AJO-DO

Voxel-based superimposition of serial craniofacial cone-beam computed tomographies for facial soft tissue assessment: Reproducibility and segmentation effects

Georgios Kanavakis,[a,b] Simeon T. Häner,[c] François Matthey,[d] and Nikolaos Gkantidis[c]
Boston, Mass, and Basel, Bern, and Neuchâtel, Switzerland

Introduction: The aim of this investigation was to evaluate the reproducibility of a voxel-based 3-dimensional superimposition method and the effect of segmentation error on determining soft tissue surface changes. **Methods:** A total of 15 pairs of serial cone-beam computed tomography images (interval: 1.69 ± 0.37 years) from growing subjects (initial age: 11.75 ± 0.59 years) were selected from an existing digital database. Each pair was superimposed on the anterior cranial base, in 3 dimensions with Dolphin 3D software (version 2.1.6079.17633; Dolphin Imaging & Management Solutions, Chatsworth, Calif). The reproducibility of superimposition outcomes and surface segmentation were tested with intra- and interoperator comparisons. **Results:** Median differences in inter- and intrarater measurements at various areas presented a range of 0.08-0.21 mm. In few instances, the differences were larger than 0.5 mm. In areas where T0-T1 changes were increased, the error did not appear to increase. However, the method error increased the farther the measurement area was from the superimposition reference structure. For individual images, the median soft tissue segmentation error ranged from 0.05 to 0.06 at various areas and in no subject exceeded 0.13 mm. **Conclusions:** The presented voxel-based superimposition method was efficient and well reproducible. The segmentation process was a minimal source of error; however, there were a few cases in which the total error was more than 0.5 mm and could be considered clinically significant. Therefore, this method can be used clinically to assess 3-dimensional soft tissue changes during orthodontic treatment in growing patients. (Am J Orthod Dentofacial Orthop 2021;159:343-51)

The paradigm shift in modern orthodontics toward more soft tissue–based diagnostics has expanded the social role of the specialty to one that manages a greater spectrum of patient needs than straight teeth. Facial and smile esthetics have gained increased attention in the professional orthodontic community and among patients whose decision to pursue treatment is primarily driven by their wish to improve their appearance.[1-3] This motive is also well supported by a large body of literature emphasizing the significant social impact of facial and smile attractiveness on various aspects of

[a]Department of Pediatric Oral Health and Orthodontics, University Center for Dental Medicine, University of Basel, Basel, Switzerland.
[b]Department of Orthodontics, Tufts University School of Dental Medicine, Boston, Mass.
[c]Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Bern, Switzerland.
[d]Private practice, Neuchâtel, Switzerland.

All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported.

This study was reviewed and approved by the Swiss Ethic Committee (protocol no. 2018-01670). All inclusion criteria were visually assessed by 2 independent researchers (S.T.H, N.G) before any data were used for the purposes of the investigation. Furthermore, the entire research methodology, including surface segmentation of cone-beam computed tomography T1, was conducted twice, by 2 independent operators (S.T.H, G.K) to test interoperator agreement. Intraoperator reliability was tested by repetition of the superimposition process for 10 sets of 3-dimensional models by 1 operator (G.K).

Address correspondence to: Georgios Kanavakis, Department of Pediatric Oral Health and Orthodontics, University Center for Dental Medicine, University of Basel, Mattenstrasse 40, Basel CH-4058, Switzerland; e-mail, georgios.kanavakis@unibas.ch or gkanavak@gmail.com.

Submitted, February 2020; revised, March 2020; accepted, April 2020.
0889-5406

https://doi.org/10.1016/j.ajodo.2020.04.022

professional and personal life and on overall well-being.[4-7]

Although facial esthetics have traditionally played an important role in treatment planning,[8,9] there is no widely used reliable tool for measuring soft tissue dimensions or soft tissue changes in clinical practice. Direct anthropometric measurements are considered reliable but are not easy to perform and are time-consuming. By contrast, 2-dimensional (2D) photography or radiography, which is used routinely, does not provide safe information on facial dimensions nor allows for reliable comparisons of different time points.[10-12] Furthermore, the dimensional reduction of a 3-dimensional (3D) object leads to a significant loss of information that limits the imaging value of 2D modalities. Three-dimensional photography seems quite promising but is not yet incorporated in clinical practice.

As a result, clinicians are mostly limited in using 2D (lateral cephalometric images) and 3D (cone-beam computed tomography [CBCT] images) radiographs to assess facial soft tissues at a certain time point. However, cephalometric assessments of soft tissue structures are prone to various sources of error[11-14]; and the availability of CBCT images requires that they are justified and have been obtained following the As Low As Diagnostically Acceptable principle. Nevertheless, if available, they depict both hard and soft tissues and also provide stable reference structures for assessing facial changes after treatment or growth.[11] Technological advances in 3D imaging techniques have extended the use of computed tomography and CBCT.[15,16] This development has increased the potential of studying the craniofacial complex and its surrounding tissues in a more thorough manner. However, this requires the adaptation of currently used diagnostic analyses into 3 dimensions. For example, the anterior cranial base, which remains stable after an early age,[17,18] has been used traditionally for studying changes on cephalometric images but can also be used reliably in 3D superimpositions.[16,19]

Superimpositions of serial 3D images are performed with landmark-based, surface-based, or voxel-based methods,[16] each of which presents its advantages and disadvantages. Landmark-based methods are highly dependent on the number of selected landmarks and can significantly become time-consuming when high accuracy is required.[19,20] Surface-based techniques are faster and more user-friendly; however, they require previous segmentation of the bony structures from the entire volume. Therefore, they are subject to the error generated by selecting different threshold values when performing this process.[21] In CBCT images, even if the same threshold is used, there is no guarantee for precise segmentation owing to the absence of correspondence of gray scale values to Hounsfield units.[22]

Voxel-based methods use the best-fit approach to superimpose original volumetric data (voxel gray scale values) and are thus not subject to segmentation error. Nevertheless, they also require bone segmentation to create a surface model that can then be used for a thorough outcome assessment. Several studies have investigated voxel-based superimposition techniques,[23,24] some of which reported the technique to be user-friendly and highly reliable.[23] The voxel-based Dolphin 3D (version 2.1.6079.17633; Dolphin Imaging & Management Solutions, Chatsworth, Calif) superimposition technique has been recently tested on a growing population, and it proved suitable for everyday use in clinical practice to evaluate hard tissue treatment outcomes and skeletal changes owing to growth.[25]

Nevertheless, CBCT volumes also include soft tissue data, and software programs allow for the construction of a surface rendering from the original volumetric information. The aim of the present investigation was to assess the reliability and reproducibility of the voxel-based Dolphin 3D superimposition technique to detect soft tissue changes in a growing patient population. The effect of soft tissue surface segmentation on the superimposition outcome was also tested.

MATERIAL AND METHODS

This prospective methodological study, using pre-existing patient data, was registered and approved by the Swiss Ethics Committee (protocol no. 2018-01670).

The electronic archives of a single orthodontic clinic, between 2008 and 2018, were reviewed. This clinic does not take CBCTs routinely on all incoming patients; therefore, all CBCTs in the archives were originally obtained in subjects for whom they were considered to be essential for proper diagnosis and treatment planning. All scans were performed with the same x-ray machine (KaVo 3D eXam; KaVo Dental, Hatfield, Pa) under the following settings: scanning area 170 mm × 232 mm; 0.4 mm^3-voxel size; 5 mA; 120 kV; scan time, 8.9 seconds; exposure time, 3.7 seconds, which allowed for lower dose scans.[26] The data were saved and exported in a Digital Imaging and Communications in Medicine format. The scans of patients with craniofacial syndromes, malformations, or severe facial asymmetries, as well as low-quality scans, were disregarded.

The final sample comprised serial CBCT images of 15 (8 males and 7 females) growing orthodontic patients. This number of scans represents all available scans in the archives that fulfilled the inclusion criteria and was considered adequate for the purpose of the study.[19]

Two time points were recorded; T0 (first CBCT volume) and T1 (second CBCT volume). The mean ages of participants were 11.75 ± 0.59 years and 13.44 ± 0.96 years at T0 and T1, respectively. Two researchers (S.T.H, N.G) visually inspected all criteria independently and agreed on them before proceeding to any data generation process.

All pairs of Digital Imaging and Communications in Medicine datasets were imported in Dolphin 3D software. Voxel-based superimpositions of each pair of CBCTs were performed on the anterior cranial base, which is a standard reference to evaluate changes in craniofacial structures over time.[17,19,23] The first CBCT volume (ie, taken at T0 and mentioned as CBCT T0 thereafter) was repositioned according to the position of the CBCT taken at T1 (named CBCT T1 thereafter). The exact process followed is displayed in Supplementary Figure 1 and has been described previously.[25] Consequently, soft tissue surfaces were extracted from the superimposed volumes through segmentation to evaluate the superimposition outcome. The soft tissue segmentation of the CBCT T0 was conducted with the automated function of Dolphin, using the same threshold for a single dataset. Therefore, the surface generated from CBCT T0 was not affected by the segmentation factor allowing for comparisons between and within operators. By contrast, the segmentation of the CBCT T1 was performed manually, through real-time visual inspection of the resulting surface with different threshold values, until the segmented surface was smooth, with minimum artifacts or holes. All segmented soft tissue surfaces were then extracted into STL format for further analyses. All aforementioned procedures were performed twice, by 2 independent operators (S.T.H, G.K) to test interrater agreement. Furthermore, 1 operator (G.K) repeated the whole superimposition process for 10 sets of 3D models to assess intraoperator agreement. To assess the effect of the segmentation process on the superimposition outcome, the same operator repeated the manual segmentation of 10 surfaces and performed a manual segmentation of 10 surfaces that were previously extracted automatically.

All generated STL files were imported and further processed with Viewbox 4 Software (version 4.1.0.1, BETA 64; dHAL software, Kifisia, Greece), which has been thoroughly tested in surface model processing.[25,27-29] To evaluate the reproducibility of the method, intra- and interoperator agreements were measured on color-coded maps by determining the mean absolute distances (MAD) between corresponding models in each set. The MADs between T0 and T1 surface models were measured on areas with a predetermined size of 100 mesh points, which were selected on the T0 models and transferred to all other models generated from the same set, for consistency reasons. The following 7 areas were selected: N-point, A-point, pogonion, zygomatic arch right and left, and gonial angle right and left (Supplementary Fig 2).

The intraoperator agreement was tested through calculating the MADs between 2 surface models both extracted from CBCT T0 at different time points, by the same operator (operator 1). T0 surface models were chosen because their segmentation was performed with an automated software function, and, thus, there was no manual threshold definition error. Similarly, the MAD between the surface models extracted by 2 operators at T0 was measured to assess interrater agreement. A MAD of 0 mm indicated perfect agreement. The intra- and interrater agreement in measured T0-T1 changes were also tested by comparing the respective MADs between registered models, on the 7 areas previously mentioned. This comparison is influenced by the segmentation error related to the manual T1 model extraction. The magnitude of this error was explored by measuring the MADs between 10 sets of repeatedly extracted surface models. Those were compared with each other, and one of them was also compared with another automatically segmented model. Any deviation between the pairs of extracted models represented the segmentation error.

Statistical analysis

For consistency reasons, the statistical methodology of the present study was identical to that applied to a previous similar study on hard tissue outcomes.[25] Statistical analysis was performed using the SPSS software (version 25.0; IBM Corp, Armonk, NY).

Raw data were tested for normality through the Shapiro-Wilk test and were not normally distributed in all cases. Thus, nonparametric statistical tests were used. Intra- and interoperator agreement on superimposition outcome was shown with box plots. Any deviation from 0 indicated a superimposition error. Differences in the amount of error among the selected areas were tested in a paired manner through the Friedman test. In the case of significant results, pairwise comparisons were performed through the Wilcoxon signed rank test.

Segmentation error was tested in the same manner as the superimposition error.

In all cases, a 2-sided significance test was carried out at an alpha level of 0.05. In the case of multiple comparisons, a Bonferroni comparison was applied to the level of significance to avoid false-positive results.

The Bland-Altman method (difference plot)[30] was also used to evaluate interoperator agreement in the detected morphologic changes. A 1-sample *t* test was used to assess if there was a systematic error between the operators.

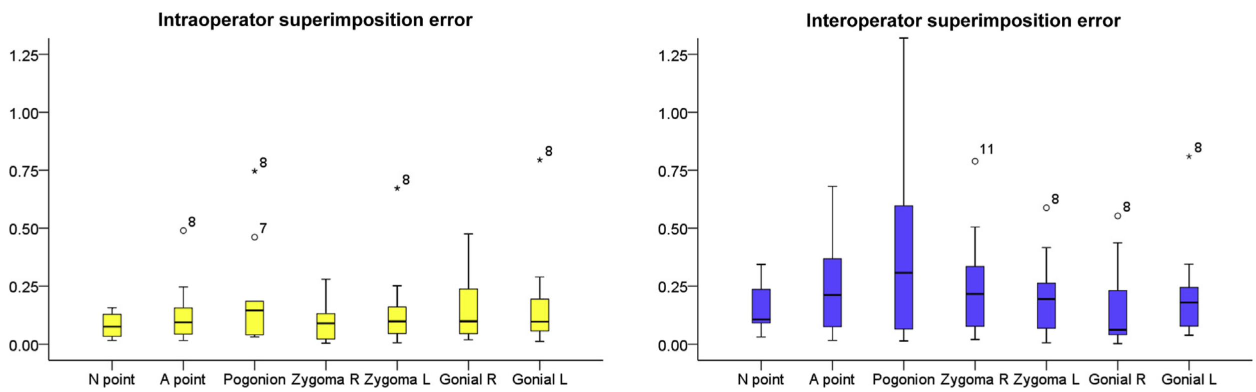


Fig 1. Box plots showing the intra- (*top*) and interoperator (*bottom*) error in millimeters, for all measurement areas. Zero indicates perfect reproducibility, whereas any deviation from 0 is considered error. The upper limit of the *black line* represents the maximum value; the lower limit, the minimum value; the *box*, the interquartile range; and the *horizontal black line,* the median value. Outliers are shown as *black dots* or *stars* in more extreme cases. No significant difference was detected between the measurement areas (Wilcoxon signed rank test). *L*, left; *R*, right.

RESULTS

The intra- and interoperator agreement on superimposition outcomes, assessed through the MAD between the relocated T0 models is shown in Figure 1. The median error ranged from 0.08 to 0.15 mm and from 0.17 to 0.21 mm for the intra- and interoperator comparisons, respectively. This amount of error was considered clinically acceptable in both cases. In a few cases, it exceeded 0.5 mm, mainly for interoperator comparisons. No significant differences were evident between the

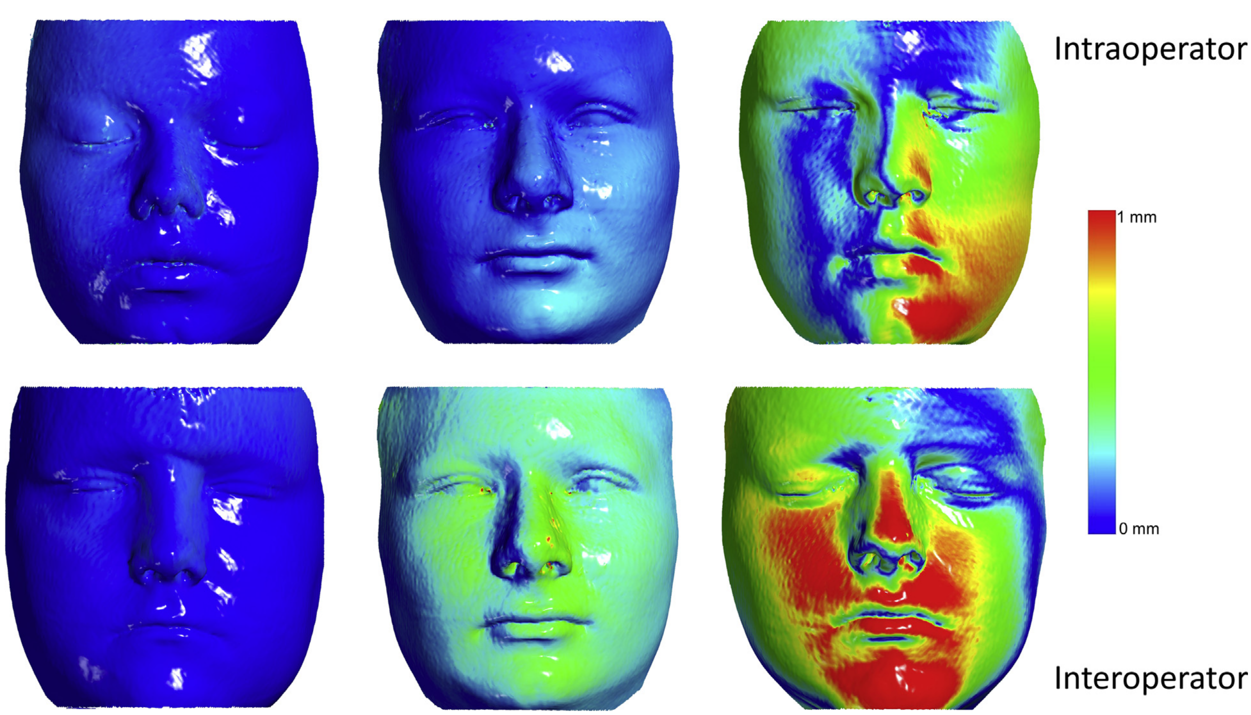


Fig 2. Intra- (*top*) and interoperator (*bottom*) differences between the T0 surfaces generated from repeated superimpositions are displayed in color-coded distance maps. The samples that presented the least (*left*), average (*middle*), and largest (*right*) differences are shown.

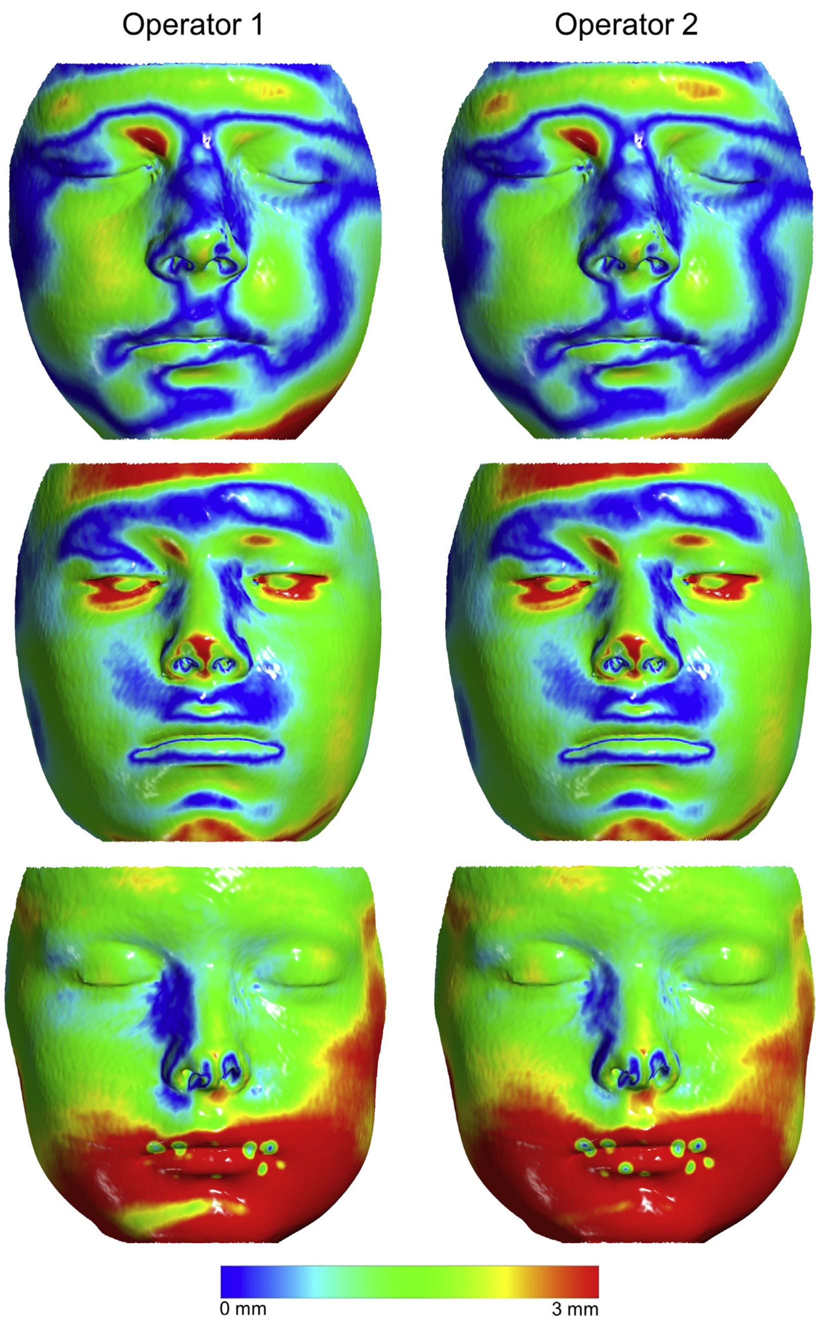


Fig 3. The detected T0-T1 changes illustrated in color-coded distance maps. The samples that presented the least (*upper*), average (*middle*), and largest (*lower*) are shown. Only minimal differences can be seen between the operators and these include the operator-dependent superimposition error plus the segmentation error of 1 surface model.

error assessed at different areas for both intra- ($P = 0.260$) and interoperator ($P = 0.095$) comparisons. In both cases, the measurements showed the highest values and the largest variance on the pogonion region, which is located at the furthest distance from the superimposition reference. Visual assessment of the respective color maps confirmed these outcomes (Fig 2). These results are free of any segmentation error, and, thus, they show the pure superimposition process error.

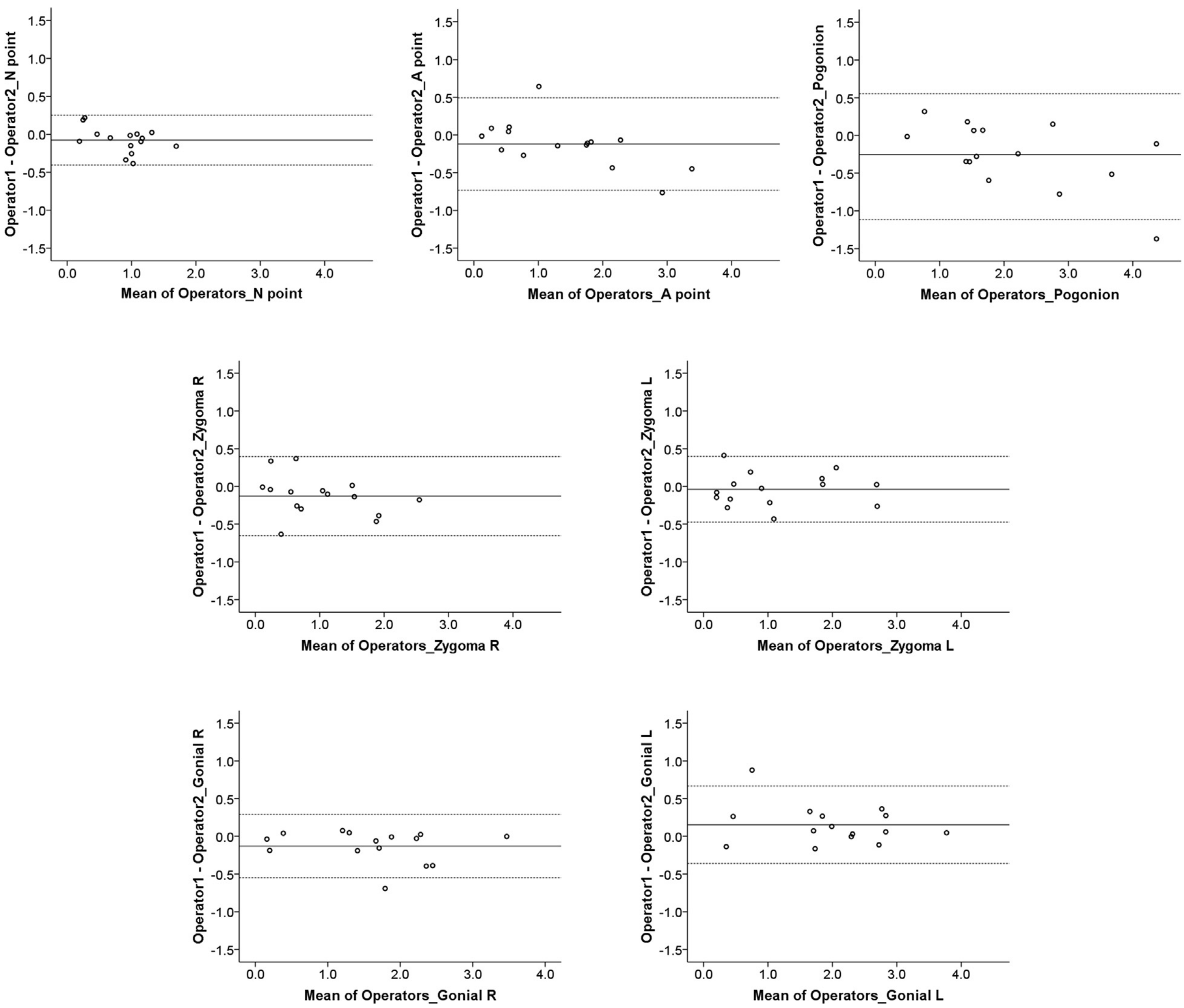


Fig 4. Bland-Altman plots on the T0-T1 changes (millimeters) detected by each operator. The continuous *horizontal line* shows the mean of the differences in the detected T0-T1 changes, and the *dashed lines* show the corresponding 95% confidence intervals. *L*, left; *R*, right.

There was no systematic difference between the 2 operators in the measured T0-T1 changes (1-sample t test, $P > 0.01$). Apart from a few cases, deviations of the individual measurements generally remained within 0.5 mm (Figs 3 and 4). Interoperator differences were not affected by the amount of the detected T0-T1 changed. At the midline, they tended to increase on increasing distance from the cranial base. These results were affected by segmentation error only regarding 1 of the 2 tested surfaces (T1).

The median error of repeated manual surface segmentation ranged between 0.05 and 0.06 mm. In all cases, the observed amount of error can be considered small. The maximum segmentation error observed in the whole sample was 0.13 mm (Fig 5). This is also evident in the respective color maps (Fig 6). Significant differences were found between different areas ($P = 0.001$), but no pairwise comparison remained significant after Bonferroni correction ($P > 0.003$).

DISCUSSION

The present investigation tested the reproducibility of a voxel-based superimposition method to detect 3D soft tissue changes owing to growth or as a result of orthodontic treatment and assessed all possible sources of potential error. The reliability of this superimposition technique has been previously reported as high.[25] The process of conducting the present study confirmed that the voxel-

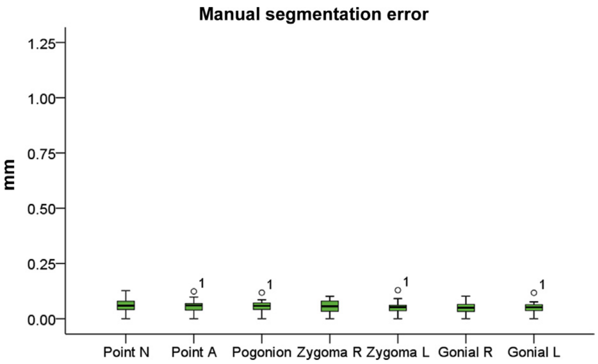


Fig 5. Box plots showing, in the y-axis (millimeters), the error between repeated manual segmentations for all measurement areas. Zero indicates perfect reproducibility, whereas any deviation from 0 is considered error. The upper limit of the *black line* represents the maximum value; the lower limit, the minimum value; the *box*, the interquartile range; and the *horizontal black line*, the median value. Outliers are shown as *black dots* or *stars* in more extreme cases. No significant difference was detected between the measurement areas (Wilcoxon signed rank test). *L*, left; *R*, right.

based Dolphin 3D superimposition method is user-friendly and fast. Other authors have also had the same experience while using this software.[23] In addition, our results indicate that the method is reproducible and can provide reliable information regarding 3D changes in soft tissues of growing subjects.

Superimposition error was less than 0.5 mm in all cases with few exceptions, whereas the segmentation error never exceeded 0.13 mm, with a median magnitude ranging between 0.05 and 0.06 mm. Previously reported data on hard tissues also presented low error values,[25] and, thus, it can be concluded that the method can be used with confidence in everyday orthodontic practice to assess hard and soft tissue changes. A more critical view on the results reveals that in selected cases, however, the superimposition error exceeded 0.5 mm. If the maximum segmentation error found in this study (ie, 0.13 mm) is added twice (once for each extracted surface file) to a potential superimposition error of 0.6 mm, then the total error would be 0.86 mm, which arguably reaches clinically significant levels. There is a 2-fold possible explanation for this; it may be related to differences in superimposition reference structures owing to anatomic changes or image imprecision, or it may also be related to Dolphin's selection tool, which is a rectangular frame, and, thus, nonosseous structures could also be included in the selection of the reference structure.

Furthermore, the total error detected in this study appeared to increase as the distance between a selected area and the superimposition reference area increased. The amount of error was larger in facial areas located more laterally and caudally to the anterior cranial base. This could be attributed to small rotations of the surface model around a center located somewhere in the reference area,[31] which manifested when a best-fit matching was performed. This was true in all cases, regardless of the operator, and is depicted on the color maps in Figure 2. Despite this observation, the error at distant surface areas still rarely exceeded 0.5 mm.

The amount of change at a certain area had no effect on the reproducibility of the superimposition outcome. The same finding was evident when assessing hard tissue changes,[25] but there the actual changes were much smaller. Here, soft tissue changes were almost double as large as the equivalent hard tissue ones and could have potentially led to higher error values. Nevertheless, this was not evident.

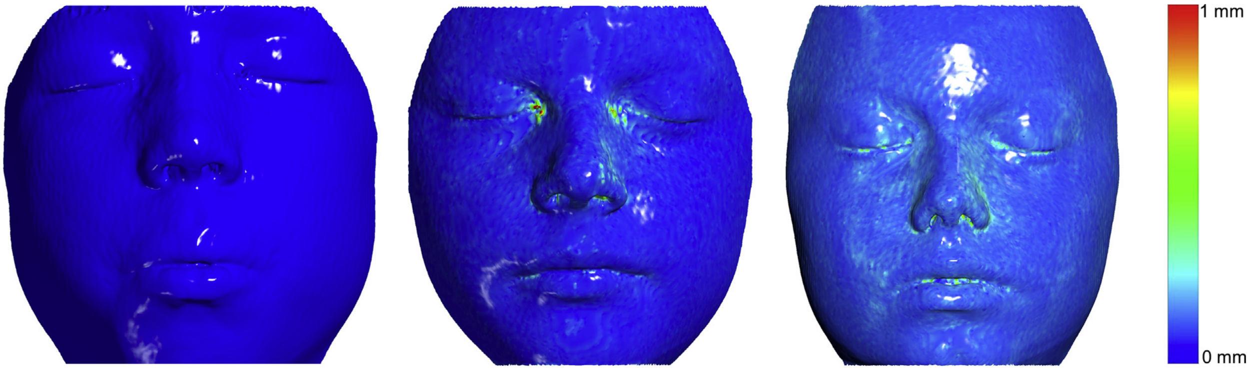


Fig 6. Color maps showing the difference between identical T0 surfaces generated from repeated manual segmentations. To illustrate the range of error the least (*left*), average (*middle*), and largest (*right*) deviating samples are presented.

The voxel-based Dolphin 3D superimposition method has previously produced promising outcomes; however, previous assessments of the method did not study individual differences in the reproducibility of the method[23] or only assessed hard tissue changes.[25] In this study, we have performed a comprehensive evaluation of the superimposition outcomes on soft tissues and also accounted for the error related to the segmentation process. The present study and the previous one on hard tissue outcomes[25] constitute a complete and thorough assessment of the voxel-based Dolphin 3D superimposition method.

Previous studies have also used the same software to measure soft tissue changes on superimposed CBCTs of surgical patients,[32] but used linear and angular measurements to test the reproducibility of their method. The identification of soft tissue landmarks on 3D images is moderately reliable[33] and when added to other sources of error might not be the optimal method for testing superimposition outcomes. Others have tested the use of a reference plane to study soft tissue changes on growing subjects and showed good reliability and reproducibility.[34] However, these researchers used not strictly defined measurement surface areas and, thereby, introduced additional error compared with the present study that used surface areas of predetermined size (100 mesh points).

Although the method we present in this study has very good reproducibility and reliability, soft tissue surfaces generated by CBCT scans do not include texture and color. This reduces the diagnostic value of these 3D images compared with the ones created through stereophotogrammetry. Therefore, attempts have been made to register 3D images taken with a 3D surface scanner on their corresponding CBCTs, which showed promising results.[35] This could certainly improve the diagnostic efficiency and expand the use of cone-beam technology in clinical practice.

Furthermore, acquiring soft tissue surface information from CBCT scans and being able to perform 3D superimpositions in the anterior cranial base require a large field of view and, thus, subject patients to higher radiation dosages.[36] To avoid large field of view scans, alternative areas of superimposition have been proposed on computed tomography scans but require further validation for use on CBCTs.[19]

In regard to the effect of the segmentation process on the superimposition outcomes, the areas mostly affected were the inner canthi, the nasal region (alar base curves and nostrils) and the stomion. This could be explained by small differences in soft tissue posturing during image acquisition or by the fine anatomy of these structures that might not be adequately captured by the imaging technique. However, this is not expected to have affected the results because we did not perform any measurements on these areas.

The effect of other factors that could have affected the superimposition outcomes, such as image quality or voxel size, was not assessed in this investigation because all images were acquired with the same machine, under the same settings. However, the used images correspond to regular image quality used for orthodontic diagnosis.

CONCLUSIONS

The voxel-based Dolphin 3D superimposition method can be used reliably to study 3D soft tissue changes on a growing population. The superimposition error was less than 0.5 mm in most cases, and the error resulting from the segmentation process was minimal. Therefore, the method is indicated for clinical use and might be a significant aid, especially in the management of complex subjects, for whom CBCT data are usually required.

CREDIT AUTHOR STATEMENT

Georgios Kanavakis was responsible for conceptualization, validation, investigation, writing-original draft, writing-review and editing, and visualization. Simeon T. Häner was responsible for validation, investigation, resources, writing-review and editing, and visualization. François Matthey was responsible for validation, resources, writing-review and editing, and visualization. Nikolaos Gkantidis was responsible for conceptualization, methodology, validation, formal analysis, writing-review and editing, visualization, and supervision.

SUPPLEMENTARY DATA

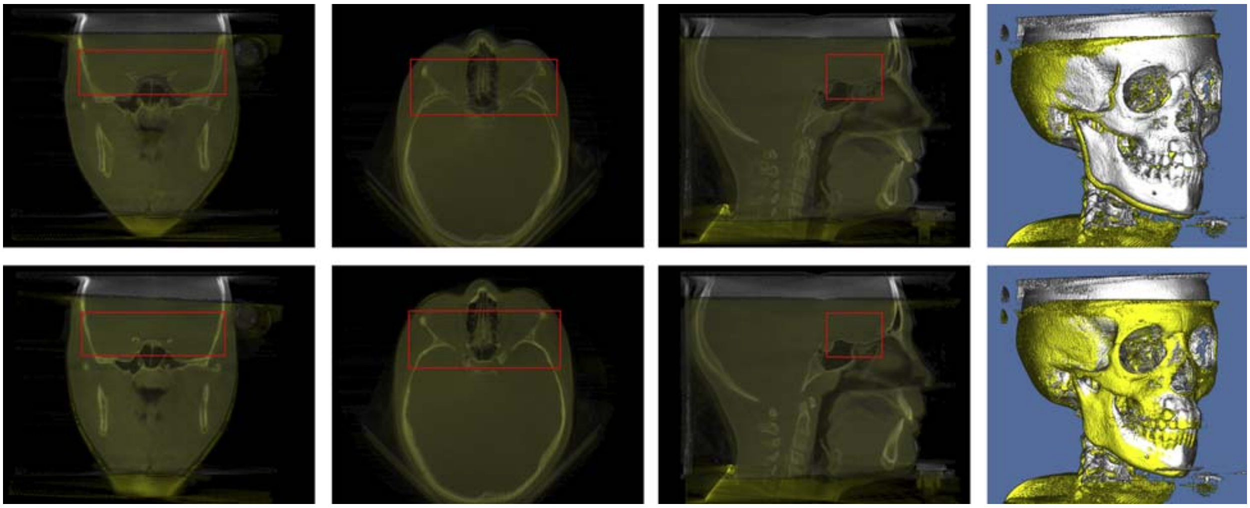
Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ajodo.2020.04.022.

REFERENCES

1. Lin F, Ren M, Yao L, He Y, Guo J, Ye Q. Psychosocial impact of dental esthetics regulates motivation to seek orthodontic treatment. Am J Orthod Dentofacial Orthop 2016;150:476-82.
2. Wedrychowska-Szulc B, Syryńska M. Patient and parent motivation for orthodontic treatment–a questionnaire study. Eur J Orthod 2010;32:447-52.
3. Pabari S, Moles DR, Cunningham SJ. Assessment of motivation and psychological characteristics of adult orthodontic patients. Am J Orthod Dentofacial Orthop 2011;140:e263-72.
4. Little AC, Jones BC, DeBruine LM. Facial attractiveness: evolutionary based research. Philos Trans R Soc Lond B Biol Sci 2011;366:1638-59.

5. Bale C, Archer J. Self-perceived attractiveness, romantic desirability and self-esteem: a mating sociometer perspective. Evol Psychol 2013;11:68-84.
6. Gangestad SW, Scheyd GJ. The evolution of human physical attractiveness. Annu Rev Anthropol 2005;34:523-48.
7. Oikawa H, Sugiura M, Sekiguchi A, Tsukiura T, Miyauchi CM, Hashimoto T, et al. Self-face evaluation and self-esteem in young females: an fMRI study using contrast effect. Neuroimage 2012;59:3668-76.
8. Arnett GW, Bergman RT. Facial keys to orthodontic diagnosis and treatment planning–part II. Am J Orthod Dentofacial Orthop 1993;103:395-411.
9. Holdaway RA. A soft-tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. Am J Orthod 1983;84:1-28.
10. Brons S, van Beusichem ME, Bronkhorst EM, Draaisma JM, Bergé SJ, Schols JG, et al. Methods to quantify soft tissue-based cranial growth and treatment outcomes in children: a systematic review. PLoS One 2014;9:e89602.
11. Halazonetis DJ. From 2-dimensional cephalograms to 3-dimensional computed tomography scans. Am J Orthod Dentofacial Orthop 2005;127:627-37.
12. Halazonetis DJ. Morphometrics for cephalometric diagnosis. Am J Orthod Dentofacial Orthop 2004;125:571-81.
13. Alkadhi RM, Finkelman MD, Trotman CA, Kanavakis G. The role of lip thickness in upper lip response to sagittal change of incisor position. Orthod Craniofac Res 2019;22:53-7.
14. Brock RA 2nd, Taylor RW, Buschang PH, Behrents RG. Ethnic differences in upper lip response to incisor retraction. Am J Orthod Dentofacial Orthop 2005;127:683-91: quiz 755.
15. Lo LJ, Weng JL, Ho CT, Lin HH. Three-dimensional region-based study on the relationship between soft and hard tissue changes after orthognathic surgery in patients with prognathism. PLoS One 2018;13:e0200589.
16. Ponce-Garcia C, Lagravere-Vich M, Cevidanes LHS, de Olivera Ruellas AC, Carey J, Flores-Mir C. Reliability of three-dimensional anterior cranial base superimposition methods for assessment of overall hard tissue changes: a systematic review. Angle Orthod 2018;88:233-45.
17. Afrand M, Ling CP, Khosrotehrani S, Flores-Mir C, Lagravère-Vich MO. Anterior cranial-base time-related changes: a systematic review. Am J Orthod Dentofacial Orthop 2014;146:21-32.e6.
18. Sgouros S, Natarajan K, Hockley AD, Goldin JH, Wake M. Skull base growth in childhood. Pediatr Neurosurg 1999;31:259-68.
19. Gkantidis N, Schauseil M, Pazera P, Zorkun B, Katsaros C, Ludwig B. Evaluation of 3-dimensional superimposition techniques on various skeletal structures of the head using surface models. PLoS One 2015;10:e0118810.
20. Becker K, Wilmes B, Grandjean C, Drescher D. Impact of manual control point selection accuracy on automated surface matching of digital dental models. Clin Oral Investig 2018;22:801-10.
21. Chang YB, Xia JJ, Yuan P, Kuo TH, Xiong Z, Gateno J, et al. 3D segmentation of maxilla in cone-beam computed tomography imaging using base invariant wavelet active shape model on customized two-manifold topology. J Xray Sci Technol 2013;21:251-82.
22. Spin-Neto R, Gotfredsen E, Wenzel A. Variation in voxel value distribution and effect of time between exposures in six CBCT units. Dentomaxillofac Radiol 2014;43:20130376.
23. Bazina M, Cevidanes L, Ruellas A, Valiathan M, Quereshy F, Syed A, et al. Precision and reliability of Dolphin 3-dimensional voxel-based superimposition. Am J Orthod Dentofacial Orthop 2018;153:599-606.
24. Nada RM, Maal TJ, Breuning KH, Bergé SJ, Mostafa YA, Kuijpers-Jagtman AM. Accuracy and reproducibility of voxel based superimposition of cone beam computed tomography models on the anterior cranial base and the zygomatic arches. PLoS One 2011;6:e16520.
25. Häner ST, Kanavakis G, Matthey F, Gkantidis N. Voxel-based superimposition of serial craniofacial CBCTs: reliability, reproducibility and segmentation effect on hard-tissue outcomes. Orthod Craniofac Res 2020;23:92-101.
26. Rottke D, Patzelt S, Poxleitner P, Schulze D. Effective dose span of ten different cone beam CT devices. Dentomaxillofac Radiol 2013;42:20120417.
27. Vasilakos G, Koniaris A, Wolf M, Halazonetis D, Gkantidis N. Early anterior crossbite correction through posterior bite opening: a 3D superimposition prospective cohort study. Eur J Orthod 2018;40:364-71.
28. Vasilakos G, Schilling R, Halazonetis D, Gkantidis N. Assessment of different techniques for 3D superimposition of serial digital maxillary dental casts on palatal structures. Sci Rep 2017;7:5838.
29. Henninger E, Vasilakos G, Halazonetis D, Gkantidis N. The effect of regular dental cast artifacts on the 3D superimposition of serial digital maxillary dental models. Sci Rep 2019;9:10501.
30. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17:571-82.
31. Baumrind S, Miller D, Molthen R. The reliability of head film measurements. 3. Tracing superimposition. Am J Orthod 1976;70:617-44.
32. 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 2019;30:525-8.
33. Baysal A, Sahan AO, Ozturk MA, Uysal T. Reproducibility and reliability of three-dimensional soft tissue landmark identification using three-dimensional stereophotogrammetry. Angle Orthod 2016;86:1004-9.
34. Brons S, Meulstee JW, Loonen TGJ, Nada RM, Kuijpers MAR, Bronkhorst EM, et al. Three-dimensional facial development of children with unilateral cleft lip and palate during the first year of life in comparison with normative average faces. PeerJ 2019;7:e7302.
35. Jayaratne YSN, McGrath CPJ, Zwahlen RA. How accurate are the fusion of cone-beam CT and 3-D stereophotographic images? PLoS One 2012;7:e49585.
36. Miracle AC, Mukherji SK. Conebeam CT of the head and neck, part 1: physical principles. AJNR Am J Neuroradiol 2009;30:1088-95.

SUPPLEMENTARY DATA



Supplementary Fig 1. Voxel-based superimposition on the anterior cranial base in Dolphin 3D Software. The *red frame* in the cranial, axial, and sagittal views represents the selected superimposition reference area, in this subject, the anterior cranial base. The frame extends from the posterior wall of the sinus frontalis (anteriorly) to the middle of sella turcica (posteriorly). The height of the box is approximately 3.5 cm, with its lower limit positioned 2-4 mm inferiorly to the lowest point of the sella turcica. Its lateral limits were extended to the lateral cranial walls. Before (*top*) and after (*bottom*) images depict the repositioning of the CBCT T0 image according to the position of CBCT T1.

Supplementary Fig 2. The colored areas represent the 7 areas (N-point, A-point, pogonion, zygoma right and left, and gonial right and left) selected for measuring the mean absolute distances between the surface models and comprising 100 mesh points each.