Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Orthodontic analysis improves the accuracy of lip morphology prediction in facial reconstruction

Nanami Aoki
aoki.nanami@tmd.ac.jp

Institute of Science Tokyo
Hajime Utsuno
Institute of Science Tokyo
Shuuji Namiki
Institute of Science Tokyo
Muhammad Garry Syahrizal Hanafi
Institute of Science Tokyo
Hitomi Nojima
Tokyo Medical and Dental University
Hisako Saitoh
Institute of Science Tokyo
Saki Minegishi
Institute of Science Tokyo
Sayaka Yamada
Institute of Science Tokyo
Hiroyuki Harada
Institute of Science Tokyo
Hideaki Kamochi
Institute of Science Tokyo
Eriko Marukawa
Institute of Science Tokyo
Koichi Sakurada
Institute of Science Tokyo

Research Article

Keywords: Lip morphology, forensic facial reconstruction, facial soft tissue thickness, multiple regression model, Northwestern analysis

Posted Date: June 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9742771/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

In forensic facial reconstruction, lip morphology is difficult to estimate because of the lack of direct bony support. Methods have been developed to evaluate not only average facial soft tissue thickness (FSTT) but also its relationship with the distance between cranial landmarks. In recent years, improved prediction accuracy of FSTT regression models using orthodontic analysis items has been reported, suggesting that these items may help improve the prediction accuracy of lip morphology. In this study, 201 multidetector computed tomography images from living individuals (training group: n = 161, test group: n = 40) were used. Horizontal and vertical distances, FSTT measurements of the lip region, and Northwestern analysis items were measured. Multiple regression models were then constructed using the training group, including models with orthodontic analysis items as explanatory variables (Orthodontic model) and models without them (Non-orthodontic model). When accuracy was validated in the test group, the developed models showed acceptable prediction accuracy and generalization performance. Furthermore, when the absolute errors calculated by each model were compared, the Orthodontic model showed significantly lower values, mainly in the items indicating the height of the vermilion lip, as well as in the FSTT of the Incisor superior ($p < 0.05$). Therefore, considering the relationship with orthodontic analysis items may contribute to improving the accuracy of lip morphology prediction.

Introduction

Forensic facial reconstruction (FFR) is a forensic anthropological technique used to estimate the facial appearance of unidentified skeletal remains and is performed as a last resort when no suitable candidate emerges despite identification efforts [1, 2]. This technique does not directly identify an individual; rather, it serves as an auxiliary method that provides clues for identification by estimating antemortem facial appearance based on the relationship between the skull and the overlying soft tissue [2]. Information is then sought using the reconstructed face, and once a potential candidate emerges, the superimposition method is used to compare the antemortem photograph with the unidentified skull to confirm identity [3]. In this process, one of the main factors determining facial contour is facial soft tissue thickness (FSTT). FSTT is defined as the distance from anthropological landmarks on the skull to the skin surface and is widely used as a fundamental indicator for morphological estimation in facial reconstruction. However, FSTT alone cannot fully reconstruct facial morphology. In particular, the lip region lacks direct bony support; therefore, estimation based solely on FSTT has limitations. Accordingly, methods have been developed to estimate lip morphology based not only on FSTT but also on the relationship between the distances between soft tissue landmarks and skull landmarks [4, 5].

In a classic work on FFR, Gerasimov pointed out that although lip soft tissue does not directly correspond to skull morphology, it is influenced by the morphological features of the dentition and jawbones [6]. However, these findings are based mainly on observation and are nonquantitative; therefore, their direct application to lip morphology estimation is limited.

Furthermore, in orthodontics, studies have also examined the relationship between the skeleton and soft tissues, and associations have been reported between changes in lip soft tissue and the movement of incisors and jawbones during treatment [7–10]. However, these studies mainly focus on changes before and after treatment and differ in purpose and approach from facial reconstruction, which aims to estimate lip morphology directly from the skull.

Importantly, in cranial measurements, the distance between landmarks alone does not adequately reflect morphological features such as the anteroposterior and vertical relationships of the teeth and jawbones, making it difficult to adequately capture lip morphology.

In recent years, regression models that incorporate orthodontic variables as independent variables have been developed to predict FSTT while accounting for morphological features of the teeth and jawbones, and improved predictive accuracy has been reported [11]. In contrast, the predictive accuracy of regression models incorporating orthodontic analysis items for distances between soft tissue landmarks associated with lip morphology has not been adequately investigated.

In orthodontics, detailed analysis items have been established for the comprehensive evaluation of craniofacial morphology, and various cephalometric analysis methods based on cephalometric radiographs are widely used [12]. These methods include

Downs analysis [13], Steiner analysis [14], Tweed analysis [15, 16], and Northwestern analysis [17, 18]. By evaluating the positional relationships between the jawbones and dentition, these methods help clarify the characteristics of maxillofacial morphology. Northwestern analysis was specifically selected for this study due to its integrative nature. While Steiner and Tweed are often tailored toward specific clinical treatment goals, Northwestern analysis integrates key elements from Downs analysis to provide a more multifaceted evaluation by assessing skeletal and dental relationships independently [18]. In the context of FFR, this distinct separation is crucial because it allows for a more granular understanding of how bone structure versus dental protrusion individually contributes to the overlying lip profile, a level of detail that is often less emphasized in other traditional analyses.

Therefore, in this study, regression models were developed to predict lip soft tissue morphology using distances between soft tissue landmarks in the lip region and FSTT as dependent variables, and distances between skull landmarks together with orthodontic analysis items (Northwestern analysis) as independent variables. The predictive accuracy of these models was then evaluated. This study aimed to clarify the usefulness of orthodontic analysis in predicting lip soft tissue morphology and to establish a more accurate method for estimating lip morphology.

Materials and methods

Samples

Multi-detector computed tomography (MDCT) images obtained for clinical purposes at the Department of Oral Surgery, Institute of Science Tokyo Hospital, were used. The sample included 201 Japanese patients aged 18–81 years (104 males, 97 females) who visited the hospital between April 2019 and August 2025. Only images with the mouth closed and in occlusion were included. Images from patients meeting any of the following criteria were excluded: facial deformities or congenital malformations; maxillofacial trauma, bone fractures, or a history thereof; a history of orthodontic treatment; minor maxillofacial surgery within the past 6 months; lack of posterior occlusal support; anterior tooth loss; or metal restorations in the anterior region causing artifacts.

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Review Committee of the Department of Dentistry, Institute of Science Tokyo (Approval No. D2018-058-08). Because this was a retrospective study, the requirement for informed consent was waived. Instead, information about the study was posted on websites, and participants were given an opportunity to opt out.

Measurements

CT viewer software OsiriX MD (version 14.0.1; Pixmeo SARL, Bernex, Switzerland) was used for measurements, and Adobe Photoshop 2026 image editing software (version 27.5.0; Adobe Inc., San Jose, CA, USA) was used to prepare explanatory images of the measured items. First, anthropological landmarks, including 18 soft tissue and 22 hard tissue landmarks, were plotted with the Frankfort Horizontal Plane (Online Resource 1) as the reference. Their definitions and abbreviations are shown in Tables 1 and 2 [12, 19–21]. Next, soft and hard tissue distances and angles were measured based on these landmarks. For soft tissues, vertical and horizontal distances (Fig. 1a) and FSTT (Fig. 1b) were measured [4, 5, 22]. For hard tissues, vertical and horizontal distances (Fig. 2) [4, 5] and Northwestern analysis items (Fig. 3) [11, 12] were measured. Definitions of the measured items are shown in Tables 3 and 4.

Table 1
Definitions of soft tissue landmarks.

Abbreviation	Landmark	Definition
Median points		
Sn	Subnasale	Median point between the lower edge of the nasal septum and the philtrum
Ss'	Subspinale (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Subnasale
Ls	Labiale superius	The most anterior point of the upper vermilion lip border
Pr'	Prosthion (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Labiale superius
Sto	Stomion	Midline point of the labial fissure when the lips are naturally closed and the teeth are in their natural positions
InS'	Incisor superior (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Stomion
Li	Labiale inferius	Midpoint of the lower vermilion border
Id'	Infradentale (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Labiale inferius
Labm	Labiomentale	The deepest point between the lower vermilion lip and the chin
Sm'	Supramentale (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Labiomentale
ChinA	Chin anterior	The most anterior point of the chin
Pg'	Pogonion (soft tissue)	Point adjusted for horizontal tilt between soft tissue and the skull midline at Chin anterior
Bilateral points (L/R)		
Sbal	Subalare	The lowest point of the nasal ala
Cph	Crista philtra	Point on the raised edge of the philtrum just before it projects into the vermilion line
Ch	Cheilion	The most lateral point at the corner of the mouth where the upper and lower lips meet

Table 2
Definitions of hard tissue landmarks.

Abbreviation	Landmark	Definition
Median points		
N	Nasion	The most anterior point of the frontonasal suture
S	Sella turcica	The center of the sella turcica
Ak	Akanthion	Tip of the anterior nasal spine
Ss	Subspinale	The deepest point below the anterior nasal spine on the midsagittal line (orthodontic point A)
Pr	Prosthion	Midline point between the central incisors at the most anterior edge of the maxillary alveolar margin
CejU1	Mid-cej U1	Midpoint between the apexes of the cementoenamel junctions of the maxillary central incisors
InS	Incisor superior	Midpoint between the mesial incisal edges of the maxillary central incisors
InI	Incisor inferior	Midpoint between the mesial incisal edges of the mandibular central incisors
Id	Infradentale	The midline point between the central incisors at the anterior upper border of the mandibular alveolar region
Sm	Supramentale	The deepest point of the groove above the mental eminence on the midsagittal line (orthodontic point B)
Pg	Pogonion	The most prominent point of the mental eminence
Gn	Gnathion	The most anterior point of the mandibular symphysis
Me	Menton	The lowest point of the mental symphysis on the midsagittal line
Bilateral points (L/R)		
CN	Nasal cavity	The most lateral point of the piriform aperture
U1d	U1 distal	Distal point of the maxillary central incisor
U3d	U3 distal	Distal point of the maxillary canine
CejU3	Cej U3	Apex of the cementoenamel junction of the maxillary canine
GoL	Gonion L	The most laterally protruding point of the angle formed by the posterior border of the mandibular ramus and the mandibular body on the left side

1. **(A)** Vertical and horizontal measurements (a-l)
2. **(B)** Facial soft tissue thickness (FSTT) (a-f)

The measurement labels correspond to those in Table 3

1. Frontal view (a-l, o)
2. Occlusal view (m, n)

The measurement labels correspond to those in Table 4

1. Items indicating jaw position (a–e)
2. Items indicating the axial relationship of the anterior teeth (f–j)

The measurement labels correspond to those in Table 4

Table 3
Definitions of soft tissue measurements.

Label[1]	Measurement	Definition
	Vertical distance	
A-a	Philtrum height	The vertical distance between Subnasale (Sn) and Labiale superius (Ls)
A-b	Total upper lip height	The vertical distance between Subnasale (Sn) and Stomion (Sto)
A-c	Total lip height	The vertical distance between Subnasale (Sn) and Labiomentale (Labm)
A-d	Lower facial height to ChinA	The vertical distance between Subnasale (Sn) and Chin anterior (ChinA)
A-e	Upper midline vermilion height	The vertical distance between Labiale superius (Ls) and Stomion (Sto)
A-f	Total vermilion height	The vertical distance between Labiale superius (Ls) and Labiale inferius (Li)
A-g	Lower vermilion height	The vertical distance between Stomion (Sto) and Labiale inferius (Li)
A-h/i	Upper paramedian vermilion height (L/R)	The vertical distance between Crista philtri (Cph) L/R and Stomion (Sto)
	Horizontal distance	
A-j	Nasal alar base width	The linear distance between both Subalare points (Sbal)
A-k	Philtrum width	The linear distance between both Crista philtri (Cph)
A-l	Mouth width	The linear distance between both Cheilion (Ch)
	FSTT[2]	
B-a	Ss–Ss'	The linear distance between Subspinale (Ss) and Subspinale (soft tissue) (Ss')
B-b	Pr–Pr'	The linear distance between Prosthion (Pr) and Prosthion (soft tissue) (Pr')
B-c	InS–InS'	The linear distance between Incisor superior (InS) and Incisor superior (soft tissue) (InS')
B-d	Id–Id'	The linear distance between Infradentale (Id) and Infradentale (soft tissue) (Id')
B-e	Sm–Sm'	The linear distance between Supramentale (Sm) and Supramentale (soft tissue) (Sm')
B-f	Pg–Pg'	The linear distance between Pogonion (Pg) and Pogonion (soft tissue) (Pg')

[1]Measurement labels correspond to those in Fig. 1.

[2]FSTT: Facial soft tissue thickness.

Table 4
Definitions of hard tissue measurements.

Label[1]	Measurement	Definition
	Vertical distance	
2-a	Ak–Pr	The vertical distance between Akanthion (Ak) and Prosthion (Pr)
2-b	Ak–InS	The vertical distance between Akanthion (Ak) and Incisor superior (InS)
2-c	Ak–Me	The vertical distance between Akanthion (Ak) and Menton (Me)
2-d	Ss–Pr	The vertical distance between Subspinale (Ss) and Prosthion (Pr)
2-e	Ss–Sm	The vertical distance between Subspinale (Ss) and Supramentale (Sm)
2-f	Pr–InS	The vertical distance between Prosthion (Pr) and Incisor superior (InS)
2-g	Pr–Id	The vertical distance between Prosthion (Pr) and Infradentale (Id)
2-h	InI–Id	The vertical distance between Incisor inferior (InI) and Infradentale (Id)
2-i	Id–Sm	The vertical distance between Infradentale (Id) and Supramentale (Sm)
2-j	Id–Me	The vertical distance between Infradentale (Id) and Menton (Me)
2-k	CejU1–InS	The vertical distance between Mid-cej U1 (CejU1) and Incisor superior (InS)
	Horizontal distance	
2-l	CN(L–R)	The linear distance between both Nasal cavity points (CN)
2-m	U1d(L–R)	The linear distance between both U1 distal points (U1d)
2-n	U3d(L–R)	The linear distance between both U3 distal points (U3d)
2-o	CejU3(L–R)	The linear distance between both Cej U3 points (CejU3)
	Northwestern analysis	
3-a	N–A–P	The angle between the lines Nasion (N)–Point A (Ss) and Point A (Ss)–Pogonion (Pg)
3-b	SNA	The angle between the lines Nasion (N)–Sella turcica (S) and Nasion (N)–Point A (Ss)
3-c	SNB	The angle between the lines Nasion (N)–Sella turcica (S) and Nasion (N)–Point B (Sm)
3-d	ANB	The angle between the lines Nasion (N)–Point A (Ss) and Nasion (N)–Point B (Sm) (ANB = SNA-SNB)
3-e	Gn–Go to S–N	The angle between the Gnathion (Gn–Gonion L (GoL) plane and the Nasion (N)-Sella turcica (S) plane
3-f	U1 to NS	The angle between the axis of the maxillary left central incisor and the Nasion (N)-Sella turcica (S) plane
3-g	L1 to Gn–Go	The angle between the axis of the left mandibular central incisor and the Gnathion (Gn)-Gonion L (GoL) plane (internal angle)
3-h	L1 to Op	The angle between the axis of the lower left central incisor and the occlusal plane (downward)
3-i	U1 to L1	The internal angle between the axis of the left mandibular central incisor and the axis of the left maxillary central incisor
3-j	InS to NP mm	The distance from Incisor superior (InS) to the Nasion (N)–Pogonion (Pg) plane

[1]Measurement labels correspond to those in Figs. 2 and 3.

In FSTT studies, soft tissue landmarks corresponding to hard tissue landmarks Ss, Pr, InS, Id, Sm, and Pg are commonly defined as Sn, Ls, Sto, Li, Labm, and Pg', respectively [11, 23–25]. However, on three-dimensional images such as MDCT, the coordinates of soft tissue landmarks derived from hard tissue landmarks do not always coincide with those based on actual soft tissue morphology. Therefore, to obtain more accurate soft tissue measurements, this study defined soft tissue landmarks based on soft tissue morphology as Sn, Ls, Sto, Li, Labm, and ChinA, whereas those based on the hard tissue landmarks used for FSTT measurements were defined as Ss', Pr', InS', Id', Sm', and Pg'.

Regarding MDCT, before mid-July 2023, the Department of Dentistry at Science Tokyo Hospital used a 64-slice detector CT system (SOMATOM Sensation 64; Siemens Healthcare, Erlangen, Germany) with a tube voltage of 120 kV, tube current of 84mA, and slice thickness of 0.6 mm. Thereafter, an 80-slice detector CT system (Aquilion Lightning Helios Edition; Canon Inc., Tokyo, Japan) was used, with a tube voltage of 120 kV, tube current of 175 mA, and slice thickness of 0.5 mm. Whyms et al. reported that various combinations of scan parameters, including differences in MDCT slice thickness, do not significantly affect linear and angular measurements derived from three-dimensional computed tomography-rendered models [26]. Accordingly, data obtained from different MDCT systems were considered comparable and were analyzed together in this study.

Analysis

Data were analyzed using Microsoft Excel (version 2603) and IBM SPSS Statistics (version 31; IBM Corp., Armonk, NY, USA). To assess measurement reproducibility, the intraclass correlation coefficient (ICC) was calculated for intra- and inter-observer errors in 20 randomly selected samples. Intra-observer errors were evaluated from two repeated measurements obtained at an interval of at least 2 weeks. The ICC criteria were based on Table 1 of Imai et al. [27]. Based on the ICC results, measurements with low reliability were excluded from subsequent analyses.

The samples were divided into a training group of 161 individuals (82 males, 79 females) and a test group of 40 individuals (22 males, 18 females). Descriptive statistical analysis was performed to confirm the age, sex, and skeletal (ANB) and dental features (InS to NP mm, U1 to L1) of all samples (total group), the training group, and the test group. In addition, the mean (standard deviation) and median (interquartile range) of the measured values were calculated, and the following analyses were conducted.

The following analyses were performed in the training group.

1. Normality assessment: The normality of the data was assessed using the Shapiro–Wilk test and Q–Q plots to determine the appropriate statistical tests for subsequent analyses.
2. Analysis of sex differences: Sex differences in the soft tissue measurement items were analyzed using the Welch's t-test. Cohen's d was calculated as the effect size to evaluate the magnitude of the sex differences. The interpretation of Cohen’s d was based on the Cohen’s criteria (|d|: 0.2 small, 0.5 medium, 0.8 large) [28]. For measurement items that did not meet the assumption of normality, the Mann–Whitney U test was additionally performed as a robustness check.
3. Analysis of differences among age groups: Age-related differences in the soft tissue measurement items were analyzed. First, the samples were divided into five age groups: 1, 18–29 years (n = 19); 2, 30–39 years (n = 33); 3, 40–49 years (n = 39); 4, 50–59 years (n = 40); and 5, 60–81 years (n = 30). To assess group differences, Welch's analysis of variance (ANOVA) was performed followed by the Games–Howell post hoc test. Furthermore, Cohen's d was calculated for adjacent age groups where age-related changes were particularly pronounced. Next, based on these results, the samples were reclassified into two groups, and Welch's t-test was performed. Cohen's d was calculated to evaluate the magnitude of the age group differences. For measurement items that did not meet the assumption of normality, the Kruskal–Wallis test with Bonferroni correction for five-group comparisons and the Mann–Whitney U test for two-group comparisons were additionally performed as a robustness check.
4. Development of multiple regression models: First, multiple regression analysis (forced entry method) was performed using soft tissue measurement items as dependent variables and hard tissue analysis items, sex, and age group as independent variables. To evaluate the effect of orthodontic analysis items, two types of multiple regression models were constructed: one including these items (Orthodontic model) and one excluding them (Non-orthodontic model). Before regression

analysis, Pearson correlation analysis was used to evaluate the relationships between the dependent and independent variables and among the independent variables themselves. Spearman correlation analysis was additionally performed to confirm the consistency of the results. The interpretation of the correlation coefficient r was based on the Cohen's criteria (|r|: 0.1 small, 0.3 medium, 0.5 large) [28]. The possibility of multicollinearity was examined from the correlation results among the independent variables, and multicollinearity was evaluated using variance inflation factor (VIF) values. Because even VIF values below 5 may affect the results, VIF values of 2 or higher were considered to indicate multicollinearity in this study [29]. Independent variables were selected based on the simple correlation coefficients with the dependent variables, together with anatomical relevance and statistical validity, such as multicollinearity, and the final models were developed. The normality and homoscedasticity of the residuals were evaluated using P–P plots and residual plots to assess the assumption of the multiple regression analysis. To verify the validity of the developed models, the adjusted coefficient of determination (adjR²) and the standard error of the estimate (SEE) were calculated.

In the test group, the accuracy of each of the two developed models was validated, their generalization performance was assessed, and their predictive accuracy was compared. For accuracy validation, the correlation coefficient between the observed and predicted values (Test r) and the mean absolute error in the test group (Test MAE) were calculated. The MAE was calculated using the following formula.

MAE = $\sum_{1}^{n} |\text{Measured value} - \text{Estimated value}| / n$

The absolute errors between the observed and predicted values were compared between the training and test groups to investigate generalization performance. The comparisons were performed using the Mann–Whitney U test. In addition, the effect size r was calculated, and its interpretation was based on Cohen's criteria (|r|: 0.1 small, 0.3 medium, 0.5 large) [28].

To evaluate differences in predictive accuracy between the two models, the differences in the coefficient of determination ($\Delta adjR^2$), Test r (ΔTest r), and Test MAE (ΔTest MAE) between the models ((Orthodontic model value) – (Non-orthodontic model value)) were calculated. ΔadjR² or ΔTest r values exhibiting positive and ΔTest MAE values exhibiting negative values indicate better accuracy of the Orthodontic model. In addition, the absolute error between the observed and predicted values was tested. The Wilcoxon signed-rank test was used to compare the models and the effect size r was calculated.

Results

Measurement reproducibility

Intra- and inter-observer ICCs generally showed good reproducibility. Detailed results are presented in Online Resource 2. Intra-observer ICC (1.1) showed reliable values for all measured items, whereas inter-observer ICC (2.1) showed a low value of 0.50 only for InI–Id. Therefore, the measured items were analyzed, excluding InI–Id.

Descriptive statistical analysis

The demographic, skeletal, and dental characteristics of the total, training, and test groups are summarized in Table 5. The distribution of age, sex, and orthodontic variables (ANB, InS to NP mm, and U1 to L1) was comparable across all groups, with an overall mean age of 46.58 ± 14.01 years. Detailed descriptive statistics, including median and interquartile range values for all measurements, are provided in Online Resource 3.

Table 5
Demographic, skeletal, and dental characteristics of the samples.

Sample group		Age (years)		ANB (°)		InS to NP mm		U1 to L1 (°)	
	n	Mean	SD[1]	Mean	SD[1]	Mean	SD[1]	Mean	SD[1]
Total group									
Total	201	46.58	14.01	3.95	2.37	9.40	3.87	129.56	12.98
Male	104	45.24	14.61	3.49	2.38	8.62	3.77	127.81	11.31
Female	97	48.01	13.26	4.44	2.28	9.43	3.99	131.43	14.39
Training group									
Total	161	46.65	13.65	3.93	2.35	9.34	3.80	129.66	12.52
Male	82	45.37	14.30	3.47	2.41	9.30	3.73	128.18	11.40
Female	79	47.97	12.90	4.41	2.19	9.37	3.89	131.21	13.48
Test group									
Total	40	46.30	15.55	4.03	2.50	9.64	4.16	129.13	14.87
Male	22	44.77	16.07	3.58	2.33	9.58	4.00	126.47	11.12
Female	18	48.17	15.13	4.59	2.67	9.72	4.48	132.39	18.29

[1]SD: Standard deviation.

Assessment of normality

The results of the Shapiro–Wilk test are shown in Online Resource 4. Mouth width, Pr–Pr', InS–InS', Sm–Sm', CejU1–InS, and L1 to Op demonstrated significant deviations from normality. Q–Q plots indicated that most measurement items nearly followed a normal distribution. Q–Q plots for the items that showed significant deviations from normality in the Shapiro–Wilk test are shown in Online Resource 5. Although some deviations were observed, the data were largely distributed along the diagonal line. Therefore, for these items, both parametric and nonparametric tests were used to confirm the robustness of the results.

Sex and age differences

The results of the sex difference analysis for soft tissue measurement items are shown in Table 6. The items indicating vermilion height did not vary significantly between sexes ($p > 0.05$). Conversely, males exhibited significantly increased dimensions across all other categories, including vertical facial heights, horizontal widths (nasal and mouth), and all FSTT measurements. The most pronounced sex differences were observed in Ss–Ss' and Pr–Pr', both of which yielded large effect sizes (Cohen's $d > 1.4$). In the multiple regression analysis, sex was converted to a dummy variable and included as an independent variable (Sex group).

For items that deviated significantly from normality, the Mann–Whitney U test was additionally performed (Online Resource 6). The results were generally consistent with those of Welch's t-test.

Table 6
Comparison of soft tissue measurements between males and females.

Measurement (mm)	Male (n = 82)		Female (n = 79)		Sex difference	
	Mean	SD[1]	Mean	SD[1]	*p*	Cohen's d [95% CI[2]]
Vertical distance						
Philtrum height	16.46	2.71	15.06	2.38	< 0.001*	0.55 [0.23, 0.86]
Total upper lip height	23.39	2.47	22.04	2.44	< 0.001*	0.55 [0.23, 0.86]
Total lip height	40.36	4.46	38.05	4.04	< 0.001*	0.54 [0.23, 0.86]
Lower facial height to ChinA	52.50	4.64	48.56	4.24	< 0.001*	0.89 [0.56, 1.21]
Upper midline vermilion height	6.92	1.80	6.97	1.39	0.826	-0.04 [-0.34, 0.28]
Total vermilion height	16.13	3.43	15.93	2.57	0.675	0.07 [-0.24, 0.38]
Lower vermilion height	9.21	1.99	8.95	1.76	0.391	0.14 [-0.17, 0.45]
Upper paramedian vermilion height L	9.13	1.69	9.07	1.34	0.791	0.04 [-0.27, 0.35]
Upper paramedian vermilion height R	9.17	1.84	9.06	1.28	0.659	0.07 [-0.24, 0.38]
Horizontal distance						
Nasal alar base width	30.39	2.74	27.86	2.72	< 0.001*	0.93 [0.60, 1.25]
Philtrum width	12.97	2.24	11.76	2.13	< 0.001*	0.55 [0.24, 0.87]
Mouth width	51.51	5.11	47.55	4.25	< 0.001*	0.84 [0.52, 1.16]
FSTT						
Ss–Ss'	12.74	1.74	10.37	1.46	< 0.001*	1.47 [1.12, 1.82]
Pr–Pr'	14.20	1.72	11.55	1.64	< 0.001*	1.58 [1.22, 1.93]
InS–InS'	5.79	1.82	5.12	1.85	0.022*	0.37 [0.05, 0.68]
Id–Id'	14.98	1.63	13.19	1.60	< 0.001*	1.11 [0.78, 1.44]
Sm–Sm'	13.32	1.79	12.30	1.84	< 0.001*	0.56 [0.25, 0.88]
Pg–Pg'	13.26	2.01	12.34	2.40	0.009*	0.42 [0.10, 0.70]

[1]SD: Standard deviation.

[2]CI: Confidence interval.

*: Statistically significant $p < 0.05$ (Welch's t test).

Furthermore, the detailed results of the comparison among the five age groups for soft tissue measurement items are shown in Online Resource 7. Initial comparison across five age groups indicated that morphological changes were most pronounced during the transition from the 40s to the 50s. This period demonstrated the highest increase in significantly altered items and moderate-to-large effect sizes (Online Resource 7 and 8). Consequently, the samples were reclassified into two groups using 50 years as the cutoff to reflect this biological threshold and ensure adequate statistical power through balanced sample sizes.

The < 50 years and ≥ 50 years groups were then compared (Table 7). Significant differences were observed between the two groups in all soft tissue measurement items except for Lower facial height to ChinA, Philtrum width, Pr–Pr', and Sm–Sm'. However, among the four items that did not show statistically significant differences, Lower facial height to ChinA (d = -0.25, |d| =

0.25) and Pr–Pr' (d = 0.25, |d| = 0.25) demonstrated slightly larger effect sizes compared to the others, although all effect sizes remained small. Significantly smaller values were observed in the ≥ 50 years group for items indicating vermilion height and FSTT excluding Pr–Pr' and Sm–Sm'. In contrast, significantly greater values were observed in the ≥ 50 years group for items indicating lip height (Philtrum height, Total upper lip height, Total lip height), Nasal alar base width, and Mouth width. The results were generally consistent with those of Welch's t-test. In the multiple regression analysis, these two age group classifications were converted to dummy variables and included as independent variables (Age group).

For items that deviated significantly from normality, the Kruskal–Wallis test with Bonferroni-adjusted post hoc comparisons and the Mann–Whitney U test was additionally performed for five-group and two-group comparisons (Online Resources 9 and 10). The results were largely consistent with those of Welch's ANOVA and Welch's t-test.

Table 7
Comparison of soft tissue measurements between < 50 years and ≥ 50 years.

Measurement (mm)	< 50 years (n = 91)		≥ 50 years (n = 70)		Age difference	
	Mean	SD[1]	Mean	SD[1]	*p*	Cohen's d [95% CI[2]]
Vertical distance						
Philtrum height	14.88	2.35	16.93	2.57	< 0.001*	-0.84 [-1.16, -0.51]
Total upper lip height	22.32	2.61	23.25	2.36	0.019*	-0.37 [-0.69, -0.06]
Total lip height	38.54	4.24	40.11	4.48	0.026*	-0.36 [-0.68, -0.05]
Lower facial height to ChinA	50.04	5.18	51.26	4.33	0.105	-0.25 [-0.57, 0.06]
Upper midline vermilion height	7.43	1.47	6.32	1.48	< 0.001*	0.75 [0.43, 1.08]
Total vermilion height	17.06	2.73	14.70	2.89	< 0.001*	0.84 [0.52, 1.17]
Lower vermilion height	9.63	1.67	8.38	1.91	< 0.001*	0.70 [0.38, 1.02]
Upper paramedian vermilion height L	9.48	1.49	8.62	1.45	< 0.001*	0.59 [0.27, 0.90]
Upper paramedian vermilion height R	9.52	1.52	8.59	1.52	< 0.001*	0.61 [0.30, 0.93]
Horizontal distance						
Nasal alar base width	28.25	2.82	30.32	2.84	< 0.001*	-0.73 [-1.05, -0.41]
Philtrum width	12.24	2.36	12.55	2.13	0.382	-0.14 [-0.45, 0.18]
Mouth width	47.92	3.81	51.70	5.75	< 0.001*	-0.80 [-1.12, -0.47]
FSTT						
Ss–Ss'	11.97	1.90	11.06	2.01	0.004*	0.47 [0.15, 0.78]
Pr–Pr'	13.13	2.27	12.60	1.94	0.112	0.25 [-0.06, 0.56]
InS–InS'	5.80	1.79	5.02	1.88	0.009*	0.43 [0.11, 0.74]
Id–Id'	14.53	1.52	13.54	2.08	0.001*	0.55 [0.24, 0.87]
Sm–Sm'	12.93	1.78	12.69	2.01	0.421	0.13 [-0.18, 0.44]
Pg–Pg'	13.20	2.21	12.31	2.22	0.013*	0.40 [0.08, 0.71]

[1]SD: Standard deviation.

[2]CI: Confidence interval.

*: Statistically significant $p < 0.05$ (Welch's t test).

Pearson correlation analysis

Pearson correlation analysis of the dependent and independent variables (Online Resource 11) revealed robust associations within vertical measurement categories. Particularly strong correlations were noted between Total upper lip height and Ak–InS (r = 0.71) and between the lower facial height to ChinA and Ak–Me (r = 0.73). Among Northwestern analysis items, InS to NP mm exhibited the strongest associations with the soft tissue items, while horizontal distance items exhibited no significant linear relationship with skeletal Northwestern variables.

According to the results of the Pearson correlation analysis among the independent variables (Online Resource 12), 13 combinations demonstrated high correlations ($|r| \geq 0.7$), indicating potential redundancy. Significant collinearity was observed between Ak–Pr and Ak–InS (r = 0.92) and between N–A–P and ANB (r = 0.95). These findings were instrumental in the selective inclusion of predictors to prevent multicollinearity within the regression models.

Furthermore, Spearman correlation analysis mirrored these trends, confirming the reliability of the observed data distributions.

Multiple regression equation development and prediction accuracy

Based on the results of sex differences, age group differences, and correlation analysis, variables showing significant associations and anatomical relevance were selected as candidate predictors for the multiple regression models. Orthodontic and Non-orthodontic models were developed by multiple regression analysis using the training group and validated using the test group. The assumption of normality and homoscedasticity of the residuals were met. The developed multiple regression equations and the results of accuracy validation are shown in Tables 8 and 9, respectively. For Mouth width and Id–Id', orthodontic analysis items were not retained as independent variables because their statistical contribution was not sufficiently supported; therefore, Orthodontic models were not constructed. In both models, Total upper lip height showed the highest coefficient of determination (Ortho: $adjR^2 = 0.635$, Non-ortho: $adjR^2 = 0.572$). In contrast, the lowest coefficient of determination was observed for Philtrum width ($adjR^2 = 0.157$) in the Orthodontic model and InS–InS' ($adjR^2 = 0.086$) in the Non-orthodontic model. Furthermore, in the accuracy validation, the correlation coefficient between the measured and predicted values was highest for Total lip height in both models (Ortho: Test r = 0.850, Non-Ortho: Test r = 0.798) and lowest for Upper paramedian vermilion height L (Ortho: Test r = 0.288, Non-ortho: Test r = 0.092). The lowest Test MAE was observed for Ss–Ss' (Ortho: 1.131 mm, Non-ortho: 1.155 mm). The highest Test MAE was observed for Lower facial height to ChinA (2.447 mm) in the Orthodontic model and Mouth width (3.126 mm) in the Non-orthodontic model.

Table 8
Multiple regression equations for the Orthodontic and Non-orthodontic models.

Measurement	Training group (n = 161)	
	Model type	Regression formula
Vertical distance		
Philtrum height	Ortho	5.141 + 0.334 * (Ak–InS) + 0.097 * (N–A–P) -1.497 * (Sex group) + 1.960 * (Age group)
	Non-ortho	3.474 + 0.410 * (Ak–InS) – 1.262 * (Sex group) + 2.022 * (Age group)
Total upper lip height	Ortho	6.917 + 0.487 * (Ak–InS) + 0.182 * (InS to NP mm) -1.129 * (Sex group) + 0.814*(Age group)
	Non-ortho	6.226 + 0.567 * (Ak–InS) – 1.074 * (Sex group) + 0.871 * (Age group)
Total lip height	Ortho	2.212 + 0.513 * (Ak–Me) + 0.200 * (InS to NP mm) + 1.185 * (Age group)
	Non-ortho	1.554 + 0.550 * (Ak–Me) + 1.244 * (Age group)
Lower facial height to ChinA	Ortho	10.932 + 0.618 * (Ak–Me) – 0.368 * (ANB) – 1.297 * (Sex group)
	Non-ortho	11.791 + 0.587 * (Ak–Me) – 1.757 * (Sex group)
Upper midline vermilion height	Ortho	3.254 + 0.115 * (Ak–InS) + 0.088 * (InS to NP mm) – 1.167 * (Age group)
	Non-ortho	2.946 + 0.153 * (Ak–InS) – 1.139 * (Age group)
Total vermilion height	Ortho	3.011 + 0.176 * (Ak–Me) + 0.240 * (InS to NP mm) – 2.562 * (Age group)
	Non-ortho	2.222 + 0.220 * (Ak–Me) – 2.492 * (Age group)
Lower vermilion height	Ortho	5.583 + 0.153 * (Ak–Pr) + 0.148 * (InS to NP mm) – 1.283 * (Age group)
	Non-ortho	5.777 + 0.220 * (Ak–Pr) – 1.211 * (Age group)
Upper paramedian vermilion height L	Ortho	4.525 + 0.140 * (Ak–InS) + 0.094 * (InS to NP mm) – 0.926 * (Age group)
	Non-ortho	4.196 + 0.180 * (Ak–InS) – 0.895 * (Age group)
Upper paramedian vermilion height R	Ortho	3.787 + 0.164 * (Ak–InS) + 0.102 * (InS to NP mm) – 1.004 * (Age group)
	Non-ortho	3.430 + 0.208 * (Ak–InS) – 0.972 * (Age group)
Horizontal distance		
Nasal alar base width	Ortho	18.480 + 0.393 * (CN(L–R)) + 0.129 * (InS to NP mm) – 2.301 * (Sex group) + 1.829 * (Age group)
	Non-ortho	13.339 + 0.135*(Ss–Sm) + 0.424*(CN(L–R))-1.988*(Sex group) + 1.709*(Age group)
Philtrum width	Ortho	-2.872 + 0.500 * (U1d(L–R)) + 0.096 * (SNB) – 0.91 * (Sex group)
	Non-ortho	3.835 + 0.557 * (U1d(L–R)) – 0.967 * (Sex group)
Mouth width	Ortho	-

Measurement		Training group (n = 161)
	Model type	Regression formula
Vertical distance		
	Non-ortho	25.619 + 0.628 * (CejU3(L–R)) − 3.266 * (Sex group) + 3.856*(Age group)
FSTT		
Ss–Ss'	Ortho	17.084–0.117 * (Ak–InS) − 0.165 * (ANB) − 2.243 * (Sex group) − 0.722 * (Age group)
	Non-ortho	18.113–0.171 * (Ak–InS) − 2.424 * (Sex group) − 0.779 * (Age group)
Pr–Pr'	Ortho	17.205 + 0.109 * (Pr–Id) − 0.148 * (Id–Me) − 0.145 * (ANB) − 2.855 * (Sex group)
	Non-ortho	16.649 + 0.130 * (Pr–Id) − 0.160 * (Id–Me) – 3.010 * (Sex group)
InS–InS'	Ortho	12.163 − 0.036 * (U1 to NS) − 0.252 * (InS to NP mm) − 0.785 * (Sex group) -0.717 * (Age group)
	Non-ortho	8.702 − 0.132 * (Pr–Id) − 0.727 * (Sex group) − 0.666 * (Age group)
Id–Id'	Ortho	-
	Non-ortho	11.663 + 0.343 * (CejU1–InS) − 1.626 * (Sex group) − 0.707 * (Age group)
Sm–Sm'	Ortho	5.529 + 0.159 * (Ss–Sm) + 0.128 * (InS to NP mm) − 0.707 * (Sex group)
	Non-ortho	5.667 + 0.185 * (Ss–Sm) − 0.646 * (Sex group)
Pg–Pg'	Ortho	5.584 + 0.152 * (Ss–Sm) + 0.167 * (InS to NP mm) − 1.105 * (Age group)
	Non-ortho	5.916 + 0.182 * (Ss–Sm) − 1.067 * (Age group)

Sex group (male = 0, female = 1), Age group (< 50 years = 0, ≥ 50 years = 1).

Table 9
Accuracy validation for Orthodontic and Non-orthodontic models.

Measurement		Training group (n = 161)				Test group (n = 40)	
	Model type	adjR² [1]	SEE[2]	*P*	Max VIF[3]	Test r[4]	Test MAE[5]
Vertical distance							
Philtrum height	Ortho	0.481	1.903	< 0.001	1.245	0.745	1.653
	Non-ortho	0.445	1.968	< 0.001	1.010	0.731	1.688
Total upper lip height	Ortho	0.635	1.533	< 0.001	1.154	0.738	1.339
	Non-ortho	0.572	1.661	< 0.001	1.010	0.665	1.577
Total lip height	Ortho	0.520	3.051	< 0.001	1.081	0.850	1.654
	Non-ortho	0.494	3.129	< 0.001	1.003	0.798	1.844
Lower facial height to ChinA	Ortho	0.579	3.148	< 0.001	1.207	0.695	2.447
	Non-ortho	0.552	3.248	< 0.001	1.130	0.717	2.320
Upper midline vermilion height	Ortho	0.240	1.369	< 0.001	1.145	0.337	1.153
	Non-ortho	0.204	1.400	< 0.001	1.001	0.159	1.234
Total vermilion height	Ortho	0.382	2.383	< 0.001	1.801	0.519	2.225
	Non-ortho	0.301	2.534	< 0.001	1.003	0.302	2.554
Lower vermilion height	Ortho	0.288	1.585	< 0.001	1.132	0.531	1.446
	Non-ortho	0.212	1.667	< 0.001	1.001	0.297	1.694
Upper paramedian vermilion height L	Ortho	0.245	1.327	< 0.001	1.145	0.288	1.252
	Non-ortho	0.202	1.364	< 0.001	1.001	0.092	1.358
Upper paramedian vermilion height R	Ortho	0.288	1.336	< 0.001	1.145	0.391	1.236
	Non-ortho	0.240	1.381	< 0.001	1.001	0.172	1.337
Horizontal distance							
Nasal alar base width	Ortho	0.378	2.368	< 0.001	1.088	0.566	2.014
	Non-ortho	0.380	2.365	< 0.001	1.147	0.562	2.014
Philtrum width	Ortho	0.157	2.077	< 0.001	1.071	0.458	1.289
	Non-ortho	0.138	2.101	< 0.001	1.037	0.459	1.289
Mouth width	Ortho	-	-	-	-	-	-
	Non-ortho	0.386	3.947	< 0.001	1.191	0.673	3.126
FSTT							
Ss–Ss'	Ortho	0.480	1.436	< 0.001	1.291	0.662	1.131
	Non-ortho	0.454	1.472	< 0.001	1.010	0.643	1.155
Pr–Pr'	Ortho	0.462	1.572	< 0.001	1.337	0.645	1.331
	Non-ortho	0.442	1.601	< 0.001	1.280	0.637	1.316
InS–InS'	Ortho	0.388	1.456	< 0.001	1.130	0.604	1.189

Measurement		Training group (n = 161)				Test group (n = 40)	
	Model type	adjR² [1]	SEE[2]	*P*	Max VIF[3]	Test r[4]	Test MAE[5]
Vertical distance							
	Non-ortho	0.086	1.780	< 0.001	1.034	0.133	1.617
Id–Id'	Ortho	-	-	-	-	-	-
	Non-ortho	0.313	1.529	< 0.001	1.164	0.509	1.226
Sm–Sm'	Ortho	0.264	1.613	< 0.001	1.115	0.658	1.122
	Non-ortho	0.204	1.678	< 0.001	1.071	0.539	1.309
Pg–Pg'	Ortho	0.199	2.017	< 0.001	1.052	0.552	1.501
	Non-ortho	0.127	2.106	< 0.001	1.016	0.473	1.783

[1]adjR[2]: Adjusted coefficient of determination.

[2]SEE: Standard error of the estimate (mm).

[3]Max VIF: Maximum variance inflation factor.

[4]Test r: Correlation coefficient between measured and predicted values in the test group.

[5]Test MAE: Mean absolute error in the test group (mm).

Generalization performance

The results for the generalization performance analysis are shown in Table 10. A significant difference in absolute error was observed between the training and test groups for Total lip height (Orthodontic model); however, the effect size was small (r = -0.161, |r| = 0.161). No significant differences were observed for the other items, and the effect sizes were small, with a maximum of |r| = 0.127.

Table 10
Generalization performance based on absolute errors between training and test groups.

Measurement		Training group	Test group	Difference in absolute error		
	Model type	Median (IQR[1])	Median (IQR[1])	Z	*p*	r^2
Vertical distance						
Philtrum height	Ortho	1.25 [0.64–2.23]	1.25 [0.74–2.59]	-0.40	0.686	-0.028
	Non-ortho	1.33 [0.73–2.37]	1.32 [0.46–2.64]	-0.11	0.915	-0.007
Total upper lip height	Ortho	0.96 [0.46–1.80]	1.07 [0.41–1.89]	-0.30	0.768	-0.021
	Non-ortho	1.00 [0.53–1.95]	1.32 [0.81–2.13]	-1.53	0.127	-0.108
Total lip height	Ortho	1.99 [0.86–3.29]	1.34 [0.72–2.56]	-2.28	0.023*	-0.161
	Non-ortho	2.23 [0.89–3.28]	1.51 [0.68–2.72]	-1.81	0.071	-0.127
Lower facial height to ChinA	Ortho	2.00 [1.20–3.50]	1.71 [0.72–2.64]	-1.21	0.227	-0.085
	Non-ortho	1.99 [0.97–3.65]	1.26 [0.62–2.77]	-1.55	0.121	-0.109
Upper midline vermilion height	Ortho	0.81 [0.42–1.48]	0.91 [0.24–2.11]	-0.06	0.954	-0.004
	Non-ortho	0.94 [0.50–1.56]	0.90 [0.27–2.23]	-0.21	0.832	-0.015
Total vermilion height	Ortho	1.62 [0.76–2.76]	1.58 [0.74–3.29]	-0.44	0.657	-0.031
	Non-ortho	1.72 [0.89–3.23]	2.19 [0.78–3.36]	-0.64	0.521	-0.045
Lower vermilion height	Ortho	1.10 [0.52–1.96]	1.14 [0.54–2.26]	-0.26	0.792	-0.019
	Non-ortho	1.36 [0.70–1.94]	1.26 [0.65–2.61]	-0.74	0.457	-0.052
Upper paramedian vermilion height L	Ortho	0.86 [0.42–1.61]	1.04 [0.47–1.95]	-0.94	0.350	-0.066
	Non-ortho	0.89 [0.49–1.57]	1.11 [0.35–1.90]	-0.91	0.362	-0.064
Upper paramedian vermilion height R	Ortho	0.83 [0.36–1.36]	0.82 [0.35–1.92]	-0.74	0.457	-0.052
	Non-ortho	0.85 [0.41–1.53]	0.97 [0.34–2.06]	-0.88	0.380	-0.062
Horizontal distance						
Nasal alar base width	Ortho	1.64 [0.72–2.74]	1.74 [0.84–2.97]	-0.95	0.343	-0.067

Measurement		Training group	Test group	Difference in absolute error		
	Model type	Median (IQR[1])	Median (IQR[1])	Z	*p*	r^2
Vertical distance						
	Non-ortho	1.56 [0.72–2.71]	1.86 [0.76–2.88]	-0.88	0.382	-0.062
Philtrum width	Ortho	1.28 [0.51–2.34]	0.75 [0.56–1.78]	-1.29	0.198	-0.091
	Non-ortho	1.34 [0.54–2.20]	1.05 [0.47–1.62]	-1.42	0.154	-0.100
Mouth width	Ortho	2.68 [0.92–4.23]	2.41 [1.35–4.06]	-	-	-
	Non-ortho	2.68 [0.92–4.23]	2.41 [1.35–4.06]	-0.07	0.942	-0.005
FSTT						
Ss–Ss'	Ortho	0.78 [0.39–1.55]	1.09 [0.57–1.56]	-0.69	0.489	-0.049
	Non-ortho	0.90 [0.41–1.69]	1.02 [0.44–1.62]	-0.42	0.677	-0.029
Pr–Pr'	Ortho	1.07 [0.55–1.67]	1.04 [0.60–1.80]	-0.22	0.825	-0.016
	Non-ortho	1.09 [0.54–1.78]	1.13 [0.44–2.05]	-0.02	0.988	-0.001
InS–InS'	Ortho	0.94 [0.41–1.65]	0.95 [0.37–1.69]	-0.02	0.988	-0.001
	Non-ortho	1.22 [0.57–2.06]	1.35 [0.77–2.34]	-1.21	0.227	-0.085
Id-Id'	Ortho	0.96 [0.46–1.59]	1.09 [0.36–1.91]	-	-	-
	Non-ortho	0.96 [0.46–1.59]	1.09 [0.36–1.91]	-0.39	0.695	-0.028
Sm–Sm'	Ortho	1.01 [0.52–1.68]	0.88 [0.46–1.51]	-0.74	0.459	-0.052
	Non-ortho	1.03 [0.45–1.75]	1.15 [0.47–1.69]	-0.07	0.944	-0.005
Pg–Pg'	Ortho	1.32 [0.66–2.29]	1.40 [0.48–2.14]	-0.22	0.825	-0.016
	Non-ortho	1.40 [0.67–2.40]	1.44 [0.78–2.43]	-0.39	0.697	-0.027

[1]IQR: Interquartile range (mm).

[2]r: Effect size; $r = Z / \sqrt{n}$, where n is the total sample size.

*: Statistically significant $p < 0.05$ (Mann–Whitney U test).

Comparison of prediction accuracy

To compare predictive accuracy, Mouth width and Id–Id', for which Orthodontic models were not constructed, were excluded. For the remaining items, $\Delta adjR^2$, ΔTest r, ΔTest MAE, the median absolute error between the observed and predicted values for each model, and the differences in these medians between the models were evaluated (Table 11).

Table 11
Comparison of prediction accuracy between Orthodontic and Non-orthodontic models.

Measurement				Orthodontic model		Non-orthodontic model		Difference in absolute error		
	ΔadjR² [1]	ΔTest r[1]	ΔTest MAE[1]	Median[2]	IQR[3]	Median[2]	IQR[3]	Z	*p*	r[4]
Vertical distance										
Philtrum height	0.036	0.014	-0.035	1.25	0.7–2.59	1.32	0.81–2.13	-0.19	0.851	-0.03
Total upper lip height	0.063	0.073	-0.238	1.07	0.41–1.89	1.32	0.81–2.13	-2.22	0.027*	-0.35
Total lip height	0.026	0.052	-0.190	1.34	0.72–2.56	1.51	0.68–2.72	-1.32	0.188	-0.21
Lower facial height to ChinA	0.027	-0.022	0.127	1.71	0.72–2.64	1.26	0.62–2.77	-0.67	0.493	-0.11
Upper midline vermilion height	0.036	0.178	-0.081	0.91	0.74–2.11	0.90	0.27–2.23	-1.61	0.107	-0.26
Total vermilion height	0.081	0.217	-0.329	1.58	0.74–3.29	2.19	0.78–3.36	-2.11	0.035*	-0.33
Lower vermilion height	0.076	0.234	-0.248	1.14	0.54–2.26	1.26	0.65–2.61	-2.51	0.012*	-0.40
Upper paramedian vermilion height L	0.043	0.196	-0.106	1.04	0.47–1.95	1.11	0.35–1.90	-2.03	0.042*	-0.32
Upper paramedian vermilion height R	0.048	0.219	-0.101	0.82	0.35–1.92	0.97	0.34–2.06	-1.69	0.090	-0.27
Horizontal distance										
Nasal alar base width	-0.002	0.004	0.000	1.74	0.84–2.97	1.86	0.76–2.88	-0.42	0.677	-0.07
Philtrum width	0.019	-0.001	0.000	0.75	0.56–1.78	1.05	0.47–1.62	-0.07	0.946	-0.01
FSTT										
Ss–Ss'	0.026	0.019	-0.024	1.09	0.57–1.56	1.02	0.44–1.62	-0.58	0.563	-0.09
Pr–Pr'	0.020	0.008	0.015	1.04	0.60–1.80	1.13	0.44–2.05	-0.15	0.882	-0.02
InS–InS'	0.302	0.471	-0.428	0.95	0.37–1.69	1.28	0.77–2.34	-2.60	0.009*	-0.41

[1]Δ values = (Orthodontic model value) - (Non-orthodontic model value) (mm)

Measurement				Orthodontic model		Non-orthodontic model		Difference in absolute error		
	ΔadjR²[1]	ΔTest r[1]	ΔTest MAE[1]	Median[2]	IQR[3]	Median[2]	IQR[3]	Z	*p*	r[4]
Vertical distance										
Sm–Sm'	0.060	0.119	-0.187	0.88	0.46–1.51	1.15	0.47–1.69	-1.94	0.053	-0.31
Pg–Pg'	0.072	0.079	-0.282	1.40	0.48–2.14	1.44	0.78–2.43	-1.91	0.056	-0.30

[1]Δ values = (Orthodontic model value) - (Non-orthodontic model value) (mm)

[2]Median: Median Absolute Error in the test group (mm).

[3]IQR: Interquartile range (mm).

[4]r: Effect size; r = Z / √n.

*: Statistically significant $p < 0.05$ (Wilcoxon signed-rank test).

ΔadjR^2 was positive for all items except Nasal alar base width (ΔadjR^2 = -0.002). Similarly, ΔTest r was positive for all items except Lower facial height to ChinA (ΔTest r = -0.022) and Philtrum width (ΔTest r = -0.001). Both ΔadjR^2 and ΔTest r were highest for InS–InS' (ΔadjR^2 = 0.302, ΔTest r = 0.471). For ΔTest MAE, positive values were observed for Lower facial height to ChinA (ΔTest MAE = 0.127) and Pr–Pr' (ΔTest MAE = 0.015), whereas no difference was observed for Nasal alar base width or Philtrum width. Among the items with negative values, InS–InS' showed the lowest value (ΔTest MAE = -0.428).

Furthermore, comparison of the absolute error between the two models showed significant differences for Total upper lip height, Total vermilion height, Lower vermilion height, Upper paramedian vermilion height L, InS–InS'. For these items, the effect size |r| ranged from 0.32 to 0.41 and the absolute errors were significantly higher in the Non-orthodontic model. Although no significant differences were observed for other items, medium effect sizes were found for Sm–Sm' (r = -0.31, |r| = 0.31) and Pg–Pg' (r = -0.30, |r| = 0.30).

Discussion

Sex and age differences

In this study, no sex differences were observed in vermilion height. This finding is generally consistent with previous studies, including Ogawa et al. [30], which described features of Japanese facial morphology, Zhao et al. [5] and Yang et al. [31], which focused on Chinese populations, and Sforza et al. [32], which focused on Italians. Thus, the influence of sex may be limited when constructing regression models in vermilion height for East Asian and Italian populations. In contrast, Houlton et al., who focused on African populations, reported significant sex differences in Total lip height and Upper midline vermilion height (corresponding to Upper lip height in that study) [33]. Therefore, depending on the ancestry group, it may be necessary to consider sex differences.

Furthermore, this study showed significant sex differences in all FSTT measurement items. However, previous studies [5, 23, 24, 34, 35] on FSTT have reported that no sex differences were observed in some regions, such as the lower lip region except in Kurkcuoglu's study [36]. One possible explanation is that FSTT in the lower face may show considerable individual variability.

In the comparison among the five age groups, differences were observed between the younger and middle-aged/older groups in items such as philtrum height, vermilion height, and mouth width. Age-related changes in lip morphology have been reported in

several previous studies. Yasumori et al. reported significant differences between Japanese women in their 20s to 30s and those aged ≥ 60 years [37]. Similarly, Chong et al. [38] found significant differences between those in their 20s to 40s and those in their 50s, which is generally consistent with the present findings. Kim et al. [39] reported findings similar to those of Chong et al., although the possibility that the sample included individuals who had undergone cosmetic surgery cannot be excluded. However, Kim et al. reported no significant difference in the item corresponding to Upper paramedian vermilion height [39]. Thus, age-related changes in lip morphology may vary according to the measurement site and method. From an aesthetic perspective, histological changes such as decreases in hyaluronic acid and collagen have been reported in the lips with aging [40]. These changes may reduce skin elasticity and volume. The cumulative effects of these age-related changes may have appeared as differences in lip morphology among age groups.

Precise age estimation in 10-year intervals using only the skull is considered challenging. Therefore, samples were reclassified into two age groups in this study. Larger effect sizes were observed between those aged ≥ 60 years for FSTT items, and participants in their 40s and 50s for other items. Although reclassification using 60 years as the cutoff was also considered, the ≥ 60 years group included only 30 samples, which could result in a bias in the sample size between groups. Accordingly, samples were reclassified using 50 years as the cutoff.

In the comparison between the two age groups, < 50 years and ≥ 50 years, significant differences were observed in Philtrum height, Total upper lip height, Nasal alar base width, and the items indicating vermilion height. A previous study by Yang et al. reported a similar trend when comparing younger and middle-aged/older groups, which was generally consistent with the present results. However, they have also reported no significant difference in the item corresponding to Upper midline vermilion height [31]. Furthermore, although significant differences were observed in multiple FSTT items in this study, Tanaka et al. found significant differences only in Ss–Ss' and Pr–Pr' among FSTT items in the lip region, suggesting that the results may differ across studies [34]. With some exceptions, it was considered necessary to consider the influence of age when developing regression models for items related to the height and width of the lip area. Regarding the influence of age on FSTT items, further investigation using a larger sample size may be necessary.

In contrast, no significant differences were observed in Philtrum width in either the two-group or five-group comparison. This finding is similar to those of previous studies by Houlton et al. [4] and Yang et al. [31]. Therefore, the influence of age may be relatively small when constructing a regression model for Philtrum width.

Pearson correlation analysis

Correlation analysis between soft tissue and hard tissue measurement items showed a tendency toward stronger correlations among vertical and horizontal distance items. This tendency was also suggested by the correlation analysis reported by Zhao et al. [5]. Regarding the Northwestern analysis variables, their contribution to regression models for items in which no clear correlations were observed is likely limited. In contrast, the other soft tissue measurement items showed moderate correlations with multiple Northwestern analysis items, particularly a stronger correlation with InS to NP mm. However, for Philtrum height, Ss–Ss', and Pr–Pr', stronger correlations were observed with N–A–P, ANB, and U1 to NS than with InS to NP mm.

Furthermore, correlation analysis among hard tissue measurement items revealed strong correlations between multiple variables. Strong correlations between independent variables can lead to multicollinearity, but they may also be advantageous [41]. For example, in this study, InS to NP mm showed correlations with many soft tissue measurement items, but this variable cannot be used for specimens with missing anterior teeth. It may be possible to predict lip morphology by developing new regression models using N–A–P, which showed strong correlations, instead of InS to NP mm.

Multiple regression equation development and predictive accuracy

For Mouth width and Id–Id', Orthodontic models that met the statistical criteria were not developed. The correlation analysis also showed no clear correlations between these items and the Northwestern analysis items, and this finding is considered consistent with that pattern.

In this study, Total upper lip height showed the highest coefficient of determination in both the Orthodontic and Non-orthodontic models. A previous study by Zhao et al. showed a similar result, but the Test MAE was lower in that study. For many other

measurement items, the Test MAE was also lower in Zhao et al.'s study; the only measurement item with significantly lower in this study was Upper paramedian vermilion height R. Zhao et al.'s study included a limited group of Chinese individuals aged 18–30 years with a body mass index (BMI) of 18.5–24.9 [5]. In contrast, this study included Japanese individuals aged 18–81 years, and BMI data were not recorded, making direct comparison impossible. Therefore, differences in population structure may have contributed to the observed differences in accuracy. In addition, in a previous study by Houlton et al. focusing on African populations, the Test MAE was lower in the present study for Mouth width but higher for Philtrum width [33]. This difference may likewise reflect differences in population structure, including ancestry-related variation. Based on these findings, although the regression models developed in this study showed acceptable predictive accuracy, differences from previous studies may reflect variation in population characteristics.

Generalization performance

When evaluating developed models, it is important to assess not only goodness of fit to the training data but also generalization performance for unknown data [42]. When comparing absolute errors between the training and test groups, a significant difference was observed only for Total lip height, suggesting the possibility of overfitting. However, the effect size was small; thus, we cannot definitively conclude that predictive performance for unknown specimens was reduced.

Comparison of prediction accuracy

For ΔTest MAE, positive values were observed for Lower facial height to ChinA and Pr–Pr', suggesting that the Non-orthodontic model may have higher predictive accuracy than the Orthodontic model. In addition, both Nasal alar base width and Philtrum width demonstrated no difference in accuracy between the Orthodontic and Non-orthodontic models. All other measurement items showed negative values, suggesting that the Orthodontic model may have higher predictive accuracy.

In addition, analysis of the absolute error in both models showed that the Orthodontic model was significantly more accurate for the five items. In these orthodontic models, InS to NP mm was involved, suggesting the strong influence of maxillary incisor protrusion. Therefore, the use of orthodontic analysis items, particularly InS to NP mm, may improve regression model accuracy, particularly for items showing significant differences.

Limitations

Facial reconstruction typically reproduces facial features with the lips closed. For this reason, this study used samples with the lips closed. In cases of excessive maxillary, mandibular, or bimaxillary protrusion, as well as excessive incisor proclination, lip incompetence is likely to occur. However, because this was a retrospective study, it was not possible to instruct participants to close their lips during MDCT imaging. Therefore, the samples in this study may have excluded cases with such excessive protrusion and/or incisor inclination due to the exclusion criteria.

Additionally, facial soft tissues are susceptible to gravity, suggesting that they may be displaced depending on body position [43–45]. Munn and Stephan reported that, in the supine position, soft tissue retracts around the nasolabial folds compared with the upright position, and this change is particularly pronounced in older individuals with reduced connective tissue adherence. Postural correction factors have also been proposed to convert supine soft tissue thickness data to upright estimates. However, these factors were not used in this study because the reliability of the improvement achieved by applying these factors to the landmarks used here has not been demonstrated [43]. Meanwhile, Nishioka et al. evaluated differences in linear facial soft tissue measurements between upright and supine positions in individuals in their 20s and those aged 60 years or older. They reported that the supine position showed significantly greater values, particularly for lower vermilion height and nose, philtrum, and lip width, with differences of 1–2 mm in both males and females in their 20s and 2–6 mm in those aged ≥ 60 years [45]. This study used supine MDCT images from individuals aged 18–81 years, which may introduce discrepancies relative to facial reconstruction performed in the upright position. Therefore, future studies should clarify postural correction factors for each measured item in each age group [11, 34].

Conclusion

This study developed regression models to predict lip morphology from hard tissue measurements and examined their predictive accuracy and generalization performance. The results showed consistent predictive accuracy for many measurement items, suggesting that these models may maintain stable accuracy even for unknown specimens. Furthermore, models including orthodontic analysis items showed higher predictive accuracy in some areas. Taken together, these regression models may be useful for estimating lip morphology from actual skulls, and considering the relationship with orthodontic analysis in specific areas may help improve facial reconstruction accuracy.

Declarations

Acknowledgments

We sincerely thank the Department of Dental Radiology, Department of Oral Surgery, and Department of Dental Implant Clinic at the Institute of Science Tokyo Hospital for granting access to data and facilities and providing the resources and support necessary for this study. We also thank Editage (www.editage.com) for English language editing.

Author Contributions

Conceptualization: Nanami Aoki, Hajime Utsuno; Methodology: Nanami Aoki, Hajime Utsuno, Shuuji Namiki, Muhammad Garry Syahrizal Hanafi; Formal analysis and investigation: Nanami Aoki, Muhammad Garry Syahrizal Hanafi; Writing–original draft preparation: Nanami Aoki, Hajime Utsuno, Koichi Sakurada, Muhammad Garry Syahrizal Hanafi; Writing–review and editing: Nanami Aoki, Hajime Utsuno, Koichi Sakurada, Muhammad Garry Syahrizal Hanafi, Shuuji Namiki; Funding acquisition: Nanami Aoki, Hajime Utsuno, Koichi Sakurada; Resources: Nanami Aoki, Hajime Utsuno, Koichi Sakurada, Shuuji Namiki, Muhammad Garry Syahrizal Hanafi, Hitomi Nojima, Hiroyuki Harada, Hideaki Kamochi, Eriko Marukawa, Hisako Saitoh, Saki Minegishi, Sayaka Yamada; Supervision: Hajime Utsuno, Koichi Sakurada

Funding

This research was supported by the Japan Science and Technology Agency (JST) Next-Generation Researchers' Challenging Research Program (JST SPRING, Grant Numbers: JPMJSP2106 and JPMJSP2180) and by a Grant-in-Aid for Scientific Research (C) (Grant Numbers: 21K06362 and 25K09797) from the Japan Society for the Promotion of Science.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

The study was conducted in accordance with the Declaration of Helsinki. This study protocol was approved by the Ethics Review Committee of the Department of Dentistry, Institute of Science Tokyo (Approval No. D2018-058-08). Because this was a retrospective study, the requirement for informed consent was waived.

Consent to participate

The requirement for informed consent was waived because this was a retrospective study. An opt-out method was implemented by posting information about this study on websites to provide participants with an opportunity to decline participation.

Consent to publish

Not applicable.

Clinical Trial Number

Not applicable. This study was a retrospective study and not a prospective clinical trial.

Data availability

The data generated and analyzed during this study are not publicly available because of privacy and ethical restrictions but are available from the corresponding author upon reasonable request. The supplementary data will be made available online upon publication.

Code Availability

Not applicable.

References

1. Wilkinson C (2010) Facial reconstruction: anatomical art or artistic anatomy? J Anat 216:235–250. https://doi.org/10.1111/j.1469-7580.2009.01182.x
2. Stephan CN (2013) Facial approximation. In: Siegel JA, Saukko PJ, Houck MM (eds) Encyclopedia of forensic sciences, 2nd edn. Elsevier, Amsterdam, pp 60–67
3. Miyasaka S (2019) The role of craniofacial reconstruction technology in crime investigation 1. Its historical context and scientific basis. Rep Natl Res Inst Police Sci 68:47–63
4. Houlton TMR, Jooste N, Steyn M (2020) Mouth width and Cupid's bow estimation in a southern African population. J Forensic Sci 65:372–379. https://doi.org/10.1111/1556-4029.14207
5. Zhao J, Ji L, Han M, Mou Q, Chu G, Chen T, Du S, Hou Y, Guo Y (2021) Lip morphology estimation models based on three-dimensional images in a modern adult population from China. Int J Legal Med 135:1887–1901. https://doi.org/10.1007/s00414-021-02559-2
6. Gerasimov MM (1971) The face finder. Hutchinson & Co, London
7. Qadeer TA, Jawaid M, Fahim MF, Habib M, Khan EB (2022) Effect of lip thickness and competency on soft-tissue changes. Am J Orthod Dentofac Orthop 162:483–490. https://doi.org/10.1016/j.ajodo.2021.05.010
8. Rains MD, Nanda R (1982) Soft-tissue changes associated with maxillary incisor retraction. Am J Orthod 81:481–488. https://doi.org/10.1016/0002-9416(82)90427-4
9. Yogosawa F (1990) Predicting soft tissue profile changes concurrent with orthodontic treatment. Angle Orthod 60:199–206. https://doi.org/10.1043/0003-3219(1990)060%3C0199:PSTPCC%3E2.0.CO;2
10. Lu W, Zhang X, Mei L, Wang P, He J, Li Y, Zhao Z (2020) Orthodontic incisor retraction caused changes in the soft tissue chin area: a retrospective study. BMC Oral Health 20:108. https://doi.org/10.1186/s12903-020-01099-2
11. Hanafi MGS, Utsuno H, Namiki S, Aoki N, Saitoh H, Minegishi S, Yamada S, Makino Y, Iwase H, Sakurada K (2025) Assessing the contribution of orthodontic profiles in predicting facial soft tissue thickness for forensic facial approximation. Int J Legal Med 139:2951–2964. https://doi.org/10.1007/s00414-025-03542-x
12. Miyashita K (1986) An atlas of roentgen anatomy and cephalometric analysis. Quintessence Publishing Co., Ltd., Tokyo
13. Downs WB (1948) Variations in facial relationships: their significance in treatment and prognosis. Am J Orthod 34:812–840. https://doi.org/10.1016/0002-9416(48)90015-3
14. Steiner CC (1953) Cephalometrics for you and me. Am J Orthod 39:729–755. https://doi.org/10.1016/0002-9416(53)90082-7
15. Tweed CH (1946) The Frankfort-mandibular plane angle in orthodontic diagnosis, classification, treatment planning, and prognosis. Am J Orthod Oral Surg 32:175–230. https://doi.org/10.1016/0096-6347(46)90001-4
16. Tweed CH (1953) Evolutionary trends in orthodontics, past, present, and future. Am J Orthod 39:81–108. https://doi.org/10.1016/0002-9416(53)90014-1
17. Graber TM (1952) New horizons in case analysis: clinical cephalometrics. Am J Orthod 38:603–624. https://doi.org/10.1016/0002-9416(52)90027-4
18. Graber TM (1954) A critical review of clinical cephalometric radiography. Am J Orthod 40:1–26. https://doi.org/10.1016/0002-9416(54)90166-9

19. Caple J, Stephan CN (2016) A standardized nomenclature for craniofacial and facial anthropometry. Int J Legal Med 130:863–879. https://doi.org/10.1007/s00414-015-1292-1
20. Editorial Committee of the Anthropology Series (ed) (1991) Anthropometry I: somatometry. Yuzankaku Publishing Co., Ltd., Tokyo
21. Editorial Committee of the Anthropology Series (ed) (1991) Anthropometry II: osteometry. Yuzankaku Publishing Co., Ltd., Tokyo
22. Sawyer AR, See M, Nduka C (2009) 3D stereophotogrammetry quantitative lip analysis. Aesthetic Plast Surg 33:497–504. https://doi.org/10.1007/s00266-008-9191-1
23. Sarilita E, Rynn C, Mossey PA, Black S, Oscandar F (2020) Facial average soft tissue depth variation based on skeletal classes in Indonesian adult population: a retrospective lateral cephalometric study. Legal Med 43:101665. https://doi.org/10.1016/j.legalmed.2019.101665
24. Utsuno H, Kageyama T, Uchida K, Kibayashi K (2014) Facial soft tissue thickness differences among three skeletal classes in Japanese population. Forensic Sci Int 236:175–180. https://doi.org/10.1016/j.forsciint.2013.12.040
25. Park E, Chang J, Park J (2023) Facial soft tissue thickness differences among three skeletal classes in Korean population using CBCT. Int J Environ Res Public Health 20:2658. https://doi.org/10.3390/ijerph20032658
26. Whyms BJ, Vorperian HK, Gentry LR, Schimek EM, Bersu ET, Chung MK (2013) The effect of computed tomographic scanner parameters and 3-dimensional volume rendering techniques on the accuracy of linear, angular, and volumetric measurements of the mandible. Oral Surg Oral Med Oral Pathol Oral Radiol 115:682–691. https://doi.org/10.1016/j.oooo.2013.02.008
27. Imai I, Shiomi T (2004) Reliability of evaluation in physical therapy studies. Rigakuryoho Kagaku 19:261–265. https://doi.org/10.1589/rika.19.261
28. Cohen J (1977) Statistical power analysis for the behavioral sciences. Academic, New York
29. Vatcheva KP, Lee M, McCormick JB, Rahbar MH (2016) Multicollinearity in regression analyses conducted in epidemiologic studies. Epidemiology 6:227. https://doi.org/10.4172/2161-1165.1000227
30. Ogawa Y, Wada B, Taniguchi K, Miyasaka S, Imaizumi K (2015) Photo anthropometric variations in Japanese facial features: establishment of large-sample standard reference data for personal identification using a three-dimensional capture system. Forensic Sci Int 257:511e1–511e9. https://doi.org/10.1016/j.forsciint.2015.07.046
31. Yang Y, Zhang M, Jin L, Long X, Zeng A, Yu N, Wang X (2023) Gender- and age-related characterization of lip morphology: a three-dimensional analysis in a Chinese population. Aesthet Surg J 43:NP990–NP1000. https://doi.org/10.1093/asj/sjad270
32. Sforza C, Grandi G, Binelli M, Dolci C, De Menezes M, Ferrario VF (2010) Age- and sex-related changes in three-dimensional lip morphology. Forensic Sci Int 200:182e1–18182 e7. https://doi.org/10.1016/j.forsciint.2010.04.050
33. Houlton TMR, Jooste N, Steyn M, Hemingway J (2022) Visualising trends in dentition to lip mouth morphology using geometric morphometrics. PLoS ONE 17:e0274127. https://doi.org/10.1371/journal.pone.0274127
34. Tanaka C, Utsuno H, Makino Y, Minegishi S, Ota J, Iwase H, Sakurada K (2020) Facial soft tissue thickness of the Japanese population determined using post mortem computed tomography images. Forensic Imaging 23:200423. https://doi.org/10.1016/j.fri.2020.200423
35. Wang J, Zhao X, Mi C, Raza I (2016) The study on facial soft tissue thickness using Han population in Xinjiang. Forensic Sci Int 266. https://doi.org/10.1016/j.forsciint.2016.04.032. :585.e1–585.e5
36. Kurkcuoglu A, Pelin C, Ozener B, Zagypan R, Sahinoglu Z, Yazici AC (2011) Facial soft tissue thickness in individuals with different occlusion pattern in adult Turkish subjects. HOMO 62:288–297. https://doi.org/10.1016/j.jchb.2011.06.001
37. Yasumori H, Tamura E, Tsukahara K, Inoue Y, Yamamoto T (2019) Age-related changes in lip morphological characteristics in Japanese women. J Soc Cosmet Chem Jpn 53:287–296. https://doi.org/10.5107/sccj.53.287
38. Chong Y, Dong R, Liu X, Wang X, Yu N, Long X (2021) Stereophotogrammetry to reveal age-related changes of labial morphology among Chinese women aging from 20 to 60. Skin Res Technol 27:41–48. https://doi.org/10.1111/srt.12906

39. Kim H, Lee M, Park SY, Kim YM, Han J, Kim E (2019) Age-related changes in lip morphological and physiological characteristics in Korean women. Skin Res Technol 25:277–282. https://doi.org/10.1111/srt.12644
40. Gomi T, Imamura T (2020) Comprehensive histological investigation of age-related changes in dermal extracellular matrix and muscle fibers in the upper lip vermilion. Int J Cosmet Sci 42:359–368. https://doi.org/10.1111/ics.12622
41. Namiki S, Utsuno H, Makino Y, Minegishi S, Toya M, Iwase H, Sakurada K (2023) Estimation of missing mandible from the skull using postmortem CT images. Legal Med 65:102321. https://doi.org/10.1016/j.legalmed.2023.102321
42. Mohri M, Rostamizadeh A, Talwalkar A (2018) Foundations of machine learning, 2nd edn. MIT Press, Cambridge
43. Munn L, Stephan CN (2018) Changes in face topography from supine-to-upright position and soft tissue correction values for craniofacial identification. Forensic Sci Int 289:40–50. https://doi.org/10.1016/j.forsciint.2018.05.016
44. Ozsoy U, Sekerci R, Ogut E (2015) Effect of sitting, standing, and supine body positions on facial soft tissue: detailed 3D analysis. Int J Oral Maxillofac Surg 44:1309–1316. https://doi.org/10.1016/j.ijom.2015.06.005
45. Nishioka H, Hishikawa M, Nagai F, Yuzuriha S (2022) Morphological changes in faces depending on posture. Facial Plast Surg 38:193–200. https://doi.org/10.1055/s-0041-1729910

Figures

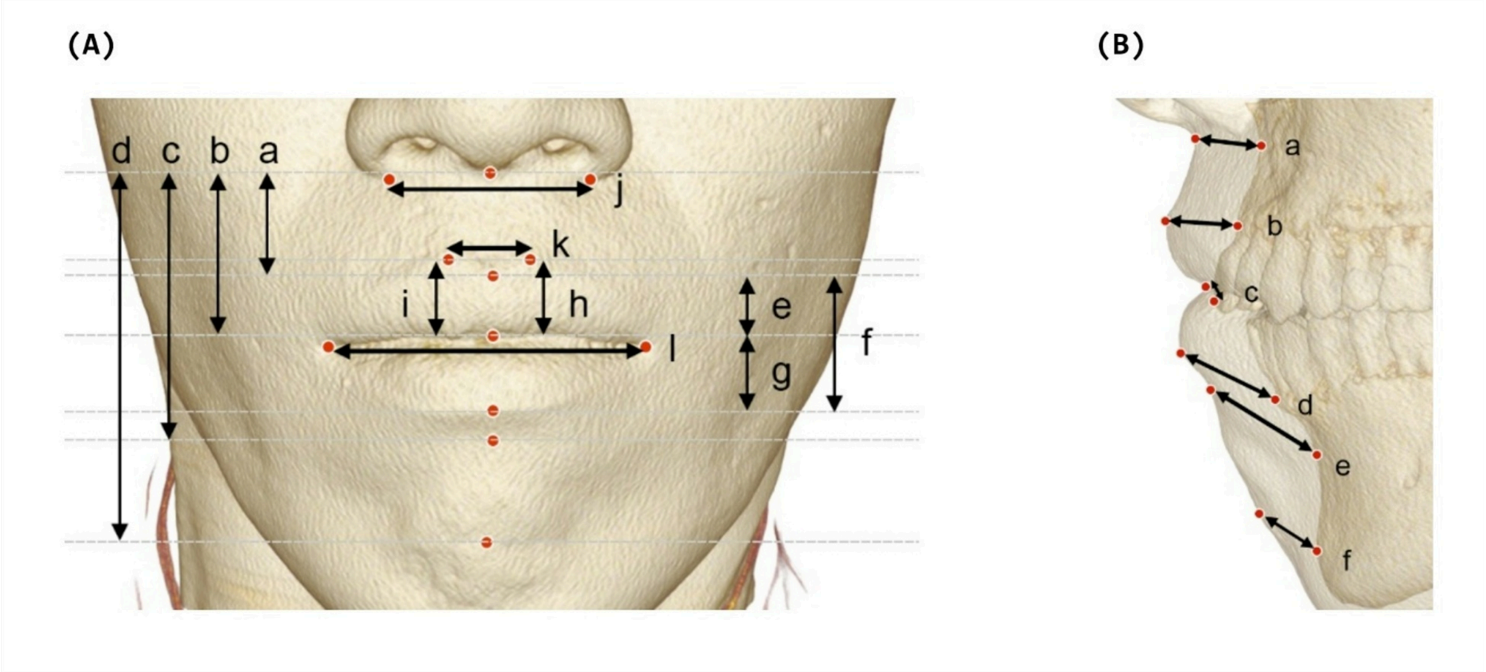


Figure 1

Soft tissue measurements

(A) Vertical and horizontal measurements (a-l)

(B) Facial soft tissue thickness (FSTT) (a-f)

The measurement labels correspond to those in Table 3

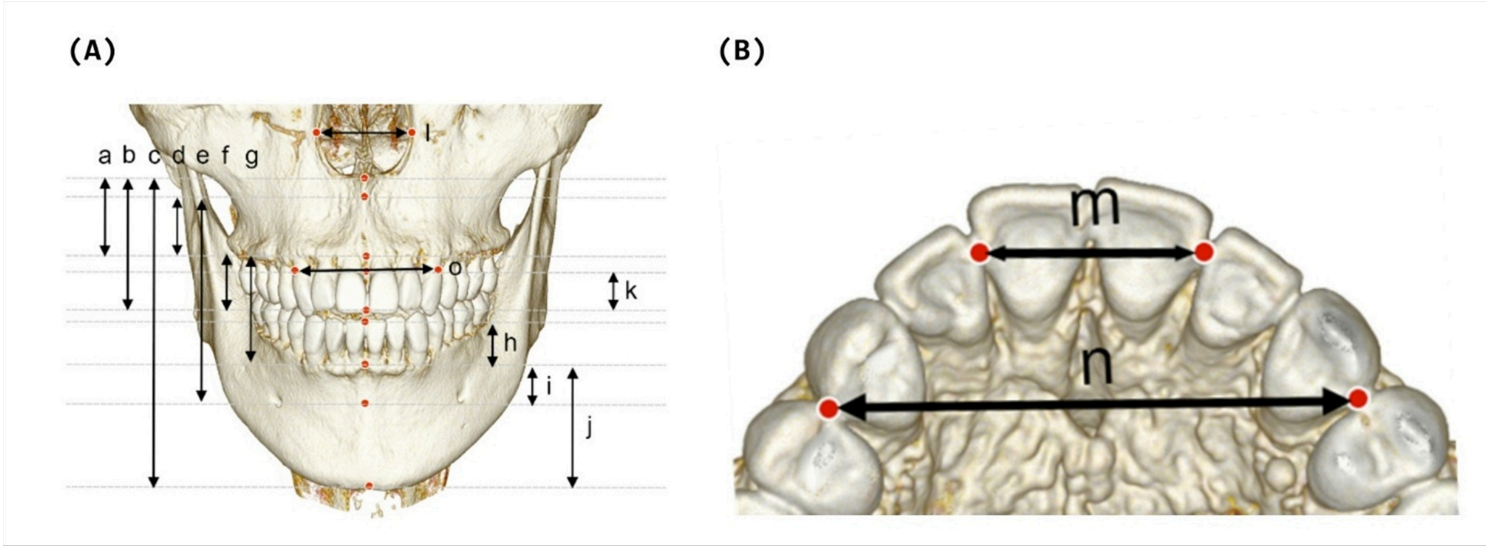


Figure 2

Hard tissue vertical and horizontal measurements

(A) Frontal view (a-l, o)

(B) Occlusal view (m, n)

The measurement labels correspond to those in Table 4

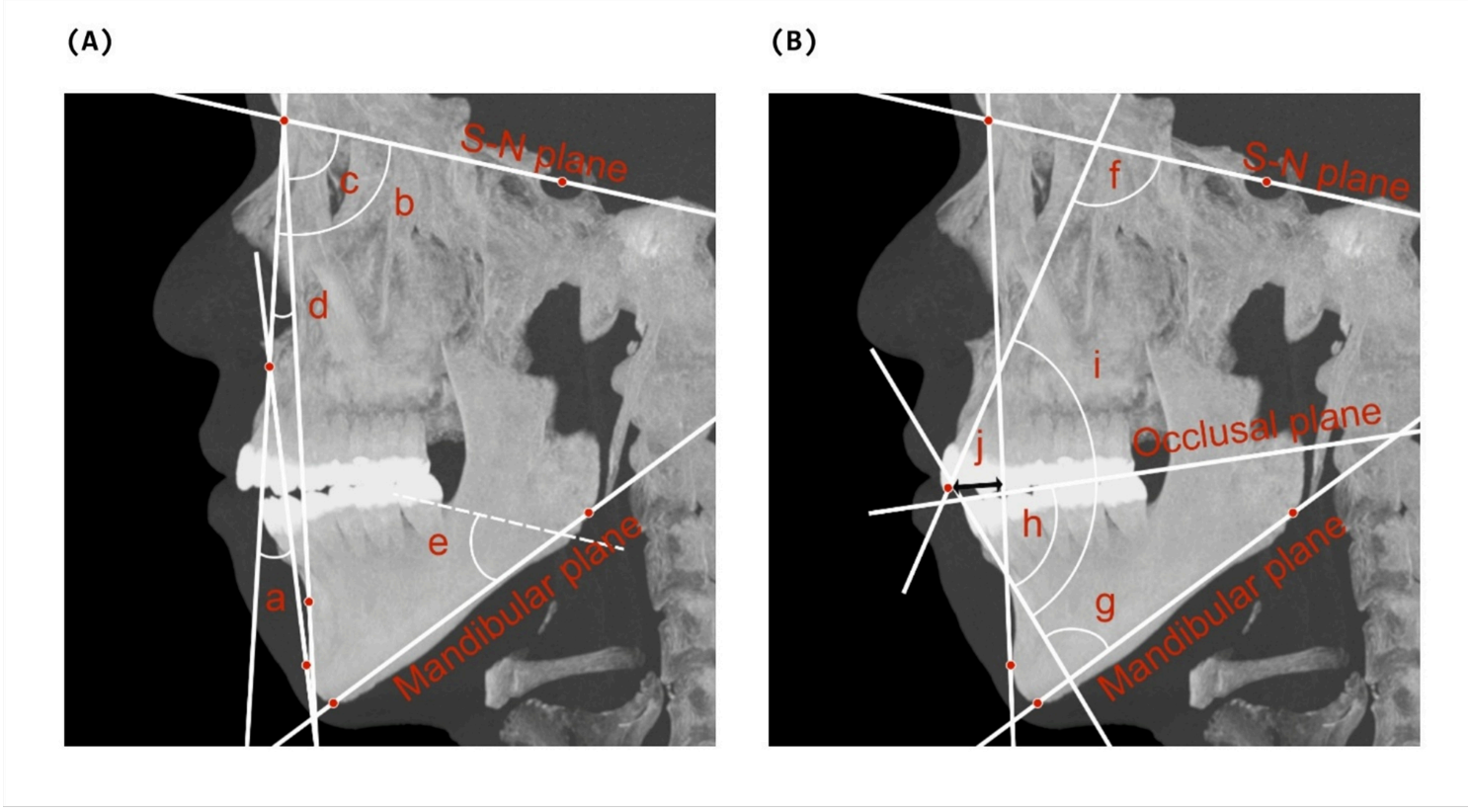


Figure 3

Northwestern analysis measurements

(A) Items indicating jaw position (a–e)

(B) Items indicating the axial relationship of the anterior teeth (f–j)

The measurement labels correspond to those in Table 4

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM.zip