



Intelligent quantitative assessment of skeletal maturation based on multi-stage model: a retrospective cone-beam CT study of cervical vertebrae

Lizhe Xie^{2,3} · Wen Tang^{1,3} · Iman Izadikhah^{1,3} · Xiaoyu Chen^{1,3} · Zhenqi Zhao⁴ · Yang Zhao⁵ · Hu Li^{1,3} · Bin Yan^{1,2,3} 

Received: 29 March 2021 / Accepted: 8 September 2021

© The Author(s), under exclusive licence to Japanese Society for Oral and Maxillofacial Radiology and Springer Nature Singapore Pte Ltd. 2021

Abstract

Objective To develop new logistic regression estimative models of the cervical vertebral maturation index (CVMI) based on cone-beam CT (CBCT)-derived parameters for intelligent evaluating skeletal maturation.

Methods From 231 CBCT volumes (age range 7–17, mean age 11.09 years), 154 were randomly selected to produce 2D sagittal projections of the second to fourth cervical vertebrae (C2–C4). From 19 quantitative parameters, significant predictors were deduced to formulate logistic models. Using the CVMI and significant predictors of 77 other subjects, performance of the models was externally examined by direct comparison and the area under the receiver operating characteristic curve (AUC). Models were modified if required, to improve their accuracy.

Results Chronological age, C3 height (H3), and ratio of posterior height to lower width of C4 (PH4/LW4) were entered as significant predictors. Accuracy of the models was acceptable (total AUC = 0.91) except for 4th and 5th stage (AUC of 0.82 and 0.83, respectively), which were mis-predicted inversely. Adjusted models were generated by bivariate logistic regression analysis and adding significant parameters (PH3/LW3 and PH4/LW4, with odds ratios of 3.308 and 3.38, respectively) from 58 subjects in 4th and 5th stages of CVMI in the model establishment group. The total AUC increased to 0.94, along with an increase in the accuracy of the latter optimized models to 77.9 and 87%, respectively.

Conclusion The new intelligent models reliably estimated skeletal maturation and can be utilized in the clinical field or machine learning-based skeletal maturation assessment.

Keywords Skeletal maturation · Cervical vertebral maturation index · Cone-beam computed tomography · Machine learning · Multi-stage model

Lizhe Xie and Wen Tang contributed equally as the first author.

✉ Bin Yan
byan@njmu.edu.cn

¹ Department of Orthodontics, The Affiliated Stomatological Hospital of Nanjing Medical University, 136 Hanzhong Street, Gulou District, Nanjing 210029, China

² Jiangsu Province Key Laboratory of Oral Diseases, Nanjing Medical University, Nanjing, China

³ Jiangsu Province Engineering Research Center of Stomatological Translational Medicine, Nanjing Medical University, 136 Hanzhong Street, Gulou District, Nanjing 210029, China

⁴ Department of Stomatology, The First People's Hospital of Nantong, Nantong, China

⁵ Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China

Introduction

Maxillofacial growth and development are determinant factors in craniofacial orthopedics, surgical treatment planning, and prognostic evaluation [1, 2]. Among methods for assessing growth and development, the method most interrelated to the craniofacial complex is the skeletal maturation which is more influenced by sexual growth rather than chronological aging. To comprehend the interrelationship between somatic growth and sexual maturity, one must remember that secretion of gonadotropin-releasing (GnRH) and sex hormones initiate puberty along with accelerating the jaws growth especially among females. This will consequently lead to a disparity between chronological and skeletal age, convincing the latter to be a better timing indicator for orthodontic and orthognathic treatment planning [3–5]. Therefore, in this research, the study sample is only composed of females.

Hand–wrist radiograph indices are extensively known as the most conventional way of skeletal maturity evaluation, though detrimentally subjecting the patients to extra dose of ionizing radiation [6]. Recently, there has been a prevailing interest in observing morphological changes in the cervical vertebrae through the lateral cephalograms as an indicator of skeletal maturity [7, 8]. Numerous authors stated that cervical vertebral maturation (CVM) can be acclaimed as an appropriate and simply effective method for assessing the mandibular growth, pubertal peak determination, and a reliable replacement for additionally taken hand–wrist radiographs while entailing fewer steps [9–11]. However, practitioner subjectivity has a great influence on the judgment of skeletal maturation [12], resulting from the visual interpretation of diagnostic parameters for determining the cervical vertebral maturity. Subsequently, novice examiners would face confusion in accurately determining skeletal development due to insufficient clinical experience and the lack of a decisive model. To minimize rater error, some studies have carried out morphological analyses of the cervical vertebrae based on parameter simplifications [13–15].

Regarding diagnostic abilities, the superimposition of the hard or soft tissues on the cervical vertebrae in cephalometric images might undesirably alter the measurement of morphological parameters, unlike the cross-sectional sagittal slices of CBCT (cone-beam CT) which is more stable [16, 17]. CBCT enacts an essential role in considerable specific circumstances, such as ectopic teeth, severe maxillofacial malformations, airway assessment, dental implant surgeries, and temporomandibular joint disorders [18–21]. In our field of work, the clinicians have been broadly discouraged to use multi-slice CT due to extra and unnecessary radiation exposure along with high cost. Not only compared to multi-slice CTs, but also compared to its own conventional models, the radiation dose and economic burden of current CBCT machines are relatively alleviating. In recent years, CBCT was undergone extensive investigations and researches leading to enormous achievements regarding low-dose and high spatial resolution technologies, offering a great potential for recruiting it as a primary diagnostic method in future [22–24].

Previously using CBCT scans [25, 26], Byun attempted to objectively explain the relationship between various influencing factors and skeletal maturation, however relying on multiple linear regression analysis. Considering the expression of the CVMI (Cervical Vertebral Maturation Index) as discrete and ordinal stages, it is an ordinal logistic model which is expected to divulge the exact maturational stage of a subject rather than its linear counterparts [27, 28], as one of the most applied techniques in machine learning, logistic regression is a predictive analytic algorithm which is used for outcome forecasting and classification based on the concept of probability in clinical

diagnosis [29]. To date, we have not come across any study applying multivariable logistic regression to create estimative analytical models based on a non-linear maturational index. Therefore, it can be anticipated that routine maturational evaluation methods can evolve to intelligent classification of CVM by generation of newly multi-class algorithms in the future [30]. Taken all together, in cases that CBCT imaging is rationally indicated for specific circumstances, utilizing multi-stage models are beneficial for CVM evaluation.

This study aimed to design a multi-stage model for intelligent skeletal maturation evaluation based on Hassel and Farman's CVMI [31] and the parameters derived from the sagittal reconstructed slices of the second, third, and fourth cervical vertebrae using CBCT images from Han Chinese girls. We hypothesize that skeletal maturation can be scored according to CVMI using the cervical parameters attained from the sagittal CBCT images.

Methods and materials

Study population

This study was approved by the Institutional ethical committee of (No. PJ2017-045-001). Pre-treatment CBCT images were collected retrospectively from the Orthodontic Department of the Affiliated Stomatological Hospital of between January and December 2018 for this study. The scans were adopted from the Orthodontic Department of the same institution that provided the approval (registered between January and December 2018) and were taken for specifically indicated clinical reasons imperative to the patient diagnosis and treatment planning (e.g. ectopic teeth, temporomandibular disorders, orthognathic surgery indicated and facially asymmetric patients). A total of 231 sets of CBCT images from females ranging in age from 7 to 17 years (mean age 11.09 years) were collected. The CBCT images were randomly divided into two groups for model development (G1, $n = 154$) and performance testing (G2, $n = 77$) (Table 1). The selection criteria were as follows: (1) no history of systemic or physiological disorders, especially skeletal deformation and diseases; (2) no history of trauma or surgery in the dentofacial region; and (3) reliable CBCT scans.

CBCT images were obtained using a NewTom 5G computed tomography system (Quantitative Radiology, Verona, Italy) with the following exposure parameters: field of view, 18×16 cm; tube voltage, 110 kV; 1–20 mA (pulsed mode); and isotropic voxel size, 0.3 mm. During the exposure, the midsagittal plane was set perpendicular to the horizontal plane, and the Frankfort plane was parallel to the ground with maximum intercuspation.

Table 1 Sample characteristics and distribution in the Model Development Group (G1, $N=154$) and the Performance Test Group (G2, $n=77$)

CVMI stage	Group (size)	Age (mean, SD)	Age range	<i>P</i> value
1	G1 ($n=22$)	10.50 ± 1.535	7–13	0.46
	G2 ($n=12$)	10.17 ± 1.94	7–13	
2	G1 ($n=29$)	11.00 ± 1.363	8–14	0.66
	G2 ($n=13$)	10.76 ± 1.42	8–12	
3	G1 ($n=23$)	11.74 ± 1.287	9–14	0.55
	G2 ($n=10$)	11.40 ± 1.51	9–14	
4	G1 ($n=32$)	12.03 ± 1.149	10–15	0.69
	G2 ($n=15$)	11.67 ± 1.18	10–14	
5	G1 ($n=26$)	13.04 ± 1.341	11–16	0.81
	G2 ($n=12$)	12.67 ± 1.23	11–17	
6	G1 ($n=22$)	14.14 ± 1.754	12–17	0.62
	G2 ($n=15$)	13.87 ± 1.41	12–16	

Significantly correlated parameters based on $P < 0.01$

Data acquisition

All CBCT records (DICOM files) were imported into Mimics software (Version 17.0; Materialise NV, Leuven, Belgium). Each image primarily was reoriented according to the Swensen protocol [32] to ensure a standardized positioning. Multiplanar reformation (MPR) mode was applied to reconstruct images in each of the three orthogonal planes. In MPR mode, the images resembled adjoining slices of the areas of interest of the cervical vertebrae visible in all three planes. According to a previous method described by Byun et al. [26], the anteroposterior axis was set in the axial view and used to define the deepest posterior point of the C2 vertebral foramen and body midpoint. In the coronal view, the vertical axis was set to pass through odontoid process midpoint. Last, The CBCT-generated sagittal slices of the cervical vertebrae were produced. Figure 1 delineates and defines all of the adopted landmarks [13], corresponding distances and ratios measured in our investigation. All of the measurements were acquired from the 154 subjects in the model establishment group by a single operator in two rounds with a 1-month interval.

Skeletal maturation level assessments

The skeletal maturation levels were estimated from the CBCT-generated lateral cephalograms [16]. The Hassel and Farman method [31] was used to rate and interpret cervical vertebral maturation in six stages according to the morphological changes of the C2–C4 (Fig. 2) [17]. The exact skeletal maturational stage of each subject in both groups was determined by three well-experienced orthodontists,

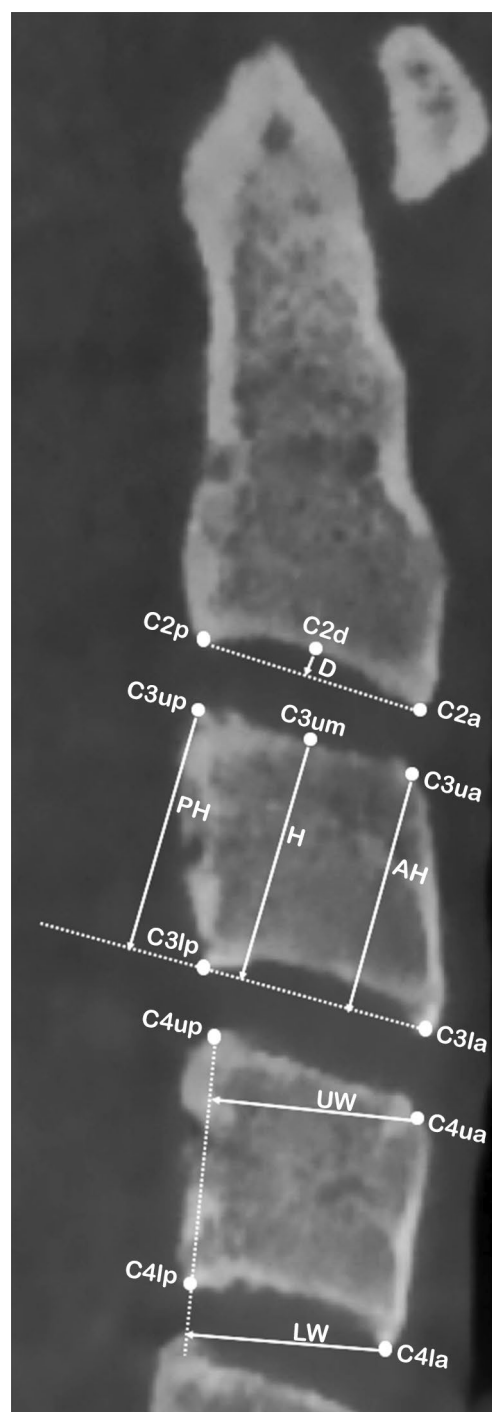


Fig. 1 Illustration of linear measurements performed on each cervical vertebra based on the following landmarks: C2a and C2p (the most anterior and posterior points on the lower border of the body of C2); C3,4la and C3,4lp (the most anterior and posterior points on the lower border of the body of C3 and C4); C3,4ua and C3,4up (the most superior points of the anterior and posterior borders of the body of C3 and C4); and C3,4um (the middle of the upper border of the body of C3 and C4). The measured parameters are D, the vertical distances between Cd and the lines connecting C1a–C1p; AH, the vertical distances between Cua and the lines connecting C1a–C1p; PH, the vertical distances between Cup and the lines connecting C1a–C1p; UW, the vertical distances between Cua and the lines connecting Cup–C1p; LW, the vertical distances between C1a and the lines connecting Cup–C1p. Including also the ratios between D/AH, PH/UW and PH/LW

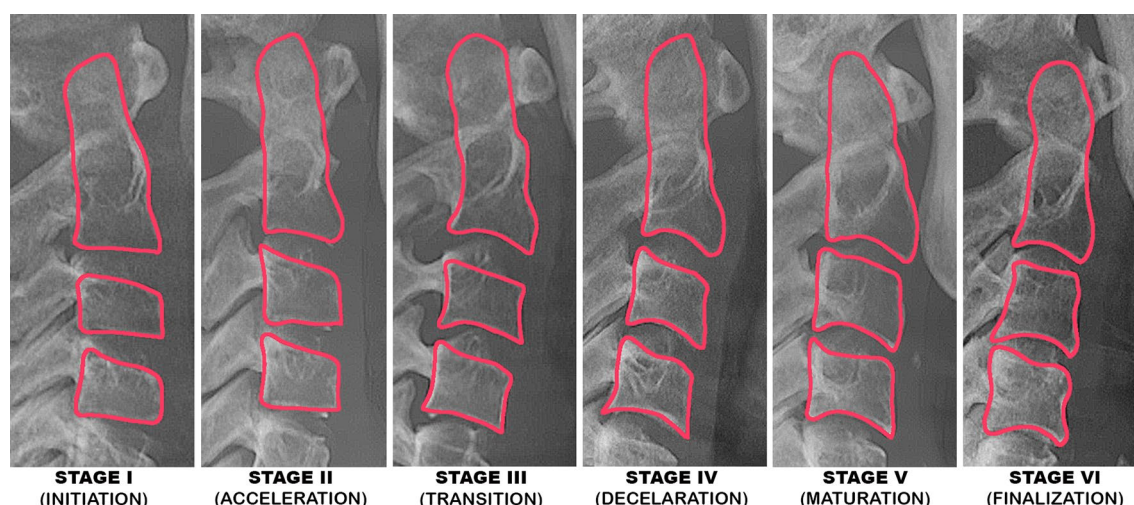


Fig. 2 Descriptions of CVM stages according to Hassel and Farman method. STAGE I (initiation); flat lower borders of C2–4 with a tapered superior border from posterior to anterior. C3 and C4 both resemble a trapezoidal shape. STAGE II (acceleration); C2 and C3 begin to develop a curved lower border, but C4 is still flat. C3 and C4 now present an almost rectangular shape. STAGE III (transition); C2 and C3 have a distinctive concave lower border but C4 is just developing to have that. The C3 and C4 shape is now fully rectan-

gular horizontal. STAGE IV (deceleration); distinct concavities are now visible in lower borders of C2–4. The shape of C3 and C4 now begin to be more of square or rectangular horizontal form. STAGE V (maturation); accentuated concavities can now be seen in C2–4 lower borders. C3 and C4 shapes are fully square. STAGE VI (finalization); deeply concaved lower borders of C2–4 with at least one or both of C3 and C4 is rectangular vertical with heights greater than widths

respectively, to provide a reference basis for assessing the accuracy of the predictive models (Table 1).

Statistical analysis

The statistical analyses were conducted using SPSS (version 19.0, SPSS, Inc., Chicago, USA), with a *P* value of 0.05 indicating a significant difference. To examine the reliability and reproducibility, the intraclass correlation coefficient (ICC) and weighted Cohen's kappa tests were performed on data from 100 computed randomly chosen samples. The descriptive statistics are displayed in Table 2. The normal distribution of the data was confirmed by Shapiro–Wilk test. With the presumption of a logistic relationship between the variables, Spearman's rank correlation coefficient was used to identify the parameters significantly associated with the CVM stage. Multivariable ordinal logistic regression was used to determine the significant independent predictors of the skeletal maturation level (as the dependent variable), followed by calculation of the odd ratio and Nagelkerke R^2 . The statistically significant parameters were selected in a step-wise manner with 0.05 and 0.05 alpha levels for entry and removal, respectively. The variance inflation factor (VIF) was calculated to monitor the severity of multicollinearity.

Model development

Logistic regression models were developed based on the proportional-odds cumulative logit model (Eq. (1)) [33].

Table 2 Spearman's rank correlation coefficients for comparisons between the G1 linear measurements and the CVM stages

Variable	Mean	SD	Correlation coefficient	<i>P</i> value
Age/year	12.86	4.022	0.750**	<0.001
C2 vertebra				
D2/mm	1.471	0.784	0.569**	<0.001
C3 vertebra				
D3/mm	1.269	0.759	0.617**	<0.001
AH3/mm	9.767	2.162	0.835**	<0.001
H3/mm	10.741	2.071	0.841**	<0.001
PH3/mm	11.553	2.145	0.832**	<0.001
UW3/mm	12.418	1.143	0.150	0.066
LW3/mm	13.685	1.253	0.103	0.210
D3/AH3	0.124	0.065	0.398**	<0.001
PH3/UW3	0.933	0.168	0.782**	<0.001
PH3/LW3	0.846	0.152	0.818**	<0.001
C4 vertebra				
D4/mm	1.157	0.796	0.620**	<0.001
AH4/mm	9.667	1.979	0.803**	<0.001
H4/mm	10.521	1.914	0.833**	<0.001
PH4/mm	11.218	1.975	0.829**	<0.001
UW4/mm	12.418	1.295	0.113	0.169
LW4/mm	13.551	1.354	0.114	0.164
D4/AH4	0.114	0.072	0.458**	<0.001
PH4/UW4	0.903	0.178	0.651**	<0.001
PH4/LW4	0.831	0.144	0.794**	<0.001

**Significantly correlated parameters based on *P* < 0.01

The new method output individual formulas for each specific stage pose the advantage of probability determination for each individual skeletal maturational stage. $P(y = j|x)$ represents the expected probability of each maturational stage in the CVMI index ($j = 1, \dots, 6$), whereas a_j denotes the coefficient value as the log-odds, $\beta(x)$ is the constant slope coefficient, and e represents Euler's number ($= 2.71828$). Equation (3) assisted us in determining the linear function of the independent variables. After calculation of the probability of each CVMI stage for each individual, the maximum resulted P_j would be indicative of his/her skeletal maturational stage:

$$P(y = j|x) = \frac{1}{1 + e^{-a_j + \beta(x)}} - \frac{1}{1 + e^{-a_{j-1} + \beta(x)}} \quad (1)$$

$$\text{logit } P_j = a + \beta(x) = g_j(x). \quad (2)$$

The performance of the models was examined by direct comparison with the actual CVMI stage (the ratio of the correctly rated samples to the total number of samples at each specific CVMI stage, expressed as a percentage) and receiver operating characteristic (ROC) curve analysis. The clinical significance level was limited to area under the ROC curve (AUC) = 0.9 and accuracy = 80% as the cut-off criteria for model optimization. To test the predictive accuracy of the to-be-designed skeletal maturation models, the test group (G2) was used. This group was subjected to measurement of the significant predictor variables and calculation of the maturational stage by the generated models. ROC curve analysis was again used to reexamine the prediction accuracy before and after model modification.

Results

Reliability analysis

For quantitative data, the ICC range of the measured cervical vertebral parameters was between 0.947 and 0.998, indicating high intra-rater reliability. The calculated mean and standard deviation (SD) of each parameter is presented in Table 2. The weighted Cohen's kappa test also showed high intra-rater reproducibility for the CVMI assessments (kappa value = 0.901).

Correlation analysis

Spearman's rank correlation coefficient analysis was applied with the CVMI stage as the dependent variable and the measured parameters as the independent variables. The 16 parameters were proven to be significantly correlated ($P < 0.05$) and included in the regression analysis (Table 2).

As revealed by multiple ordinal logistic regression (Table 3), out of the 16 independent parameters, only chronological age, H3 and PH4/LW4 exhibited a significant correlation with the cervical vertebral maturational stage ($P < 0.05$). The statistically significant parameters were selected in a stepwise manner with 0.05 and 0.05 alpha levels for manual entry and removal, respectively. The odds ratios were 9.459 for chronological age, 5.918 for H3, and 15.674 for PH4/LW4 (Nagelkerke $R^2 = 0.729$).

Model development

Model formulation

Chronological age, H3, and PH4/LW4 were introduced for constructing the models. Equation (3) allowed us to calculate $\beta(x)$ for each of the subjects as a function of their significant predictor variables. The final model for the prediction of each stage was determined as Eq. (5), which results in six P_j ($j = 1, \dots, 6$) values ranging between 0 and 1, indicative of each stage probability. After calculating each stage probability, the largest estimated P_j value among the six results designated the current skeletal maturation stage for each individual (VIF < 5 ; tolerance > 2 , indicating no collinearity between parameters).

$$\beta(x) = 2.247x_1 + 1.778x_2 + 2.752x_3 \quad (3)$$

x_1 : Age x_2 : H3 x_3 : PH4/LW4

$$\begin{aligned} \text{logit } P_1 &= 3.545 + \beta(x) = g_1(x) \\ \text{logit } P_2 &= 1.948 + \beta(x) = g_2(x) \\ \text{logit } P_3 &= 1.334 + \beta(x) = g_3(x) \\ \text{logit } P_4 &= -1.393 + \beta(x) = g_4(x) \\ \text{logit } P_5 &= -3.761 + \beta(x) = g_5(x) \end{aligned} \quad (4)$$

$$\begin{aligned} P(y = 1|x) &= \frac{1}{1 + e^{g_1(x)}} \\ P(y = 2|x) &= \frac{1}{1 + e^{g_2(x)}} - \frac{1}{1 + e^{g_1(x)}} \\ P(y = 3|x) &= \frac{1}{1 + e^{g_3(x)}} - \frac{1}{1 + e^{g_2(x)}} \\ P(y = 4|x) &= \frac{1}{1 + e^{g_4(x)}} - \frac{1}{1 + e^{g_3(x)}} \\ P(y = 5|x) &= \frac{1}{1 + e^{g_5(x)}} - \frac{1}{1 + e^{g_4(x)}} \\ P(y = 6|x) &= 1 - \frac{1}{1 + e^{g_5(x)}} \end{aligned} \quad (5)$$

$y = \text{argmax } P_j.$

To make the final decision about the skeletal maturational stage, the highest value ($\text{argmax } P_j$) will be the most possible CVMI stage (y).

Table 3 Multivariate associations determined by ordinal logistic regression

	<i>B</i>	Standard error of <i>B</i>	Wald statistic	Sig	Confidence interval (95%)		Odds ratio
					Minimum	Maximum	
<i>Thresholds</i>							
[CVMI=1]	− 3.545	0.462	58.828	0.000*	− 4.451	− 2.639	
[CVMI=2]	− 1.948	0.359	29.442	0.000*	− 2.651	− 1.244	
[CVMI=3]	− 1.334	0.336	15.802	0.000*	− 1.992	− 0.676	
[CVMI=4]	1.393	0.353	15.529	0.000*	0.7	2.086	
[CVMI=5]	3.761	0.507	54.922	0.000*	2.766	4.755	
<i>Predictors</i>							
Chronological age	2.247	0.74	9.218	0.002 [†]	0.796	3.697	9.459
D2	− 0.374	0.297	1.578	0.209	− 0.956	0.209	−
D3	− 1.718	2.438	0.496	0.481	− 6.497	3.061	−
D4	0.368	2.336	0.025	0.875	− 4.209	4.946	−
AH3	0.589	0.479	1.511	0.219	− 0.350	1.529	−
AH4	0.006	0.513	0.000	0.991	− 0.999	1.011	−
H3	1.778	0.57	9.724	0.002 [†]	0.66	2.895	5.918
H4	− 0.022	0.630	0.001	0.972	− 1.258	1.214	−
PH3	− 1.071	1.539	0.484	0.487	− 4.088	1.946	−
PH4	1.942	1.507	1.660	0.198	− 1.012	4.896	−
UW3	1.195	1.887	0.401	0.527	− 2.504	4.894	−
UW4	− 0.467	0.404	1.336	0.248	− 1.259	0.325	−
LW3	0.663	1.669	0.158	0.691	− 2.609	3.934	−
LW4	− 1.747	1.236	1.997	0.158	− 4.170	0.676	−
D3/AH3	21.941	22.445	0.956	0.328	− 22.050	65.933	−
D4/AH4	− 2.251	21.712	0.011	0.917	− 44.806	40.303	−
PH3/UW3	10.913	24.980	0.191	0.662	− 38.047	59.873	−
PH4/UW4	0.015	1.683	0.000	0.993	− 3.284	3.313	−
PH3/LW3	13.635	27.326	0.249	0.618	− 39.923	67.194	−
PH4/LW4	2.752	0.78	12.464	0.000 [†]	1.224	4.281	15.674

B: Standardized regression coefficient, *significant predictors based on $P < 0.05$, [†]entered significant predictor variables

Performance analysis

G2 ($n = 77$) was employed to evaluate the performance accuracy of the predictive models. The initial performance assessment yielded a total accuracy of 77.9%. As displayed in Fig. 3A, the 4th and 5th stage estimates were remarkably different from the corresponding Hassel and Farman CVMI stages (66.7% and 58.3%, respectively). Predictions from 40% of the subjects with CVMI stages 4 and 5 were mispredicted inversely. As depicted by the ROC curve (Fig. 3B), the original total AUC was equal to 0.91. All of the models had $AUC > 0.9$ except for the models predicting stages 2 ($AUC = 0.83$), 4 ($AUC = 0.82$), and 5 ($AUC = 0.83$).

Model updating

Model optimization

To improve the accuracy, adjusted models were established for precisely distinguishing the 4th and 5th stages based on the measurements from 58 subjects in the model development group who were determined to be at those CVMI stages. A bivariate ordinal logistic regression model was applied to analyze the effects of the independent variables on the skeletal maturation level assessment with the same entry and removal protocol as before. As depicted in Table 4, PH3/LW3 and PH4/LW4, with odds ratios of 3.308 (95% CI 1.148–9.529)

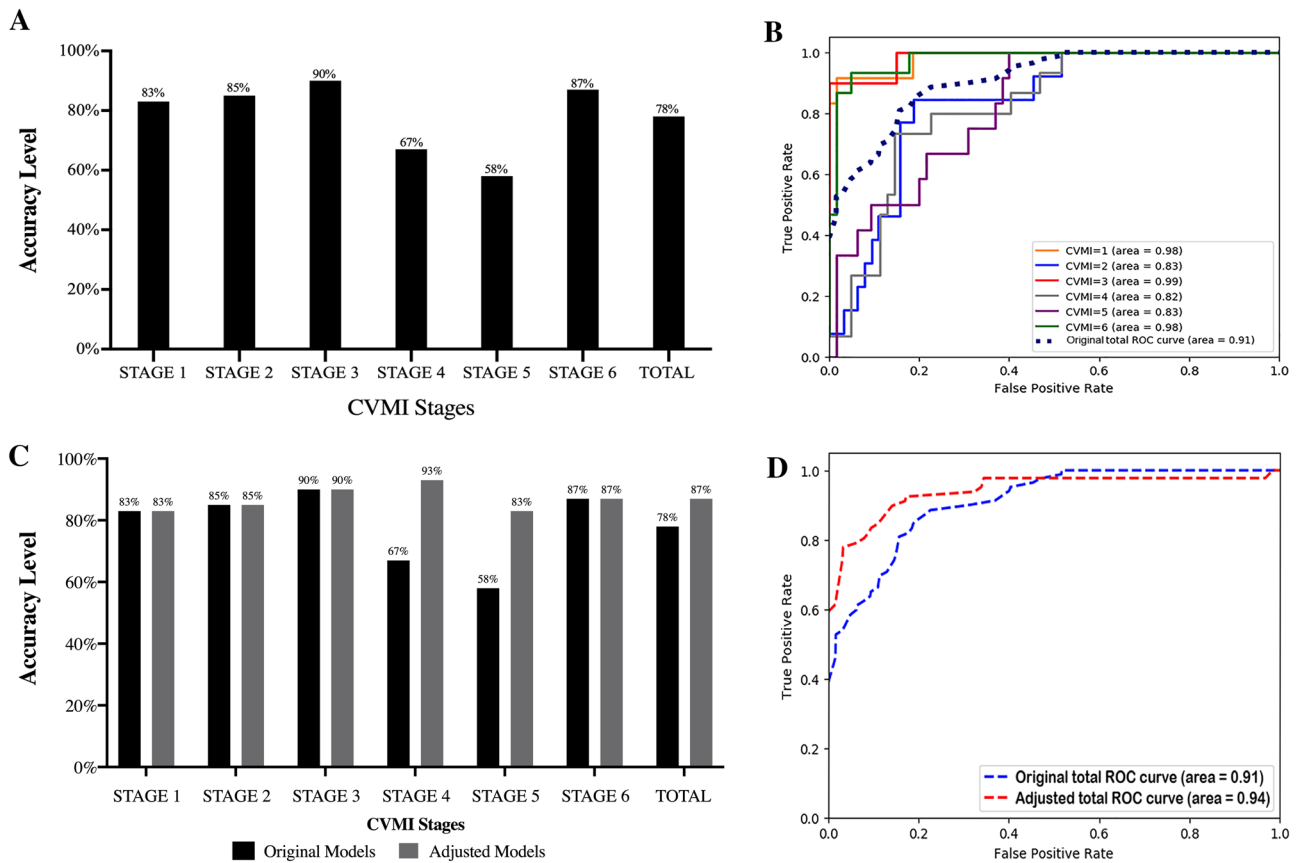


Fig. 3 Prediction accuracy of original and adjusted models. **A** Bar chart: comparison of original models' accuracy for different stages. **B** Receiver operating characteristic curve: separate ROC curves for each

original model and the total performance. **C** Comparison of model accuracy for different stages before and after optimization. **D** ROC curves for total model performance before and after adjustment

Table 4 Secondary parameter estimates and bivariate associations determined by ordinal logistic regression

	Parameter	<i>B</i>	Standard error of <i>B</i>	Wald statistic	Sig.	Odds ratio	95% CI for odds ratio	
							Lower	upper
Location	PH3/LW3	1.196	0.54	4.912	0.027*	3.308	1.148	9.529
	PH4/LW4	1.218	0.525	5.372	0.020*	3.38	1.207	9.466
	Constant	− 0.328	0.37	0.789	0.374	0.72		

B: Standardized regression coefficient, *significant predictors based on $P < 0.05$

and 3.38 (95% CI 1.207–9.466), respectively, exhibited statistically significant correlations ($P < 0.05$) and were employed to generate two adjusted logistic regression equations (Nagelkerke $R^2 = 0.86$). The process of generating the adjusted logistic regression models is shown below:

$$\beta'(x) = 1.196x'_1 + 1.218x'_2 \quad (6)$$

x'_1 : PH3/LW3 x'_2 : PH4/LW4

$$\text{logit } P'_4 = 0.328 + \beta'(x) = g'_4(x) \quad (7)$$

$$P'(y = 4|x) = \frac{e^{g'_4(x)}}{1 + e^{g'_4(x)}} \quad (8)$$

$$P'(y = 5|x) = \frac{1}{1 + e^{g'_4(x)}}.$$

Performance revaluation

Following adjustment of the equations, the accuracy-test was repeated (for the 27 samples of stage 4th and 5th in G2), and an increase in the predictability of the two stages

was observed with the new equations (93.30% and 83.30%, respectively). Additionally, the total accuracy was enhanced as high as 87% after the adjustments (Fig. 3C). The AUC based on the ROC results for the total predictions was improved to 0.94 after reexamining the performance of the models (Fig. 3D).

Discussion

Morphological changes of the cervical vertebrae in different imaging modalities are widely considered helpful indicators of skeletal development. Although CVM method is recently being more acknowledged in clinical practice, still literature shows important shortcomings, such as insufficiency in growth spurts identification, inadequate validity and reproducibility, and subjective technique sensitivity, which influences the results [34]. However, many have verified the CVM method to be in good agreement with hand–wrist method; conforming it to be used as gold standard in quantitative analyses or as skeletal maturity indicator specially when hand–wrist images are not accessible [8, 17, 30, 35]. Nevertheless, our principal focus was not to generalize or evaluate the reliability and clinical applicability of the CVM technique but to construct and introduce a novel logistic regression model detecting the CVM stage. Among methods for assessing growth and development, the method most interrelated to the craniofacial complex is the skeletal maturation which is more influenced by sexual growth rather than chronological aging. Due to the effects of gender, socioeconomic, and ethnic factors on skeletal growth and development [36, 37], we only included images from Han Chinese girls. To eradicate the risk of outcome disparity and intra-observant bias, we recruited three different raters to visually assess the skeletal maturation level.

Generally, the quantification of cervical vertebral parameters has been noticeably used to design predictive models, although limited to continuous linear approaches [12, 13, 26]. Baccetti et al. showed that morphological changes in the C2, C3, and C4 vertebrae were sufficient to evaluate bone maturity [38]. Consistent with previous studies [12, 13, 26, 39], our analysis also demonstrates the effective association of the height of the C3 vertebral body (H3), the height-to-width ratio of the C4 vertebral body (PH4/LW4), and chronological age with skeletal maturation ($OR > 1$, $P < 0.05$).

With the significant predictors, a multi-stage model was established based on multivariable logistic regression to estimate skeletal maturation. Meanwhile, the logistic regression model presents the advantages of its linear regression counterpart in nature and can properly and simultaneously reflect the essential relationship between the significant cervical vertebral parameters and each CVM stage.

Performance analysis of the original models indicated that it was more difficult to precisely discriminate stages 4 and 5 from each other (Fig. 3A, B). According to the ROC analysis, the larger the AUC, the greater the performance accuracy. Generally, the total AUC indicated satisfactory performance. However, distinguishing the stages, the AUCs corresponding to the performance of the models predicting the second, fourth and fifth stages were lower than those of the others (Fig. 3A). Hence, optimization of the less accurate models was performed. Regarding the second stage model, despite the lower AUC (area = 0.83) than that of the 1st-, 3rd-, and 6th-stage models, the accuracy was still as high as 84.6%. On the other hand, it was a predicament and near to impossible to discover with which of the other stages stage two was mispredicted. Our analysis revealed that integration of the C3 aspect ratio (PH3/LW3) into the models increased the distinguishability between the Deceleration and Maturation stages. We speculate that the configurational similarities between C3 and C4 induced the original inaccuracy in the skeletal maturation estimation (Fig. 2). This was confirmed previously in another study in Korean girls, which demonstrated indistinguishable incremental curves between the total length and height of C3 and C4 [26]. On the other hand, we hypothesize that the skeletal growth of C3 during puberty not being as extensive or apparent as that of the other vertebrae is another reason for the low effect in the models [40]. Therefore, optimization of the models for the Deceleration and Maturation stages was carried out by bivariable ordinal logistic regression. The prediction provided by the new adjusted models was more simplified, precise, and objective. Additionally, both the total accuracy and the AUC were improved (Fig. 3C, D), indicating that improvement in the performance of the models was accomplished by adding a new parameter for the 4th and 5th stages. The logical process behind optimizing the models provides an opportunity to rate skeletal maturation similar to clinical diagnostic cognition.

Recently, an increasing trend and evolution have been witnessed in the replacement of manual diagnostic radiography with computerized and digitized methods [41–44]. Newly developed softwares in dentistry can assist clinicians in the decision-making process with a great potential to gain substantial momentum in the future. Reports are indicative of reliable linear and angular cephalometric measurements achieved via different digital methods in contrast to conventional tracing [45, 46]. In terms of practical application, our proposed models can be widely used not only in determining the timing of maxillofacial orthodontic orthopedic and orthognathic surgery, but also in forensic medicine and pediatric endocrinology, thereby allowing more accurate outcomes with a simplified, comprehensible calculation process. With advancements in computerized automatic image recognition and landmark detection, the CVM stage can be

obtained automatically and programmatically by the integration of our method. We believe that after further external validation in different populations, our introduced model can be helpful in clinical practice by prompt and accurate skeletal development rating on account of the widespread use of digitalized radiology.

Compared to deep learning methods, the application of logistic regression was more effective since we had well-structured data with an explainable algorithm [47]. Amasya et al. recently developed a deep learning-based model for CVM analysis. Although, the performance of their Artificial Neural Network (ANN) model was reported close to human observers in CVM analysis, still it was declared insufficient to be trusted blindly requiring human interaction in the clinical decision-making process. Such these predicaments arise when the ANN model or generally deep learning technique continuously optimizes model parameters by labeling a large amount of data without marking the actual significant parameters in the analyzed image. This learning process is like a “black box” in medicine and could lead to problems with accountability and legal consequences [48].

Our presented CBCT-based model is prospective and will have a great contribution to clinical dentistry in the future. At the same time, the workload of clinicians will be reduced, the evaluation results will be more objective, and the evaluation consistency will be satisfying. With the development of technology and decreased radiation dose, CBCT has a great potential to be recruited as the primary diagnostic method in future [22]. Therefore, by increasing the CBCT utilization in clinical fields our model also would be more appreciated accordingly. Generally, the establishment of skeletal maturation norms by regression models requires a vast range of data from samples with various demographic features, such as size and ethnicity, which is not feasible via CBCT considering its limited clinical indications. We expect that the course of cervical vertebral ossification is susceptible to regional and ethnical diversities which requires additional studies to establish externally applicable models. Therefore, the differences between our sample and the general population should not be overlooked. Nevertheless, the presented modeling algorithm provided new grounds for the prediction of cervical vertebral maturation in other ethnic groups and also for other clinical multi-stage problems. However, utilizing the CVM method with intelligent algorithms is still early to be considered an absolute faultless way to precisely grade the individual's skeletal maturation and requires clinician supervision. Although this study benefited from the utilization of CBCT reconstructions of the cervical vertebrae, we cannot suggest a clinical indication solely for skeletal maturational estimation. Further adjustment of the other stage estimative models rather than only the 4th- and 5th-stage models is recommended for future studies.

This study presents newly established logistic models for skeletal maturation estimation based on CBCT-generated parameters. After adjustment of the 4th and 5th CVM-stage models, skeletal maturation could be assessed precisely by rating the probability of each maturational stage for each individual. The formulated cervical vertebral-based models can be used as an evaluation index for the growth and development of the craniomaxillofacial region in Han Chinese girls, providing a reference for clinical orthopedics. These new models can be further verified and cross-checked to be incorporated by computerized methods for the objective assessment of skeletal age. Furthermore, this method can provide new schemes for multi-stage modeling of other medical conditions in clinical fields.

Author contributions LX and WT designed the study and performed the experiments and conceptualizing the study. Data collection and manuscript writing was done by II, WT, and XC. ZZ and YZ carried out statistical analysis and interpretation of the results. HL and BY conceived and supervised the whole project along with funding acquisition. All authors have read and approved the final manuscript.

Funding This study is supported by National Natural Science Foundation of China (Grant no. 82071143 and no. 81571005), Natural Science Foundation of Jiangsu province (BK20180670), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD, 2018-87), and the Key Medical Research Projects of Jiangsu Health Commission (ZDA2020003).

Declarations

Conflict of interest The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

Ethics approval This research study was conducted retrospectively from data obtained for clinical purposes. We consulted extensively with the IRB of Nanjing Medical University who granted our ethical approval for our study (PJ2017-045-001). This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

1. Manlove AE, Romeo G, Venugopalan SR. Craniofacial growth: current theories and influence on management. *Oral Maxillofac Surg Clin N Am*. 2020;32:167–75. <https://doi.org/10.1016/j.coms.2020.01.007>.
2. Lee YS, Choi SH, Kim KH, Hwang CJ. Evaluation of skeletal maturity in the cervical vertebrae and hand-wrist in relation to vertical facial types. *Korean J Orthod*. 2019;49:319–25. <https://doi.org/10.4041/kjod.2019.49.5.319>.
3. Demirjian A, Buschang PH, Tanguay R, Patterson DK. Interrelationships among measures of somatic, skeletal, dental, and sexual maturity. *Am J Orthod*. 1985;88:433–8. [https://doi.org/10.1016/0002-9416\(85\)90070-3](https://doi.org/10.1016/0002-9416(85)90070-3).
4. Morris JM, Park JH. Correlation of dental maturity with skeletal maturity from radiographic assessment: a review. *J Clin Pediatr*

- Dent. 2012;36:309–14. <https://doi.org/10.17796/jcpd.36.3.1403471880013622>.
5. Fishman LS. Chronological versus skeletal age, an evaluation of craniofacial growth. *Angle Orthod.* 1979;49:181–9. [https://doi.org/10.1043/0003-3219\(1979\)049%3c0181:CVSAAE%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(1979)049%3c0181:CVSAAE%3e2.0.CO;2).
6. Gandini P, Mancini M, Andreani F. A comparison of hand-wrist bone and cervical vertebral analyses in measuring skeletal maturation. *Angle Orthod.* 2006;76:984–9. <https://doi.org/10.2319/070605-217>.
7. McNamara JA, Franchi L. The cervical vertebral maturation method: a user's guide. *Angle Orthod.* 2018;88:133–43. <https://doi.org/10.2319/111517-787.1>.
8. Szemraj A, Wojtaszek-Słomińska A, Racka-Pilszak B. Is the cervical vertebral maturation (CVM) method effective enough to replace the hand-wrist maturation (HWM) method in determining skeletal maturation?—a systematic review. *Eur J Radiol.* 2018;102:125–8. <https://doi.org/10.1016/j.ejrad.2018.03.012>.
9. Franchi L, Baccetti T, McNamara JA. Mandibular growth as related to cervical vertebral maturation and body height. *Am J Orthod Dentofac Orthop.* 2000;118:335–40. <https://doi.org/10.1067/mod.2000.107009>.
10. Cunha AC, Cevidanes LHS, Sant'Anna EF, et al. Staging hand-wrist and cervical vertebrae images: a comparison of reproducibility. *Dentomaxillofac Radiol.* 2018;47:20170301. <https://doi.org/10.1259/dmfr.20170301>.
11. Sohrabi A, Babay Ahari S, Moslemzadeh H, et al. The reliability of clinical decisions based on the cervical vertebrae maturation staging method. *Eur J Orthod.* 2016;38:8–12. <https://doi.org/10.1093/ejo/cjv030>.
12. Mito T, Sato K, Mitani H. Cervical vertebral bone age in girls. *Am J Orthod Dentofac Orthop.* 2002;122:380–5. <https://doi.org/10.1067/mod.2002.126896>.
13. Chen L, Liu J, Xu T, et al. Quantitative skeletal evaluation based on cervical vertebral maturation: a longitudinal study of adolescents with normal occlusion. *Int J Oral Maxillofac Surg.* 2010;39:653–9. <https://doi.org/10.1016/j.ijom.2010.03.026>.
14. Chen F, Terada K, Hanada K. A new method of predicting mandibular length increment on the basis of cervical vertebrae. *Angle Orthod.* 2004;74:630–4. [https://doi.org/10.1043/0003-3219\(2004\)074%3c0630:ANMOPM%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(2004)074%3c0630:ANMOPM%3e2.0.CO;2).
15. Chatzigianni A, Halazonetis DJ. Geometric morphometric evaluation of cervical vertebrae shape and its relationship to skeletal maturation. *Am J Orthod Dentofac Orthop.* 2009;136:481–9. <https://doi.org/10.1016/j.ajodo.2009.04.017>.
16. Echevarría-Sánchez G, Arriola-Guillén LE, Malpartida-Carrillo V, et al. Reliability of cephalograms derived of cone beam computed tomography versus lateral cephalograms to estimate cervical vertebrae maturity in a Peruvian population: a retrospective study. *Int Orthod.* 2020;18:258–65. <https://doi.org/10.1016/j.ortho.2020.01.001>.
17. Tekin A, Cesur Aydın K. Comparative determination of skeletal maturity by hand–wrist radiograph, cephalometric radiograph and cone beam computed tomography. *Oral Radiol.* 2019. <https://doi.org/10.1007/s11282-019-00408-y>.
18. Tadinada A, Schneider S, Yadav S. Role of cone beam computed tomography in contemporary orthodontics. *Semin Orthod.* 2018;24:407–15. <https://doi.org/10.1053/j.sodo.2018.10.005>.
19. Kapila SD, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. *Dentomaxillofac Radiol.* 2015;44:20140282. <https://doi.org/10.1259/dmfr.20140282>.
20. Obelenis Ryan DP, Bianchi J, Ignácio J, et al. Cone-beam computed tomography airway measurements: can we trust them? *Am J Orthod Dentofac Orthop.* 2019;156:53–60. <https://doi.org/10.1016/j.ajodo.2018.07.024>.
21. Hodges RJ, Atchison KA, White SC. Impact of cone-beam computed tomography on orthodontic diagnosis and treatment planning. *Am J Orthod Dentofac Orthop.* 2013;143:665–74. <https://doi.org/10.1016/j.ajodo.2012.12.011>.
22. Dai X, Bai J, Liu T, Xie L. Limited-view cone-beam CT reconstruction based on an adversarial autoencoder network with joint loss. *IEEE Access.* 2019;7:7104–16. <https://doi.org/10.1109/ACCESS.2018.2890135>.
23. Jain S, Choudhary K, Nagi R, et al. New evolution of cone-beam computed tomography in dentistry: combining digital technologies. *Imaging Sci Dent.* 2019;49:179–90. <https://doi.org/10.5624/isd.2019.49.3.179>.
24. Pauwels R, Jacobs R, Bosmans H, Schulze R. Future prospects for dental cone beam CT imaging. *Imaging Med.* 2012;4:551–63. <https://doi.org/10.2217/iim.12.45>.
25. Byun BR, Il KY, Yamaguchi T, et al. Quantitative skeletal maturation estimation using cone-beam computed tomography-generated cervical vertebral images: a pilot study in 5- to 18-year-old Japanese children. *Clin Oral Investig.* 2015;19:2133–40. <https://doi.org/10.1007/s00784-015-1415-6>.
26. Byun B-R, Kim Y-I, Yamaguchi T, et al. Quantitative assessment of cervical vertebral maturation using cone beam computed tomography in Korean girls. *Comput Math Methods Med.* 2015;2015: 405912. <https://doi.org/10.1155/2015/405912>.
27. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Linear and logistic regression analysis. *Kidney Int.* 2008;73:806–10. <https://doi.org/10.1038/sj.ki.5002787>.
28. Draper NR, Smith H. *Applied regression analysis.* New York: Wiley; 1998.
29. Ray S. A quick review of machine learning algorithms. In: 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon). IEEE; 2019. p. 35–9. <https://doi.org/10.1109/COMITCon.2019.8862451>.
30. Amasya H, Cesur E, Yıldırım D, Orhan K. Validation of cervical vertebral maturation stages: artificial intelligence vs human observer visual analysis. *Am J Orthod Dentofac Orthop.* 2020;158:e173–9. <https://doi.org/10.1016/j.ajodo.2020.08.014>.
31. Hassel B, Farman AG. Skeletal maturation evaluation using cervical vertebrae. *Am J Orthod Dentofac Orthop.* 1995;107:58–66. [https://doi.org/10.1016/S0889-5406\(95\)70157-5](https://doi.org/10.1016/S0889-5406(95)70157-5).
32. Swennen GRJ, Schutyser F, Hausamen J-E. *Three-dimensional cephalometry: a color atlas and manual.* Berlin: Springer; 2006.
33. Grilli L, Rampichini C. *Encyclopedia of quality of life and well-being research.* Dordrecht: Springer; 2014.
34. Morris KM, Fields HW, Beck FM, Kim DG. Diagnostic testing of cervical vertebral maturation staging: an independent assessment. *Am J Orthod Dentofac Orthop.* 2019;156:626–32. <https://doi.org/10.1016/j.ajodo.2018.11.016>.
35. Uysal T, Ramoglu SI, Basciftci FA, Sari Z. Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist: is there a relationship? *Am J Orthod Dentofac Orthop.* 2006;130:622–8. <https://doi.org/10.1016/j.ajodo.2005.01.031>.
36. Cericato GO, Bittencourt MAV, Paranhos LR. Validity of the assessment method of skeletal maturation by cervical vertebrae: a systematic review and meta-analysis. *Dentomaxillofac Radiol.* 2015;44:20140270. <https://doi.org/10.1259/dmfr.20140270>.
37. Kang ST, Choi SH, Kim KH, Hwang CJ. Evaluation of cephalometric characteristics and skeletal maturation of the cervical vertebrae and hand-wrist in girls with central precocious puberty. *Korean J Orthod.* 2020;50:181–7. <https://doi.org/10.4041/kjod.2020.50.3.181>.
38. Baccetti T, Franchi L, McNamara JA. An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. *Angle Orthod.* 2002;72:316–23. [https://doi.org/10.1043/0003-3219\(2002\)072%3c0316:AIVOTC%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(2002)072%3c0316:AIVOTC%3e2.0.CO;2).

39. Ramírez-Velásquez M, Vilorio-Ávila TJ, Rodríguez DA, et al. Maturation of cervical vertebrae and chronological age in children and adolescents. *Acta Odontol Latinoam*. 2018;31:125–30.
40. Yoganandan N, Pintar FA, Lew SM, et al. Quantitative analyses of pediatric cervical spine ossification patterns using computed tomography. *Ann Adv Automot Med Assoc Adv Automot Med Annu Sci Conf*. 2011;55:159–68.
41. Debelmas A, Ketoff S, Lanciaux S, et al. Reproducibility assessment of Delaire cephalometric analysis using reconstructions from computed tomography. *J Stomatol Oral Maxillofac Surg*. 2020;121:35–9. <https://doi.org/10.1016/j.jormas.2019.04.008>.
42. İzgi E, Pekiner FN. Comparative evaluation of conventional and OnyxCeph™ dental software measurements on cephalometric radiography. *Turk J Orthod*. 2019;32:87–95. <https://doi.org/10.5152/TurkJOrthod.2019.18038>.
43. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics: evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. *J Orofac Orthop*. 2020;81:52–68. <https://doi.org/10.1007/s00056-019-00203-8>.
44. Dot G, Rafflenbeul F, Arbotto M, et al. Accuracy and reliability of automatic three-dimensional cephalometric landmarking. *Int J Oral Maxillofac Surg*. 2020;49:1367–78. <https://doi.org/10.1016/j.ijom.2020.02.015>.
45. AlBarakati SF, Kula KS, Ghoneima AA. The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. *Dentomaxillofac Radiol*. 2012;41:11–7. <https://doi.org/10.1259/dmfr/37010910>.
46. Goracci C, Ferrari M. Reproducibility of measurements in tablet-assisted, PC-aided, and manual cephalometric analysis. *Angle Orthod*. 2014;84:437–42. <https://doi.org/10.2319/061513-451.1>.
47. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. *Radiographics*. 2017;37:2113–31. <https://doi.org/10.1148/rg.2017170077>.
48. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. <https://doi.org/10.1016/j.media.2017.07.005>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.