TURKISH JOURNAL OF
ORTHODONTICS

DOI: 10.4274/TurkJOrthod.2025.2024.30

Original Article

Prediction of Skeletal Age Through Cervical Vertebral Measurements Using Different Machine Learning Regression Methods

İrem Yılmaz[1], Merve Gonca[2]

[1]Private Practice, İstanbul, Türkiye
[2]Eskişehir Osmangazi University Faculty of Dentistry, Department of Orthodontics, Eskişehir, Türkiye

Cite this article as: Yılmaz İ, Gonca M. Prediction of skeletal age through cervical vertebral measurements using different machine learning regression methods *Turk J Orthod.* 2025; 38(1): 36-48

Main Points

- The least absolute shrinkage and selection operator regression exhibited the highest predictive accuracy in estimating vertebral skeletal age.
- Vertebral depth of concavities emerged as a significant predictor of skeletal age in both sexes.
- Vertebral skeletal age estimation did not demonstrate a clinical advantage over chronological age.
- Vertebral skeletal age estimation showed greater variability in boys than in girls, indicating lower consistency with hand-wrist skeletal age assessment.

ABSTRACT

Objective: To compare skeletal ages determined using three different regression methods from measurements made on cervical vertebrae from lateral cephalometric radiographs (LCRs) with the skeletal age determined from hand-wrist radiographs (HWRs).

Methods: LCRs and HWRs of 794 individuals (329 boys, 465 girls) aged 7-18 years were examined. The hand-wrist skeletal age of the participants was determined using the Greulich-Pyle (GP) atlas. Forty-four linear and nine angular morphometric measurements in the C2-C5 vertebrae were made in LCRs. Vertebral skeletal age (VSA) was determined in both sexes using Ridge, the least absolute shrinkage and selection operator (LASSO), and ElasticNet regression methods. The study results were evaluated using R2 (explainability power). Bland-Altman analysis was performed to determine the consistency of chronologic age (CA), GP age, and VSAs.

Results: LASSO regression showed the highest explainability power for VSA, with boys at 0.783 and girls at 0.741. In both sexes, the vertebral depth of concavities had high beta coefficients, and the posterior height of C3 vertebrae (TVup-TVlp) had the highest beta coefficient in boys in LASSO regression. The width of the limits of agreement in both CA and VSA graphs of GP age was wider in boys than in girls. The width of the limits of agreement of CA-VSAs was wider in girls than in boys.

Conclusion: Although high R2 values were obtained, VSA showed no superiority over CA in the assessment of skeletal age, and no significant clinical advantage was observed. For the Turkish population, using GP age may be more accurate for determining skeletal age in orthodontic treatment planning.

Keywords: Orthodontics, hand-wrist radiograph, cephalometry, cervical vertebrae, machine learning

Corresponding author: Merve Gonca, **e-mail:** mervegonca@gmail.com
Received: February 14, 2024 **Accepted:** February 09, 2025 **Epub:** 19.03.2025 **Publication Date:** 27.03.2025

INTRODUCTION

Assessing growth potential during pre-adolescence and adolescence is crucial, and various indicators such as body height and weight, sexual maturation, chronologic age (CA), dental development, and skeletal development can be used to identify growth stages. The identification of the growth and development stage of an individual has a significant impact on the diagnosis, treatment planning, and treatment outcome of orthodontic treatment. Although CA is commonly used, it may not always be a reliable indicator for growth stages due to variations in the timing, velocity, and duration of growth among individuals.[1] Skeletal age is commonly evaluated in orthodontics via hand-wrist radiographs (HWRs) or lateral cephalometric radiographs (LCRs).[2]

The Greulich-Pyle (GP) atlas is commonly used to determine patients' skeletal age by evaluating the maturation of the hand and wrist bones. The main use of skeletal age in orthodontic treatment is the determination of the timing of orthopedic treatment or the confirmation of the end of growth.[3] HWRs are considered the gold standard; the other most commonly used method for evaluating skeletal maturity in orthodontics is cervical vertebral maturation (CVM), which is based on assessing the maturation stage of the cervical vertebrae.[4,5] It is often suggested that HWRs in orthodontics should be limited to cases where the information obtained is considered essential for treatment planning and cannot be obtained by other means, given the importance of minimizing radiographic exposure.[6]

To objectify skeletal age assessment and make it more efficient, many artificial intelligence (AI) systems have been developed to increase diagnostic accuracy mostly via HWRs.[7] Due to the significant correlation between hand-wrist bone and CVM, most AI studies have focused on classifying developmental phases and comparing AI-based classifications with human diagnoses. However, skeletal age estimation has not been thoroughly studied. The clinical application of these studies was limited because they focus on evaluating success metrics rather than automated systems.[8,9] To address this gap, this study aims to evaluate cervical vertebrae maturity using a quantitative method of morphologic changes.

Regression-based methods determine how independent factors affect a dependent variable by identifying a non-deterministic function representing the independent variables' effect on the dependent variable's mean. While regression procedures are straightforward, they require a suitable model for data fitting. Predictions can be made by applying the parameters obtained in a clinical application into the regression formula.[10] Ridge, The least absolute shrinkage and selection operator (LASSO), and ElasticNet are regression models are commonly used in multiple linear regression problems to prevent overfitting. Optimizing the selection of the proper technique and fine-tuning the hyperparameters via cross-validation is essential for constructing a model that effectively manages bias and variance, thereby enhancing predicted accuracy.[11]

The explainability power (R^2) provides valuable information regarding the degree to which the analyzed data can understand the dependent variable. The higher the R^2 value, the higher the capacity of the obtained data to describe the dependent variable.[12-14] The predominant methodology in the scholarly literature for estimating skeletal age through vertebral parameters involved stepwise regression analysis.[15-18] To our knowledge, no previous study in the literature includes a quantitative approach with AI regression methods to determine skeletal age through LCRs.

Although correlation analysis can compare actual and regression-predicted skeletal age studies, it only evaluates the connection between variables, not their differences.[15,17] The Bland-Altman analysis offers an alternative approach by quantifying the agreement between two quantitative measures by calculating the mean difference and agreement limits. However, Bland-Altman plots only depict the range of agreement without indicating whether it is acceptable. Acceptable limits must be determined based on predefined clinical requirements, biologic considerations, or other relevant goals.[19] Also, there is limited research explicitly addressesing R^2 in skeletal age determination using vertebral measurements and assessing the compatibility and repeatability [vertebral skeletal age (VSA) -GP age] of this method through Bland-Altman analysis.[20]

The aim of this study was to develop a predictive model of VSA by using Ridge, LASSO, and ElasticNet regression models.

The null hypothesis of the study was that there would be no significant difference between the vertebral age prediction models developed using Ridge, LASSO, and ElasticNet regression.

METHODS

Study Design

The study received ethical approval from the Research Ethics Committee of Recep Tayyip Erdoğan University (date: 02.02.2023 and protocol number: 33) and involved a retrospective analysis of LCRs and HWRs from patients referred for orthodontic treatment at the Department of Orthodontics, Faculty of Dentistry at Recep Tayyip Erdoğan University. The study was conducted in accordance with the applicable ethical principles of the World Medical Association Declaration of Helsinki of 1964 and later versions.[21] Informed written consent forms, which included the use of patient records in scientific studies, were obtained from all patients at the beginning of treatment. Patients who met specific criteria were included in the study, including individuals of Turkish ethnicity, between the ages of 7-18 years, with good quality LCRs and HWRs, normal growth and development, no systemic disease, no congenital deformities, no bone syndromes, no previous hand-wrist injury, and good nutrition without serious illness. LCRs and HWRs were taken on the same day and all LCRs included

in the study were of sufficient quality, with a clear view of the cervical spine (C2-C5).

The LCRs and HWRs were acquired using a Planmeca Promax 2D S2 imaging unit (Planmeca Oy; Helsinki, Finland) with specific exposure parameters (66 kVp, 10 mA, 10.5 s in LCRs, 60 kVp, 4 mA, and 10.5 s in HWRs). During LCR acquisition, ear rods, and nasal support were used to stabilize the head, and the Frankfort horizontal plane was set parallel to the floor. HWRs were obtained using a specific focus-to-film distance of 170 cm and 30° angulation of the thumb to allow for the depiction of the sesamoid bone.

In the sample size calculation performed considering the number of independent variables as 53, the adjusted R^2=0.686 result in the study of Varshosaz et al.[15], 95% confidence (1-α), 95% test strength (1-β), and f2=2.185 effect size, the minimum required number of samples was determined as 69.

A total of 1257 individuals' LCRs and HWRs were reviewed, and radiographs from 463 individuals who did not meet the inclusion criteria were excluded from the study. We analyzed 794 sets of radiographs (LCR and HWR) of untreated subjects (329 boys, 465 girls) and identified 27 cervical vertebral reference points (Figure 1) for the analysis and obtained 44 linear and nine angular morphometric measurements (Figures 2 and 3), which were located in the C2-C5 vertebrae. The GP age was determined using the HWR images.

All LCRs were calibrated using a 45-mm-long bar, and linear and angular measurements were performed by an orthodontist with 4 years of orthodontic clinical experience using AudaxCeph version 4.2.0.3101 software. To assess the intra-rater and inter-rater agreement, a random sample of 393 LCRs and HWRs was chosen. The measurements were repeated after 1 month by the same orthodontist with 4 years of clinical experience to determine intra-rater reproducibility. Another orthodontist with 10 years of clinical experience performed the measurements to evaluate inter-rater reliability. The intraclass correlation coefficient (ICC) was used to assess the measurement error.

Regression Methods

Ridge, LASSO, and ElasticNet are all regression models used in multiple linear regression problems to prevent overfitting. Choosing the appropriate method and tuning the hyperparameters through cross-validation are crucial for building a model that balances bias and variance, thus improving predictive performance.[12,14]

Multicollinearity occurs when independent variables in a regression model are highly correlated, making it difficult to determine each variable's effect. This issue can be detected using the variance inflation factor (VIF) and tolerance values. A VIF above 10 or a tolerance below 0.2 indicates multicollinearity. Regularization techniques such as Ridge, LASSO, and ElasticNet address multicollinearity by adding a penalty term to the regression model, which helps shrink the coefficients of correlated variables.

Ridge regression incorporates an L2 penalty, the sum of squared coefficients, into the loss function. This technique is particularly effective when dealing with many small and approximately equal coefficients because it distributes the values evenly among correlated variables. By using the lambda (λ) parameter, Ridge regression controls the strength of the L2 regularization.

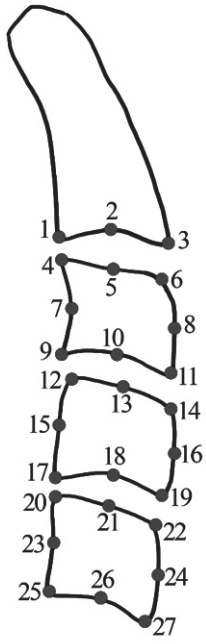

Figure 1. 1.SVp: The most posterior point of the lower edge of the 2nd cervical vertebra, 2.SVd: The deepest point of the concavity at the lower edge of the 2nd cervical vertebra, 3.SVa: The most anterior point of the lower edge of the 2nd cervical vertebra, 4.TVup: The most posterior point of the upper edge of the 3rd cervical vertebra, 5.TVum: Midpoint of the upper edge of the 3rd cervical vertebra, 6.TVua: The most anterior point of the upper edge of the 3rd cervical vertebra, 7.TVpm: Midpoint of the posterior border of the 3rd cervical vertebra, 8.TVam: Midpoint of the anterior edge of the 3rd cervical vertebra, 9.TVlp: The most posterior point of the lower border of the 3rd cervical vertebra, 10.TVd: The deepest point of the concavity at the lower edge of the 3rd cervical vertebra, 11.TVla: The most anterior point of the lower border of the 3rd cervical vertebra, 12.FVup: The most posterior point of the upper edge of the 4th cervical vertebra, 13.FVum: Midpoint of the upper edge of the 4th cervical vertebra, 14.FVua: The most anterior point of the upper edge of the 4th cervical vertebra, 15.FVpm: Midpoint of the posterior edge of 4th cervical vertebra, 16.FVam: The midpoint of the anterior edge of the 4th cervical vertebra, 17.FVlp: The most posterior point of the lower edge of the 4th cervical vertebra, 18.FVd: The deepest point of the concavity at the lower edge of the 4th cervical vertebra, 19.FVla: The most anterior point of the lower edge of the 4th cervical vertebra, 20.FiVup: The most posterior point of the upper edge of the 5th cervical vertebra, 21.FiVum: Midpoint of the upper edge of the 5th cervical vertebra, 22.FiVua: The most anterior point of the upper edge of the 5th cervical vertebra, 23.FiVpm: Midpoint of the posterior edge of 5th cervical vertebra, 24.FiVam: The midpoint of the anterior border of the 5th cervical vertebra, 25.FiVlp: The most posterior point of the lower border of the 5th cervical vertebra, 26.FiVd: The deepest point of the concavity at the lower edge of the 5th cervical vertebra, 27.FiVla: The most anterior point of the lower border of the 5th cervical vertebra

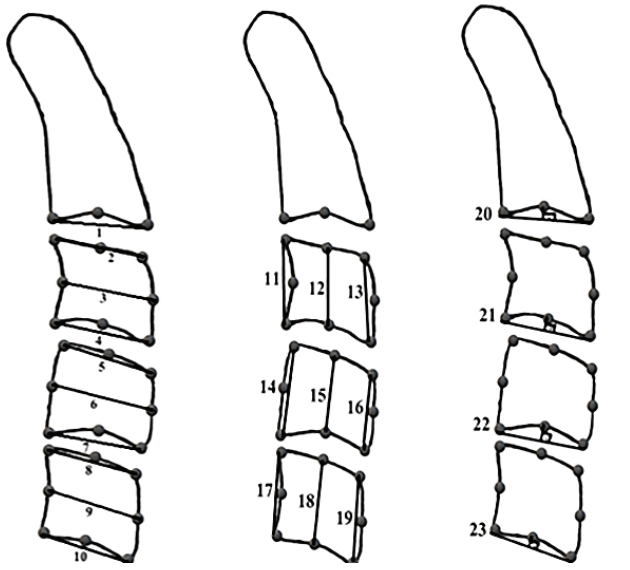

Figure 2. 1st SVp-SVa, 2nd TVup-TVua, 3rd TVpm-TVam, 4th TVlp-TVla, 5th FVup-FVua, 6th FVpm-FVam, 7th FVlp-FVla, 8th FiVup-FiVua, 9th FiVpm-FiVam, 10th FiVlp-FiVla, 11th TVup-TVlp, 12th TVum-TVd, 13th TVua-TVla, 14th FVup-FVlp, 15th FVum-FVd, 16th FVua-FVla, 17th FiVup-FiVlp, 18th FiVum-FiVd, 19th FiVua-FiVla, 20th SVD, 21st TVD, 22nd FVD, 23rd FiVD

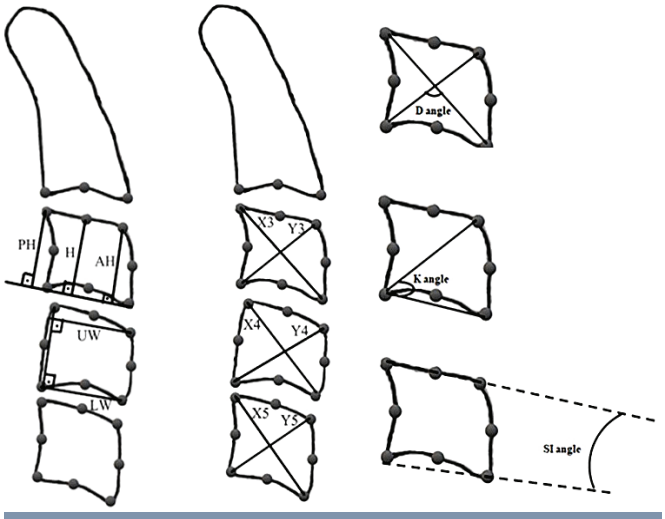

Figure 3. PH3, H3, AH3, PH4, H4, AH4, PH5, H5, AH5, UW3, LW3, UW4, LW4, UW5, LW5, X3, Y3, X4, Y4, X5, Y5, D3 angle, D4 angle, D5 angle, K3, K4, K5, SI3, SI4, SI5

This regularization term penalizes large coefficients, thereby reducing their variance without eliminating any variables, and mitigating multicollinearity in the model.[12] LASSO regression applies an L1 penalty, which is the sum of the absolute values of coefficients. This approach can shrink some coefficients to zero, effectively performing variable selection by eliminating less important predictors. This makes it particularly useful when only a few predictors are expected to be significant. LASSO regression uses the lambda (λ) parameter to control the strength of the L1 regularization, which penalizes the absolute values of the coefficients and enables variable selection by shrinking some coefficients to zero.[13] ElasticNet combines both L1 and L2 penalties, offering a balance between Ridge and LASSO regressions. This approach is advantageous when multiple correlated predictors are present, and some need to be eliminated. ElasticNet regression uses lambda (λ) and alpha (α) parameters. Lambda (λ) controls the overall strength of the regularization, and alpha (α) determines the mix between L1 and L2 regularization. When alpha is 0, ElasticNet behaves like Ridge regression; when alpha is 1, it behaves like LASSO regression. Values between 0 and 1 provide a balance between the two methods. The optimal values of the regularization parameters in Ridge, LASSO, and ElasticNet regression are determined by minimizing the mean squared error.[14]

The performance of these models is typically evaluated using metrics such as R^2 and the Akaike information criterion (AIC) (to measure the model's fit and complexity).[12,14] Cross-validation is a method used to evaluate the performance of machine-learning models. Among various methods, k-fold cross-validation is the most widely used. The dataset is divided into k parts, and each of the k parts is used separately as the test dataset, and the remaining dataset is used as the training dataset. This process is repeated k times, and the mean of the test errors obtained each time is used to predict the model's performance. K-fold cross-validation method ensures that all the samples in the dataset are used to train the model. After k cross-validation, the mean error is calculated for the training and test data and expresses how much the predicted values deviate from the actual values. A lower mean error value means a better fit and more accurate predictions. Cross-validation, especially k-fold cross-validation, is often used to tune the hyperparameters (lambda and alpha), ensuring that the model generalizes well to new data.[12,14]

Statistical Analysis

Statistical analysis was performed using the Eviews v12 program (IHS Markit Ltd, London, UK). Descriptive statistics were calculated as mean, standard deviation, median, minimum/maximum (min./max.), Kurtosis, and Skewness. Vertebral morphometric measurements were included to generate a calculated VSA. The ENET-ElasticNet regularization method was used for estimating skeletal age. Estimation was made using Ridge, LASSO, and ElasticNet regression models included in the method. Lambda hyperparameter was used in Ridge and LASSO methods and the optimal lambda value was determined according to the min./max. ratio (0.0001) according to the minimum mean square error within 50 periods. In ElasticNet regression, both lambda and alpha editing parameters were used and the alpha value was automatically taken as 0.5. Bland-Altman analysis was used to assess the agreement among different methods of age estimation, including the GP age, VSA (Ridge, LASSO, ElasticNet), and CA. Limits of agreement were identified.

RESULTS

Measurement Error

The intra-rater and inter-rater agreements were estimated using the intra-class correlation coefficient (ICC) and were found to be excellent for all vertebral measurements (ICC ≥0.977, and ICC ≥0.960, respectively). Both intra-observer and inter-observer agreements of GP skeletal age were 0.997 (95% confidence interval: 0.996 to 0.997) with excellent reliability.

The First Phase of the Regression Methods

The descriptive statistics in the study are demonstrated for each sex (Table 1). Independent variables with VIF values greater than 10 are shown in bold (Table 2). Our study was conducted separately for both girls and boys.

Statistical analysis consisted of two parts. In the first part, all independent variables (vertebral measurements) were evaluated. The target variable was GP age. To obtain the VSA, three regression methods were used. In the second part, the analyses were repeated with the variables with the highest beta coefficients obtained from each regression model.

In the initial phase of the statistical analysis, the lambda values were chosen based on minimum mean square error values. The beta coefficients and lambda values for each regression model were determined separately for boys and girls, and the results are presented in Table 3. The R^2 values obtained in the first stage of the statistical analysis were between 0.799 and 0.804.

In boys, all variables except one in the Ridge and ElasticNet regressions and 15 variables in the LASSO regression had non-zero beta coefficients. In girls, all variables in the Ridge and ElasticNet regressions and all variables except 11 in the LASSO regression had non-zero beta coefficients.

The Second Phase of the Regression Methods

Due to the high number of independent variables (n=53) statistically evaluated in our study, in the second part of the analysis, the beta coefficients were examined to determine which variables had the greatest impact on the regression models and to select the most important variables for clinical applicability. Separate analyses were conducted for boys and girls, and the eight variables with the highest coefficients in each regression model were chosen.

For both girls and boys, eight measurements with the highest coefficients were selected in each regression model, and a total of 24 measurements were determined. In boys, for the elimination of 24 measurements selected for the second part of the statistical analysis, the first three measurements (SVD, FiVD, FVD) were common to all three regression models and had the highest beta coefficients, and PH3, TVD, TVup-TVlp, and Y3 measurements, which were common to all three models, were selected. In addition, UW3, which was common to ElasticNet and Ridge regressions was selected. For boys, the selected measurements were SVD, FiVD, FVD, PH3, TVD, TVup-TVlp, Y3, and UW3 (Figure 4a, Table 4). In girls, for the elimination of 24

Table 1. The descriptive statistics for each sex

	Boys=329		Girls=465	
Measurements	Mean	Standard deviation	Mean	Standard deviation
Skeletal age	12.9	3.1	13.6	2.7
SVp-SVa	13.2	1.5	11.8	1.1
TVup-TVua	13.1	1.6	11.9	1.3
TVpm-TVam	13.8	1.6	12.9	1.3
TVlp-TVla	13.9	1.4	12.7	1.2
FVup-FVua	13.5	1.7	12.4	1.3
FVpm-FVam	13.7	1.7	12.7	1.4
FVlp-FVla	14.0	1.7	12.8	1.3
FiVup-FiVua	13.6	2.0	12.4	1.5
FiVpm-FiVam	13.8	1.9	12.7	1.4
FiVlp-FiVla	14.4	1.9	13.2	1.4
TVup-TVlp	10.8	2.5	11.2	2.0
TVum-TVd	9.4	2.2	9.9	1.7
TVua-TVla	9.0	2.6	10.0	2.4
FVup-FVlp	10.6	2.4	10.9	2.0
FVum-FVd	9.2	1.9	9.6	1.7
FVua-FVla	8.5	2.2	9.3	2.2
FiVup-FiVlp	10.4	2.4	10.8	2.1
FiVum-FiVd	9.3	1.9	9.5	1.6
FiVua-FiVla	8.3	2.1	9.1	2.0
SVD	1.2	0.7	1.4	0.6
TVD	0.9	0.7	1.1	0.7
FVD	0.8	0.6	1.0	0.6
FiVD	0.7	0.6	0.9	0.6
X3	18.7	2.6	18.1	2.0
Y3	14.5	2.1	14.1	1.7
X4	18.4	2.6	17.8	2.1
Y4	14.8	2.2	14.4	1.8
X5	18.3	2.7	17.6	2.1
Y5	15.2	2.4	14.7	1.8
D3 angle	109.2	12.1	99.6	11.7
D4 angle	111.7	10.6	103.4	10.9
D5 angle	113.7	10.0	105.2	10.4
K3 angle	36.4	8.2	42.9	8.4
K4 angle	33.8	6.6	39.3	7.3
K5 angle	32.2	6.1	37.3	6.5
SI3 angle	8.7	5.3	6.9	5.7
SI4 angle	9.6	4.9	8.0	4.8
SI5 angle	9.5	4.8	8.5	4.5
PH3	10.6	2.5	11.0	2.0
H3	10.1	2.6	10.8	2.0
AH3	8.7	2.6	9.6	2.3
PH4	10.5	2.4	10.8	2.0
H4	9.7	2.3	10.4	2.0
AH4	8.2	2.2	9.1	2.1
PH5	10.4	2.4	10.7	2.1
H5	9.7	2.2	10.2	1.9
AH5	8.1	2.1	8.9	2.0
UW3	12.5	1.5	11.4	1.2
LW3	13.8	1.4	12.5	1.2
UW4	12.9	1.7	11.9	1.3
LW4	13.8	1.7	12.7	1.3
UW5	13.2	1.9	12.1	1.4
LW5	14.3	1.9	13.1	1.5

Table 2. Tolerance and VIF values in boys and girls

Measurements	Boys		Girls	
	Tolerance	VIF	Tolerance	VIF
FVup-FVlp	0	**26202.8**	0	**13112.1**
AH3	0	**3996.27**	0.001	**1840.53**
AH4	0	**3782.6**	0	**5082.87**
AH5	0	**2757.9**	0	**2051.75**
SI3 angle	0.003	**324.895**	0.004	**284.422**
SI4 angle	0.004	**226.826**	0.004	**250.702**
SI5 angle	0.008	**133.063**	0.006	**178.519**
D3 angle	0	**2681.55**	0	**2326.59**
D4 angle	0	**2017.23**	0	**2474.24**
D5 angle	0.001	**1416.93**	0.001	**1722.65**
FiVD	0.184	5.429	0.138	7.267
FiVlp-FiVla	0.001	**1481.23**	0.001	**947.5**
FiVpm-FiVam	0.046	**21.886**	0.092	**10.886**
FiVua-FiVla	0	**2072.05**	0.001	**1479**
FiVum-FiVd	0.026	**37.796**	0.026	**38.321**
FiVup-FiVlp	0	**10946.8**	0	**34200**
FiVup-FiVua	0.001	**670.271**	0.001	**850.398**
FVD	0.137	7.278	0.105	9.536
FVlp-FVla	0.001	**1441.84**	0.001	**810.977**
FVpm-FVam	0.049	**20.232**	0.066	**15.11**
FVua-FVla	0	**3653.2**	0	**4776.87**
FVum-FVd	0.019	**51.932**	0.019	**51.533**
FVup-FVua	0.001	**1114.8**	0.001	**747.306**
H3	0.007	**144.1**	0.008	**127.789**
H4	0.01	**103.722**	0.009	**109.27**
H5	0.013	**75.204**	0.012	**80.036**
K3 angle	0	**2676.69**	0	**2289.14**
K4 angle	0.001	**1527.84**	0.001	**1992.53**
K5 angle	0.001	**917.457**	0.001	**1358.53**
LW3	0.002	**578.674**	0.001	**772.953**
LW4	0.001	**1185.02**	0.002	**632.613**
LW5	0.001	**1133.61**	0.001	**844.341**
PH3	0	**18165.8**	0.001	**1177.72**
PH4	0.001	**1125.44**	0.001	**1337.63**
PH5	0.001	**1018.7**	0.001	**847.303**
SVD	0.353	2.835	0.405	2.47
SVp-SVa	0.311	3.216	0.395	2.529
TVD	0.091	**10.972**	0.084	**11.935**
TVlp-TVla	0.002	**635.172**	0.001	**973.426**
TVpm-TVam	0.058	**17.144**	0.091	**10.958**
TVua-TVla	0	**3760.9**	0.001	**1771.12**
TVum-TVd	0.014	**71.786**	0.014	**69.296**
TVup-TVlp	0.001	**1308.24**	0	**8472.63**
TVup-TVua	0.001	**731.716**	0.002	**481.131**
UW3	0.002	**504.403**	0.003	**368.668**
UW4	0.001	**952.503**	0.002	**652.179**
UW5	0.001	**685.526**	0.001	**741.698**
X3	0.001	**1729.68**	0.001	**979.453**
X4	0.001	**1753.84**	0.001	**1404.84**
X5	0.001	**1302.94**	0.001	**1081.78**
Y3	0.002	**492.202**	0.003	**330.541**
Y4	0.001	**682.317**	0.002	**615.488**
Y5	0.002	**572.438**	0.002	**521.382**

Independent variables with a VIF value of 10 or more are demonstrated in bold
VIF: The variance inflation factor

measurements selected for the second part of the statistical analysis, SVD, FVD, TVum-TVd, TVpm-TVam, UW4, and Y3, which are common to all three regression models and have high beta coefficients were selected.

In addition, FiVD, which is common to Ridge and ElasticNet regression, and UW5, which is common to LASSO and Ridge regression, were selected. For girls, the selected measurements were SVD, FVD, TVum-TVd, TVpm-TVam, UW4, Y3, FiVD, and UW5 (Figure 4b, Table 4). The lambda values and beta coefficients were recalculated based on new datasets created separately for each sex.

In the second phase of the statistical analysis, the lambda values were chosen based on minimum mean square error values. The minimum mean square error for boys was obtained at lambda values 0.0762, 0.000148, and 0.002344 for the Ridge, LASSO, and ElasticNet, respectively. For girls, the minimum mean square error was obtained at lambda values of 0.04913915, 0.00003718, and 0.00000113 for the Ridge, LASSO, and ElasticNet regression, respectively.

The R^2 values obtained in the second stage of the statistical analysis were between 0.740 and 0.783. The highest R^2 in both boys and girls was obtained using LASSO regression (respectively, R^2=0.783, 0.741) (Table 5), and the performance of each regression model was assessed using 10-fold cross-validation.

The means and errors for both the training and test datasets from the initial and second parts of the analyses are presented in Table 6.

Vertebral skeletal age formulas obtained in each regression model in boys:

Ridge regression: VSA=0.318*FVD + 0.561*FiVD + 0.307*PH3 + 0.487*SVD - 0.059*TVD + 0.33*TVup-TVlp + 0.025*UW3 + 0.252*Y3 + 0.889

LASSO regression: VSA=0.185*FVD + 0.534*FiVD + 0.019*PH3 + 0.448*SVD + 0*TVD + 0.647*TVup-TVlp + 0*UW3 + 0.259*Y3 + 0.868

ElasticNet regression: VSA=0.323*FVD + 0.564*FiVD + 0.306*PH3 + 0.483*SVD – 0.048*TVD + 0.326*TVup-TVlp + 0.031*UW3 + 0.249*Y3 + 0.906

Vertebral skeletal age formulas obtained in each regression model in girls:

Ridge regression: VSA=0.528*FiVD + 0.909*FVD + 0.638*SVD + 0.023*TVpm-Tvam + 0.508*TVum-TVd -0.138*UW4 - 0.064*UW5 + 0.456*Y3 + 1.988

LASSO regression: VSA=0.481*FiVD + 0.935*FVD + 0.614*SVD + 0*TVpm-Tvam + 0.513*TVum-TVd -0.149*UW4 - 0.065*UW5 + 0.494*Y3 + 1.892

Table 3. The model results obtained according to the minimum mean square error as a result of the Ridge, LASSO, and ElasticNet regression models in boys and girls in the first part of the analysis

	Boys			Girls		
	Ridge	LASSO	ElasticNet	Ridge	LASSO	ElasticNet
Lambda	0.809	0.000398	0.000027	0.05002	0.000012	0.0000012
SVp-SVa	-0.01	0	-0.005	0.029	0.023	0.019
TVup-TVua	0.04	0.032	0.039	-0.058	-0.026	-0.035
TVpm-TVam	-0.001	0	0.003	0.164	0.205	0.137
TVlp-TVla	0.024	0	0.023	-0.07	-0.083	-0.065
FVup-FVua	0.015	0.001	0.017	0.104	0.104	0.084
FVpm-FVam	0.028	0.014	0.027	-0.062	-0.087	-0.04
FVlp-FVla	0.012	0	0.012	-0.056	-0.046	-0.067
FiVup-FiVua	0.018	0	0.018	0.05	0.014	0.046
FiVpm-FiVam	0	0	0	0.098	0.1	0.08
FiVlp-FiVla	-0.006	0	-0.002	-0.05	-0.075	-0.047
TVup-TVlp	0.045	0.072	0.043	0.051	0.021	0.054
TVum-TVd	0.047	0.057	0.045	-0.173	-0.259	-0.11
TVua-TVla	0.025	0.014	0.025	-0.007	0	0.005
FVup-FVlp	0.038	0.061	0.037	0.042	0.01	0.043
FVum-FVd	0.023	0.008	0.025	-0.038	-0.014	-0.021
FVua-FVla	0.03	0.014	0.031	-0.032	-0.036	-0.004
FiVup-FiVlp	0.035	0.051	0.033	0.009	0	0.015
FiVum-FiVd	0.012	0	0.016	0.022	0	0.011
FiVua-FiVla	0.025	0.008	0.026	-0.014	0	0.002
SVD	0.234	0.307	0.215	0.211	0.186	0.218
TVD	0.103	0.06	0.107	0.046	-0.016	0.106
FVD	0.146	0.123	0.144	0.374	0.418	0.362
FiVD	0.183	0.178	0.178	0.112	0.06	0.116
X3	0.023	0.032	0.024	-0.033	-0.029	-0.022
Y3	0.052	0.072	0.048	0.11	0.139	0.098
X4	0.021	0.03	0.022	-0.044	-0.048	-0.026
Y4	0.029	0.027	0.028	0.063	0.036	0.062
X5	0.019	0.022	0.019	-0.017	0	-0.01
Y5	0.017	0.007	0.018	0.069	0.059	0.06
D3 angle	-0.01	-0.012	-0.01	-0.024	-0.039	-0.021
D4 angle	-0.012	-0.014	-0.012	-0.029	-0.045	-0.023
D5 angle	-0.011	-0.012	-0.011	-0.023	-0.028	-0.018
K3 angle	0.008	0	0.008	0.017	0.012	0.014
K4 angle	0.013	0.003	0.013	0.031	0.03	0.023
K5 angle	0.011	0	0.011	0.025	0.022	0.02
SI3 angle	0.012	0	0.011	0.014	0.012	0.011
SI4 angle	0.013	0	0.013	-0.002	0.002	-0.001
SI5 angle	0.02	0.011	0.019	-0.007	-0.002	-0.006
PH3	0.047	0.07	0.045	0.072	0.076	0.071
H3	0.039	0.048	0.036	0.095	0.116	0.064
AH3	0.03	0.021	0.029	0.029	0	0.03
PH4	0.039	0.052	0.038	0.036	0	0.041
H4	0.031	0.032	0.031	0.091	0.085	0.071
AH4	0.034	0.033	0.034	0.001	0	0.016
PH5	0.035	0.043	0.034	0.016	0	0.02
H5	0.033	0.037	0.032	-0.014	-0.001	0.002
AH5	0.029	0.017	0.029	0.002	0	0.014
UW3	0.049	0.046	0.047	0.099	0.068	0.076
LW3	0.025	0	0.023	-0.019	0	-0.021
UW4	0.023	0.014	0.023	0.157	0.228	0.128
LW4	0.015	0.001	0.015	-0.052	-0.025	-0.06
UW5	0.018	0.001	0.018	0.11	0.165	0.091
LW5	-0.006	0	-0.002	-0.027	-0.005	-0.032
c	2.869	5.309	2.756	8.661	12.86	7.636
df	53	53	53	53	53	53
L1 Norm	4.693	6.964	4.53	11.861	15.918	10.475
R-squared	0.801	0.801	0.8	0.802	0.804	0.799
AIC	0.521	0.522	0.522	0.426	0.424	0.429

LASSO: Least absolute shrinkage and selection operator, AIC: Akaike information criterion

Table 4. Definitions of vertebral measurements used for the regression formula

Vertebral measurements	Definitions
TVup-TVlp[a]	The vertical distance between the uppermost and lowest points of the posterior edge of the 3rd cervical vertebra.
FiVD[ab]	The vertical distance from the deepest point of the 5th cervical vertebra to the plane between the most anterior and posterior points of its lower edge
SVD[ab]	The vertical distance from the deepest point of the 2nd cervical vertebra to the plane between the most anterior and posterior points of its lower edge
Y3[ab]	The most anterior point of the upper border of the 3rd cervical vertebra and the most posterior point of the lower border.
FVD[ab]	The vertical distance from the deepest point of the 4th cervical vertebra to the plane between the most anterior and posterior points of its lower edge
PH3[a]	The distance of the perpendicular from the highest point of the posterior edge of the 3rd cervical vertebra to the plane formed by the most anterior and most posterior points of the lower edge
UW3[a]	The horizontal distance of the perpendicular descending from the highest point of the anterior edge to the plane formed between the upper and lower points of the posterior edge of the 3rd cervical vertebra.
TVD[a]	The vertical distance from the deepest point of the 3rd cervical vertebra to the plane between the most anterior and posterior points of its lower edge
TVum-TVd[b]	The vertical distance between the uppermost and lowest points of the medial edge of the 3rd cervical vertebra.
UW4[b]	The horizontal distance of the perpendicular descending from the highest point of the anterior edge to the plane formed between the upper and lower points of the posterior edge of the 4th cervical vertebra.
UW5[b]	The horizontal distance of the perpendicular descending from the highest point of the anterior edge to the plane formed between the upper and lower points of the posterior edge of the 5th cervical vertebra.
TVpm-TVam[b]	The horizontal distance between the midpoints of the anterior and posterior edges of the 3rd cervical vertebra.

[a]Only boys; [b]Only girls; [ab]Both boys and girls

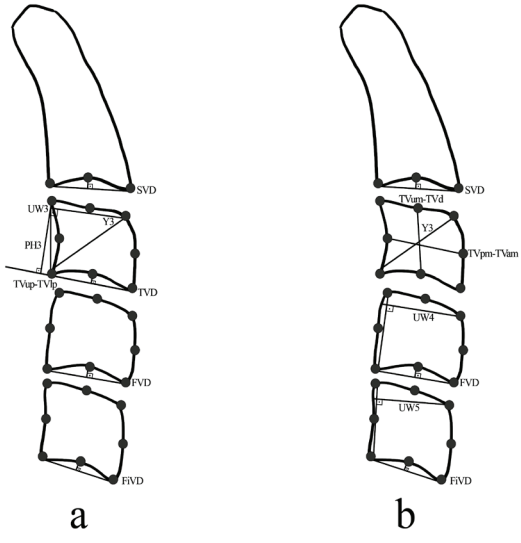

Figure 4. a) The vertebral measurements used for the second phase of the statistical analysis in boys, **b**) The vertebral measurements used for the second phase of the statistical analysis in girls

ElasticNet regression: VSA=0.543*FiVD + 0.891*FVD + 0.642*SVD + 0.033*TVpm-Tvam + 0.498*TVum-TVd - 0.126*UW4 - 0.055*UW5 + 0.435*Y3 + 1.985

The highest power of explainability was obtained using LASSO regression for both girls and boys (Table 5).

Bland-Altman Analysis

Figures 5 and 6 display the Bland-Altman plots illustrating the consistency of inter-age measurements for boys and girls, respectively, including CA, GP age, Ridge regression age, LASSO regression age, and ElasticNet age. The plots depict a solid line indicating zero bias, the middle-dashed line represents the bias, and the outer dashed lines define the limits of agreement.

DISCUSSION

This study identified key findings in skeletal age prediction using Ridge, LASSO, and ElasticNet regression models. Among these, LASSO regression demonstrated the highest R^2 values (0.783 in boys and 0.741 in girls). Additionally, in both sexes, the vertebral depth of concavities exhibited high beta coefficients, highlighting their significance in skeletal age estimation. The Bland-Altman analysis indicated that the limits of agreement for GP age with CA and VSA were wider in boys than in girls, whereas the limits of agreement between CA and VSA were wider in girls than in boys.

Furthermore, although LASSO exhibited the highest R^2, the observed differences in predictive accuracy among Ridge, LASSO, and ElasticNet regression models suggest that the assumption of equal model performance does not hold. The performance variations among models differed, leading to the rejection of the null hypothesis (H_0), which stated that there would be no difference between VSA prediction models developed using Ridge, LASSO, and ElasticNet regression.

Morphologic changes in the cervical vertebrae are considered useful indicators of skeletal development, although the CVM method has some limitations, such as subjectivity and inadequate validity and reproducibility.[22] We attempted to overcome these restrictions by assessing VSA using morphometric measurements. CVM and hand-wrist methods may be consistent,[9,23] making them reliable skeletal maturity indicators, especially when HWR images are unavailable.[24]

The sample sizes in the literature for skeletal age estimation from vertebral measurements varied from 66 to 958 individuals. Our study sample size was larger than in many studies in the literature, except for Roman et al.'s[24] study.[15-17,25]

There are noticeable differences between boys and girls in the timing of the growth spurt (pre-peak, peak, and post-peak). Hägg and Taranger[26] reported that pubertal growth spurts begin on average at the age of 10 years in girls and 12 years in

Table 5. The model results obtained according to the minimum mean square error as a result of the Ridge, LASSO, and ElasticNet regression models in boys and girls in the second part of the analysis

Boys	Ridge	LASSO	ElasticNet	Girls	Ridge	LASSO	ElasticNet
Lambda	0.0762	0.000148	0.002344	Lambda	0.04913915	0.00003718	0.00000113
TVup-TVlp	0.330	0.647	0.326	FVD	0.909	0.935	0.891
FiVD	0.561	0.534	0.564	SVD	0.638	0.614	0.642
SVD	0.487	0.448	0.483	FiVD	0.528	0.481	0.543
Y3	0.252	0.259	0.249	TVum-TVd	0.508	0.513	0.498
FVD	0.318	0.185	0.323	Y3	0.456	0.494	0.435
PH3	0.307	0.019	0.306	UW4	-0.138	-0.149	-0.126
UW3	0.025	0.000	0.031	UW5	-0.064	-0.065	-0.055
TVD	-0.059	0.000	-0.048	TVpm-TVam	0.023	0.000	0.033
C	0.889	0.868	0.906	C	1.988	1.892	1.985
df	8	6	8	df	8	7	8
L1 Norm	3.229	2.959	3.236	L1 Norm	5.251	5.143	5.207
R-squared	0.782	**0.783**	0.781	R-squared	0.74	**0.741**	0.74
AIC	0.267	0.254	0.267	AIC	0.294	0.289	0.295

LASSO: Least absolute shrinkage and selection operator, AIC: Akaike information criterion

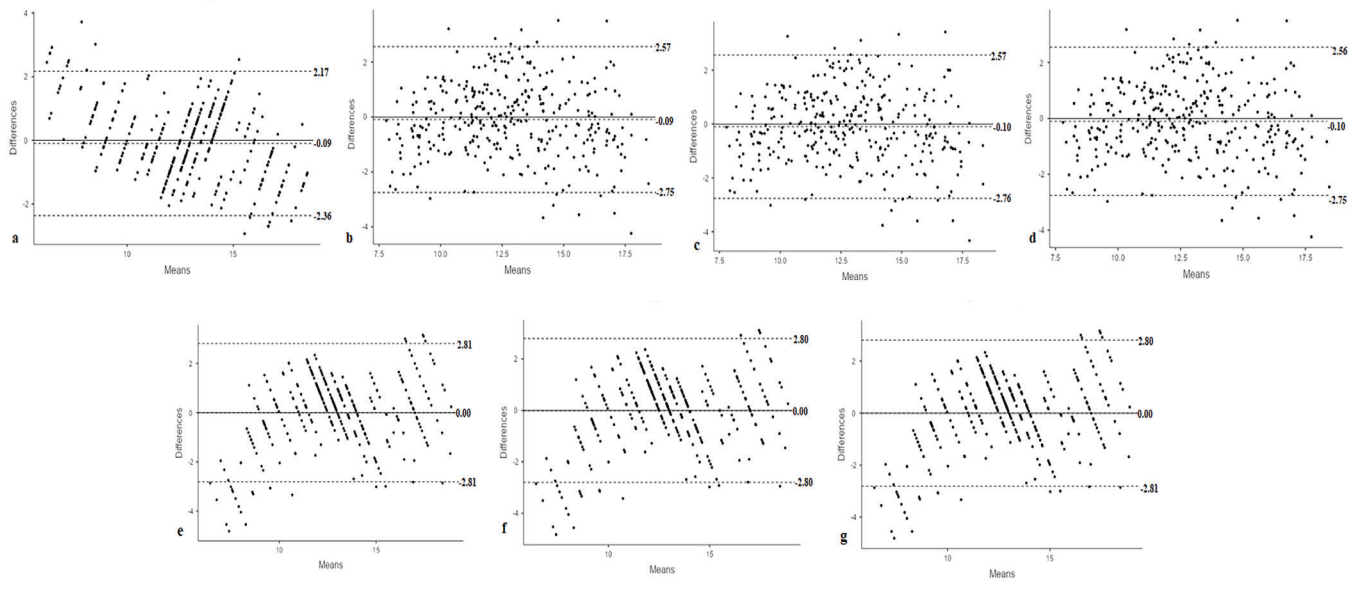

Figure 5. The X-axis represents the means of the 1st and 2nd measurements. The Y-axis represents the differences between the 1st and 2nd measurements. The solid line in the purple area indicates zero bias. The dashed middle line defines bias. The dashed outer lines define the limits of agreement. **a)** The Bland-Altman plot of the consistency between chronologic age (CA) and Greulich-Pyle (GP) age in boys. **b)** The Bland-Altman plot of the consistency between chronologic age (CA) and Ridge regression age in boys. **c)** The Bland-Altman plot of the consistency between chronologic age (CA) and LASSO regression age in boys. **d)** The Bland-Altman plot of the consistency between chronologic age (CA) and ElasticNet regression age in boys. **e)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and Ridge regression age in boys. **f)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and LASSO regression age in boys. **g)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and ElasticNet regression age in boys

boys. Fishman[27] also reported that the pubertal growth spurt ended at the age of 14.77 years in girls and 16.4 years in boys. In the present study, VSA was determined separately in boys and girls because the difference in growth and development between the sexes is often considered important.[24,26,27]

Previous studies examined C2-C5,[9,28,29] C2-C4,[4,8,30] and C3-C4[15-17,25] vertebrae for estimating skeletal age and maturation from cervical vertebrae. In our study, we focused on evaluating C2-C5 vertebrae.

The age range of the sample of our study (7-18 years) was wider than in Caldas et al.'s[25] study (7-15.9 years), Mito et al.'s[17] study (7-14.9 years), and Alhadlaq and Al-Maflehi's[16] study (10-15 years).

Caldas et al.[25] reported that the anterior (TVua-TVla), median (TVum-TVd), and posterior (TVup-TVlp) heights of the C3 vertebrae increased between 10 and 13 years, and the anterior (FVua-FVla), median (FVum-FVd), and posterior (FVup-FVlp) heights of the C4 vertebrae increased between the ages of

Table 6. Means and errors of test and training set in the first and second analyses for boys and girls

SEX (first and second analysis)	Regression model	Lambda	Test Set means	Train set means	Test set errors	Train set errors
Male (first analysis)	**Ridge**	0.809	2.016	1.849	0.159	0.017
	LASSO	0.000398	2.048	1.849	0.164	0.016
	ElasticNet	0.000027	2.018	1.847	0.159	0.017
Male (second analysis)	**Ridge**	0.07621	2.134	2.039	0.149	0.016
	LASSO	0.000148	2.132	2.031	0.152	0.017
	ElasticNet	0.000002	2.135	2.038	0.149	0.016
Female (first analysis)	**Ridge**	0.05002	1.613	1.456	0.121	0.013
	LASSO	0.000012	1.636	1.437	0.17	0.018
	ElasticNet	0.0000012	1.633	1.467	0.17	0.018
Female (second analysis)	**Ridge**	0.04913915	2.016	1.915	0.192	0.021
	LASSO	0.00003718	2.018	1.911	0.191	0.021
	ElasticNet	0.00000113	2.016	1.918	0.192	0.021

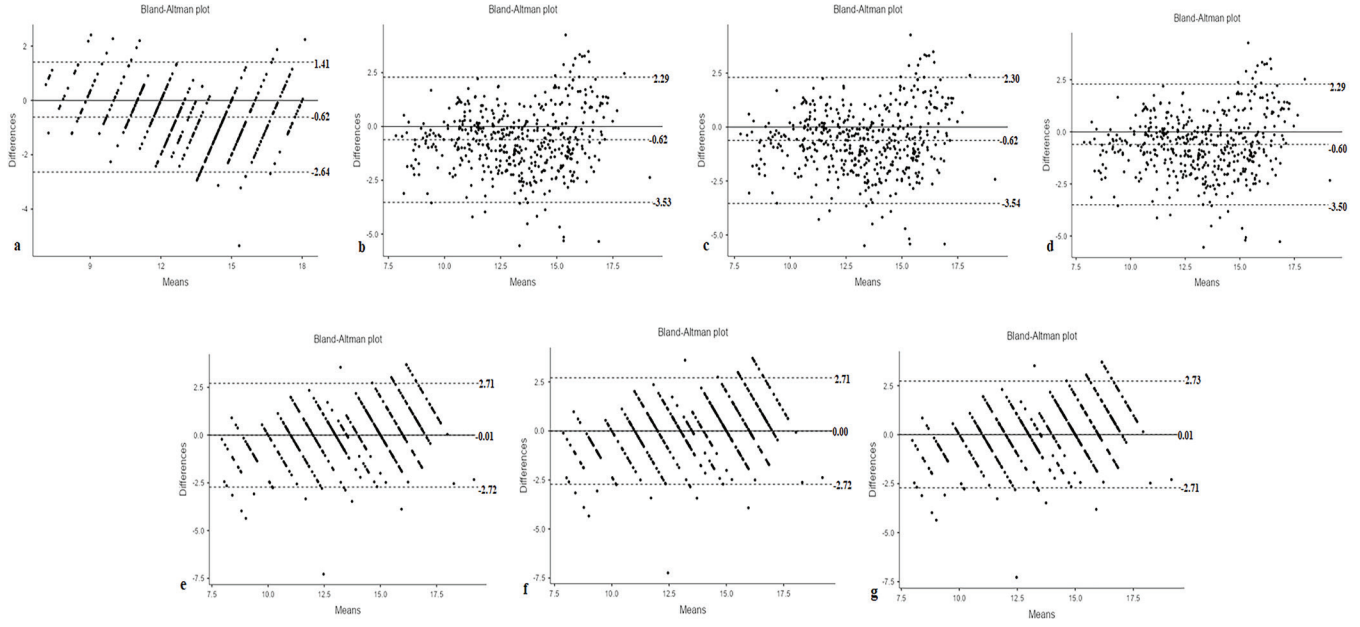

Figure 6. The X-axis represents the means of the 1st and 2nd measurements. The Y-axis represents the differences between the 1st and 2nd measurements. The solid line in the purple area indicates zero bias. The dashed middle line defines bias. The dashed outer lines define the limits of agreement. **a)** The Bland-Altman plot of the consistency between chronological age (CA) and Greulich-Pyle (GP) age in girls. **b)** The Bland-Altman plot of the consistency between the chronologic age (CA) and Ridge regression age in girls. **c)** The Bland-Altman plot of the consistency between chronologic age (CA) and LASSO regression age in girls. **d)** The Bland-Altman plot of the consistency between chronologic age (CA) and ElasticNet regression age in girls. **e)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and Ridge regression age in girls. **f)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and LASSO regression age in girls. **g)** The Bland-Altman plot of the consistency between Greulich-Pyle (GP) age and ElasticNet regression age in girls

11-13 years in girls. In addition, they reported that the anterior (Tvua-Tvla), median (Tvum-TVd), and posterior (Tvup-TVlp) heights and median width (TVpm-Tvam) of the C3 vertebrae increased between 12 and 15 years, but no significant changes were observed in the C4 vertebral measurements in boys.[25] Mito et al.[17] reported that the anterior, median, and posterior heights of the C3 and C4 vertebrae increased rapidly from age 10 to 13 years in girls.

Alhadlaq and Al-Maflehi[16] reported an increase in the heights of the C3 and C4 vertebrae between 10-15 years, but the median width did not change in this period in boys. In the present study, the median height of the C3 vertebrae (TVum-TVd) in girls and the posterior height of the C3 vertebrae (TVup-TVlp) in boys had high beta coefficients, and the coefficients of C3 height measurements were high. However, the concavity depth of all vertebrae may have been more pronounced than C4 height measurements due to the wider age range compared to other studies,[16,17,25] and the higher number of independent variables. Roman et al.[24] found that the most influential variable in determining the vertebral maturation period was the vertebral depth of concavity.

Likewise, concavity depth at the lower border of C4 (FVd) and C3 (TVd) vertebrae in girls and concavity depth at the lower border of C5 (FiVD) and C2 (SVD) vertebrae in boys were found to be the most influential variables in skeletal age estimation.

Generally, stepwise regression has been used in studies to obtain VSA.[15-18] Varshosaz et al.[15] reported that the anterior length of the fourth vertebrae was the most important variable for determining skeletal age by performing a stepwise multivariable regression analysis. The focus of the present study was to introduce different regression models for detecting VSA. The power of explainability in their study was R^2=0.686, whereas, in our study, it was R^2=0.741 in girls and R^2=0.783 in boys.[15] Although both studies were conducted in similar age groups, our study provided separate evaluations for boys and girls. Difference in variables, sample size, ethnic differences, and the use of different regression models may have influenced the results.

Although many studies have reported that evaluating cervical vertebrae with morphologic and morphometric methods yields successful results in skeletal age estimation,[16,17,23-25,29,31] Beit et al.[20] reported that methods based on vertebral morphology were insufficient for estimating skeletal age. In addition to the ratio measurements in their study, the SI angle, which was also included in our study, was included. When the first part of the statistical analysis was examined in our study, the beta coefficient of the SI angle was found to be low, likewise in the study of Beit et al.[20]. Thus, the SI angle was excluded from the second part of the statistical analysis. The explanatory power of this study model (R^2=0.783 for boys, R^2=0.741 for girls) was found to be higher than for Beit et al.[20] (R^2=0.693 for boys and R^2=0.671 for girls). Although our R^2 values are higher than those in the studies by Varshoaz et al.[15] and Beit et al.[20], the clinical advantage was insufficient to predict the skeletal age.

It is important to evaluate the differences between the two methods to assess their compatibility and reproducibility. Bland-Altman analysis was used to examine the agreement between GP age, CA, and VSA.

Varshosaz et al.[15] evaluated the relationship using the correlation method and stated that LCRs are useful for skeletal age estimation and might be an alternative to HWRs, with the advantage of radiation reduction. In the study of Beit et al.[20], the limit of agreement between CA and GP skeletal age (in boys ULA: 2.1, LLA: -1.7; in girls ULA: 2.2, LLA: -1.2) was found to be better than in our study (in boys ULA: 2.17, LLA: -2.36, in girls ULA: 1.41, LLA: -2.64). They reported that the agreement between CA and GP age was higher than the agreement between GP age and VSA in both sexes.[20] In our study, in both CA and VSA (Ridge, LASSO, Elastic Net) graphs of GP age, the width of the limits of agreement was wider in boys than in girls (Figures 5a, e, f, g, 6a, e, f, g). The width of the limits of agreement of CA-VSA (Ridge, LASSO, ElasticNet) was wider in girls than in boys (Figures 5b, c, d, 6b, c, d). Similar to our findings, by comparing GP age with VSA and CA, Beit et al.[20] reported that VSA was not superior to CA. Therefore, differences in interpretation based on statistical analysis methods are important.

In studies performed to obtain VSA, ratio or angular measurements were generally used.[15-17,20,25] In the present study, only linear and angular measurements were included. Although image magnification was mentioned as a disadvantage in the use of linear measurements,[16] the power of explainability in our study was higher than the ratio measurements used in other studies.[15,20]

Circumpubertal growth differences are more closely related to skeletal age than CA. Variations in the maturation stage are closely associated with changes in when and how much growth happens. Comprehending the development of the oro-facial region is crucial for orthodontic therapy. Determining skeletal age is important in creating effective orthodontic treatment plans because patients grow at different times, durations, and velocities. Orthodontic treatment for growth modification requires proper patient selection, appliance prescription, and compliance. Clinical decisions involving extra-oral traction forces, functional appliances, extraction vs. non-extraction therapy, or orthognathic surgery are primarily based on growth considerations.[32,33]

The methods mentioned in our study have provided useful but limited information on determining the timing of orthopedic treatment or confirming the end of growth. Clinicians should know the average differences between chronologic and skeletal ages for each sex and identify ages when there is good concordance or within clinically acceptable limits of treatment or purpose. Suri et al.[32] reported that a 0.5-year difference between skeletal and CA was acceptable in clinical practice. Despite observing high R^2 values, no significant clinical advantage was observed when comparing it with CA in the present study.

Study Limitations

Skeletal age is influenced by ethnic factors.[34] To avoid ethnic influences on skeletal growth and development, only individuals of Turkish ethnicity were included in this study. Although GP atlas assessment has been reported to exhibit minimal inter-observer and intra-observer discrepancies, it should be noted that this evaluation is inherently subjective.[35]

Future studies should be conducted using a group-based approach, employing larger sample sizes and encompassing diverse age ranges within the groups. Variations in vertebral maturations may exhibit dissimilarities across distinct age cohorts. Evaluations can be made about which vertebral variables play a more important role in different age groups.[16,17,25]

This study had several strengths. First, with a sample size of 794 individuals (329 boys, 465 girls), it included a larger dataset than many previous studies evaluating skeletal age through cervical vertebrae measurements, except for Roman et al.'s[24] study.[15-17,25] Second, by incorporating multiple regression models (Ridge, LASSO, and ElasticNet), this study enabled a comparative assessment of different predictive methodologies, providing insights into their strengths and weaknesses. Additionally, the Bland-Altman analysis enhanced reliability by quantifying the agreement between VSA and GP skeletal ages, thereby improving the interpretability of the findings. However, some limitations should be acknowledged. The retrospective design and the the inclusion of only a single ethnic group may limit the generalizability of the results. Future research should incorporate longitudinal data, investigate the influence of ethnic variability on skeletal age prediction, and validate findings using external datasets to improve model robustness and clinical applicability.

CONCLUSION

In our study, the difference in skeletal age estimation was greater than 0.5 years, which does not provide enough information in clinical practice. Relying on VSA alone to determine the skeletal age of individuals within the Turkish population is insufficient for determining the timing of orthopedic treatment or confirming the end of growth.

Ethics

Ethics Committee Approval: The study received ethical approval from the Research Ethics Committee of Recep Tayyip Erdoğan University (date: 02.02.2023 and protocol number: 33).

Informed Consent: Informed written consent forms, which included the use of patient records in scientific studies, were obtained from all patients at the beginning of treatment.

Footnotes

Author Contributions: İ.Y., M.G.; Concept - İ.Y., M.G.; Design - İ.Y., M.G.; Data Collection and/or Processing - İ.Y., M.G.; Analysis and/or Interpretation - İ.Y., M.G.; Literature Search - İ.Y., M.G.; Writing - İ.Y., M.G.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Szemraj A, Wojtaszek-Slominska 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;102:125-128. [CrossRef]
2. Chandrasekar R, Chandrasekhar S, Sundari KKS, Ravi P. Development and validation of a formula for objective assessment of cervical vertebral bone age. *Prog Orthod.* 2020;21(1):38. [CrossRef]
3. Verma D, Peltomaki T, Jager A. Reliability of growth prediction with hand-wrist radiographs. *Eur J Orthod.* 2009;31(4):438-442. [CrossRef]
4. Kok H, Acilar AM, Izgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod.* 2019;20(1):41. [CrossRef]
5. Lucchese A, Bondemark L, Farronato M, et al. Efficacy of the cervical vertebral maturation method: a systematic review. *Turk J Orthod.* 2022;35(1):55-66. [CrossRef]
6. Patcas R, Signorelli L, Peltomäki T, Schätzle M. Is the use of the cervical vertebrae maturation method justified to determine skeletal age? A comparison of radiation dose of two strategies for skeletal age estimation. *Eur J Orthod.* 2013;35(5):604-609. [CrossRef]
7. Booz C, Yel I, Wichmann JL, et al. Artificial intelligence in bone age assessment: accuracy and efficiency of a novel fully automated algorithm compared to the Greulich-Pyle method. *Eur Radiol Exp.* 2020;4(1):6. [CrossRef]
8. Kim DW, Kim J, Kim T, et al. Prediction of hand-wrist maturation stages based on cervical vertebrae images using artificial intelligence. *Orthod Craniofac Res.* 2021;24(Suppl 2):68-75. [CrossRef]
9. Kok H, Izgi MS, Acilar AM. Determination of growth and development periods in orthodontics with artificial neural network. *Orthod Craniofac Res.* 2021;24(Suppl 2):76-83. [CrossRef]
10. Dallora AL, Anderberg P, Kvist O, et al. Bone age assessment with various machine learning techniques: a systematic literature review and meta-analysis. *PLoS One.* 2019;14(7):e0220242. [CrossRef]
11. Krishnan NMA, Kodamana H, Bhattoo R. Model refinement. Machine learning for materials discovery: numerical recipes and practical applications. Cham: Springer International Publishing; 2024. p. 131-143. [CrossRef]
12. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction: Springer; 2009. [CrossRef]
13. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc B: Stat Methodol.* 1996;58(1):267-288. [CrossRef]
14. Zou HH. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B.* 2005;67(2):301-320. [CrossRef]
15. Varshosaz M, Ehsani S, Nouri M, Tavakoli MA. Bone age estimation by cervical vertebral dimensions in lateral cephalometry. *Prog Orthod.* 2012;13(2):126-131. [CrossRef]
16. Alhadlaq AM, NS Al-Maflehi. New model for cervical vertebral bone age estimation in boys. *King Saud Univ J Dent Sci.* 2013;4(1):1-5. [CrossRef]
17. Mito T, Sato K, Mitani H. Cervical vertebral bone age in girls. *Am J Orthod Dentofacial Orthop.* 2002;122(4):380-385. [CrossRef]

18. Chen F, Terada K, Hanada K. A special method of predicting mandibular growth potential for Class III malocclusion. *Angle Orthod.* 2005;75(2):191-195. [CrossRef]
19. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb).* 2015;25(2):141-151. [CrossRef]
20. Beit P, Peltomaki T, Schatzle M, Signorelli L, Patcas R. Evaluating the agreement of skeletal age assessment based on hand-wrist and cervical vertebrae radiography. *Am J Orthod Dentofacial Orthop.* 2013;144(6):838-847. [CrossRef]
21. Association WM. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bull World Health Organ.* 2001;79(4):373. [CrossRef]
22. Hussain U, Ul Hassan F, Kamran MA, et al. Inter-observer and intra-observer agreement of cervical vertebral maturation staging: A systematic review and meta-analysis. *Int Orthod.* 2024;22(3):100874. [CrossRef]
23. Chen LL, Xu TM, Jiang JH, Zhang XZ, Lin JX. Quantitative cervical vertebral maturation assessment in adolescents with normal occlusion: a mixed longitudinal study. *Am J Orthod Dentofacial Orthop.* 2008;134(6):720.e1-720.e7; discussion 720-1. [CrossRef]
24. Roman SP PJ, Oteo MD, Nevado E. Skeletal maturation determined by cervical vertebrae development. *Eur J Orthod.* 2002;24(3):303-311. [CrossRef]
25. Caldas Mde P, Ambrosano GM, Haiter Neto F. New formula to objectively evaluate skeletal maturation using lateral cephalometric radiographs. *Braz Oral Res.* 2007;21(4):330-335. [CrossRef]
26. U. Hägg, J Taranger. Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt. *Acta Odontol Scand.* 1980;38(3):187-200. [CrossRef]
27. Fishman LS. Chronological versus skeletal age, an evaluation of craniofacial growth. *Angle Orthod.* 1979;49(3):181-189. [CrossRef]
28. Kok H, Izgi MS, Acilar AM. Evaluation of the artificial neural network and naive bayes models trained with vertebra ratios for growth and development determination. *Turk J Orthod.* 2021;34(1):2-9. [CrossRef]
29. Chen L, Liu J, Xu T, Long X, Lin J. Quantitative skeletal evaluation based on cervical vertebral maturation: a longitudinal study of adolescents with normal occlusion. *Int J Oral Maxillofac Surg.* 2010;39(7):653-659. [CrossRef]
30. Xie L, Tang W, Izadikhah I, et al. Intelligent quantitative assessment of skeletal maturation based on multi-stage model: a retrospective cone-beam CT study of cervical vertebrae. *Oral Radiol.* 2022;38(3):378-388. [CrossRef]
31. Gandini P, Mancini M, Andreani F. A comparison of hand-wrist bone and cervical vertebral analyses in measuring skeletal maturation. *Angle Orthod.* 2006;76(6):984-989. [CrossRef]
32. Suri S, Prasad C, Tompson B, Lou W. Longitudinal comparison of skeletal age determined by the Greulich and Pyle method and chronologic age in normally growing children, and clinical interpretations for orthodontics. *Am J Orthod Dentofacial Orthop.* 2013;143(1):50-60. [CrossRef]
33. Hashim HA, Mansoor H, Mohamed MHH. Assessment of Skeletal Age Using Hand-Wrist Radiographs following Bjork System. *J Int Soc Prev Community Dent.* 2018;8(6):482-487. [CrossRef]
34. 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(1):228-235. [CrossRef]
35. Dahlberg PS, Mosdol A, Ding Y, et al. A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. *Eur Radiol.* 2019;29(6):2936-2948. [CrossRef]