



OPEN

# Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients

Chihiro Tanikawa<sup>1,2,3✉</sup> & Takashi Yamashiro<sup>1</sup>

From a socio-psychological standpoint, improving the morphology of the facial soft-tissues is regarded as an important therapeutic goal in modern orthodontic treatment. Currently, many of the algorithms used in commercially available software programs that are said to provide the function of performing profile prediction are based on the false assumption that the amount of movement of hard-tissue and soft-tissue has a proportional relationship. The specification of the proportionality constant value depends on the operator, and there is little evidence to support the validity of the prediction result. Thus, the present study attempted to develop artificial intelligence (AI) systems that predict the three-dimensional (3-D) facial morphology after orthognathic surgery and orthodontic treatment based on the results of previous treatment. This was a retrospective study in a secondary adult care setting. A total of 137 patients who underwent orthognathic surgery ( $n = 72$ ) and orthodontic treatment with four premolar extraction ( $n = 65$ ) were enrolled. Lateral cephalograms and 3-D facial images were obtained before and after treatment. We have developed two AI systems to predict facial morphology after orthognathic surgery (System S) and orthodontic treatment (System E) using landmark-based geometric morphometric methods together with deep learning methods; where cephalometric changes during treatment and the coordinate values of the faces before treatment were employed as predictive variables. Eleven-fold cross-validation showed that the average system errors were 0.94 mm and 0.69 mm for systems S and E, respectively. The total success rates, when success was defined by a system error of  $< 1$  mm, were 54% and 98% for systems S and E, respectively. The total success rates when success was defined by a system error of  $< 2$  mm were both 100%. AI systems to predict facial morphology after treatment were therefore confirmed to be clinically acceptable.

The face plays an important role as a means of nonverbal communication in the transmission of emotions and thoughts in social life. Further, the degree of recognition of an individual's face by others has a significant socio-psychological impact on the acceptance of the individual in their community. If there is dysmorphology of the maxillofacial region, such as cleft palate, or morphological distortion of the face due to jaw deformity, it may cause serious socio-psychological maladaptation. From this socio-psychological standpoint, improving the morphology of facial soft-tissues is regarded as an important therapeutic goal in modern orthodontic treatment<sup>1</sup>. When planning treatment for patients with malocclusion, the prediction of profile and/or facial changes that will occur after treatment is of the utmost importance when judging whether a treatment method is appropriate (i.e., tooth extraction or non-extraction, treatment with or without orthognathic surgery).

<sup>1</sup>Department of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Osaka University, Suita, Osaka, Japan. <sup>2</sup>Center for Advanced Medical Engineering and Informatics, Osaka University, Suita, Osaka, Japan. <sup>3</sup>Institute for Datability Science, Osaka University, Suita, Osaka, Japan. ✉email: ctanika@dent.osaka-u.ac.jp

Thus far, many studies have tried to predict facial profiles after orthodontic treatment and have examined the facial changes according to orthodontic treatment (i.e., four premolar extraction<sup>2–11</sup>, non-extraction treatment<sup>7,8,10–12</sup>, orthognathic surgery<sup>13–15</sup>, or the use of a functional appliance<sup>16–18</sup>). As for the profile changes that occur following four-premolar extraction, most studies have evaluated the relationship between incisor movement and soft-tissue profiles on cephalograms<sup>2–6</sup>. Previous studies examining factors associated with two-dimensional (2-D) profile changes after orthodontic treatment have revealed a proportional relationship between the amount of movement of hard and soft-tissues. However, the value of the proportionality constant in the lip region has been shown to vary greatly depending on the case<sup>19–21</sup>. Currently, many of the algorithms used in commercially available software programs that are said to provide the function of performing 2-D profile prediction are based on the false assumption that the amount of movement of hard-tissue and soft-tissue has a proportional relationship. The specification of the proportionality constant value depends on the operator, and there is little evidence regarding the validity of the prediction result. Furthermore, regarding the pattern matching technology used, the specific algorithm has not been clarified, and it is unclear whether a valid prediction has been made<sup>22</sup>. According to studies<sup>22–26</sup> that compared results predicted with such software programs to actual treatment results, it is reported that it is still difficult to predict the facial morphology before and after orthodontic treatment.

Recently, three-dimensional (3-D) stereophotogrammetry of faces has emerged in dentistry and orthopedics as a method of recording 3-D facial topography<sup>27–31</sup>. Using this technique, 3-D images are acquired by combining photographs captured from various angles with synchronous digital cameras. The values recorded by 3-D systems are reported to be accurate and reliable for clinical use<sup>27,28</sup>. Since 3-D data include a greater amount of information than two-dimensional data, it was assumed that 3-D images are useful for creating a system that predicts the post-treatment facial shape. Although commercial 3-D simulation software programs exist (e.g., Dolphin system), they are only for prediction after orthognathic surgery and internal calculations have not been clarified; thus it is unclear whether their calculations are based on evidence. Studies that measured the ability to accurately predict 3-D soft-tissue changes after Le Fort I osteotomy and/or sagittal split osteotomy using the Dolphin 3-D software program<sup>32,33</sup> showed limited accuracy, especially in the upper lip and base of the nose. There is a great need for patients to be able to check the post-treatment face in 3-D before treatment; thus, a 3-D facial simulation system that can predict 3-D faces after orthognathic surgery or orthodontic treatment with four premolar extraction would be a clinically important tool.

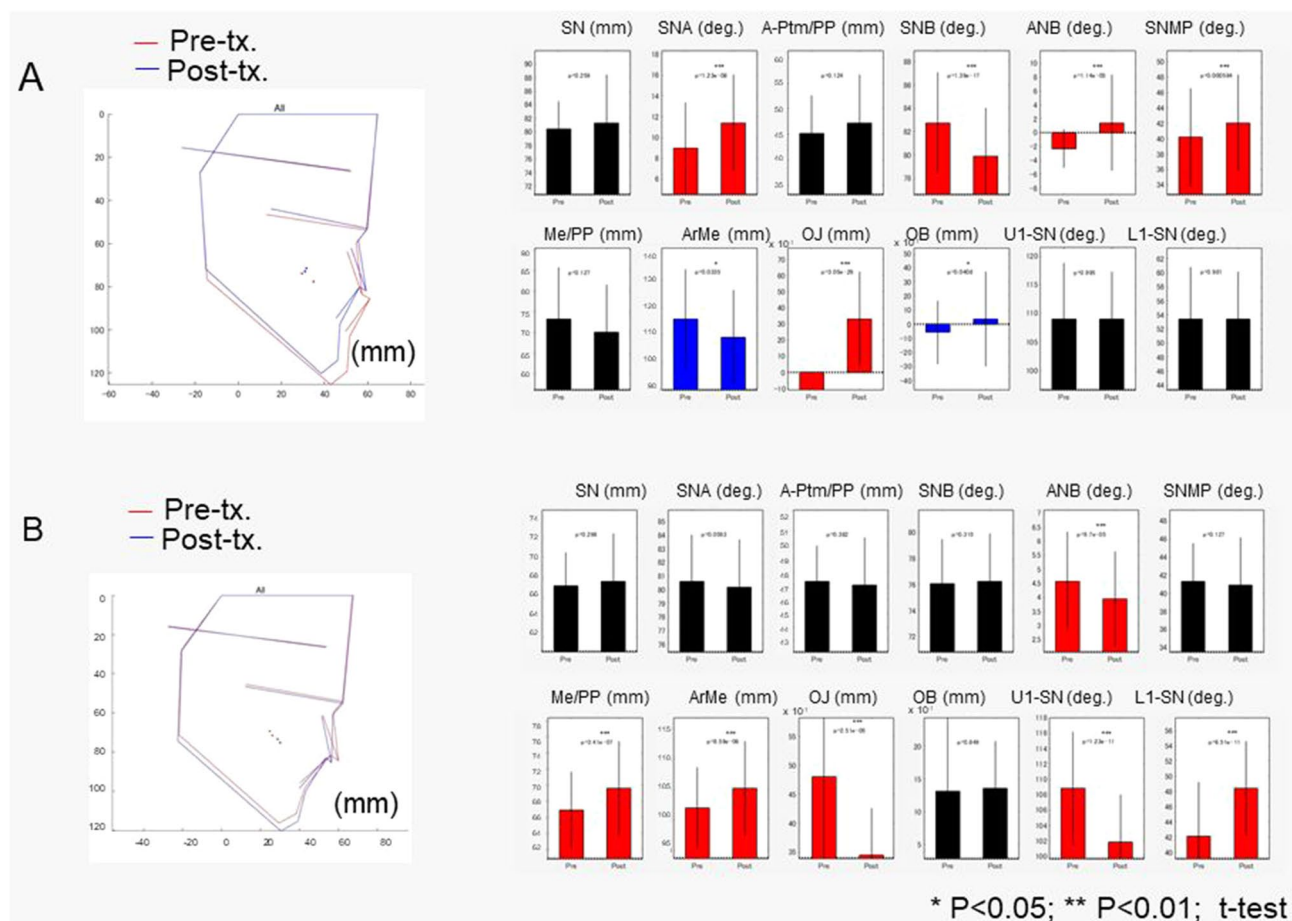
On the other hand, with recent computation advances, two important methods that are related to 3-D prediction have become available in machine learning, that is, landmark-based geometric morphometric methods (GMMs) and deep learning. GMMs that have revealed some statistical variation in the shape and size of target objects (phenotypic variation) have recently emerged. In developmental biology, GMMs use homologous landmarks, which can be defined as precise locations on biological specimens that hold some functional, structural, developmental, or evolutionary significance, and which are directly comparable between specimens<sup>34</sup>. A key concept of a GMM is based on the fact that the morphology can be systematically mapped, often within a “morphospace”, with the use of these landmarks. Morphospaces are maps that show how shapes are defined by quantitative traits. GMMs rely on the superimposition of landmark coordinate data to place individuals in a common morphospace. As deep learning should use the same number of inputs, we hypothesized that this GMM would be useful, together with deep learning methods, for developing artificial intelligence (AI) systems that predict the 3-D facial topography after orthognathic surgery and fixed edgewise orthodontic treatment.

Thus, the present study aimed to develop AI systems that predict the 3-D facial topography after orthognathic surgery and fixed edgewise orthodontic treatment based on the combination of GMM and deep learning. The usefulness of the combination of GMM and deep learning was also discussed when predicting the 3-D shape changes in relation to treatment in the present study.

## Results

**Hard-tissue and soft-tissue changes after treatment.** The changes of the hard and soft-tissue after treatment are shown in Figs. 1, 2 and 3. In the Surgery group, the maxillary positions were repositioned anteriorly by the average maxillary length of 2 mm, and the mandibular positions were repositioned posteriorly by two-jaw surgeries by the average mandibular effective length of 5 mm, leading to a significant increase in overjet with average changes of – 1 mm to + 3 mm. As the facial soft-tissue changes, the nasal alar and mid-facial cheek moved laterally by an average of 2 mm and the corner of the lip was moved mesially by an average of 1 mm (x-axis). As for the vertical changes (y-axis), the nose and chin were repositioned into superior positions by average distances of 2 mm and 3 mm, respectively, whereas the cheek was moved downward by the average distances of 1–2 mm. With regard to the antero-posterior changes (z-axis), the midface was moved forward by an average distance of 2 to 3 mm and the chin was moved backward by an average distance of 4 to 5 mm. Since the backward movement of the mandible was 5 mm, the soft-tissue movement corresponded well to the hard-tissue in the antero-posterior direction on the midline; however, the lateral and vertical changes seem to be difficult to predict from the assumption that the amount of movement of hard and soft-tissue has a proportional relationship.

In the Extraction group, the maxillary incisors were inclined lingually by an average of 6.5 degrees and the tip of the maxillary incisors was repositioned lingually by an average of 5 mm. The mandibular incisors were inclined lingually by an average of 6.2 degrees, and repositioned lingually by an average of 2.5 mm. As a result, the overjet was significantly decreased from + 4.8 mm to an average of + 3.3 mm. The patients showed significant growth in the mandibular effective length (5 mm) and lower facial height (3 mm), leading to a significant 0.5 degrees decrease in the A-N-B angle. As the facial soft-tissue changes, the facial width was increased laterally by an average of 2 mm (x-axis). As for the vertical changes (y-axis), the forehead was moved upward by 2 mm, the chin moved downward by 3 mm in average, which were corresponded to the facial growth. Regarding the antero-posterior changes (z-axis), the cheek, nose, and chin moved forward by 2 to 3 mm. The upper and lower



**Figure 1.** The average cephalometric changes in the Surgery (A) and Extraction (B) groups. The left column indicated the average cephalometric profilogram for pre-treatment (red) and post-treatment (blue). The right column shows the results of the comparison of cephalometric variables between pre-treatment and post-treatment. \* Blue bars indicate significant differences ( $p < 0.05$ ); \*\*\* red bars indicated significant differences ( $p < 0.001$ ). SN indicates the distance between Sella and Nasion; SNA, Sella-Nasion-Point A angle; A-Ptm/PP, Point A to Ptm distance projected on the palatal plane; SNB, Sella-Nasion-Point B angle; ANB, Point A-Nasion-Point B angle; SNMP, angle formed by SN and the mandibular plane; Me/PP, anterior lower face height; Ar-Me, distance between Articulare and Menton; OJ, overjet; OB, overbite; U1-SN, angle formed by upper incisors and the SN plane; L1-SN, angle formed by lower incisors and the SN plane.

lips were moved backward by average distances of 2 mm and 3 mm, respectively. As the upper and lower incisors positions were posteriorly moved by 5 mm and 2.5 mm, this upper and lower lip movement showed that the assumption that the amount of movement of hard-tissue and soft-tissue has a proportional relationship that is not applicable to post-treatment facial form.

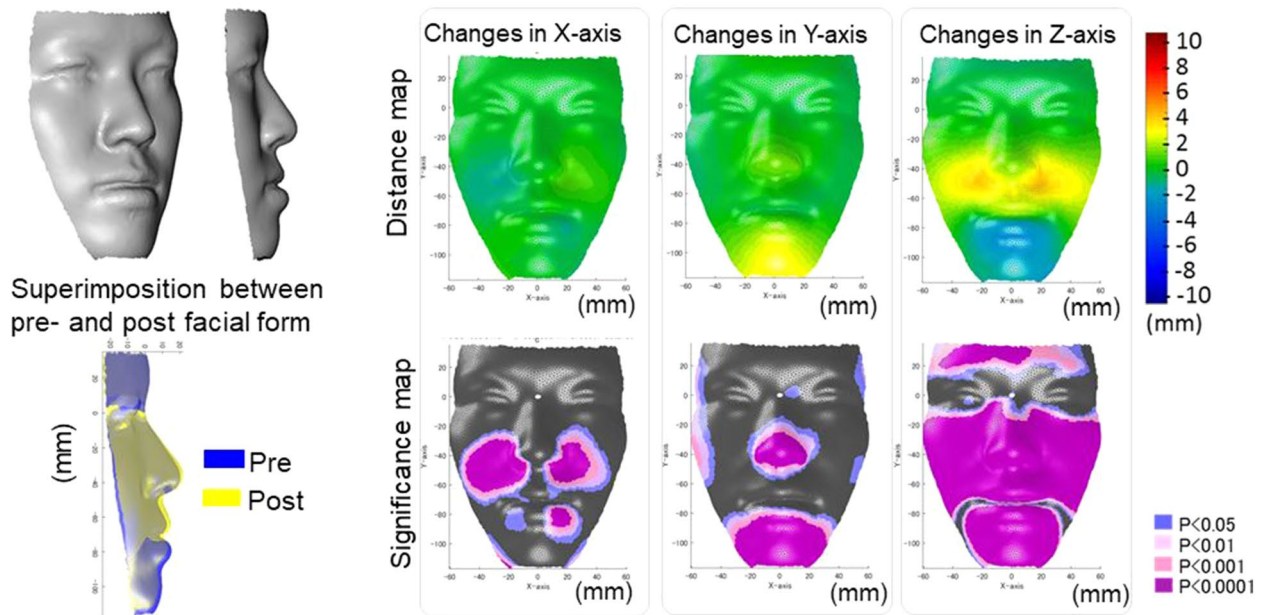
**The development of the AI systems.** Our system to predict facial morphology after treatment was successfully developed with a combination of GMM and deep learning (Fig. 4). The prediction errors were demonstrated in Fig. 5. The system error was  $0.89 \pm 0.30$  mm (S) and  $0.69 \pm 0.18$  mm (E; Table 1). Maximum errors were observed in the nasal alar (S), the chin (S), the corner of the mouth (S), and the lower lip (S and E; Figs. 5 and 6). The total success rate of  $< 1$  mm was 82% (74% [S] and 92% [E]), while the total success rate of  $< 2$  mm was 100% (Fig. 7).

## Discussion

In the present study, we successfully developed two AI systems that predict the 3-D facial shape after orthodontic treatment as well as orthognathic surgery by the combination of deep learning and a GMM, which was a novel method for predicting 3-D shapes. Because this method is versatile, the present study suggests possibilities for developing systems that predicts any 3-D shape (e.g., breast shape) after any intervention or change (e.g., aging, cancer surgery, cosmetic surgery) once data are stored.

The proposed AI system to predict the facial form after orthognathic surgery (System S) showed an error of  $0.89 \pm 0.30$  mm, which was small enough for clinical application. It is difficult to compare the present system directly with several commercial 3-D systems that are available for orthognathic surgery (e.g., the Dolphin system, Mimics, or other systems). However, studies<sup>32,33</sup> demonstrated that the Dolphin 3-D software program

## Averaged facial form pre-tx. (n = 72)



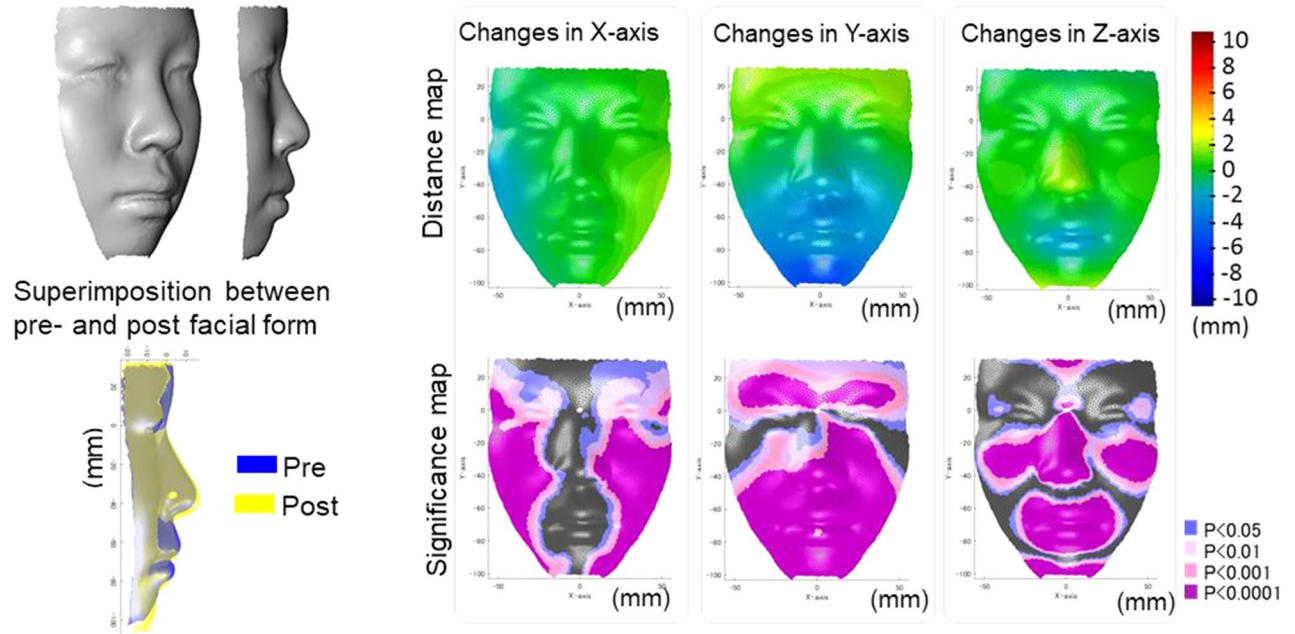
**Figure 2.** The average actual facial changes in the Surgery group for pre-treatment (top-left) and the superimposition of the actual facial changes between pre-treatment (blue) and post-treatment (yellow). For the significance probability maps, significant changes between pre-treatment and post-treatment are indicated as follows: blue,  $P < 0.05$ ; pale pink,  $P < 0.01$ ; dark pink,  $P < 0.001$ ; purple,  $P < 0.0001$ . For the distance maps, yellow indicates that the treatment changed from pre-treatment to post-treatment (post-treatment minus pre-treatment) is a positive value, whereas blue indicates that the change had a negative value. Thus, in the horizontal direction in the Surgery group, yellow in the left-half face indicates lateral displacement and blue in the right-half face also indicates lateral displacement after treatment. In the vertical direction, blue indicates a downward displacement after treatment in the Surgery group and yellow indicates upward displacement after treatment. In the anteroposterior direction, yellow indicates a protrusive movement, whereas blue indicates a retrusive movement after treatment. The figures were created using a custom MATLAB software (MATLAB 2020a, The Math Works, Natick, MA, USA).

had limited accuracy, especially in the upper lip and base of the nose. In a previous study<sup>32</sup>, the mean linear prediction error was  $2.91 \pm 2.16$  mm and the mean error at the nasolabial angle was  $8.1 \pm 5.68$  degrees. Notably, in most cases, the cheek around the nasal alar and the upper lip showed errors of  $> 5$  mm. In contrast, a study using Dolphin 3-D software program showed errors that were, overall, clinically acceptable; however, the largest difference was also observed in the perioral area<sup>35</sup>. As for the facial changes after maxillary advancement osteotomy<sup>36</sup>, the differences in the mean absolute distance between the actual soft-tissue and the prediction were significantly smaller than 3 mm for all segmented anatomical areas, and ranged from 0.65 mm (chin) to 1.17 mm (upper lip). A study<sup>37</sup> examining the changes of the malar-midfacial region after LeFort I osteotomy indicated that maxillary advancement leads to a more pronounced shifting of the soft-tissues in the malar-midfacial area in comparison to the upper lip.

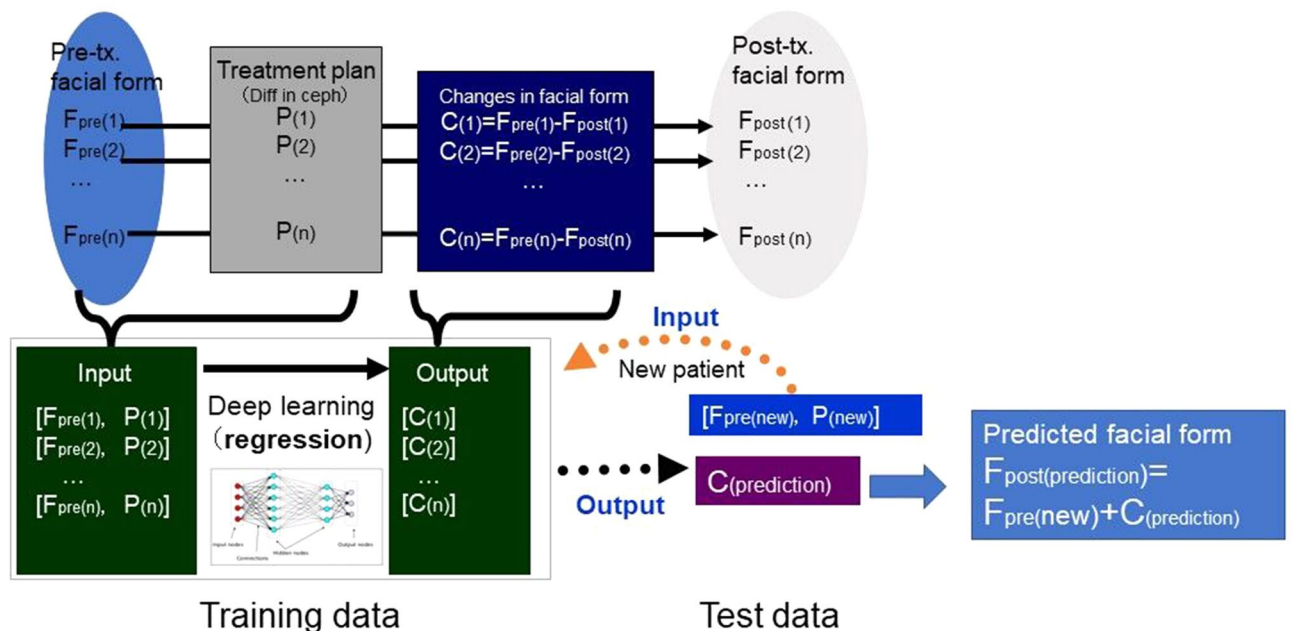
There were also several attempts that studies that have aimed to predict 3-D soft-tissue outcomes based on the soft-tissue density on cone beam computed tomography<sup>38–40</sup>. Further, novel methods have been considered, such as the finite element model<sup>40</sup> or anatomically driven soft-tissue simulation using a muscle template<sup>41</sup>. In an attempt to find a linear relationship between the soft-tissue and hard-tissue changes, it was impossible to find any specific proportional constant values in mandibular landmarks, and a wide range of relationships were observed between the soft-tissue and hard-tissue changes<sup>38–40</sup>. The nonlinear response of maxillary soft-tissues to underlying hard-tissue change in the nasal base and maxillary spline indicated the need for the use of nonlinear ratios for selected maxillary soft-tissue to hard-tissue movements in a simulation software program<sup>39</sup>. In the present study, we created an AI system that predicted the facial form after orthognathic surgery using nonlinear ratios and succeeded in predicting the post-treatment facial form, including malar-midfacial area.

The nasal alar, the chin, and the corner of the mouth showed errors of approximately 1.3 mm, whereas the lower lip showed an error of approximately 0.8 mm on average in system S. When compared to the commercial software program<sup>32,37</sup> which showed errors of  $> 5$  mm in the cheek around the nasal alar, the cheek was an area that commonly showed errors among the systems. When observing the soft-tissue changes in the present study, the nasal alar and mid-facial cheek were expanded laterally, moved downward, and moved forward, which was a new result that the previous studies did not find. These soft-tissue changes were considered to be one of the reasons why the cheek showed greater errors among the systems. However, our system well for this complicated movement and showed smaller errors (1.3 mm) in comparison to the commercial software program ( $> 5$  mm).

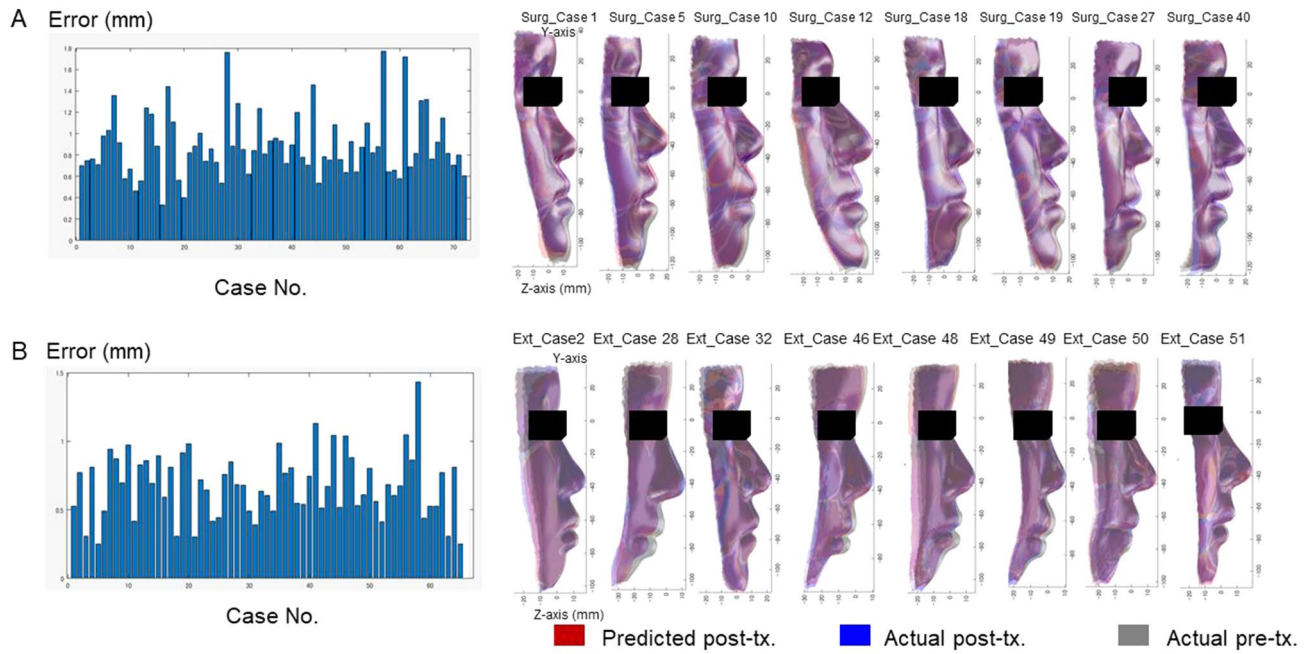
Averaged facial form pre-tx. (n = 65)



**Figure 3.** The average actual facial changes in the Extraction group for pre-treatment (top-left) and the superimposition of the actual facial changes between pre-treatment (blue) and post-treatment (yellow). For the significance probability maps, significant changes between pre-treatment and post-treatment are indicated as follows: blue,  $P < 0.05$ ; pale pink,  $P < 0.01$ ; dark pink,  $P < 0.001$ ; purple,  $P < 0.0001$ . For the distance maps, yellow indicates that the treatment changes from pre-treatment to post-treatment is a positive value, whereas blue indicates that the change is a negative value. Thus, in the horizontal direction in the Extraction group, yellow in the left-half face indicates lateral displacement and blue in the right-half face also indicates lateral displacement after treatment. In the vertical direction, blue indicates downward displacement after treatment in the Extraction group and yellow indicates upward displacement after treatment. In the anteroposterior direction, yellow indicates a protrusive movement, whereas blue indicates a retrusive movement after treatment. The figures were created using a custom MATLAB software (MATLAB 2020a, The Math Works, Natick, MA, USA).



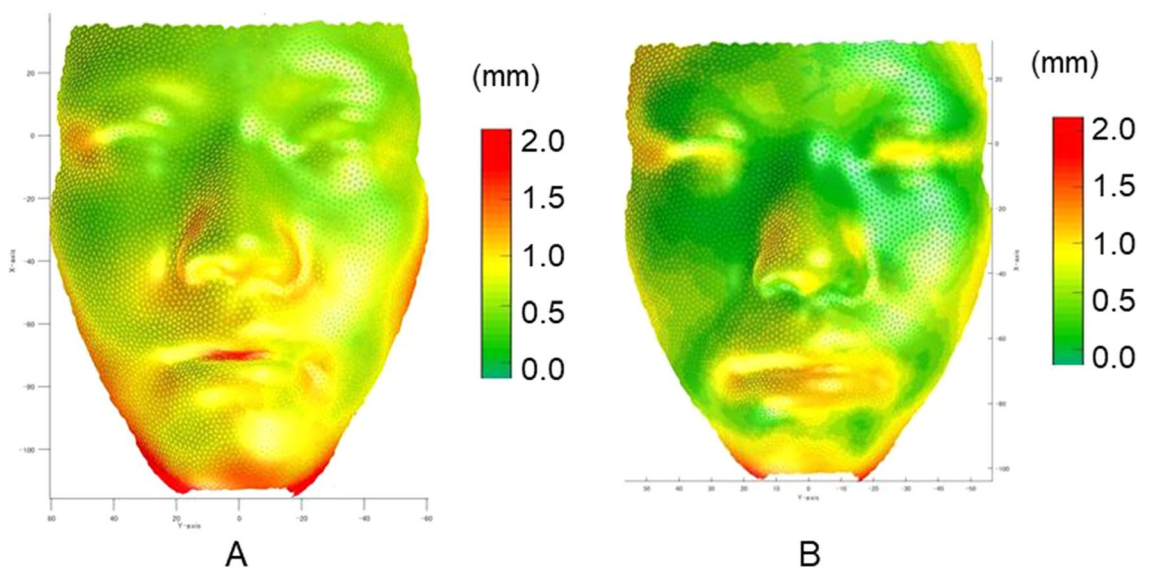
**Figure 4.** The mathematical model used to develop the artificial intelligence (AI) systems.



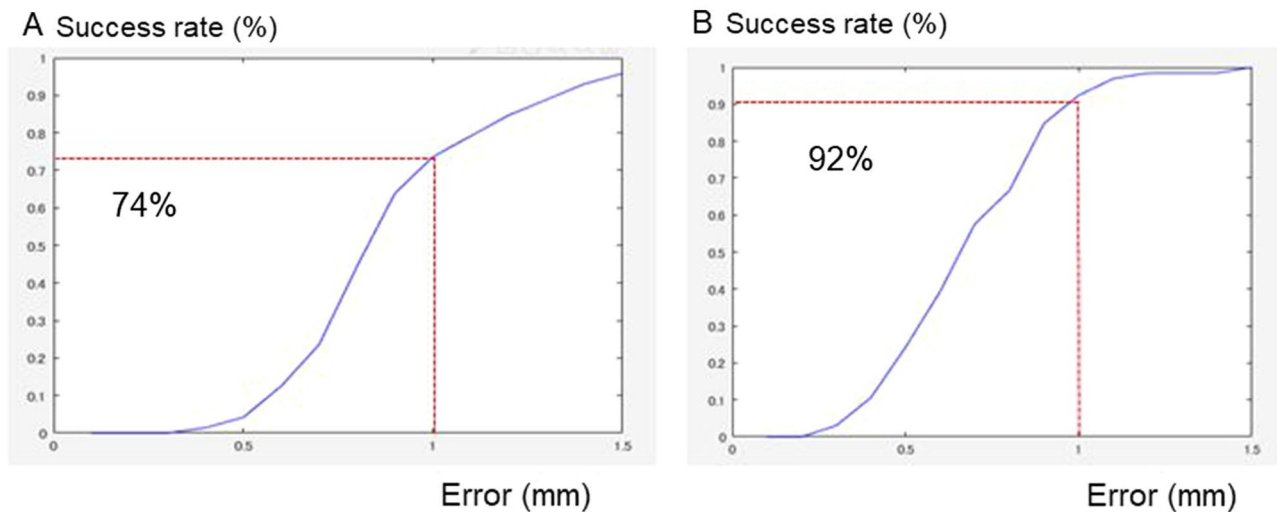
**Figure 5.** Average error for each patient with Systems S (A) and E (B). In the left column, the x-axis indicates the case number and the y-axis indicates the prediction error, which was defined as the difference between the predicted and actual post-treatment surface values. The right column shows the predicted face (red), actual post-treatment face (blue), and actual pre-treatment face (grey) for the eight representative cases in each group.

(unit, mm)	Total error	Variation among the points (AveEachPc)			Variation among the patients (AveEachPt)		
		S.D	Min	Max	S.D	Min	Max
System S	0.89	0.36	0.42	4.02	0.30	0.33	1.77
System E	0.69	0.22	0.24	1.77	0.18	0.30	1.02

**Table 1.** Average error, standard deviation (S.D.), minimum and maximum values for Systems S and E. For the definition of the AveEachPc and AveEachPt, please see Supplementary Table S1.



**Figure 6.** Average error (absolute differences in Z-axis between the predicted and actual facial form at post-treatment) with Systems S (A) and E (B).



**Figure 7.** Success rate for each threshold between 0.1 mm and 1.5 mm for Systems S (A) and E (B).

Since the mid-facial deficiency or cheek concavity is an important factor when considering the advancement of the maxilla by LeFort I, the present study is useful when deciding the amount of advancement.

The corner of the mouth is another area that has not previously been confirmed. In the present study the corner of the mouth showed significant mesial movement after orthognathic surgery, which means that the mouth width became narrower. This may be related to the assumption that backward movement of the lower lip may reduce the lip prominence from the axial view, which may reduce the width of the mouth. The prediction for the corner of the mouth also showed a greater error, which means that the hard-tissue and soft-tissue movement in the corner of the mouth is complex.

Because there were no 3-D systems that could predict the 3-D facial form after orthodontic treatment with four premolar extraction, it was difficult to compare the present AI system (System E) to previous studies. However, the 2-D Dolphin system<sup>26</sup> showed a mean error of 2.02 mm (tip of the nose) to 4.13 mm (soft-tissue gnathion) in the 2-D system; thus, the error of the present system, which was 0.69 mm, was considered to be small enough for clinical application. When the clinically acceptable prediction error was set as 2 mm, a previous study using 2-D systems showed success rates of 47.71% and 54.90%<sup>23</sup>; in contrast, the present system showed a success rate of 100%.

Regarding the facial portions showing prediction errors in system E, the lower lip showed a greater error (approximately 1.1 mm) in comparison to the upper lip (0.8 mm) in the present study. This result partially corresponded to the prediction error observed in a 2-D system<sup>26</sup>. The 2-D system<sup>26</sup> showed an error of 1.59 mm for the upper lip, while the error of the lower lip was 2.77 mm. In contrast, another study<sup>19</sup> showed that changes in the lower lip in response to orthodontic tooth movement were more predictable than those of the upper lip. Another study examining the lip changes in relation to the lip strain, which were expressed as lip thickness<sup>42</sup>, showed a strong correlation found between osseous changes and soft-tissue changes in subjects with thin lips ( $r = 0.9$ ), whereas no significant correlation was found in subjects with thick lips ( $r = 0.48\text{--}0.67$ ). This result indicated that pre-treatment functional morphological structures influence lip changes. It has been shown that the value of the proportionality constant in the lip region varies greatly depending on the case. A study examining two different types of malocclusion (i.e., maxillary protrusion and bimaxillary protrusion cases) showed that the cases with maxillary protrusion showed slightly less soft-tissue change in comparison to cases with bimaxillary protrusion<sup>5</sup>. The relationship between hard-tissue and soft-tissue movement is too complex to examine in a relatively simplistic way<sup>20</sup>. Because the method that the present system employed involved prediction based on past cases in patients with a similar facial form, we considered that the present system could predict more complicated movement in comparison to the conventional assumption that the proportional relationship between hard-tissue and soft-tissues can be applied to all patients.

The treatment results that we employed to develop the AI system were obtained from the records of two hospitals that had similar treatment procedures; thus, the system can mimic treatment results from these hospitals. However, it is unknown whether these systems can be applied in other hospitals or clinics. Further studies are needed to validate these results.

## Conclusion

The combination of GMM and deep learning were confirmed to be useful for the development of AI systems that predict the 3-D facial topography after orthognathic surgery (S) and fixed edgewise orthodontic treatment with four premolar extraction (E). The system error for two systems was  $0.94 \pm 0.43$  mm (S) and  $0.69 \pm 0.28$  mm (E), respectively. Maximum errors of approximately 0.8–1.2 mm were observed in the nasal ala, chin, corner of the mouth (S) and the lower lip (E). Although these difficult areas were clinically important regions, the total success rate at < 1 mm was 81% (54% [S] and 98% [E]); and that at < 2 mm was 100%, which indicates the possibility of applying the system in the clinical setting.

## Material and methods

**Subjects.** This was a retrospective study in a secondary adult care setting. A total of 137 pre-treatment and post-treatment records from patients who underwent orthognathic surgery (Surgery group;  $n=72$ ; mean age = 23.5 years) and those who underwent orthodontic treatment with premolar extraction (Extraction group;  $n=65$ ; mean age = 15.6 years) were enrolled. The inclusion criteria for the Surgery group were as follows: pre-treatment patients who visited the orthodontic department of Osaka University Dental Hospital, during 2011–2015; age 15–37 years; negative overjet; no severe mandibular deviation of  $>5$  mm; planned orthognathic surgery with LeFort I osteotomy and bilateral sagittal split osteotomy; no facial paralysis; body mass index, 18.50–24.99; and no maxillofacial plastic surgery in the previous six months. The inclusion criteria for the Extraction group were as follows: pre-treatment patients who visited the orthodontic department of Osaka University Dental Hospital and Kanomi Dental Hospital, during 2011–2015; age 12–37 years; positive overjet; no severe mandibular deviation of  $>5$  mm; planned edgewise orthodontic treatment with four premolar extraction; no facial paralysis; body mass index, 18.50–24.99; and no maxillofacial plastic surgery in the previous six months. A written informed consent form was distributed to and signed by all participants. Informed consent was approved by the Research Ethics Committee, Osaka University Dental Hospital (project ID: H25-E37-1). The treatment outcome was assessed in 2014–2020, after treatment was completed. The sample size was decided based on practical grounds (existing study cohort). There were no missing data on the predictors or outcomes. All experiments were performed in accordance with the relevant guidelines and regulations, and all experimental protocols were approved by the Research Ethics Committee, Osaka University Dental Hospital. Further, informed consent was obtained to publish the images in an online open access publication.

**Data acquisition.** Facial images of the subjects at rest were recorded once with a 3-D image capturing device (3-DMDcranial System, 3-DMD, Atlanta, GA, USA) with a natural head position without head support. 3-D images were acquired after treatment during the retention period.

**Data processing. Landmark identification.** For each 3-D facial image, the positions of 18 landmarks (nasion, pronasale, subnasale, labiale superius, stomion, labiale inferius, submentale, pogonion, porion [right and left], exocanthion [right and left], endocanthion [right and left], alar curvature [right and left], and cheilion [right and left]<sup>43</sup>) were identified and digitized using a commercial software program (Face Rugle, Medic Engineering Co., Kyoto, Japan). Facial images were standardized with a common coordinate system [Supplementary Fig S1]<sup>44</sup>.

**Wire mesh fitting.** For each participant at each time point, a landmark-based GMM analysis<sup>45</sup> was performed using the HBM software program (National Institute of Advanced Industrial Science and Technology, Japan). This analysis generated a fitted mesh consisting of a set of 6017 points as nodes (Supplementary Fig. S2).

**Morphological differences between pre-treatment and post-treatment.** To determine morphological characteristics of the facial soft-tissue surface changes during treatment in the Surgery and Extraction groups, the coordinate values of each node of the wire mesh on the facial surface on the  $x$ ,  $y$ , and  $z$  axes were statistically analyzed. A one-sample  $t$ -test was performed to compare each axis in the two subject groups. To visualize the differences between pre-treatment and post-treatment in the two subject groups, the results were represented as a color map showing the  $P$  values for the comparison between the two subject groups (the significance probability map) and a color map representing the differences between two subject groups (hereafter referred to as the distance map)<sup>45</sup>.  $P$  values of  $<0.05$  were considered to indicate statistical significance.

**AI systems.** Based on the deep learning method, we developed two AI systems (S and E) to predict changes in the coordinate values of the semi-landmarks due to orthognathic surgery (Surgery group) and orthodontic treatment with four premolar extraction (Extraction Group), respectively. As predictive variables for the AI system, the  $x$ -,  $y$ - and  $z$ -coordinate values of the 6017 3-D facial semi-landmarks for the pre-treatment,  $x$ - and  $y$ -coordinate values of 27 cephalometric landmarks (Table 1), and  $x$ - and  $y$ -coordinate changes during the treatment of 16 cephalometric landmarks (Table 1) were employed (i.e.,  $6,017 \times 3 + 27 \times 2 + 16 \times 2 = 18,137$  dimensions). The AI system's predicted outcome was set as the predicted change for each semi-landmark on the 3-D face ( $6,017 \times 3 = 18,051$  dimensions). The predicted post-treatment facial morphology was calculated as the sum of the coordinate value on the face before treatment and the predicted change for each semi-landmark (Fig. 4).

**Structure of deep neural networks.** The number of neurons in the input layer is the same as the number of features in our data. In the present study, 18,137 neurons were set as input layers. The 18,051 neurons were set for the regression output layer, since the expected output was the predicted change for each semi-landmark on the 3-D face. The system used two dense layers and one dropout layer, in which dense layers were the most common type of layer, where each of its neurons was connected to the neurons of the previous and next layer. Each dense layer has an activation function that determines the output of its neurons based on the inputs and weights of the synapses. In the present study, Rectified Linear Unit (ReLU) activation was performed for each dense layer. Dropout layers were also utilized for regularization, which randomly dropped some of the input units to 0, which helps to reduce the chance of overfitting the neural network. In our system, the dropout layer was set as 30%. As the optimized layer, which is the optimizing algorithm that helps achieve better results for the loss function, the Adam optimizer was used for learning. The mean squared error was set as the loss function when compiling. The system was trained with 30 epochs using keras and TensorFlow libraries in Python, where



we intentionally stopped the training process when the validation loss did not improve for certain epochs. Here, to tune the model's hyperparameters, we used records from five patients who were excluded from our samples (two upper premolar extraction and one jaw surgery). As a result, the same hyperparameters were used for both systems. Calculation was conducted using an Octopus supercomputer at Osaka University. Once the system was trained, the predicted change for each semi-landmark on the 3-D face for the new patients was estimated using the trained neural networks.

**Evaluation.** The performance of each AI system was evaluated using 11-fold cross-validation. Namely, the whole data were randomly split into eleven sets (or folds). The following procedure was conducted for each of the eleven folds: a model was trained using the 10 folds as training data; the resulting model was validated with the remaining data; and the system error for all patients was computed in the eleven folds. Here, the system error was determined as the difference between actual post- and predicted post-treatment coordinates of the semi-landmarks in the Z-axis (Table 1). Furthermore, the total success rate was examined, wherein successful cases were defined as those with an average error of < 1 mm or < 2 mm.

Received: 24 March 2021; Accepted: 13 July 2021

Published online: 04 August 2021

## References

1. Proffit, W. R. & Fields, H. W. *Contemporary Orthodontics* 3rd edn. (Mosby, 2000).
2. Chaconas, S. J. & Bartroff, J. D. Prediction of normal soft tissue facial changes. *Angle Orthod.* **45**, 12–25. [https://doi.org/10.1043/0003-3219\(1975\)045%3c0012:PONSTP%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(1975)045%3c0012:PONSTP%3e2.0.CO;2) (1975).
3. Droboczy, O. B. & Smith, R. J. Changes in facial profile during orthodontic treatment with extraction of four first premolars. *Am. J. Orthod. Dentofacial. Orthop.* **95**, 220–230. [https://doi.org/10.1016/0889-5406\(89\)90052-8](https://doi.org/10.1016/0889-5406(89)90052-8) (1989).
4. Diels, R. M., Kalra, V., DeLoach, N. Jr., Powers, M. & Nelson, S. S. Changes in soft tissue profile of African-Americans following extraction treatment. *Angle Orthod.* **65**, 285–292. [https://doi.org/10.1043/0003-3219\(1995\)065%3c0285:CISTPO%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(1995)065%3c0285:CISTPO%3e2.0.CO;2) (1995).
5. Yogosawa, F. 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](https://doi.org/10.1043/0003-3219(1990)060%3c0199:PSTPCC%3e2.0.CO;2) (1990).
6. Tan, T. J. Profile changes following orthodontic correction of bimaxillary protrusion with a preadjusted edgewise appliance. *Int. J. Adult Orthodon. Orthognath. Surg.* **11**, 239–251 (1996).
7. Saelens, N. A. & De Smit, A. A. Therapeutic changes in extraction versus non-extraction orthodontic treatment. *Eur. J. Orthod.* **20**, 225–236. <https://doi.org/10.1093/ejo/20.3.225> (1998).
8. Ismail, S. F., Moss, J. P. & Hennessy, R. Three-dimensional assessment of the effects of extraction and nonextraction orthodontic treatment on the face. *Am. J. Orthod. Dentofac. Orthop.* **121**, 244–256. <https://doi.org/10.1067/mod.2002.121010> (2002).
9. Basciftici, F. A., Uysal, T., Buyukerkmen, A. & Demir, A. The influence of extraction treatment on Holdaway soft-tissue measurements. *Angle Orthod.* **74**, 167–173. [https://doi.org/10.1043/0003-3219\(2004\)074%3c0167:TIOETO%3e2.0.CO;2](https://doi.org/10.1043/0003-3219(2004)074%3c0167:TIOETO%3e2.0.CO;2) (2004).
10. Erdinc, A. E., Nanda, R. S. & Dandajena, T. C. Profile changes of patients treated with and without premolar extractions. *Am. J. Orthod. Dentofac. Orthop.* **132**, 324–331. <https://doi.org/10.1016/j.ajodo.2005.08.045> (2007).
11. Freitas, B. V., Rodrigues, V. P., Rodrigues, M. E., de Melo, H. V. F. & Dos Santos, P. C. F. Soft tissue facial profile changes after orthodontic treatment with or without tooth extractions in Class I malocclusion patients: A comparative study. *J. Oral Biol. Craniofac. Res.* **9**, 172–176. <https://doi.org/10.1016/j.jobcr.2018.07.003> (2019).
12. Young, T. M. & Smith, R. J. Effects of orthodontics on the facial profile: A comparison of changes during nonextraction and four premolar extraction treatment. *Am. J. Orthod. Dentofac. Orthop.* **103**, 452–458. [https://doi.org/10.1016/S0889-5406\(05\)81796-2](https://doi.org/10.1016/S0889-5406(05)81796-2) (1993).
13. Kolokitha, O. E., Athanasiou, A. E. & Tuncay, O. C. Validity of computerized predictions of dentoskeletal and soft tissue profile changes after mandibular setback and maxillary impaction osteotomies. *Int. J. Adult Orthodon. Orthognath. Surg.* **11**, 137–154 (1996).
14. Tsang, S. T., McFadden, L. R., Wiltshire, W. A., Pershad, N. & Baker, A. B. Profile changes in orthodontic patients treated with mandibular advancement surgery. *Am. J. Orthod. Dentofac. Orthop.* **135**, 66–72. <https://doi.org/10.1016/j.ajodo.2007.01.033> (2009).
15. Betts, N. J. & Dowd, K. F. Soft tissue changes associated with orthognathic surgery. *Atlas Oral Maxillofac. Surg. Clin. N. Am.* **8**, 13–38 (2000).
16. Cozza, P., De Toffol, L. & Colagrossi, S. Dentoskeletal effects and facial profile changes during activator therapy. *Eur. J. Orthod.* **26**, 293–302. <https://doi.org/10.1093/ejo/26.3.293> (2004).
17. Idris, G., Hajeer, M. Y. & Al-Jundi, A. Soft- and hard-tissue changes following treatment of Class II division 1 malocclusion with Activator versus Trainer: A randomized controlled trial. *Eur. J. Orthod.* **41**, 21–28. <https://doi.org/10.1093/ejo/cjy014> (2019).
18. Kilic, N., Celikoglu, M. & Oktay, H. Effects of the functional regulator III on profile changes in subjects with maxillary deficiency. *Eur. J. Orthod.* **32**, 729–734. <https://doi.org/10.1093/ejo/cjq011> (2010).
19. Talass, M. F., Talass, L. & Baker, R. C. Soft-tissue profile changes resulting from retraction of maxillary incisors. *Am. J. Orthod. Dentofac. Orthop.* **91**, 385–394. [https://doi.org/10.1016/0889-5406\(87\)90391-x](https://doi.org/10.1016/0889-5406(87)90391-x) (1987).
20. Battagel, J. M. The relationship between hard and soft tissue changes following treatment of Class II division 1 malocclusions using Edgewise and Frankel appliance techniques. *Eur. J. Orthod.* **12**, 154–165. <https://doi.org/10.1093/ejo/12.2.154> (1990).
21. Jin, S. C., Kasai, K., Iwasawa, T. & Kanazawa, E. Lip form responses to changes in maxillary incisor position. *J. Nihon Univ. Sch. Dent.* **38**, 146–154. <https://doi.org/10.2334/josnusd1959.38.146> (1996).
22. Konstantos, K. A., O'Reilly, M. T. & Close, J. The validity of the prediction of Soft Tissue profile changes after LeFort I osteotomy using the dentofacial planner (computer software). *Am. J. Orthod. Dentofac. Orthop.* **105**, 241–249. [https://doi.org/10.1016/S0889-5406\(94\)70117-2](https://doi.org/10.1016/S0889-5406(94)70117-2) (1994).
23. Peterman, R. J., Jiang, S., Johe, R. & Mukherjee, P. M. Accuracy of Dolphin visual treatment objective (VTO) prediction software on class III patients treated with maxillary advancement and mandibular setback. *Prog. Orthod.* **17**, 19. <https://doi.org/10.1186/s40510-016-0132-2> (2016).
24. Kaipatur, N. R. & Flores-Mir, C. Accuracy of computer programs in predicting orthognathic surgery soft tissue response. *J. Oral Maxillofac. Surg.* **67**, 751–759. <https://doi.org/10.1016/j.joms.2008.11.006> (2009).
25. Smith, J. D., Thomas, P. M. & Proffit, W. R. A comparison of current prediction imaging programs. *Am. J. Orthod. Dentofac. Orthop.* **125**, 527–536. <https://doi.org/10.1016/S0889540604001210> (2004).

26. Zhang, X. *et al.* Accuracy of computer-aided prediction in soft tissue changes after orthodontic treatment. *Am. J. Orthod. Dentofac. Orthop.* **156**, 823–831. <https://doi.org/10.1016/j.ajodo.2018.11.021> (2019).
27. Kuijpers, M. A., Chiu, Y. T., Nada, R. M., Carels, C. E. & Fudalej, P. S. Three-dimensional imaging methods for quantitative analysis of facial soft tissues and skeletal morphology in patients with orofacial clefts: a systematic review. *PLoS ONE* **9**, e93442. <https://doi.org/10.1371/journal.pone.0093442> (2014).
28. Metzler, P. *et al.* Validity of the 3D VECTRA photogrammetric surface imaging system for cranio-maxillofacial anthropometric measurements. *Oral Maxillofac. Surg.* **18**, 297–304. <https://doi.org/10.1007/s10006-013-0404-7> (2014).
29. Tanikawa, C., Akcam, M. O. & Takada, K. Quantifying faces three-dimensionally in orthodontic practice. *J. Cranio Maxill Surg.* **47**, 867–875. <https://doi.org/10.1016/j.jcms.2019.02.012> (2019).
30. Tanikawa, C. *et al.* Functional decline in facial expression generation in older women: A cross-sectional study using three-dimensional morphometry. *PLoS ONE* **14**, e0219451. <https://doi.org/10.1371/journal.pone.0219451> (2019).
31. Tanikawa, C. Facial morphospace: A clinical quantitative analysis of the three-dimensional face in patients with cleft lip and palate. *Plast. Aesthet. Res.* **7**, 48. <https://doi.org/10.20517/2347-9264.2020.136> (2020).
32. Resnick, C. M., Dang, R. R., Glick, S. J. & Padwa, B. L. Accuracy of three-dimensional soft tissue prediction for Le Fort I osteotomy using Dolphin 3D software: A pilot study. *Int. J. Oral Maxillofac. Surg.* **46**, 289–295. <https://doi.org/10.1016/j.ijom.2016.10.016> (2017).
33. Elshebiny, T., Morcos, S., Mohammad, A., Quereshy, F. & Valiathan, M. Accuracy of three-dimensional soft tissue prediction in orthognathic cases using dolphin three-dimensional software. *J. Craniofac. Surg.* **30**, 525–528. <https://doi.org/10.1097/SCS.0000000000005037> (2019).
34. Bookstein, F. L. Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Med. Image Anal.* **1**, 225–243. [https://doi.org/10.1016/s1361-8415\(97\)85012-8](https://doi.org/10.1016/s1361-8415(97)85012-8) (1997).
35. Schendel, S. A., Jacobson, R. & Khalessi, S. 3-Dimensional facial simulation in orthognathic surgery: Is it accurate?. *J. Oral Maxillofac. Surg.* **71**, 1406–1414. <https://doi.org/10.1016/j.joms.2013.02.010> (2013).
36. Ullah, R., Turner, P. J. & Khambay, B. S. Accuracy of three-dimensional soft tissue predictions in orthognathic surgery after Le Fort I advancement osteotomies. *Br. J. Oral Maxillofac. Surg.* **53**, 153–157. <https://doi.org/10.1016/j.bjoms.2014.11.001> (2015).
37. Nkenke, E. *et al.* Three-dimensional analysis of changes of the malar-midfacial region after LeFort I osteotomy and maxillary advancement. *Oral Maxillofac. Surg.* **12**, 5–12. <https://doi.org/10.1007/s10006-008-0094-8> (2008).
38. Mankad, B., Cisneros, G. J., Freeman, K. & Eisig, S. B. Prediction accuracy of soft tissue profile in orthognathic surgery. *Int. J. Adult Orthodon. Orthognath. Surg.* **14**, 19–26 (1999).
39. Chew, M. T., Sandham, A. & Wong, H. B. Evaluation of the linearity of soft- to hard-tissue movement after orthognathic surgery. *Am. J. Orthod. Dentofac. Orthop.* **134**, 665–670. <https://doi.org/10.1016/j.ajodo.2006.12.017> (2008).
40. Wang, S. & Yang, J. An improved finite element model for craniofacial surgery simulation. *Int. J. Comput. Assist. Radiol. Surg.* **4**, 579–587. <https://doi.org/10.1007/s11548-009-0373-3> (2009).
41. Kim, H., Jurgens, P., Nolte, L. P. & Reyes, M. Anatomically-driven soft-tissue simulation strategy for cranio-maxillofacial surgery using facial muscle template model. *Lect. Notes Comput. Sci.* **6361**, 61 (2010).
42. Oliver, B. M. The influence of lip thickness and strain on upper lip response to incisor retraction. *Am. J. Orthod.* **82**, 141–149. [https://doi.org/10.1016/0002-9416\(82\)90492-4](https://doi.org/10.1016/0002-9416(82)90492-4) (1982).
43. Farkas, L. G. *Anthropometry of the Head and Face* (Raven Press, 1994).
44. Tanikawa, C., Zere, E. & Takada, K. Sexual dimorphism in the facial morphology of adult humans: A three-dimensional analysis. *Homo* **67**, 23–49. <https://doi.org/10.1016/j.jchb.2015.10.001> (2016).
45. Kono, K., Tanikawa, C., Yanagita, T., Kamioka, H. & Yamashiro, T. A novel method to detect 3D mandibular changes related to soft-diet feeding. *Front. Physiol.* **8**, 567. <https://doi.org/10.3389/fphys.2017.00567> (2017).

## Acknowledgements

This work was partially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Nos. 22792048 and 25862008) and Japan Agency for Medical Research and Development (AMED) (Grant No. A100). We declare that the JSPS and AMED had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

C.T. conceived the study and were in charge of overall direction and planning. C.T. also designed the model and the computational framework, carried out the data acquisition, and analyzed the data and performed the calculations. C.T. and T.Y. discussed the results and contributed to the final manuscript.

## Competing interests

C.T. have applied a patent for this technology, and T.Y. declares that there is no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-95002-w>.

**Correspondence** and requests for materials should be addressed to C.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021