	12.00 ± 1.44	
CVM IV	Male	28	14.38 ± 0.98	0.158
	Female	109	14.17 ± 1.25	
CVM V	Male	5	17.00 ± 0.84	0.007
	Female	15	16.00 ± 0.90	

p value from Mann-Whitney U test.

CVM: Cervical vertebral maturation; BA: Bone age; SD: Standard deviation.

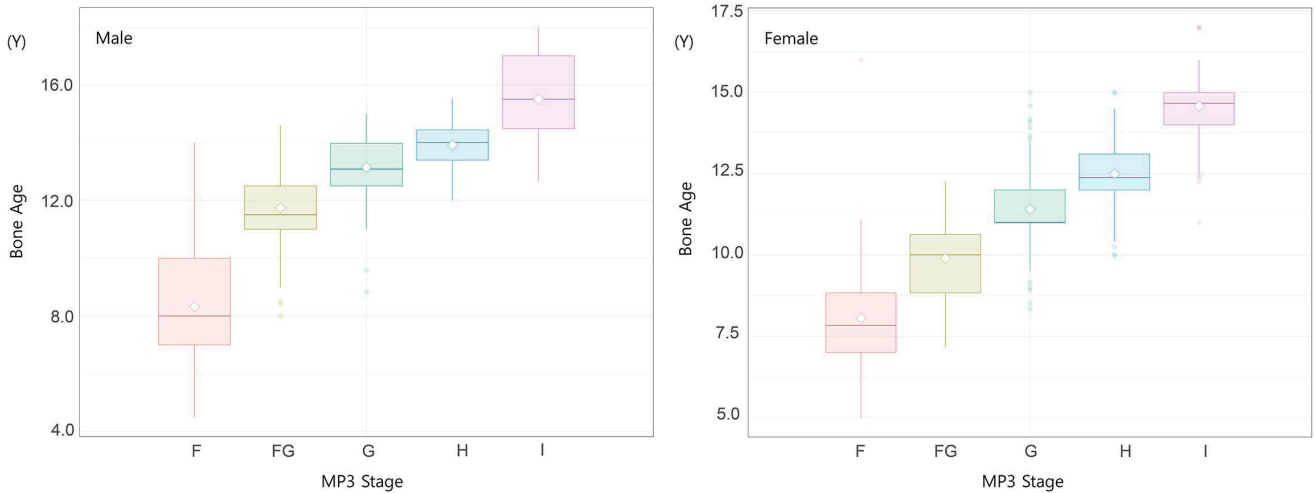


Fig. 6. Bone age distribution according to cervical vertebral maturation stage in males and females. MP3: Middle phalanx of the third finger.

Table 5. The mean bone age and standard deviation according to the middle phalanx of the third finger stage

Stage	Gender	n	BA	<i>p</i>
			Mean \pm SD (year)	
MP3 F	Male	1183	8.33 \pm 1.74	< 0.0001
	Female	789	8.06 \pm 1.13	
MP3 FG	Male	170	11.74 \pm 1.16	< 0.0001
	Female	303	9.90 \pm 0.95	
MP3 G	Male	102	13.14 \pm 1.07	< 0.0001
	Female	240	11.40 \pm 1.09	
MP3 H	Male	52	13.92 \pm 0.78	< 0.0001
	Female	136	12.50 \pm 1.03	
MP3 I	Male	22	15.50 \pm 1.53	0.013
	Female	121	14.58 \pm 1.11	

p value from Mann-Whitney U test.

MP3: Middle phalanx of the third finger; BA: Bone age; SD: Standard deviation.

4. CVM 단계에 따른 MP3 단계 분포 및 골령과의 상관관계

성별에 따른 CVM 각 단계에 해당하는 MP3 단계의 전체 분포는 다음과 같았다(Table 6, 7). 남자의 CVM I 단계에서는 MP3 - F, FG, G 단계가 분포했고, 여자는 MP3 - F, FG, G, H 단계가 분포했다. CVM II 단계에서 남자는 MP3 - F, FG, G, H 단계가 분포했고, 여자는 MP3 모든 단계가 분포했다. CVM III 단계에서는 남자, 여자 모두 MP3 전 단계에 걸쳐 분포했

다. CVM VI 단계에서 남자는 MP3 - FG, G, H, I 단계에 분포했고, 여자는 MP3 - G, H, I 단계에 걸쳐 분포했다. CVM V 단계에서 남자는 MP3 - I 단계, 여자는 MP3 - H, I 단계에 분포했다. CVM과 골령의 상관성은 남자에서 0.68, 여자에서 0.74였고, MP3와 골령에서 상관성은 남자에서 0.69, 여자에서 0.86으로 CVM과 MP3 모두 여자에서 골령과 상관관계가 다소 높음을 보였다. CVM과 MP3의 상관성은 0.68, 남자에서 0.67, 여자에서 0.72였다(Table 8, $p < 0.0001$).

Table 6. Distribution of the cervical vertebral maturation and the middle phalanx of the third finger stages in male

MP3 CVM	F	FG	G	H	I		<i>r</i>	<i>p</i>
I	976 (82.5%)	51 (30.0%)	4 (3.9%)	-	-			
II	168 (14.2%)	77 (45.3%)	24 (23.5%)	7 (13.5%)	-			
III	39 (3.3%)	41 (24.1%)	70 (68.6%)	33 (63.5%)	6 (27.3%)		0.67	< 0.0001
IV	-	1 (0.6%)	4 (3.9%)	12 (23.1%)	11 (50.0%)			
V	-	-	-	-	5 (22.7%)			
Total	1183	170	102	52	22	1529		

p value from Spearman's rank-order test.

CVM: Cervical vertebral maturation; MP3: Middle phalanx of the third finger.

Table 7. Distribution of the cervical vertebral maturation and the middle phalanx of the third finger stages in female

MP3 CVM	F	FG	G	H	I		<i>r</i>	<i>p</i>
I	681 (86.3%)	184 (60.7%)	60 (25.0%)	6 (4.4%)	-			
II	101 (12.8%)	102 (33.7%)	73 (30.4%)	24 (17.6%)	4 (3.3%)			
III	7 (0.9%)	17 (5.6%)	97 (40.4%)	72 (52.9%)	38 (31.4%)		0.72	< 0.0001
IV	-	-	10 (4.2%)	32 (23.5%)	66 (54.5%)			
V	-	-	-	2 (1.5%)	13 (10.7%)			
Total	789	303	240	136	121	1589		

p value from Spearman's rank-order test.

CVM: Cervical vertebral maturation; MP3: Middle phalanx of the third finger.

Table 8. Spearman rank correlation coefficient between cervical vertebral maturation, the middle phalanx of the third finger, and bone age in both genders

	Male		Female		Total	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
CVM and BA	0.68	< 0.0001	0.74	< 0.0001	0.71	< 0.0001
MP3 and BA	0.69	< 0.0001	0.86	< 0.0001	0.75	< 0.0001
CVM and MP3	0.67	< 0.0001	0.72	< 0.0001	0.68	< 0.0001

p value from Spearman's rank-order test.

CVM: Cervical vertebral maturation; MP3: Middle phalanx of the third finger; BA: Bone age.

총괄 및 고찰

성장은 성인에서는 관찰할 수 없는 소아 청소년기에 나타나 는 특징적인 과정이다. 성장은 계속해서 변화하는 과정이며 신체적 성장뿐만 아니라 사회적, 정신적인 상태와도 관련이 있다 [2]. 또한 소아의 신체 발달 및 사춘기 발현 시기, 치아 발달 정도 등은 연대 연령보다는 골 성숙도와 연관성이 깊다[1]. 골령은 개인의 성적 성숙도나 신체의 성장 정도를 반영하며 현재 아

동의 연령대에 해당하는 신체 성장을 평균과 비교하여 성장 잠재력을 파악하고 예상 성인 키를 예측하는 데 도움을 줄 수 있으며, 소아의 성장과 관련된 질환을 진단하고 관리하는 데 도움을 줄 수 있다[37]. 따라서 정확한 골령을 평가하여 적절한 시기에 필요한 치료를 시행하는 것이 중요하다.

수완부골을 이용해 골 성숙도를 평가하는 기존의 방법 중에서 가장 대표적인 것으로 GP 방법과 TW 방법이 있다. 이번 연구에서 사용된 GP 방법은 1930년대 미국의 소수 중산층의 원

쪽 손목의 표본 방사선 사진을 이용해 연령대별로 나열된 GP 도감과 비교하여 골 성숙도가 표본 방사선과 가장 근접한 골령을 측정하는 방법이다. GP 방법은 골령을 측정하는데 비교적 간편하고 측정이 빠르지만, 표본 방사선 사진이 1개씩만 존재하며, 방사선 사진이 6개월 또는 1년 간격으로 나뉘어 있으므로 세밀한 골령을 평가하는데 TW 방법보다 정확도가 떨어질 수 있다[16]. 하지만 기존의 여러 연구를 통해 두 방법을 통한 골령 평가에서 큰 차이가 없다는 것이 확인되었고, TW 방법에 비해 GP 방법은 짧은 시간 내 골령 평가가 가능하므로 임상에서 더 활용도가 높다[16,19,38]. King 등[38]의 연구에 의하면 TW2 방법과 GP 방법으로 골령 평가에 필요한 평균 시간은 각각 7.9 분, 1.4분으로 GP 방법이 훨씬 시간이 적게 걸리기 때문에 임상적으로 사용하기에 더 유용하다는 장점이 있다.

골령을 평가하는 기존의 방법들은 대부분 수동으로 골령을 평가했는데, 문제는 관찰자 간, 관찰자 내의 평가의 다양성이 존재하는 것이다[37]. 더욱 정확한 골령 평가를 위해서는 평가자의 다양성을 최소화해야 하며, 판독의 정확성도 있어야 한다. 이러한 판독의 다양성을 최소화하고, 술자의 보다 정확한 판독을 위해 진단의 보조 역할로 인공지능이 널리 이용되고 있다[39]. 이번 연구에서 골령을 평가하기 위해 사용된 프로그램은 2018년 5월 식품의약품안전처의 허가를 받은 국내 최초의 인공지능 기반의 의료기기로서 심층학습을 기반으로 한 최초의 자동화된 골령 평가 프로그램이다. 이 프로그램은 GP 방법을 기반으로 입력된 수완부 방사선 사진의 골령과 가장 유사한 골령을 자동으로 평가한다. 인공지능을 이용한 골령 평가에 대한 정확도를 평가한 Kim 등[30]의 연구에 따르면 이 프로그램을 이용한 골령 평가에서는 가장 가능성이 높은 추정 뼈 나이 3개와 각 확률률 제시하는데, 전체 일치율은 93%였다. 확률적으로 가장 가능성이 높은 1순위 골령을 수동으로 평가한 기준 골령과의 일치도를 평가하기 위해 Root Mean Square Error (RMSE)를 사용했고, 값이 낮을수록 일치도가 높음을 뜻하며 그 값은 0.60년이었다. 수동으로 평가한 경우와 인공지능의 도움을 받은 경우 골령의 일치도 평가에서는 인공지능 프로그램의 도움을 받은 경우 검토자 2명 모두 각각 9.5%, 8% 증가한 판독의 일치율을 보였다. 골령 간의 일치도를 평가하기 위한 RMSE 역시 검토자 2명 모두 0.11년, 0.09년 감소하였고 이를 통해 판독의 정확도가 증가함을 확인할 수 있었다. 또한 판독 시간에 있어 200개의 수완부 방사선 사진의 골령을 평가하는 데 인공지능 프로그램의 도움을 받은 경우 수동으로 평가를 시행한 것보

다 검토자 2명 모두에서 골령 판독 시간을 평균 29% 정도 단축했고 임상에서 더욱 빠르게 적용할 수 있었다. Lea 등[40]의 국내 아동을 대상으로 한 골령 평가에서 이번 연구에 사용된 인공지능 프로그램을 사용하여 연대 연령을 기준으로 인공지능을 이용한 골령과 검토자가 평가한 골령 간의 일치도 평가에서 인공지능을 이용한 경우 RMSE 차이가 1.6개월로 인공지능에서 더 높은 일치도를 보였다. Hwang 등[41]의 연구에서도 연대 연령을 기준으로 했을 때 인공지능을 이용해 평가한 골령과 2명의 검토자가 추정한 골령 사이의 RMSE 차이가 인공지능을 이용한 골령 평가에서 더 낮은 값을 보여 수동 평가에 비해 인공지능의 평가가 우수함을 확인할 수 있었다. 이를 통해 인공지능을 이용한 평가가 수동으로 평가한 것보다 정확하며 술자의 진단 효율성에 기여함을 알 수 있었다.

GP 방법을 이용한 골령 평가 방법의 신뢰도를 평가한 Kim 등[42]의 연구에 따르면, GP, TW3, Korean standard bone age chart 방법을 비교하였을 때 각각의 방법으로 추정된 골령과 연대 연령 간의 유의한 상관관계가 있었으며($r = 0.87 - 0.90$), 이는 인공지능을 이용해 평가한 골령과 연대 연령 간에 높은 상관관계를 보인 이번 연구와도 유사한 결과를 나타냈다. 또한 관찰자 간의 평균 골령 차이를 비교했을 때도 3가지 방법 모두 ICC 0.95 이상으로 높은 신뢰도를 보였다. Kim 등[16]의 연구에서도 국내 아동의 골령 평가 시 GP 방법과 TW3 방법을 비교하였고, 두 방법을 통한 골령 간에 강한 상관관계가 있음을 확인할 수 있었다($r = 0.94$). 또한 관찰자 간, 관찰자 내의 판독에 대한 재현성에서 유의한 차이가 없어 단시간 내에 사용해야 하는 외래에서는 오히려 GP 방법이 더욱 편리할 수 있음을 언급하였다. 이를 통해 이번 연구에서 사용된 인공지능 프로그램은 국내 아동을 대상으로 한 골령 평가에서 정확도 및 신뢰성이 임상에서 충분히 적용할 만한 가치가 있으며, 인공지능 프로그램의 골령 평가방식은 GP 방법을 이용한 골령 평가도 국내 아동에 적용 가능한 신뢰할 만한 방법으로 판단된다.

골령과 연대 연령은 일반적으로 거의 일치하거나 6개월 이내의 차이가 있는 것을 정상 범주로 본다[3]. 골령과 연대 연령이 차이 나는 이유는 골령 평가에 있어 골 성숙 속도의 개인차, 골령 측정 방법 자체의 오차, 측정자 간의 오차 등이 원인으로 작용할 수 있다[43]. 이번 연구에서 인공지능을 통해 평가한 골령은 남녀 모두 연대 연령과 높은 상관관계를 보였다. 국내 아동을 대상으로 하여 기존의 GP 방법을 통해 골령을 평가한 Kim 등[16]의 연구에 의하면 남자 7 - 15세, 여자 7 - 14세에서 남자

의 연대 연령과 골령은 각각 10.9 ± 1.7 세, 10.61 ± 2.27 세였고, 여자는 10.5 ± 1.6 세, 10.69 ± 2.21 세였다. GP 방법으로 평가한 골령과 연대 연령 간에 상관성은 0.91로 강한 상관관계를 보였다. Oh 등[44]은 기존의 GP 방법을 이용한 골령 평가와 연대 연령 간의 비교에서 8 - 15세의 연대 연령 평균은 11.4 ± 2.2 세, 골령 11.9 ± 2.8 세로 골령이 다소 높게 평가되는 경향이 있으며, 특히 여자에게서 더 높게 평가된다고 하였고 이는 이번 연구 결과와 유사함을 보였다. 하지만 앞선 여러 연구의 결과와 비교했을 때 대상자의 수 및 연령대의 차이가 있기 때문에 시대에 따라 변하는 국내 아동의 연대 연령과 골령을 정확히 파악하기는 어렵지만, GP 방법 기반의 인공지능을 이용해 평가한 골령과 연대 연령 간에 높은 상관관계를 확인할 수 있었고, 남자보다 여자에서 연대 연령에 비해 골령이 다소 높은 경향성을 확인할 수 있었다. 따라서 보다 정확한 비교와 평가를 위해서는 전국적으로 다기관 대상자 모집을 통한 연구가 필요할 것으로 생각된다. 골령은 사춘기 발현 시기와 성장 완료 시 예상 성인 키와 밀접한 관련이 있고, 임상적으로 저신장(short stature)을 판단하는데 유용한 요소다[1]. 따라서 골령 평가를 통해 성장장애를 의심할 수 있는 능력을 갖추어 조기 치료를 의뢰하는 것도 소아 청소년의 성장을 다루는 치과의사의 역할 중 하나라고 사료된다.

국내 아동에 대해 GP 방법을 이용한 골령 평가의 적용 가능성을 확인하기 위해 아시아 인종에 대한 여러 기존 선행 연구를 비교한 결과 동일한 인종 간에서도 국가, 지역에 따라 적용 가능성에 일부 차이는 존재했지만, 남녀 모두에서 적용 가능한 평가 방법임을 알 수 있었다. Chiang 등[45]의 2005년 타이완의 7 - 19세 대상의 남녀를 평가한 연구에서는 일부 연령대에서 골령과 연대 연령의 차이가 1년 이상 나는 것을 통해 GP 방법의 적용을 위해선 개선이 필요함을 언급하였다. 하지만 2012년 Soudack 등[21]의 0 - 18세 이스라엘 남녀 대상의 연구에서는 GP 방법을 통한 골령과 연대 연령 간에 유의한 차이가 없었다. 다만 남자에서 15세 미만의 연령대의 골령이 연대 연령보다 낮게 평가되고 특히 6 - 10세의 연령대에서 낮은 점이 있었지만, 정상 범주 내의 수준이라 남녀 모두에서 GP 방법의 적용에 문제가 없음을 언급하였다. 이는 남자에서 평균 연대 연령과 평균 골령사이의 유의한 차이가 있는 이번 연구 결과와 유사하였고, 이러한 결과는 아시아인 남자에서 특정 연령 구간에서 골령이 낮게 평가되는 기존의 연구[46]와도 유사한 경향을 보였다. 2015년 Patel 등[47]의 연구에서도 인도 서부 지역의 6 - 16

세 남녀를 대상으로 한 골령 평가에서 연대 연령과 유의한 차이가 없었고 남녀 모두에서 적용할 수 있으며 특히 여자에서 더욱 적용 가능성이 높다고 하였다. 이는 여자에서 연대 연령과 골령의 평균에 유의한 차이가 없음을 보인 이번 연구 결과와 일치했다. 국내 아동을 대상으로 한 2015년 Kim 등[42]의 연구에서는 7 - 12세 남녀에서 연대 연령과 기존의 GP 방법으로 평가한 골령은 남자 10.09 ± 1.81 세, 9.61 ± 2.36 세였고, 여자 9.63 ± 1.77 세, 9.85 ± 1.98 세로 평가 간 유의한 차이가 없었으며 다른 골령 평가 방법과 비교해도 남녀 모두에서 적용 가능한 평가 방법임을 확인할 수 있었다. 이를 통해 GP 방법을 이용해 평가한 골령의 적용에는 동일한 인종에서도 지역에 따른 차이가 존재했지만, 이번 연구에서 인공지능을 이용해 평가한 골령과 기존의 방법으로 평가한 골령에서 나타난 유사한 경향성을 통해 국내 아동의 골령 평가에 적용할 수 있는 유용한 방법으로 판단되며, 보다 정확한 평가를 위해서는 이에 관한 추가적인 연구와 검증이 필요할 것으로 보인다.

이번 연구에서 사용된 CVM 단계는 Baccetti 등[35]이 제시한 방법으로 경추골 성숙단계를 총 5단계로 나누었다. 경추골 성숙단계는 사춘기 최대 성장 및 하악골 성장 평가에 효율적인 방법이며, 사춘기 최대 성장이 나타나는 시기는 CVM II와 III 단계로 알려져 있다. 이번 연구 결과에 따른 최대성장기에 해당하는 골령은 CVM II 단계에서 남, 녀 각각 11.00 ± 1.81 세, 10.00 ± 1.49 세였고, CVM III 단계에서는 13.00 ± 1.46 세, 12.00 ± 1.44 세로 나타났다. GP 방법으로 평가한 골령과 CVM 단계를 분석한 2007년 Al-Hadlaq 등[48]은 10 - 15세 사우디 남자 아동을 대상으로 한 연구에서 최대성장기에 해당하는 CVM II, III 단계의 골령은 11.323 ± 0.8902 세, 13.016 ± 0.5984 세로 이번 연구의 국내 남자 아동의 결과와 비교 시 유사하였다. 2017년 한국 소아 청소년 성장 도표[49]에 따라 사춘기 최대 성장이 나타나는 연령대인 남자 11.91 - 12.83세, 여자 9.41 - 10.83세와 비교 시 비슷한 골령 범위는 남녀 모두 CVM II, III 단계에 해당했다. 하지만 골령의 범위가 남자의 경우 CVM III 단계에 가까웠고, 여자의 경우 CVM II 단계에 가까워 여자에서 남자보다 빠른 단계에서 사춘기 최대 성장이 나타남을 예상할 수 있었다.

이번 연구에 사용된 MP3 방법은 Hägg와 Taranger[36]가 제안한 방법을 사용했고, Madhu 등[50]의 연구에 따라 크게 최대 성장기 전, 최대 성장기, 최대 성장기 후로 나눌 수 있다. 최대 성장기 전에 해당하는 단계는 F, FG 단계로 F 단계는 성장

개시기, FG 단계는 최대 성장기 1년 전에 해당하는 성장 가속기로 본다. 최대 성장기는 G 단계로 최대 성장기 또는 1년 후로 본다. 최대 성장기 후는 H, I 단계로 H 단계는 성장 하강기, I 단계는 성장 종료로 볼 수 있다. 이번 연구에서 나타난 MP3 단계별 골령의 분포로 보면 최대 성장기 전부터 최대 성장기 후까지 모든 골령이 남자보다 여자에게서 낮게 나왔다. 이는 여자에서 골 성숙이 더 빠르게 진행되는 것으로 해석할 수 있으며 선행 연구 결과와도 일치함을 알 수 있었다[51,52]. 2017년 한국 소아 청소년 성장 도표[49]에 따른 사춘기 최대 성장이 나타나는 연령은 남자에서 11.91 - 12.83세, 여자에서 9.41 - 10.83세였고, 이는 이번 연구 결과 사춘기 최대 성장기에 해당하는 G 단계 보다는 오히려 성장 가속기에 해당하는 FG 단계의 골령과 유사하였다. 2019년 Lee와 Mah[53]의 연구에서는 MP3 단계의 골령을 평가하기 위해 1997년 한국 소아 청소년 3,407명을 대상으로 TW2-20 방법을 이용해 제작된 표준 골령 도표를 이용하였는데[54], GP 방법을 기반으로 한 인공지능 프로그램을 이용해 골령을 평가한 이번 연구 결과와 비교했을 때 MP3 단계의 골령이 이전 연구에 비해 대체로 낮아졌음을 알 수 있었다. 연구 대상자의 수와 연령 분포의 차이가 있지만, TW2-20 방법과 비교 시에도 GP 방법의 골령 평가는 큰 차이가 없다는 기존 연구 결과[43]를 통해 시대가 변함에 따라 소아 청소년의 골 성숙이 빠르게 진행되며 사춘기 최대 성장이 빠른 시기에 나타날 것으로 예상할 수 있다. 이러한 결과는 과거에 비해 소아 청소년의 생활환경 및 식습관 변화로 인해 성장 속도가 빨라졌기 때문으로 사료된다.

기존의 여러 연구를 통해 경추골과 수완부골 간에는 높은 상관관계가 있었고, 각각의 방법 역시 골 성숙도 평가에서 신뢰할 만한 방법임을 알 수 있었다[31-33]. 이번 연구에서 최대성장기에 해당하는 MP3 - G 단계에서 남녀 모두 가장 높은 비율을 차지한 경추골 성숙도는 CVM III 단계였다. Hegde 등[55]의 연구에서는 최대성장기에 해당하는 MP3 - G 단계에서 가장 높은 비율을 차지한 CVM 단계는 II, III 단계로 이는 이번 연구 결과와 유사했다. Wong 등[32]의 연구에서 CVM의 최대성장기는 CVM III 단계로 이에 해당하는 MP3 단계는 대부분 FG 단계로 이번 연구의 CVM III 단계에서 가장 높은 비율을 차지한 MP3 - G 단계와는 약간의 차이를 보였다. 또한 CVM과 MP3 단계 간의 Spearman 상관관계수는 남자 0.9521, 여자 0.9408 였는데, 이번 연구에서는 남자 0.67, 여자 0.72 로 기존 연구에 비해 다소 낮았다. 이는 CVM과 MP3 단계에서 성장 완료시기에 해당

하는 조사 대상자 수가 다른 단계에 비해 상대적으로 부족했기 때문이라고 생각된다. 하지만 Prion과 Haerling[56]의 연구에서 제시된 상관관계수 0.6 이상일 경우 강한 상관관계가 있다는 것을 고려한다면 CVM과 MP3 방법 간에는 높은 상관관계가 있다고 볼 수 있다. CVM과 MP3 방법 간의 높은 상관관계와 최대성장기에 해당하는 CVM과 MP3 단계에서 골령이 비슷한 점을 고려한다면 두 가지 방법 중 한 가지만을 이용하더라도 골령 평가가 가능할 것으로 예측되며, 방사선 노출량을 줄일 수 있는 측면에서도 유용할 것으로 보인다.

그러나 이 연구에는 몇 가지 한계점이 존재한다. 먼저 각 CVM 단계와 MP3 단계에서 조사 대상의 분포가 특정 단계에 밀집된 점이다. 대학병원의 소아치과에 내원하는 환자의 특성상 연령대가 제한적이기 때문에 성장 완료시기에 다다른 조사 대상을 확보하는 데 어려움이 있었다. 이에 따라 대상자 수가 적은 CVM과 MP3의 특정 단계에 해당하는 평균 골령을 현시대의 국내 아동의 평균 골령으로 결론 내리기에 한계가 있다. 따라서 향후 연구에서는 골령 평가가 필요한 조사 대상이 내원하는 연관된 과에서 다양한 정보를 통합하여 많은 조사 대상을 확보하는 것이 필요하다고 생각된다. 다음으로 이번 연구에서 골령을 측정하는 데 사용된 프로그램의 GP 방법은 골령 평가의 기준으로 삼은 당시의 소아 청소년의 수완부 골이 현재와 시대적 차이가 있을 뿐만 아니라 인종의 차이도 있다. 또한 시대에 따라 변화하는 아동의 성장 속도를 반영하지는 못하고 있다. 소아 청소년의 골 성숙 정도는 유전, 생활 습관 및 환경에 영향을 받고 시대에 따라 빠르게 변화하기 때문에 국가별, 인구 개별집단에 따라 다르게 나타날 수 있다. 따라서 동일한 평가 방법을 적용했을 때 골령 평가에서 오차가 발생할 수 있다. 마지막으로 이 프로그램을 통해 측정된 골령이 평가 대상의 골령을 대표하는 값으로 볼 수 있는가에 대한 문제가 있다. 이번 연구에 사용한 프로그램은 심층학습을 기반으로 한 것으로, Hwang 등[41]의 연구에 의하면 프로그램을 이용한 골령 평가에 있어 구조적 편향(Systemic bias)이 존재하기 때문에 술자는 프로그램을 활용할 때 이에 대한 한계점을 인식하고 있어야 한다고 언급했다. 이를 개선하기 위해서 다양한 기관의 더욱 많은 자료를 확보하고 국내 소아 청소년의 정확한 골령을 평가하기 위한 프로그램의 검증을 통한다면 국내 아동을 대표할 수 있는 골령 평가가 가능할 것으로 보인다. 추후 국내 소아 청소년의 자료를 이용하여 인공지능 프로그램을 활용한 국내 소아 청소년의 골령 평가에 관한 추가적인 연구와 검증이 필요할 것으로 사료된다.

결론

이번 연구는 GP 방법을 기반으로 한 인공지능 프로그램을 활용하여 자동으로 골령을 평가하고, 국내 아동의 CVM과 MP3 각 단계에 해당하는 평균 골령을 평가하였다. CVM 단계의 최대 성장기로 예상되는 II, III 단계에 해당하는 남자의 골령은 11.00 ± 1.81 세, 13.00 ± 1.46 세였고, 여자의 골령은 10.00 ± 1.49 세, 12.00 ± 1.44 세였다. MP3 단계에서 최대 성장기로 예상되는 단계는 G 단계로 이 단계의 골령은 남자 13.14 ± 1.07 세, 여자 11.40 ± 1.09 세로 나타났다.

소아 청소년의 최대성장기를 예측하는 것은 부정교합의 치료 시기를 결정하는데 중요한 요인이며, CVM과 MP3 단계의 최대성장기에 해당하는 평균 골령을 활용한다면 더욱 정확한 성장 평가가 가능할 것으로 생각된다. 또한 인공지능을 이용한 골령 평가는 높은 정확도와 신뢰도를 바탕으로 술자의 진단 보조 역할로 충분한 활용 가치가 있음을 확인할 수 있었고, 이를 활용해 골령을 평가한다면 임상에서 더욱 편리하고 빠르게 신뢰할 수 있는 골령 예측이 가능할 것으로 사료된다.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2020R1G1A1011629).

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

References

1. Lee KH : Growth assessment and diagnosis of growth disorders in childhood. *Clin Exp Pediatr*, 46:1171-1177, 2003.
2. Sung IK : Monitoring growth in childhood: practical clinical guide. *J Korean Med Assoc*, 52:211-224, 2009.
3. Kim DH : Assessment of Bone Age During Pubertal Age. *J Korean Soc Pediatr Endocrinol*, 16:135-138, 2011.

4. Pancherz H, Hägg U : Dentofacial orthopedics in relation to somatic maturation: An analysis of 70 consecutive cases treated with the Herbst appliance. *Am J Orthod*, 88:273-287, 1985.
5. Hassel B, Farman AG : Skeletal maturation evaluation using cervical vertebrae. *Am J Orthod Dentofacial Orthop*, 107:58-66, 1995.
6. Fishman LS : Maturational patterns and prediction during adolescence. *Angle Orthod*, 57:178-193, 1987.
7. Björk A, Helm S : Prediction of the age of maximum puberal growth in body height. *Angle Orthod*, 37:134-143, 1967.
8. Leite HR, O'Reilly MT, Close JM : Skeletal age assessment using the first, second, and third fingers of the hand. *Am J Orthod Dentofacial Orthop*, 92:492-498, 1987.
9. Lilliequist B, Lundberg M : Skeletal and tooth development. A methodologic investigation. *Acta Radiol Diagn*, 11:97-112, 1971.
10. Flores-Mir C, Burgess CA, Champneyn M, Jensen RJ, Pitcher MR, Major PW : Correlation of skeletal maturation stages determined by cervical vertebrae and hand-wrist evaluations. *Angle Orthod*, 76:1-5, 2006.
11. Gandini P, Mancini M, Andreani F : A comparison of hand-wrist bone and cervical vertebral analyses in measuring skeletal maturation. *Angle Orthod*, 76:984-989, 2006.
12. Stiehl J, Müller B, Dibbets J : The development of the cervical vertebrae as an indicator of skeletal maturity: comparison with the classic method of hand-wrist radiograph. *J Orofac Orthop*, 70:327-335, 2009.
13. Román PS, Palma JC, Oteo MD, Nevado E : Skeletal maturation determined by cervical vertebrae development. *Eur J Orthod*, 24:303-311, 2002.
14. Garn SM : Radiographic atlas of skeletal development of the hand and wrist. *Am J Hum Genet*, 11:282-283, 1959.
15. Chapman SM : Ossification of the adductor sesamoid and the adolescent growth spurt. *Angle Orthod*, 42:236-244, 1972.
16. Kim SY, Oh YJ, Shin JY, Rhie YJ, Lee KH : Comparison of the Greulich-Pyle and Tanner Whitehouse (TW3)

- Methods in Bone Age Assessment. *J Korean Soc Pediatr Endocrinol*, 13:50-55, 2008.
17. Lee JS : The application of TW3 method for prediction about bone age in hand AP image of children. *J Korean Soc Radiol*, 9:349-356, 2015.
 18. Buckler JM : How to make the most of bone ages. *Arch Dis Child*, 58:761-763, 1983.
 19. Kim SY, Yang SW : Assessment of Bone Age: A comparison of the Greulich Pyle Method to the Tanner Whitehouse Method. *J Korean Endocr Soc*, 13:198-204, 1998.
 20. Milner GR, Levick RK, Kay R : Assessment of bone age: a comparison of the Greulich and Pyle, and the Tanner and Whitehouse methods. *Clin Radiol*, 37:119-121, 1986.
 21. Soudack M, Ben-Shlush A, Jacobson J, Raviv-Zilka L, Eshed I, Hamiel O : Bone age in the 21st century: is Greulich and Pyle's atlas accurate for Israeli children? *Pediatr Radiol*, 42:343-348, 2012.
 22. Cantekin K, Celikoglu M, Miloglu O, Dane A, Erdem A : Bone age assessment: the applicability of the Greulich-Pyle method in Eastern Turkish children. *J Forensic Sci*, 57:679-682, 2012.
 23. 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*, 57:21-24, 2013.
 24. Dahlberg PS, Mosdøl A, Ding Y, Bleka Ø, Rolseth V, Straumann GH, Skjerven-Martinsen M, Delaveris GJM, Vist GE : A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. *Eur Radiol*, 29:2936-2948, 2019.
 25. Lee JH, Kim DH, Jeong SN, Choi SH : Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*, 77:106-111, 2018.
 26. Arijji Y, Yanashita Y, Kutsuna S, Muramatsu C, Fukuda M, Kise Y, Nozawa M, Kuwada C, Fujita H, Katsumata A, Arijji E : Automatic detection and classification of radiolucent lesions in the mandible on panoramic radiographs using a deep learning object detection technique. *Oral Surg Oral Med Oral Pathol Oral Radiol*, 128:424-430, 2019.
 27. Russel S, Norvig P : Artificial intelligence: a modern approach, 3rd ed. Pearson Prentice Hall, Saddle River, 1-5, 2010.
 28. LeCun Y, Bengio Y, Hinton G : Deep learning. *Nature*, 521:436-444, 2015.
 29. Lee BD, Lee MS : Automated bone age assessment using artificial intelligence: the future of bone age assessment. *Korean J Radiol*, 22:792-800, 2021.
 30. Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, Kim S : Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency. *AJR Am J Roentgenol*, 209:1374-1380, 2017.
 31. Perinetti G, Perillo L, Franchi L, Di Lenarda R, Conterdo L : Maturation of the middle phalanx of the third finger and cervical vertebrae: a comparative and diagnostic agreement study. *Orthod Craniofac Res*, 17:270-279, 2014.
 32. Wong RWK, Alkhal HA, Rabie ABM : Use of cervical vertebral maturation to determine skeletal age. *Am J Orthod Dentofacial Orthop*, 136:484.E1-E6, 2009.
 33. Kasimoglu Y, Marsan G, Gencay K : Skeletal Maturity Prediction Using Radiographs of the Medial Phalanx of the Third Finger and Cervical Vertebrae. *Int J Med Invest*, 9:42-49, 2020.
 34. Jeon JY, Kim CS, Kim JS, Choi SH : Correlation and Correspondence between Skeletal Maturation Indicators in Hand-Wrist and Cervical Vertebra Analyses and Skeletal Maturity Score in Korean Adolescents. *Children*, 8:910, 2021.
 35. Baccetti T, Franchi L, McNamara Jr. JA : An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. *Angle Orthod*, 72:316-323, 2002.
 36. Hägg U, Taranger J : Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt. *Acta Odontol Scand*, 38:187-200, 1980.
 37. Satoh M : Bone age: assessment methods and clinical applications. *Clin Pediatr Endocrinol*, 24:143-152, 2015.

38. King DG, Steventon DM, O'Sullivan MP, Cook AM, Hornsby VP, Jefferson IG, King PR : Reproducibility of bone ages when performed by radiology registrars: an audit of Tanner and Whitehouse II versus Greulich and Pyle methods. *Br J Radiol*, 67:848-851, 1994.
39. Li D, Pehrson LM, Lauridsen CA, Tøttrup L, Fraccaro M, Elliott D, Zajac HD, Darkner S, Carlsen JF, Nielsen MB : The added effect of artificial intelligence on physicians' performance in detecting thoracic pathologies on CT and chest X-ray: A systematic review. *Diagnostics*, 11:2206, 2021.
40. Lea WW, Hong SJ, Nam HK, Kang WY, Yang ZP, Noh EJ : External validation of deep learning-based bone-age software: a preliminary study with real world data. *Sci Rep*, 12:1232, 2022.
41. Hwang J, Yoon HM, Hwang JY, Kim PH, Bak B, Bae BU, Sung J, Kim HJ, Jung AY, Cho YA, Lee JS : Re-Assessment of Applicability of Greulich and Pyle-Based Bone Age to Korean Children Using Manual and Deep Learning-Based Automated Method. *Yonsei Med J*, 63:683-691, 2022.
42. 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*, 16:201-205, 2015.
43. Cole AJ, Webb L, Cole TJ : Bone age estimation: a comparison of methods. *Br J Radiol*, 61:683-686, 1988.
44. Oh Y, Lee R, Kim HS : Evaluation of skeletal maturity score for Korean children and the standard for comparison of bone age and chronological age in normal children. *J Pediatr Endocrinol Metab*, 25:279-284, 2012.
45. Chiang KH, Chou ASB, Yen PS, Ling CM, Lin C, Lee CC, Chang PY : The reliability of using Greulich-Pyle method to determine children's bone age in Taiwan. *Tzu Chi Med J*, 17:417-420, 2005.
46. Alshamrani K, Messina F, Offiah AC : Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. *Eur Radiol*, 29:2910-2923, 2019.
47. Patel PS, Chaudhary AR, Dudhia BB, Bhatia PV, Soni NC, Jani YV : Accuracy of two dental and one skeletal age estimation methods in 6-16 year old Gujarati children. *J Forensic Dent Sci*, 7:18-27, 2015.
48. Al-Hadlaq A, Al-Qarni M, Al-Kahtani A, Al-Obaid A : Comparative study between hand-wrist method and cervical vertebral maturation method for evaluation of skeletal maturity in Saudi boys. *Pak Oral Dent J*, 27:187-192, 2007.
49. Kim JH, Yun S, Hwang SS, Shim JO, Chae HW, Lee YJ, Lee JH, Kim SC, Lim D, Yang SW, Oh K, Moon JS; Committee for the Development of Growth Standards for Korean Children and Adolescents; Committee for School Health and Public Health Statistics, the Korean Pediatric Society; Division of Health and Nutrition Survey, Korea Centers for Disease Control and Prevention : The 2017 Korean National Growth Charts for children and adolescents: development, improvement, and prospects. *Korean J Pediatr*, 61:135-149, 2018.
50. Madhu S, Hegde AM, Munshi AK : The developmental stages of the middle phalanx of the third finger (MP3): a sole indicator in assessing the skeletal maturity? *J Clin Pediatr Dent*, 27:149-156, 2003.
51. Krailassiri S, Anuwongnukroh N, Dechkunakorn S : Relationships between dental calcification stages and skeletal maturity indicators in Thai individuals. *Angle Orthod*, 72:155-166, 2002.
52. Yoo HK, Ra JY, Lee JW : Skeletal Maturity Evaluation using Maxillary Canine Development in Growing Children. *J Korean Acad Pediatr Dent*, 46:247-254, 2019.
53. Lee YJ, Mah YJ : Skeletal Age Assessment of SMI and MP3 Stages to Predict the Pubertal Growth Spurt. *J Korean Acad Pediatr Dent*, 46:233-238, 2019.
54. Yeon KM : Standard bone-age of infants and children in Korea. *J Korean Med Sci*, 12:9-16, 1997.
55. Hegde DY, Baliga S, Yeluri R, Munshi AK : Digital radiograph of the middle phalanx of the third finger (MP3) region as a tool for skeletal maturity assessment. *Indian J Dent Res*, 23:447-453, 2012.
56. Prion S, Haerling KA : Making sense of methods and measurement: Spearman-rho ranked-order correlation coefficient. *Clin Simul Nurs*, 10:535-536, 2014.

골 성숙도 단계의 골령 평가를 위한 Greulich-Pyle 방법을 이용한 인공지능 프로그램의 활용

김지훈¹ · 서혜준^{1,2} · 박소영³ · 이은경^{1,3} · 정태성^{1,3} · 남옥형^{4,5} · 최성철^{4,5} · 신종현^{1,3}

¹부산대학교 치의학전문대학원 소아치과학교실 및 치의생명과학연구소

²울산대학교 치과병원

³부산대학교 치과병원 소아치과 및 치의학연구소

⁴경희대학교 치과병원 소아치과

⁵경희대학교 치의학전문대학원 소아치과학교실

이 연구의 목적은 Greulich-Pyle (GP) 방법을 기반으로 한 인공지능 프로그램을 이용해 골령을 측정하고 경추골 성숙도(Cervical vertebral maturation, CVM)와 중지 중절골 성숙도(Middle phalanx of the third finger, MP3) 각 단계에 해당하는 골령을 파악하는 것이다. 연구는 2013년부터 2021년까지 경희대학교와 부산대학교 치과병원 소아치과에 내원한 총 3,118명을 대상으로 하였다. CVM은 Baccetti 분류에 따라 5단계로 나누었고, MP3는 Hägg와 Taranger의 방법에 따라 5단계로 나누었다. 골령은 GP 방법 기반의 인공지능 프로그램을 통해 평가하였다. 최대 성장기의 CVM 단계는 II, III로 CVM II의 평균 골령은 남자 11.00 ± 1.81세, 여자 10.00 ± 1.49세였고, III는 남자 13.00 ± 1.46세, 여자 12.00 ± 1.44세였다($p < 0.0001$). MP3 최대 성장기는 G 단계로 평균 골령은 남자 13.14 ± 1.07세, 여자 11.40 ± 1.09세였다($p < 0.0001$). 인공지능을 통한 골령 평가는 임상적 활용 가치가 있으며 신속하고 정확한 진단이 가능할 것으로 예상된다. [J Korean Acad Pediatr Dent 2023;50(1):89-103]

원고접수일 2022년 10월 7일
 원고최종수정일 2022년 11월 29일
 원고채택일 2022년 11월 30일

© 2023 대한소아치과학회
 © 이 글은 크리에이티브 커먼즈 코리아
 저작자표시-비영리 4.0 대한민국
 라이선스에 따라 이용하실 수 있습니다.

교신저자 신종현
 (50612) 경상남도 양산시 물금읍 금오로 20 부산대학교 치의학전문대학원 소아치과학교실
 Tel: 055-360-5183 / Fax: 055-360-5174 / E-mail: jonghyuns@pusan.ac.kr