Measurement Error and Reliability of three available 3D Superimposition Methods in Growing Patients

by

Cecilia Ponce-Garcia

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Medical Sciences – Orthodontics
University of Alberta

ABSTRACT

Introduction: Cone-Beam Computed Tomography (CBCT) images can be superimposed, allowing a three-dimensional (3D) evaluation of growth changes and treatment effects. One of the main challenges of 3D superimposition of serial images is to understand that linear/angular measurements in two-dimensional (2D) and 3D images are not directly applicable, due to differences in size, shape and relative spatial location of skeletal, dental and soft tissue between the two imaging systems. Most of the limitations of 3D superimposition techniques are related to imaging and landmark identification errors and software/hardware related errors.

Objective: The aims of this research are to determine and compare the intra-rater reliability generated by three 3D cephalometric superimposition methods, and to determine and compare the changes observed in treated cases by the three methods over an average period of approximately 24 months using CBCT images acquired from an iCAT machine.

Methods: Thirty-six growing individuals (11-14 years old) were selected from patients that received orthodontic treatment at the University of Alberta. Before and after treatment (on average 24 months apart) CBCTs were analyzed using three superimposition methods (two voxel-based and one landmark-base). The superimposed scans with the two voxel-based methods were then used to construct surface models and quantify differences using the open-source SlicerCMF software, while the distances in the landmark-derived method were calculated manually through a formula using Excel. 3D linear and angular measurements computed with each method were then compared.

Results: Repeated measurements with each method separately presented good to excellent

intraclass correlation coefficient (ICC). Reproducibility of the quantitative assessments among the three methods had a smaller agreement with a wide range of confidence interval.

Conclusions: The findings of the research indicate good to excellent intra-examiner reliability of the three 3D superimposition methods when assessed individually. However, when assessing reliability among the three methods, the ICC demonstrated less powerful agreement with a wide range of confidence interval. ICC values were the lowest when compared the landmark-based method and the voxel-based (CMFreg/Slicer and Dolphin) methods. Moderate to excellent agreement was observed for the intra-examiner reliability when compared the voxel-based methods against each other. The landmark-based method generated the highest measurement error among the three methods.

PREFACE

This thesis is an original work by Cecilia Ponce-Garcia. Ethics approval was previously obtained by the Institutional Health Research Ethics Board at the University of Alberta, No. Pro00013379.

Chapter 2 of the thesis has been published as Cecilia Ponce-Garcia; Manuel Lagravere-Vich; Lucia Helena Soares Cevidanes; Antonio Carlos de Olivera Ruellas; Jason Carey; Carlos Flores-Mir. Reliability of three-dimensional anterior cranial base superimposition methods for assessment of overall hard tissue changes: A systematic review. *Angle Orthod. 2018; 88:233–245*. I was responsible for the article collection and analysis along with the manuscript composition. M. Lagravere assisted with the article review as the second reviewer, contributed to manuscript edits, and assisted with manuscript edits and conceptualization. L. Cevidanes, A. Ruellas and J. Carey assisted with manuscript edits. Flores-Mir was the supervisor of this chapter and was involved in the conceptualization and manuscript editing.

The data collection and statistical analysis in chapters 3 and 4 are my original collaborative work, as well as the general introduction in chapter 1 and the concluding analysis in chapter 5.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Manuel Lagravere for the continuous guidance and support throughout the past 3 years. His knowledge and accessibility helped me in all the time of research and writing of this thesis and made the whole experience easier.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Carlos Flores-Mir, Dr. Lucia Cevidanes and Dr. Jason Carey, for their insightful comments and perspective while serving on my supervising committee.

My sincere thanks also go to Dr. Antonio Ruellas for his guidance and training of all open-source image analysis procedures for orientation, segmentation, voxel-based registration, visualization of 3D semi-transparencies and color-coded maps, the quantitative approaches for comparison of all 3 methods. His precious time and expertise has been invaluable for this research. Without his unconditional support it would not be possible to conduct this project.

Finally, I must express my very profound gratitude to my parents and to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

General Introduction

1.1 Background

Monitoring treatment progress and outcomes is pivotal to patient care.[1] Therefore, an important part of orthodontic treatment involves the study of longitudinal changes induced by growth and treatment in the dentofacial complex.[2, 3] Superimposing tracings of serial lateral cephalograms has facilitated knowledge about normal craniofacial growth and development as well as knowledge about the treatment effects produced by various orthodontic, orthopedic, and surgical procedures.[2, 4] A reference system is required for a superimposition to be able to determine exactly what changes occurred. Such references must be consistently visible in the cephalograms of the individual, and they must be stable within the time frame of the observation period.[2, 5]

Several studies [6-12] have proposed the use of the anterior cranial base since there is little or no growth after 7-8 years of age when the spheno-ethmoidal synchondrosis ceases to grow. After that time a number of structures especially those associated with neural tissues remain stable and can be relied upon for superimposition.[1]

The most important limitation of conventional cephalometric superimposition is that three-dimensional (3D) changes are measured in two dimensions (2D).

Since its introduction into dentistry in 1998, Cone-Beam Computed Tomography (CBCT) has been increasingly utilized for orthodontic diagnosis, treatment planning, and research to overcome the limitations of 2D radiographic views.[13, 14] When computed tomography was initially introduced into the dental field, it was not preferred for orthodontic diagnosis due to the high ionizing radiation dose, especially when considering the young

populations that usually receive orthodontic treatment in whom cellular growth and organ development is associated with multiplied radiosensitivity of tissues.[15] The technology has been evolving ever since, resulting in a reduction in ionizing radiation dose and a continuous decrease in the cost of the CBCT systems, so they have become more widely used to visualize and assess the craniofacial complex in 3D.[16]

The introduction of CBCT also offered new insights into the changes induced by normal growth, and orthodontic treatment. It provides improved data of growth in the three planes of space of the entire skull and the maxillofacial complex.[2, 17] Currently, three overall approaches have been published and used for analyzing 3D craniofacial anatomy and changes due to growth and treatment outcomes: (1) voxel-based, (2) landmark-based, and (3) surface-based.[5, 14, 18-24]

Most of the limitations of these 3D superimposition techniques are related to imaging and landmark identification flaws and software/hardware related errors. In addition, most of the methods that are currently being used in clinical settings are time-consuming. Thus, the establishment of a reliable and simple system to analyze the craniofacial complex and provide clinicians with new possibilities in determining the structural changes produced by growth and treatment in children and adolescents is needed.

1.2 Study Objectives

The main objectives of this thesis are to:

1. Determine and compare the intra-rater reliability generated by three 3D cephalometric superimposition methods over an average period of 24 months using CBCT images acquired from a specific scan, the iCAT machine.

2. Determine and compare the changes observed by three 3D cephalometric superimposition methods over an average period of 24 months using CBCT images.

1.3 References

1. American Board of Orthodontics 2D Cranial Base Superimposition. https://www.americanboardortho.com. Accessed on March 10, 2018.

2. Duterloo H, Planché P. Handbook of cephalometric superimposition Hanover Park, IL : Quintessence Pub., c2011.; 2011.

3. De Clerck HJ, Nguyen T, Koerich L, Cevidanes L. Three-dimensional assessment of mandibular and glenoid fossa changes after bone-anchored Class III intermaxillary traction. Am J Orthod Dentofacial Orthop 2012;142:25-31.

4. Gu Yan, Jr MJ. Cephalometric Superimpositions. A Comparison of Anatomical and Metallic Implant Methods. Angle Orthodontist 2008;78(6):967-76.

5. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

6. Björk A. The Use of Metallic Implants in the Study of Facial Growth in Children : Method and Application. Am. J. Phys. Anthrop 1968;29:243-54.

7. Melsen B. The cranial base. Acta Odontol Scand 1974;32 (Suppl 62):86-101.

8. Melsen B, Melsen F. The postnatal development of the palatomaxillary region studied on human autopsy material. Am J Orthod 1982;82:329-42.

9. Cevidanes LH, Motta A, Proffit WR, Ackerman JL, Stynere M. Cranial base superimposition for 3-dimensional evaluation of soft-tissue changes. AJODO 2010;137(4 Suppl):S120-S29.

10. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.

11. Ghafari J, Engel FE, Laster LL. Cephalometric superimposition on the cranial base: a review and a comparison of four methods. Am J Orthod Dentofacial Orthop 1987;91(5):403-13.

12. Afrand M. Anterior and middle cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2015.

13. Kapila S, Conley R, Harrell Jr W. The current status of cone beam computed tomography imaging in orthodontics. Dentomaxillofacial Radiology 2011;40:24-34.

14. Kapila S, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. Dentomaxillofacial Radiology 2015;44(1):20140282.

15. Cevidanes L, Benavides Erika, Ludlow J, Ruellas A. Orthodontic diagnosis and treatment planning with cone-beam computed tomography imaging. In: Graber, editor. Orthodontics Current Principles and Techniques. USA: Elsevier; 2017.

16. Kaygısız E, Tortop T. Cone Beam Computed Tomography in Orthodontics. Intechopen 2017;DOI: 10.5772/intechopen.68555.

17. Lenza MA, Carvalho AA, Lenza EB, et al. Radiographic evaluation of orthodontic treatment by means of four different cephalometric superimposition methods. Dental Press J. Orthod 2015;20(3):29-36.

18. Park JH, Tai K, Owtad P. 3-Dimensional Cone-Beam Computed Tomography Superimposition: A review. Seminars in Orthodontics 2015;21(4):263-73.

19. Cevidanes LH, Bailey L'TJ, Tucker SF, Styner MA, et al. Three-dimensional cone-beam computed tomography for assessment of mandibular changes after orthognathic surgery. AJODO 2007;131:44-50.

20. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

21. Weissheimer A, Menezes L, Koerich L, Pham J, Cevidanes L. Fast three-dimensional superimposition of cone beam computed tomography for orthopaedics and orthognathic surgery evaluation. Int. J. Oral Maxillofac. Surg. 2015;44:1188-96.

22. Lagravère M, Secanell M, Major P, Carey J. Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;111:771-77.

23. Gkantidis N, Schauseil M, Pazera P, Zorkun B, et al. Evaluation of 3-Dimensional Superimposition Techniques on Various Skeletal Structures of the Head Using Surface Models. PLoS ONE 2015;10(2):1-20.

24. Almukhtar A, Ju X, Khambay B, McDonald J, Ayoub A. Comparison of the accuracy of voxel based registration and surface based registration for 3D assessment of surgical change following orthognathic surgery. PLoS ONE 2014;9(4):e93402.

CHAPTER 2

Reliability of Three-Dimensional Anterior Cranial Base Superimposition Methods for Assessment of Overall Hard Tissue Changes: A Systematic Review

Cecilia Ponce-Garcia; Manuel Lagravère-Vich; Lucia Helena Soares Cevidanes; Antonio Carlos de Olivera Ruellas; Jason Carey; Carlos Flores-Mir. Reliability of Three-Dimensional Anterior Cranial Base Superimposition Methods for Assessment of Overall Hard Tissue Changes: A Systematic Review. Angle Orthod. 2018; 88:233–245

2.1 Introduction

Superimposition of cephalometric headfilms taken at defined intervals is used by researchers and clinicians to help in orthodontic diagnosis, treatment planning and obtaining a general view of growth changes and treatment outcomes in the dentofacial complex.[20, 24, 25]

Conventional lateral cephalometric radiographs have proven to be an invaluable part of initial and final orthodontic records to quantify and determine craniofacial growth changes and effects of orthodontic treatment.[17, 24] However, two-dimensional (2D) cephalometric radiographs suffer from a number of inherent flaws, such as errors generated because of inadequate patient's head position, alignment of imaging device, inherent geometric distortions, and differential magnification created by projection distance and beam divergence.[5, 13, 24, 26-28]

Over the past decade, craniofacial three-dimensional (3D) digital records have become increasingly popular among orthodontists as the specialty progresses towards a 3D virtual representation of the patient for diagnosis, treatment planning, and surgical simulation. The advanced imaging capabilities of Cone-Beam Computed Tomography (CBCT) are depicted through 3D cephalometric analysis, temporomandibular joint visualization, and 3D evaluation of dental anomalies, to name only a few.[17, 29] A single scan provides an overlap-free 3D visualization of different components of the skull, enables volumetric measurements and also allows detailed assessment of the maxillofacial structures in variable thickness of axial, coronal and sagittal slices, providing real measurements with no magnification.[13, 30]

Nowadays, similar to 2D cephalometric tracings, CBCT images can be superimposed, allowing a 3D evaluation of growth changes, treatment effects, and stability over a certain time

interval, through registration points, angles, shapes, and volumes.[31-33] One of the main challenges of 3D superimposition of serial images is to understand that linear/angular measurements in 2D and 3D images are not directly comparable, due to differences in size, shape and relative spatial location of skeletal, dental and soft tissue between the two imaging systems.[17, 34]

Three general methods of 3D cephalometric superimposition have been published and used for clinical diagnosis and assessment of orthodontic treatment outcomes: (1) voxel-based,[9, 17-20, 35] (2) landmark-based,[5, 21] and (3) surface-based.[17, 22]

A review addressing the 3D CBCT superimposition methods was published in 2015.[17] Although it discusses the three main techniques, it focuses mainly on their clinical applications, benefits, and limitations. It did not consider the measuring capabilities of any of those methods. To our knowledge, no systematic review has been specifically conducted to investigate the reliability of these 3D superimposition methods when assessing changes in craniofacial hard tissues. Without an in-depth understanding of the measurement properties of each method, its indiscriminate use should be questioned, as treatment decisions/assessments would not have been based on sound superimposition evidence.

The purpose of this systematic review was to synthesize the available literature concerning the reliability of 3D superimposition methods to evaluate craniofacial hard tissues changes.

2.2 Materials and Methods

This systematic review followed, whenever applicable, the Preferred Reporting Items for Systematic Reviews and Meta-Analysis checklist.[36]

Protocol and Registration

The study protocol was not registered in advance.

2.2.1 Eligibility criteria

The following selection criteria were applied for the review.

- Study design: clinical trials, cohort, case-control and cross-sectional studies that evaluated the reliability, repeatability or reproducibility of 3D superimposition methods on the anterior cranial base were included. No restrictions were applied regarding language or year of publication.

- Exclusion Criteria: review articles, meeting abstracts, book chapters, case reports, editorial letters and personal opinions were ruled out from the review.

Information sources and search strategy

A systematic search of four electronic databases (Embase, Medline via OVID, Web of Science and SCOPUS) was carried. All searches were inclusive until December 2016. The search strategy was designed with the assistance of a health science senior librarian. We selected appropriate truncation and word combination and adjusted them for each database search. Keywords used in the search and combination of terms per database can be found in Table 2.1.

The EndNote Basic software (Thompson Reuters, New York, NY) was used to manage the references, and RefWorks software was used to remove duplicates. No restrictions were applied to electronic database searches.

Table 2.1. Database Search Strategy

Database	Search Strategy
WEB OF SCIENCE (1945-August, 2016)	(#1 TOPIC: cone beam computed tomography OR #2 TOPIC: CBCT) AND ((#3 TOPIC: three-dimensional cephalometric superimposition OR #4 TOPIC: three dimensional cephalometric superimposition OR #5 TOPIC: 3-dimensional cephalometric superimposition OR #6 TOPIC: 3D cephalometric superimposition) OR (#7 TOPIC: three-dimensional superimposition OR #8 TOPIC: three dimensional superimposition OR #9 TOPIC: 3-dimensional superimposition OR #10 TOPIC: 3D superimposition) OR (#11 TOPIC: craniofacial three-dimensional superimposition OR #12 TOPIC: craniofacial three dimensional superimposition OR #13 TOPIC: craniofacial 3-dimensional superimposition OR #14 TOPIC: craniofacial 3D superimposition) OR (#15 TOPIC: three-dimensional superimposition methods OR #16 TOPIC: three dimensional superimposition methods OR #17 TOPIC: 3-dimensional superimposition methods OR #18 TOPIC: 3D superimposition methods) OR (#19 TOPIC: three-dimensional cephalometric superimposition methods OR #20 TOPIC: three dimensional cephalometric superimposition methods OR #21 TOPIC: 3-dimensional cephalometric superimposition methods OR #22 TOPIC: 3D cephalometric superimposition methods))
EMBASE (1974-August, 2016)	(cone beam computer tomography.mp. OR craniofacial three dimensional imaging.mp. OR CBCT) AND (3D cephalometric superimposition.mp. OR three-dimensional cephalometric superimposition.mp. OR three dimensional cephalometric superimposition.mp. OR three-dimensional superimposition.mp. OR three dimensional superimposition.mp. 3D superimposition OR craniofacial superimposition.mp.)
MEDLINE (1966-August, 2016)	(cone beam computer tomography.mp. OR craniofacial three dimensional imaging.mp. OR CBCT) AND (3D cephalometric superimposition.mp. OR three-dimensional cephalometric superimposition.mp. OR three dimensional cephalometric superimposition.mp. OR three-dimensional superimposition.mp. OR three dimensional superimposition.mp. 3D superimposition OR craniofacial superimposition.mp.)
SCOPUS (Until August, 2016)	((TITLE-ABS-KEY (cone beam computed tomography)) OR (TITLE-ABS-KEY (cbct))) AND (((TITLE-ABS-KEY (3d cephalometric superimposition)) OR (TITLE-ABS-KEY (three dimensional cephalometric superimposition)) OR (TITLE-ABS-KEY (three-dimensional cephalometric superimposition))) OR ((TITLE-ABS-KEY (3d superimposition)) OR (TITLE-ABS-KEY (three dimensional superimposition)) OR (TITLE-ABS-KEY (three-dimensional superimposition))) OR ((TITLE-ABS-KEY (three-dimensional superimposition methods)) OR (TITLE-ABS-KEY (three-dimensional cephalometric superimposition methods)) OR (TITLE-ABS-KEY (three dimensional cephalometric superimposition methods)) OR (TITLE-ABS-KEY (3d cephalometric superimposition methods)) OR (TITLE-ABS-KEY (3d superimposition methods)) OR (TITLE-ABS-KEY (three dimensional superimposition methods))) OR ((TITLE-ABS-KEY (craniofacial three dimensional superimposition)) OR (TITLE-ABS-KEY (craniofacial three-dimensional superimposition)) OR (TITLE-ABS-KEY (craniofacial 3d superimposition))))
Limitation: human subjects.	

2.2.2 Study Selection

The relevant articles were selected through a two-phase process. In Phase 1, two authors (CPG and ML) independently reviewed the titles and abstracts of all references. If there were more than one publication from the same study, the one with the most comprehensive information was selected. In case the abstract was judged to contain insufficient information for a final decision of inclusion or exclusion, the full article was obtained and reviewed before a decision was made. Additionally, a hand search was also conducted over the reference lists of

the included publications for potentially important studies that could have been missed through the electronic search, and a limited gray literature search of the first two hundred results was performed using Google Scholar.

In Phase 2, full texts of potentially relevant abstracts were retrieved, reviewed and screened by the same two reviewers according to the same selection criteria to confirm once again while considering the full manuscript the final selection. Any disagreement was settled by means of discussion until a mutual consensus was reached.

2.2.3 Data collection process and data items

Data were extracted from each of the selected studies including authors, year of publication, aim of the study, study design, sample size, superimposition method, statistical analysis as related to reliability, repeatability or reproducibility and outcomes using a developed standardized data collection form based on the Cochrane Consumers and Communication Review.[37] One reviewer (CPG) collected the required information from the selected articles. The second reviewer (ML) crosschecked the gathered data and confirmed its accuracy. Once again, any disagreement in either phase was resolved by consensus.

2.2.4 Risk of bias in individual studies

The Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist[38] a standardized tool for assessing the methodological quality of studies that evaluates measurement properties was used for quality assessment of included studies. This checklist contains nine boxes, each one dealing with a determined measurement property. Each box contains 5–18 items that are used to assess whether a study on a specific measurement

property meets the requirements for good methodological quality. The boxes used in this current review were reliability, measurement error, criterion validity, and the one assessing the generalizability of the study results on measurement properties. Each measurement property was evaluated separately. The scoring system determined an overall methodological quality score per measurement property for every included study. Four response options for each item of the COSMIN checklist were used as reported by Terwee et al,[39] representing excellent, good, fair, and poor methodological quality. Subsequently, a methodological quality score per box was obtained by taking the lowest rating of any item in a box (‘‘worst score counts’’). A poor score on any item was thus considered to represent a fatal flaw. In the scoring system, items 1 and 2 (on the number of missing items and how missing items are handled) were scored less strictly than the other items as this information is often not reported in articles. In all boxes, a small sample size was considered poor methodological quality. Disagreements between the reviewers in relation to quality assessment were resolved by means of discussion, and the third reviewer (CFM) made a final decision if consensus was not reached by the first 2 reviewers.

2.2.5 Synthesis of studies

Due to the nature of the question and the available data, a meta-analysis was not possible. Included studies assessed reliability of measurements form different craniofacial anatomical regions.

2.3 Results

2.3.1 Study Selection

A flow chart of the selection process of articles included in this study is outlined in Figure 2.1. A total of 254 manuscripts were selected for a Phase 1 assessment. Thereafter a total of 219 studies were excluded following abstract/title assessment. Only 35 references and 1 additional study found through a manual search (from the reference list) were subsequently selected and received full-text reading (Phase 2). From the total full-text articles retrieved and reviewed, 30 studies were later excluded; therefore, only 6 studies fulfilled the criteria to be included in this review.

2.3.2 Study characteristics

All the selected studies for the qualitative synthesis were published in English. Sample size ranged from 3 to 18 subjects and included adults undergoing orthognathic surgery and children undergoing either orthopedic treatment with miniplates for class III correction or maxillary expansion (RME). All of the studies used pre and post images to perform the 3D superimposition method to assess treatment or growth changes. Four imaging systems were used, iCAT 3D (3 out of 6 studies),[19, 20, 31] NewTom 9000,[40] NewTom 3G[5] and Philips Brilliance 16 CT.[22] Large field of view was used in all of the studies and voxel size ranged from 0.25 to 0.8. For voxel-based registration methods Insight SNAP, MIRIT, VALMET, ITK-Snap, CFM, MAXILIM, and OnDemand 3D software systems were used, on the other hand, AVIZO and OSIRIX were used for landmark-based and surface-based methods respectively. Four out of six studies tested the voxel-based registration;[19, 20, 31, 40] one assessed the landmark-based method,[5] and one assessed the surface-based method.[22] An abbreviated summary of the descriptive characteristics of the included articles is provided in Table 2.2.

2.3.3 Risk of bias within studies

The methodological quality scores based on COSMIN checklist when assessing reliability were evaluated to have high risk of bias (poor methodological quality) in all included studies. Similarly, when evaluating measurement error and validity, five studies scored poor methodological quality except for one study on each measurement property that obtained fair methodological quality. The critical appraisal details about each of the items and the evaluation criteria are described in Tables 2.3-2.6

The main methodological limitations associated with the high risk of bias in the included studies when assessing reliability, measurement error and validity were related to sample size, critical flaws in the study design and the lack of detailed statistical methods. None of the studies stated the power analysis and/or the justification of the sample size, also leading to a high risk of bias.

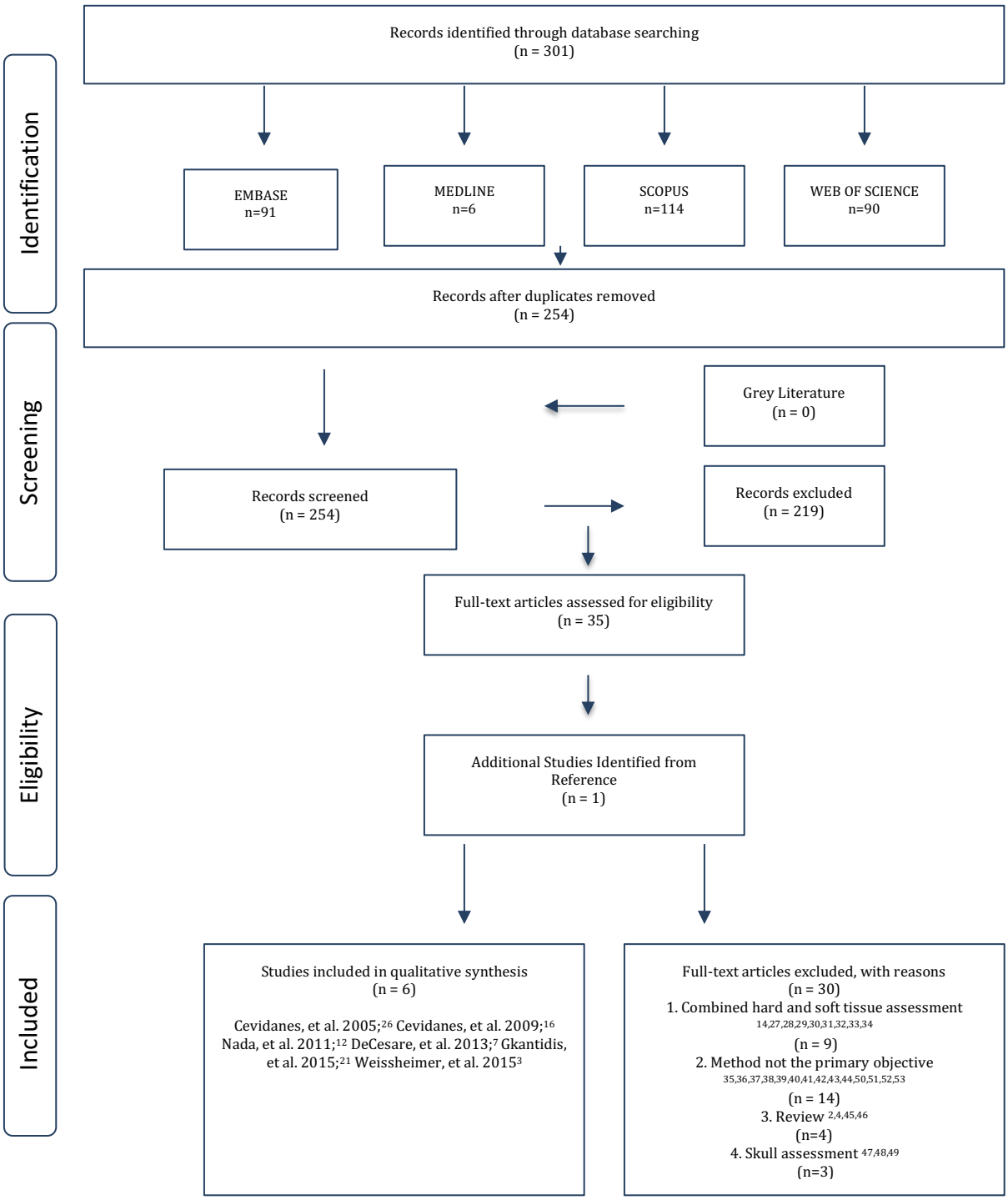


Fig 2.1. Flow diagram of data search according to PRISMA

Table 2.2 Summary of descriptive characteristics of finally selected studies

Author/ Year	Aim of Study	Participants Demographics and Characteristics		Method Details						Main Results
		Age	**Sample Size/Study population/Time between Scan**	**Imaging system used to acquired Scan**	**Field of View and Voxel Size**	**Software used**	**Superimposition Method**	**Time to complete superimposition**	**Statistics Analysis**	**Outcomes (*P*-value) or (OR; 95%CI)**
Cevidanes et al 2005[26]	Evaluate the registration of 3D models from CBCT images taken before and after orthognathic surgery for the assessment of mandibular anatomy and position	20.6 ± 5.2 years	10 patients Maxillary surgery at the UNC Scans taken before and 1 week after orthognathic surgery	NewTom 9000	FOV: 23 x 23cm Voxel size: 0.58 x 0.58 x 0.6mm	Insight SNAP: regional semiautomatic segmentation MIRIT: computing fully automated registration VALMET: comparison of 3D models	Voxel-based	Not mentioned	Inter-observer reliability	Mandibular rami surface: Inward (0.26 SD 0.12) Outward (0.24 SD 0.10) Posterior border of the mandibular ramus: Inward (0.15 SD 0.07) Outward (0.11 SD 0.04) Condyles: Inward (0.05 SD 0.02) Outward (0.13 SD 0.08) *P*-value <0.001
Cevidanes et al 2009[16]	Evaluate a new method for superimposition of 3D models of growing subjects	11.4 years	3 patients Orthopedic treatment with miniplates Scans taken before and after treatment (about 1 year follow-up)	iCAT 3D Imaging System	FOV: 22 x 16 cm Voxel size: 0.5mm	ITK-Snap: 3D surface models Imagine: computing the rigid registration CFM: overlaying 3D surface models	Voxel-based	Not mentioned	Not mentioned	Inter-examiner range of measurement across anatomic regions was equal or less than 0.5mm, which are clinically insignificant.
Nada et al 2011[1241]	Evaluate accuracy and reproducibility of a semi-automated voxel based image registration technique for the superimposition of 3D CBCT models on two different regions, the anterior cranial base and the zygomatic arches	26 ± 9 years	16 adult patients Two surgical interventions at Radboud University Nijmegen Medical Centre Scans taken prior to treatment and before the second orthognathic surgery - average 18 (±4.6) months later	iCAT 3D Imaging System	FOV: 22 x 16 cm Voxel size: 0.4mm	MAXILIM	Voxel-based	30-40 minutes	Intra-observer and inter-observer reliability using the Pearson correlation coefficient for the mean distances at 4 anatomical regions following the first and second superimpositions. Paired-sample t-test to compare the means of corresponding measurements following registration on the anterior cranial base and the left zygomatic arch.	The correlation coefficients registered on the anterior cranial base ranged between 0.53 and 0.94 for the mean distances at the 4 regions and between 0.24 and 0.71 for the mean distances at the 4 anatomic regions registered on the zygomatic arches The *P*-values ranged between 0.001 and 0.025 and were statistically significant for the 4 regions.
DeCesare et al 2013[7]	Minimize errors that occur when using a four vs. six-landmark superimpositioning method in the cranial base to define the co-ordinate system.	Not mentioned	10 patients Maxillary expansion clinical trial Scans were taken 12 months apart	NewTom 3G	FOV: 12” Voxel size: 0.25mm	AVIZO	Point-based	Not mentioned	Intrareliability values were determined using intra-class correlation coefficient for all four landmarks, repeating the process three times for each image.	When analyzing real patient data, it was found that the 6-point correction algorithm reduced errors between images and increased intrapoint reliability. This method demonstrated greater reliability and reproducibility than the previous 4-point correction algorithm.
Gkantidis et al 2015[21]	Test the applicability, accuracy, precision, and reproducibility of various 3D superimposition techniques for radiographic data, transformed to triangulated surface data.	16.2 (range: 15.1, 22.9) years	8 young adult patients Rapid maxillary expansion performed by a mini-implant supported device Scans acquired just before placement of the appliance and at the end of the activation period at a median of 15 days later	Philips Brilliance 16 CT Scanner	FOV: 21 x 21 x 12cm Voxel size: 0.8mm	OSIRIX	Surface-based	25 minutes	Differences in the measured variables were evaluated using permutational multivariate analysis of variance (MANOVA) with factorial mixed effects models. In all cases, a two-sided significance test was carried out at an alpha level of 0.05. The level of significance used for the study was set at 0.05. Bonferroni correction was applied for pairwise a posteriori multiple comparison tests.	Pairwise a posteriori tests between superimposition techniques showed that all techniques differed from each other (*P*<0.005). The AC + F technique was the most accurate (D<0.17 mm), as expected, followed by AC and BZ superimpositions that presented similar level of accuracy (D<0.5 mm).
Weissheimer et al 2015[3]	The aim of this study was to validate a method for fast three-dimensional (3D) superimposition of cone beam computed tomography (CBCT) in growing patients and adults (surgical cases).	11.4 ± 1 years (pre-treatment scans) 26.3 ± 5.7 years (non-growing adults) 9.5 ± 1.8 years (growing patients)	18 patients total. 10 patients, saved as a reoriented volume, and then superimposed on the original image. 4 non-growing and 4 growing Scans taken pre- and 1 year post-orthognathic surgery for adult sample and post rapid palatal expansion for growing sample	iCAT 3D Imaging System	FOV: Large Voxel size: 0.25mm	OnDemand3D: automatic voxel-based rigid registration ITK-Snap: automatic segmentation STL to SGI Inventor 2.0: convert files from STL to IV CMF app: provided closest point colour maps between registered 3D surface models	Voxel-based	10-15 seconds	Not mentioned	The quantification of the superimposition errors by colour-coded surface distances revealed that distances in the anterior cranial base between registered surface models were less than 0.5 mm for most regions for both growing patients and adults.

2.3.4 Results of individual studies

For better interpretation results were separated according to the superimposition techniques: voxel-based, landmark-based and surface-based. Pre- and post- treatment images were registered on the anterior cranial base surface on all studies[5, 19, 20, 22, 31, 40] and one was also superimposed on the left zygomatic arch.[31]

Voxel-based method:

Four studies tested this method. A first study carried out by Cevidanes et al.,[40] assessed inter-observer reproducibility in a subset of ten CBCT scans (before and after treatment) of five patients undergoing orthognathic surgery using three observers. They showed the similarity between the 3D colour-coded maps and that pre- to post- surgery surface distance measurements differed amongst the three observers by no more than 0.26 mm (maximal error measured as displacement at the mandibular rami surface). The average inward displacement for all surfaces (mandibular rami, posterior border of the mandibular ramus and condyles) was smaller than the image spatial resolution of 0.6mm. The one-sample t-test *P*-values were statistically significant at all surfaces, despite the small values of displacements that were observed.

A second study[31] assessed the voxel-based method on pairs of CBCT scans of sixteen adult patients. The mean absolute distances between the two 3D images were calculated in four different regions (cranial base, forehead, right and left zygomatic arches). Results showed small inter-observer variability when 3D models construction and superimposition procedure was repeated by a second observer. Mean differences between superimpositions performed by the first and second observer were 0.01mm (CI95% -0.03,0.05) for the forehead region, -0.07mm

(CI95% -0.13,-0.003) for the right zygomatic arch and -0.01mm (CI95% -0.09,0.07) for the left zygomatic arch. The correlation coefficient between the repeated superimpositions (Intra-observer repeatability) ranged from 0.53 – 0.94.

A third study[20] using a sample of CBCT images of eighteen patients assessed, also, the voxel-based method. Ten patients were used as a reference standard, reorienting the spatial position of the pretreatment CBCT volume, and then superimposing on the original image and eight patients (four non-growing and four growing) pre- and post-treatment superimposed images. Results showed that the surface distance error was less than 0.25 mm for the sample that tested the superimpositions of CBCT images with a 1-year interval for growing patients treated with RME. Similarly, the adult sample, which underwent orthognathic surgery, revealed that discrepancies in the anterior cranial base between the registered surface models were less than 0.5 mm for most regions.

A fourth study[19] assessed the voxel-based method in growing subjects. Three observers were trained for analysis of CBCT images using two images not included in the study. After calibration, each observer examined pre- and post- treatment CBCT scans of three growing patients. The inter-examiner range of measurements across anatomic regions was equal or less than 0.5 mm.

Landmark-based registration:

DeCesare et al.[5] assessed the six-landmark superimposition method when defining the co-ordinate system using data from ten growing patients with scans taken with 12 months apart. To ensure reproducibility, each patient's data set was corrected ten times with an algorithm, and

the average correction factors were calculated. The average error for each control point was obtained by first finding the distance between a control point and each of Auditory External Meatus Left (AEML), Auditory External Meatus Right (AEMR) and Dorsum of Foramen Magnum (DFM) in the first image for each patient. Those same distances were then found in the second image. The error reported was the absolute value of the difference in distances between points calculated for the first image and the second image. The results showed high inter-test reproducibility and great consistency between trials. The average error seen in the distances between the first image and the second uncorrected image was 1.64 mm. This average error dropped to 1.24 mm when corrected with the 6-point algorithm.

Surface-based registration:

Gkantidis et al.[22] in a sample of 8 non-growing orthodontic patients treated with rapid maxillary expansion tested five surface-based 3D superimposition techniques (3P: three-point registration; AC: anterior cranial base; AC+ F: anterior cranial base + foramen magnum; BZ: both zygomatic arches; 1Z: one zygomatic arch). Results showed that all techniques differed from each other (P<0.005), except for the AC and BZ superimpositions (P=0.43) using CT scans. The AC + F was the most accurate technique (P=0.07), followed by AC and BZ superimpositions that presented similar level of accuracy, but approximately four times reduced relatively to AC + F. 3P and 1Z superimpositions were the least accurate. The reproducibility and precision of all techniques were also acceptable since there were no significant differences between repeated measurements and among examiners on measured structural changes (AC + F: examiner 1: 0.11mm (CI95% 0.09, 0.17); examiner 2: 0.07mm (CI95% 0.04, 0.09); examiner 3: 0.09mm (CI95% 0.04, 0.14).

2.3.5 Synthesis of results

All included studies reported adequate reliability and acceptable measurement error of all 3D superimposition methods. Nevertheless, the quality of the studies was consistently poor. Hence it is unknown how the poor quality of evidence influenced the results.

2.3.6 Risk of bias across studies

The main methodological limitations across studies were related to small sample size, different age groups, treatment type, flaws in the study design and lack of a detailed description of statistical analysis. All of the identified studies used pre and post scans. However, some included children and others adults; either underwent rapid maxillary expansion, maxillary protraction or orthognathic surgery, some of them gave detailed information about the treatment and changes expected due to the procedure, others not; some evaluated inter and intra-observer reliability, others either inter-observer reliability or intra-observer reliability. Intraclass correlation coefficient (ICC) was used only in two studies and standard error of measurement (SEM) described in detailed only in one study.

2.3.7 Additional analysis

All the articles used different anatomical regions to assess reliability and measurement error; this made the application of a meta-analysis questionable.

Table 2.3. COSMIN Risk of Bias Assessment - Reliability

	Cevidanes et al[26]				*Cevidanes et al*[16]				*Nada et al*[12]				*DeCesare et al*[7]				*Gkantidis et al*[21]				*Weissheimer et al*[3]			
Box B. Reliability: relative measures (including test-retest reliability, inter-rater reliability and intra-rater reliability)	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P
Design requirements																								
Was the percentage of missing items given?			✓				✓				✓				✓				✓				✓	
Was there a description of how missing items were handled?			✓				✓				✓				✓				✓				✓	
Was the sample size included in the analysis adequate?				✓				✓		✓						✓				✓		✓		
Were at least two measurements available?	✓				✓				✓				✓				✓				✓			
Were the administrations independent?	✓				✓				✓				✓				✓							✓
Was the time interval stated?			✓				✓		✓						✓		✓				✓			
Were patients stable in the interim period on the construct to be measured?			✓				✓				✓				✓				✓				✓	
Was the time interval appropriate?			✓				✓		✓						✓		✓				✓			
Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions			✓				✓				✓				✓				✓					✓
Were there any important flaws in the design or methods of the study?				✓				✓			✓					✓		✓						✓
Statistical Methods																								
For continuous scores: Was an intra class correlation coefficient (ICC) calculated?				✓				✓	✓							✓	✓							✓
For dichotomous/nominal/ordinal scores: Was kappa calculated?				✓				✓				✓				✓				✓				✓
For ordinal scores: Was a weighted kappa calculated?				✓				✓				✓				✓				✓				✓
For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic				✓				✓				✓				✓	✓							✓
Score	***Poor***				***Poor***				***Poor***				***Poor***				***Poor***				***Poor***			

COSMIN[24] box with 4-point scale for methodological quality: E: Excellent, G: Good, F: Fair, P: Poor
A methodological quality score per box was obtained by taking the lowest rating of any item in a box (‘‘worst score counts’’). A poor score on any item was thus considered to represent a fatal flaw.[25]

Table 2.4. COSMIN Risk of Bias Assessment – Measurement Error

	Cevidanes et al[26]				*Cevidanes et al*[16]				*Nada et al*[12]				*DeCesare et al*[7]				*Gkantidis et al*[21]				*Weissheimer et al*[3]			
Box C. Measurement Error: absolute measures	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P
Design requirements																								
Was the percentage of missing items given?			✓				✓				✓				✓				✓				✓	
Was there a description of how missing items were handled?			✓				✓				✓				✓				✓				✓	
Was the sample size included in the analysis adequate?				✓				✓		✓						✓				✓		✓		
Were at least two measurements available?	✓				✓				✓				✓				✓				✓			
Were the administrations independent?	✓				✓				✓				✓				✓							✓
Was the time interval stated?			✓				✓		✓						✓		✓				✓			
Were patients stable in the interim period on the construct to be measured?			✓				✓				✓				✓				✓				✓	
Was the time interval appropriate?	✓				✓				✓				✓				✓				✓			
Were the test conditions similar for both measurements? e.g. type of administration, environment, instructions			✓				✓				✓				✓				✓					✓
Were there any important flaws in the design or methods of the study?				✓				✓			✓					✓		✓						✓
For CTT: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?			✓				✓			✓					✓		✓						✓	
Score	***Poor***				***Poor***				***Fair***				***Poor***				***Poor***				***Poor***			

COSMIN[24] box with 4-point scale for methodological quality: E: Excellent, G: Good, F: Fair, P: Poor
A methodological quality score per box was obtained by taking the lowest rating of any item in a box (‘‘worst score counts’’). A poor score on any item was thus considered to represent a fatal flaw.[25]

Table 2.5. COSMIN Risk of Bias Assessment - Validity

	Cevidanes et al[26]				*Cevidanes et al*[16]				*Nada et al*[12]				*DeCesare et al*[7]				*Gkantidis et al*[21]				*Weissheimer et al*[3]			
Box H. Criterion Validity	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P
Design requirements																								
Was the percentage of missing items given?			✓				✓				✓				✓				✓				✓	
Was there a description of how missing items were handled?			✓				✓				✓				✓				✓				✓	
Was the sample size included in the analysis adequate?				✓				✓		✓						✓				✓		✓		
Can the criterion used or employed be considered as a reasonable 'gold standard'?				✓				✓				✓				✓			✓			✓		
Were there any important flaws in the design or methods of the study?			✓				✓			✓					✓			✓					✓	
Statistical methods																								
For continuous scores: Were correlations, or the area under the receiver-operating curve calculated?			✓				✓				✓				✓				✓				✓	
For dichotomous scores: Were sensitivity and specificity determined?			✓				✓				✓				✓				✓				✓	
Score	***Poor***				***Poor***				***Poor***				***Poor***				***Poor***				***Fair***			

COSMIN[24] box with 4-point scale for methodological quality: E: Excellent, G: Good, F: Fair, P: Poor
A methodological quality score per box was obtained by taking the lowest rating of any item in a box (''worst score counts''). A poor score on any item was thus considered to represent a fatal flaw.[25]

Table 2.6 COSMIN Risk of Bias Assessment - Generalizability

	Cevidanes et al[26]				*Cevidanes et al*[16]				*Nada et al*[12]				*DeCesare et al*[7]				*Gkantidis et al*[21]				*Weissheimer et al*[3]			
Generalizability box ***Was the sample in which the HR-PRO instrument was evaluated adequately described? In terms of:***	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P	E	G	F	P
Median or mean age (with standard deviation or range)?				✓				✓	✓							✓	✓				✓			
Distribution of sex?				✓				✓				✓				✓	✓							✓
Important disease characteristics (e.g. severity, status, duration) and description of treatment?		✓				✓				✓						✓	✓					✓		
Setting(s) in which the study was conducted? e.g. general population, primary care or hospital/rehabilitation care		✓					✓			✓				✓					✓			✓		
Countries in which the study was conducted?			✓				✓		✓						✓				✓					✓
Language in which the HR-PRO instrument was evaluated?			✓				✓					✓			✓				✓					✓
Was the method used to select patients adequately described? e.g. convenience, consecutive, or random			✓				✓				✓				✓				✓				✓	
Was the percentage of missing responses (response rate) acceptable?			✓				✓				✓				✓				✓				✓	

COSMIN[24] box with 4-point scale for methodological quality: E: Excellent, G: Good, F: Fair, P: Poor
Generalizability box: used to extract data on the characteristics of the study population and sampling procedure, no scoring system used.

2.4 Discussion

2.4.1 Summary of evidence

In this systematic review, the available evidence concerning the reliability of the 3D superimposition methods when assessing changes in craniofacial hard tissues was investigated. Although, all included studies in this review reported acceptable reliability of all three methods, the quality of evidence was low. Therefore, any reported conclusions are not to be supported with a high level of certainty.

CBCT is nowadays a well-established diagnostic tool for the 3D assessment of growth and/or treatment changes on craniofacial structures. However, it is important to understand the challenges, since 3D superimposition is much more complicated than 2D superimposition. The difficulties assessing the reliability of 3D superimpositions are not only due to registration issues but also due to the choice of regions to test the reproducibility of the superimposition, with landmark locations in various anatomic surfaces in the three planes of space.[42]

To be suitable for routine application in medical image processing a superimposition method should be able to precisely register and aid understanding of the changes due to growth or treatment relative to the structures of reference. The image analysis procedures include 3D construction, registration, superimposition, and quantification of changes.

In the voxel-based registration method, all these steps are automated, which may allow image analysis procedures independent of observer errors. The application of this method has

been widely described in the literature to assess changes after orthognathic surgery and orthopedic treatment.[9, 25, 43-45]

Cevidanes et al.[40] introduced this method into dentistry. This first study used this method to assess mandibular anatomy and position before and after maxillary advancement. They applied distance measurement to quantify mandibular rotations and displacement. Results showed that the inter-observer errors had a range of 0.26 mm. Similarly, Nada et al.[31] used the voxel-based image registration method to test the reliability and measurement error of CBCT superimposition on the anterior cranial base in adult patients who underwent combined surgical orthodontic treatment. Authors reported small differences within 0.5mm, which was considered to be clinically insignificant. It was also mentioned that the registration of the superimposed scans on the zygomatic arch could be contemplated as an alternative to the anterior cranial base when using smaller FOV scans, in non-growing patients. However, it is important to be aware that the regions used for quantification of error in this study were closer to the region of reference reducing the magnitude of error. It is known that the further the region of interest is in regards to the superimposition structures the larger the theoretical error of measurement. Weissheimer et al.,[20] also tested the voxel-based method, and their results revealed that distances in the anterior cranial base between registered surface scans were <0.5 mm for most regions, indicating reliable superimposition. Nevertheless, the statistical analysis was not reported, whence, their conclusion should be taken warily.

When comparing the reliability of the voxel-based and the surface-based method Almukhtar et al.,[23] using pre and post CBCT images of orthognathic surgical patients, reported no significant statistical difference between the two methods, although voxel-based registration

was associated with less variability. The high variability in the surface-based method could have been due to the extra step involving 3D model rendering that this registration required generating a 3D surface mesh model, which may have introduced a possible source of error. As it was also reported by Kang et al.,[46] when compared four different software packages to produce the 3D surface meshes. They found that all four software programs generated reasonable similar meshing accuracies for clinic use. However, there were statistically significant differences at all anatomic regions between them, revealing that there was an inherent range of error in the CT image-based meshing process and highlighted that precautions should be taken in selecting the appropriate software and/or anatomic regions to avoid potential error in specific clinical applications.

When using the landmark-based method, the main drawback is that it requires landmark registration, which can increase the risk of observer-dependent errors. However, as previous studies have reported images can offer consistent and reproducible data if protocols for operator training and calibration are followed.[47, 48] This method utilizes a reference point for 3D cephalometric analysis with CBCT.[49, 50] Lagravere et al.[27] evaluated the potential errors associated with superimposition of serial CBCT images. They utilized reference planes based on cranial base landmarks using a sensitivity analysis. DeCesare et al. later optimized this analysis using a 6-point correction algorithm. This optimized method added two extra landmarks foramen ovale right and left (FOR and FOL), which were shown to decrease the envelope of error when determining the co-ordinate system and increase the intrapoint reliability when comparing images.

Assessing reproducibility is in the same way relevant for a 3D superimposition method to be used in research and clinical settings. Cevidanes et al.[19] using the voxel-based method for superimposition in growing patients assessed the reproducibility of the method, although in a small sample. They analyzed before and after treatment CBCT scans of only three growing patients, who had orthopedic treatment with miniplates as a treatment to correct a Class III malocclusion. The changes with growth and treatment were measured on the 3D models constructed by three examiners. They reported an inter-observer range of measurements across anatomic regions equal or less than 0.5 mm, concluding that these variations are clinically insignificant; therefore, the technique provided a reproducible 3D assessment of growing patients. Comparable reproducibility results were reported by Nada et al.,[31] using the voxel-based method and Gkantidis et al.,[22] using the surface-based method when repeating the superimpositions on the anterior cranial base. Even though the last study used CT scans instead, authors reported that the validity of the proposed superimposition method does not anticipate a substantial change when applied to CBCT images. Although CBCT images have to some extent lower segmentation accuracy when compared to CT, anatomical landmarks and models are generated in a reliable and clinically applicable way.[36]

2.5 Conclusions

- Findings from most of the studies included in the current review suggest that all three methods for 3D superimposition provide an acceptable level of reliability when assessing changes in craniofacial hard tissues.

- However, due to a low methodological quality of the identified evidence, the overall results should be considered cautiously.

- In addition, even though the 3D superimposition methods are more convenient for craniofacial assessment than conventional 2D methods, up to date, no studies have utilized a gold standard to determine the real accuracy of any of these methods.

2.6 References

1. American Board of Orthodontics 2D Cranial Base Superimposition. https://www.americanboardortho.com. Accessed on March 10, 2018.

2. Duterloo H, Planché P. Handbook of cephalometric superimposition Hanover Park, IL : Quintessence Pub., c2011.; 2011.

3. De Clerck HJ, Nguyen T, Koerich L, Cevidanes L. Three-dimensional assessment of mandibular and glenoid fossa changes after bone-anchored Class III intermaxillary traction. Am J Orthod Dentofacial Orthop 2012;142:25-31.

4. Gu Yan, Jr MJ. Cephalometric Superimpositions. A Comparison of Anatomical and Metallic Implant Methods. Angle Orthodontist 2008;78(6):967-76.

5. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

6. Björk A. The Use of Metallic Implants in the Study of Facial Growth in Children : Method and Application. Am. J. Phys. Anthrop 1968;29:243-54.

7. Melsen B. The cranial base. Acta Odontol Scand 1974;32 (Suppl 62):86-101.

8. Melsen B, Melsen F. The postnatal development of the palatomaxillary region studied on human autopsy material. Am J Orthod 1982;82:329-42.

9. Cevidanes LH, Motta A, Proffit WR, Ackerman JL, Stynere M. Cranial base superimposition for 3-dimensional evaluation of soft-tissue changes. AJODO 2010;137(4 Suppl):S120-S29.

10. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.

11. Ghafari J, Engel FE, Laster LL. Cephalometric superimposition on the cranial base: a review and a comparison of four methods. Am J Orthod Dentofacial Orthop 1987;91(5):403-13.

12. Afrand M. Anterior and middle cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2015.

13. Kapila S, Conley R, Harrell Jr W. The current status of cone beam computed tomography imaging in orthodontics. Dentomaxillofacial Radiology 2011;40:24-34.

14. Kapila S, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. Dentomaxillofacial Radiology 2015;44(1):20140282.

15. Kaygısız E, Tortop T. Cone Beam Computed Tomography in Orthodontics. Intechopen 2017;DOI: 10.5772/intechopen.68555.

16. Lenza MA, Carvalho AA, Lenza EB, et al. Radiographic evaluation of orthodontic treatment by means of four different cephalometric superimposition methods. Dental Press J. Orthod 2015;20(3):29-36.

17. Park JH, Tai K, Owtad P. 3-Dimensional Cone-Beam Computed Tomography Superimposition: A review. Seminars in Orthodontics 2015;21(4):263-73.

18. Cevidanes LH, Bailey L'TJ, Tucker SF, Styner MA, et al. Three-dimensional cone-beam computed tomography for assessment of mandibular changes after orthognathic surgery. AJODO 2007;131:44-50.

19. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

20. Weissheimer A, Menezes L, Koerich L, Pham J, Cevidanes L. Fast three-dimensional superimposition of cone beam computed tomography for orthopaedics and orthognathic surgery evaluation. Int. J. Oral Maxillofac. Surg. 2015;44:1188-96.

21. Lagravère M, Secanell M, Major P, Carey J. Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;111:771-77.

22. Gkantidis N, Schauseil M, Pazera P, Zorkun B, et al. Evaluation of 3-Dimensional Superimposition Techniques on Various Skeletal Structures of the Head Using Surface Models. PLoS ONE 2015;10(2):1-20.

23. Almukhtar A, Ju X, Khambay B, McDonald J, Ayoub A. Comparison of the accuracy of voxel based registration and surface based registration for 3D assessment of surgical change following orthognathic surgery. PLoS ONE 2014;9(4):e93402.

24. Jacobson A, Jacobson R. Radiographic Cephalometry. Second Edition ed: Quintessence; 2006.

25. Cevidanes LH, Styner MA, Proffit WR. Image analysis and superimposition of 3-dimensional cone-beam computed tomography models. AJODO 2006;129:611-18.

26. Hatcher D. Maxillofacial imaging. In McNeill C, ed: Science and practice of occlusion. Chicago: Quintessence Pub. Co.; 1997.

27. Lagravère M, Major P, Carey J. Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images. Dentomaxillofacial Radiology 2010;39:400-08.

28. Adams G, Gansky S, Miller A, et al. Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls. Am J Orthod Dentofacial Orthop 2004;126:397-409.

29. Redmond R, Huang J, Bumann A, Mah J. Three-Dimensional Radiographic Analysis in Orthodontics. JCO 2005;XXXIX(7):421-28.

30. Angelopoulos C. Cone Beam Tomography Imaging Anatomy of the Maxillofacial Region. Dent Clin North Am 2008;52:731-52.

31. Nada R, Maal T, Breuning K, Berge S, et al. 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(2):e16520.

32. Grauer D, Cevidanes L, Proffit W. Working with DICOM craniofacial images. AJODO 2009;136:460-70.

33. Terajima M, Yanagita N, Ozeki K, et al. Three dimensional analysis system for orthognathic surgery patients with jaw deformities. AJODO 2008;134:100-11.

34. Naji P, Alsufyani N, Lagravere M. Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist 2014;84:762-72.

35. Viola P, Wells W. Alignment by maximization of mutual information. Paper presented at: Fifth International Conference on Computer Vision, 1995.

36. Moher D, Shamseer L, Clarke M, Ghersi D, et al. Preferred Reporting Items for systematic Reviews and Meta-Analyses: The PRISMA Statement Systematic Reviews 2015;4(1):3-9.

37. Cochrane Consumers and Communication Group resources for authors. Data template extraction. Available from: https://cccrg.cochrane.org/author-resources; 2013. Accessed on May 7, 2016.

38. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual. Available from: http://www.cosmin.nl/cosmin_checklist.html; 2012. Accessed on March 3, 2017.

39. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality Life Research 2012;21:651-57.

40. Cevidanes LH, Bailey LJ, Tucker Jr. GR, Styner MA, et al. Superimposition of 3D cone-beam CT models of orthognathic surgery patients. Dentomaxillofacial Radiology 2005;34:369-75.

41. Beebe D, Rausch J, Byars K, Lanphear B, Yolton K. Persistent Snoring in Preschool Children: Predictors and Behavioral and Developmental Correlates. Pediatrics 2012;130:382-89.

42. Ruellas AC, Tonello C, Alonso N, et al. Author's response. Am J Orthod Dentofacial Orthop 2016;150(3):398-400.

43. Heymann G, Cevidanes L, Cornelis M, De Clerck H, et al. Three-dimensional analysis of maxillary protraction with intermaxillary elastics to miniplates. AJODO 2010;137:274-84.

44. Ribeiro Carvalho F, Cevidanes LH, Motta AT, Oliveira Almeida MA, Phillips C. Three-dimensional assessment of mandibular advancment 1 year after surgery. AJODO 2010;137(4):S53.e1-S53.e12.

45. Toyama Hino C, Cevidanes L, Nguyen T, De Clerck H, et al. Three-dimensional analysis of maxiallry changes associated with facemask and rapid maxillary expansion compared with bone anchored maxillary protraction. Am J Orthod Dentofacial Orthop 2013;144:705-14.

46. Kang AH, Kim MK, Kim HJ, Zhengguo P, Lee SH. Accuracy assessment of image-based surface meshing for volumetric computed tomography images in the craniofacial region. Journal of Craniofacial Surgery 2014;25:2051-55.

47. Almeida MA, Cevidanes L, Phillips C, Motta A, et al. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009;107(2):256-65.

48. Lagravère MO, Low C, Flores-Mir C, Carey J, et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. AJODO 2010;137(5):598-604.

49. Lagravère M, Major P. Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography. AJODO 2005;128:657-60.

50. Lagravère M, Hansen L, Harzer W, Major P. Plane orientation for standardization in 3-dimensional cephalometric analysis with computerized tomography imaging. AJODO 2006;129:601-04.

CHAPTER 3

Intra-examiner Reliability of Three 3D Superimposition Methods

3.1 Introduction:

Superimposition of cephalometric images is a universal method used by orthodontists for aiding in diagnosis and treatment planning, as well as for evaluating growth and treatment outcomes in the craniofacial complex in individual patients.[1, 2] A common reference is required for a superimposition to be able to determine what changes occurred. Such references must be consistently visible and relatively easy to identify in the cephalograms of the individual, and they must be stable within the time frame of the observation period.[2, 3] Based on evidence obtained through scientific research, it is known that the anterior cranial base grows rapidly in early post-natal life, reaching 90% of its final size by 4-5 years of age.[4, 5] Thus, the anterior cranial base is considered a relatively stable reference structure for cephalometric superimposition.[1, 2, 6-8] In addition, references must satisfy two conditions: validity and reliability. Validity, the predominant prerequisite, requires that the reference represents an anatomical entity and that the superimposition technique depict as close as possible the actual changes that occurred. Reliability, on the other hand, relates to precision and reproducibility.[2, 9-11]

In 2008, Leonardi *et al*[12] stated that the inconsistency in landmark identification was still a considerable source of random errors both in manual cephalometric and in computer-aided digital cephalometric analysis.[13-15] In fact, variability in landmark identification has been determined to be fivefold greater than measurement variability.[16, 17] Even though, it is widely accepted that CBCT images provide orthodontist and clinician in general with the ability to observe the inner depths of the craniofacial complex slice by slice through computer reconstruction of radiographic images, landmark identification errors are not exempt when

assessing 3D scans. Although reliability and accuracy of landmark placement in 3D has been tested,[18-24] it is important to understand that when two coordinates are used, the landmarks only need to be identified reliably in two planes, but when three coordinates are used, as in the case of using CBCT images an additional source of error, from the z-axis, may also contribute to the overall error.[25]

Many types of superimposing methods have been used for 2D lateral cephalograms. However, 2D imaging does not fully represent a 3D structure, because much of the information is lost when 3D structures are depicted as 2D images.[26-28] Thus, while 2D cephalometric superimposition is the conventional method used to evaluate craniofacial growth and treatment outcomes, superimposition of CBCT scans, nowadays, allows a 3D visualization of these effects. Similar to cephalometric tracings, 3D models constructed from CBCT scans could be superimposed manually by registering common stable landmarks or by best fit of stable anatomical regions.[29-31]

Three general methods of 3D cephalometric superimposition are well-published and used for clinical diagnosis and assessment of orthodontic treatment outcomes: (1) voxel-based, (2) point-based, and (3) surface-based. For overall superimposition, these methods use parts of the anterior cranial base, as a reference structure for CBCT superimposition, a structure known to have completed most of its growth before the adolescent growth spurt, therefore making it a quite stable reference structure for superimposition.[32, 33]

The voxel-based image registration method,[34-37] has been reported to be an accurate and reproducible automated technique whereby CBCT scans are superimposed by comparing the grey values in a defined volume of interest in two scans to compute the rotation and translation

required to align the two datasets[31, 34, 38, 39] This method attempts to match the cranial bases voxel by voxel between two CBCT images taken at two time-points [pre-treatment/initial status (T1) and post-treatment/growth changes (T2)] from the same subject and then computes the spatial difference between all other data.[34-37] The most popular voxel-based registration method is maximization of mutual information (MMI).[40]

The point-based method works by aligning a series of points to be able to get corresponding points as close as possible. The point-based method minimizes the least square errors of the corresponding points between T1 and T2. The most used point-based method is the landmark-derived plane method (LMD), which uses several identifiable landmarks, such as foramina located in the cranial base, to align the three planes based on a 3D Cartesian coordinate system.[38, 41]

The third method is the surface-based that works by creating a surface from a similar region in both T1 and T2 images. The most common surface-based method is the iterative closest point (ICP); in this technique, an operator manually defines a certain area on the surface of the CBCT scans. Then the software automatically matches and registers the identical landmarks of the selected domains on the two scans and completes the superimposition process.[31, 42]

Most of the limitations of 3D superimposition techniques are related to imaging and landmark identification flaws and software/hardware related errors. In addition, most of the methods that are currently being used in clinical settings are quite time-consuming. Thus, the establishment of a precise, reliable and efficient system to analyze images produced by 3D imaging is needed. Therefore, this research will analyze two voxel-based [CMFreg

(Craniomaxillofacial registration) and Dolphin] and one point-based (LMD) superimposition methods. Surface-method will not be included due to lack of software and training availability. The voxel-based and the landmark-based methods have been previously validated, therefore, this study will evaluate and verify the reliability and tendency to measurement error of the three methods when aligning the pre and post-growth/treatment images to provide clinicians with new possibilities in determining with higher reproducibility the structural changes produced by growth and treatment effects in children and adolescents.

3.2 Material and Methods:

3.2.1 Study population:

A retrospective, observational longitudinal study was carried out on individuals that received comprehensive orthodontic treatment at the University of Alberta. Thirty-six patients with available pre- and post-treatment CBCT were selected from a population of teenagers from 11 to 14 years.

A total of eighty-two patients who had the required full FOV CBCT images generated between 2010 and 2015 were initially selected. Thirty-four individuals with a CVM stage above 5 at pre-treatment and twelve patients with either pre or post-treatment CBCT images with lack of complete cranial base visibility were excluded. Thus, a final sample of thirty-six patients was included.

The mean age of patients at the time of the initial CBCT was 12.4 ± 0.9 years. The mean age at final CBCT was 14.3 ± 0.8 years. The sample included seventeen males and nineteen females.

The interval between T1 and T2 ranged from 22 - 25 months apart. Fourteen patients presented Class I malocclusion, eight patients presented mild Class II malocclusion and fourteen patients presented mild Class III malocclusion. All patients received a non-extraction treatment and included rapid maxillary expansion, full fixed appliances, and intermaxillary elastics.

Patients with any known syndromic anomalies were excluded as well as CBCT imaging taken with machines other than iCAT New Generation Volumetric Scanner.

This study only analyzed gathered data from patients that participated in clinical trials. No additional imaging was requested for these patients. Ethics approval was previously obtained by the Institutional Health Research Ethics Board at the University of Alberta.

3.2.2 Data Collection:

CBCT volumetric data were taken using the iCAT New Generation Volumetric Scanner at 120 kV, 5mA, and 8.9sec. Images were obtained and converted to Digital Imaging and Communications in Medicine (DICOM) format using the iCAT software with a voxel size of 0.3mm. Utilizing AVIZO (Visualization Sciences Group, Burlington, MA), ITK-Snap (Penn Image Computing and Science Laboratory at the University of Pennsylvania), 3D SLICER (Open Platform for the Medical Image Computing Community) and DOLPHIN software, the DICOM format images were rendered into a volumetric image. Sagittal, axial and coronal volumetric slices, as well as the 3D image reconstructions, were used to determine the landmark positions, required for all three methods.

Analysis of the images was carried out by one researcher using the respective superimposition techniques (CMFreg/Slicer, Dolphin and landmark-derived). Extensive training

was required prior to superimposing with each method. CMFreg/slicer demanded more training time due to multiple steps, followed by Dolphin and the landmark-derived method. Intra-observer reliability within each method was done using ten images and two repetitions each, with each measurement trial being at least 1 week apart. Reliability among the three methods was performed using the complete sample. Fourteen landmarks, used in previous studies, were marked on three-dimensional images at T1 and T2 with each of the three methods to assess reliability. (Tables 3.1 and 3.2)

Table 3.1 Landmark Definition

Maxilla	
ANS	The tip of the bony anterior nasal spine, in the median plane
APoint	The point at the deepest midline concavity on the maxilla between the anterior nasal spine and prosthion
PNS	The intersection of a continuation of the anterior wall of the pterygopalatine fossa and the floor of the nose, marking the dorsal limit of the maxilla
OrR	The lowest point in the inferior margin of the right orbit
OrL	The lowest point in the inferior margin of the left orbit
IE #11	Incisal edge of the upper right central incisor
Apex #11	Apex of the upper right central incisor
Mandible	
Me	Menton – The most inferior midline point on the mandibular symphysis
BPoint	The point at the deepest midline concavity on the mandibular symphysis between infradentale and pogonion
GoR	Constructed point of intersection of the right ramus and the mandibular plane
GoL	Constructed point of intersection of the left ramus and the mandibular plane
IE #41	Incisal edge of the lower right central incisor
Apex #41	Apex of the lower right central incisor
Pg	Pogonion – The most anterior of the bony chin in the median plane

Table 3.2 Linear Measurements	
Maxilla	
ANS T2-T1	Distance between ANS pre and post treatment
APoint T2-T1	Distance between A-point pre and post treatment
PNS T2-T1	Distance between PNS pre and post treatment
OrR T2-T1	Distance between OrR pre and post treatment
OrL T2-T1	Distance between OrL pre and post treatment
IE #11 T2-T1	Distance between IE #11 pre and post treatment
Apex #11 T2-T1	Distance between apex #11 pre and post treatment
Mandible	
Me T2-T1	Distance between Me pre and post treatment
BPoint T2-T1	Distance between BPoint pre and post treatment
GoR T2-T1	Distance between GoR pre and post treatment
GoL T2-T1	Distance between GoL pre and post treatment
IE #41 T2-T1	Distance between IE #41 pre and post treatment
Apex #41 T2-T1	Distance between apex #41 pre and post treatment
Pg T2-T1	Distance between Pg pre and post treatment

Voxel-based CMFreg/Slicer Method:

This method uses two different open-source programs ITK-Snap (http://www.itksnap.org) and 3D Slicer (http://www.slicer.org). Using ITK-Snap software program (version 2.0.0) pre-treatment and post-treatment DICOM files were opened and converted to GIPL (Guys Imaging Processing Lab) format for easy processing. Segmentations then were created using the GIPL.GZ files for both pre and post treatment scans using the

intensity segmenter function on 3D Slicer software program (version 4.7.0) to construct 3D volumetric label maps. The same range intensity level was considered adequate for the whole sample, mainly due to the same machine and machine settings used for all scans and similar bone density among the subjects. Once the segmentations were completed and saved, ITK-Snap was used to segment the area of the cranial base to be used as a reference for the superimposition using semi-automatic segmentation. Both scan and segmentation were downloaded so the area of the cranial base to be segmented could be manually "painted" using the software. The software then automatically removed the lower density, and the non-painted areas, leaving only the cranial base. At this point a complete T1 scan, a T1 segmented cranial base, a complete T2 scan, and a T2 segmented cranial base were available. The software then combined each image with its respective cranial base.

Then, surface models were created using the pre-treatment scan and segmentation in 3D Slicer to re-orient the head to establish a common coordinate system across subjects for group comparisons. Using the Transform function pre and post-treatment images were reoriented on the sagittal plane utilizing Foramen Magnum, Crista Galli and Glabella, on the vertical plane using Frankfort horizontal (Porion-Orbitale) and on the transverse plane using Porion to Porion as references, as suggested by Ruellas.[43]

Once the head orientation step was completed, the post-treatment image was manually approximated in relation to pre-treatment image using 3D Slicer, a matrix for cranial base approximation and post-treatment cranial base approximation scan and segmentation files were created. Then using ITK-Snap, the post-treatment mask was cropped before the cranial base registration.

The registration (superimposition) of the post-treatment image upon the pre-treatment image was carried out on the segmented cranial base, using the craniomaxillofacial (CMF) tool and the setting growing rigid automatic registration in 3D Slicer. During the superimposition, T2 was reoriented guided by the best fit of the outlines of the anterior cranial base and automatically superimposed on a static T1, creating a registered T2 surface model.

Once the superimposition was completed, the pre-treatment scan and segmentation, as well as the registered post-treatment scan and segmentation, were landmarked using ITK-Snap. Fourteen 3D landmarks were identified using the three views (axial, sagittal and coronal) for consistency of landmark location.

After placing the defined landmarks to pre and post-images, 3D surface models were created using 3D Slicer for all the levels used in ITK-Snap. These models were utilized to measure the absolute differences between the pre and post-treatment images by applying the Q3DC function (Quantification of directional changes in each plane of the three planes of space). 3D linear distances between T1 and T2 of corresponding landmarks were quantified in the transversal (x-axis), antero-posterior (y-axis) and vertical (z-axis) direction. (Fig. 3.1 and Appendix 3.5)

Landmark-derived Method

Using AVIZO software, the DICOM files were rendered into a volumetric image using 512 x 512 matrices giving a range of 400–420 DICOM slices. Sagittal, axial and coronal multiplanar slices, as well as the 3D image reconstructions, were used to determine the position of the seven landmarks used to superimpose the pre and post-treatment images. The principal

investigator marked the landmarks using the virtual spherical marker of 0.5mm diameter in the x, y and z-axes. The center of the spherical marker is considered as the position of the landmark by the software, therefore the size of the spherical marker does not affect the position of the landmark.

Given the coordinates of three reference landmarks for a plane, 3D visualization software can compute the plane; however, entering the three-point coordinates usually is a time-consuming repetitive manual process. A similar argument applies to determine the perpendicular distance. In order to resolve this issue, this study reproduced the mathematic procedure in Microsoft Excel. This allowed the reference planes and perpendicular distances to be automatically calculated whenever the landmark coordinates were updated.

Four landmarks were required to define a 3D anatomical reference co-ordinate system. The left and right auditory external meatus (AEML and AEMR, respectively) and the dorsum foramen magnum (DFM) were selected as suggested by previous research. The fourth point, ELSA, defined as the midpoint between the left and right foramen spinosum[19] was selected as the origin of the new Cartesian co-ordinate system. From the origin, 3D positional co-ordinates for the AEML, AEMR and DFM were determined.[41]

The optimization formulation used in this study was the 6-point algorithm, that not only optimizes the location of the same three points (i.e. AEML, AEMR and DFM) as used in the 4-point algorithm but also includes both foramen ovale (right and left) in each image.[22, 44] The addition of two extra landmarks (FOR and FOL) in the optimization analysis was shown to reduce the envelope of error when determining the co-ordinate system.[41]

Once all final coordinates were obtained in AVIZO software, the dataset was exported into a spreadsheet database (Excel 2007, Microsoft, Redmond, Wash). Matlab (Matrix Laboratory) software was then used to generate the optimization.

Once data was optimized, linear distances between the 3D coordinates were calculated using the following Euclidean distance formula:

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2}$$

With this formula, *d* represented the distance (mm) between two anatomic landmarks, while (X1, Y1, Z1) and (X2, Y2, Z2) are the respective coordinates of any two given landmarks of 3D interest. Each landmark was included in multiple linear measurements of different orientations to be able to assess all dimensions (superior-inferior, anterior-posterior, right-left). (Fig. 3.2 and Appendix 3.5)

Voxel-based – Dolphin Method

For each patient, T1 and T2 CBCT images were approximated using 4 landmarks located at the right and left frontozygomatic sutures and the right and left mental foramen and superimposed on the cranial base using voxel-based superimposition tool in Dolphin 3D (Chatsworth, CA -version 11.8.06.15 premium). The area of the cranial base used for superimposition was defined by a red box in the three different multiplanar views (axial, sagittal and coronal). The superimposition was achieved by moving the T2 image in relation to the T1

image creating a registered T2 image. No head orientation procedure was performed, as Dolphin software does not have the tool. (Appendices 3.6.8 - 3.6.10)

Then the slice views (axial, sagittal and coronal) were used to confirm the precision of Dolphin 3D superimposition. Once this step was completed, the registered post-treatment scans were exported as DICOM files and opened in ITK-Snap software to convert them into GIPL format similar to the procedure done with the CMFreg/Slicer method. 3D slicer was then used to segment the whole skull using Intensity Segmenter tool, with the same intensity level for all cases to remove any potential error due to the segmentation process. Thus, a surface model of post-treatment segmentation was created for each particular patient. Then pre and post-treatment images were ready for landmarking using ITK-Snap.

After placing the defined landmarks to pre and post-treatment images, 3D surface models were created using 3D Slicer for all the levels used in ITK-Snap. These models were utilized to measure the absolute differences between the pre and post-treatment images by applying the Q3DC function (Quantification of directional changes in each plane of the three planes of space). 3D linear distances between T1 and T2 of corresponding landmarks were quantified in the transversal (x-axis), antero-posterior (y-axis) and vertical (z-axis) direction. (Fig. 3.3 and Appendix 3.5)

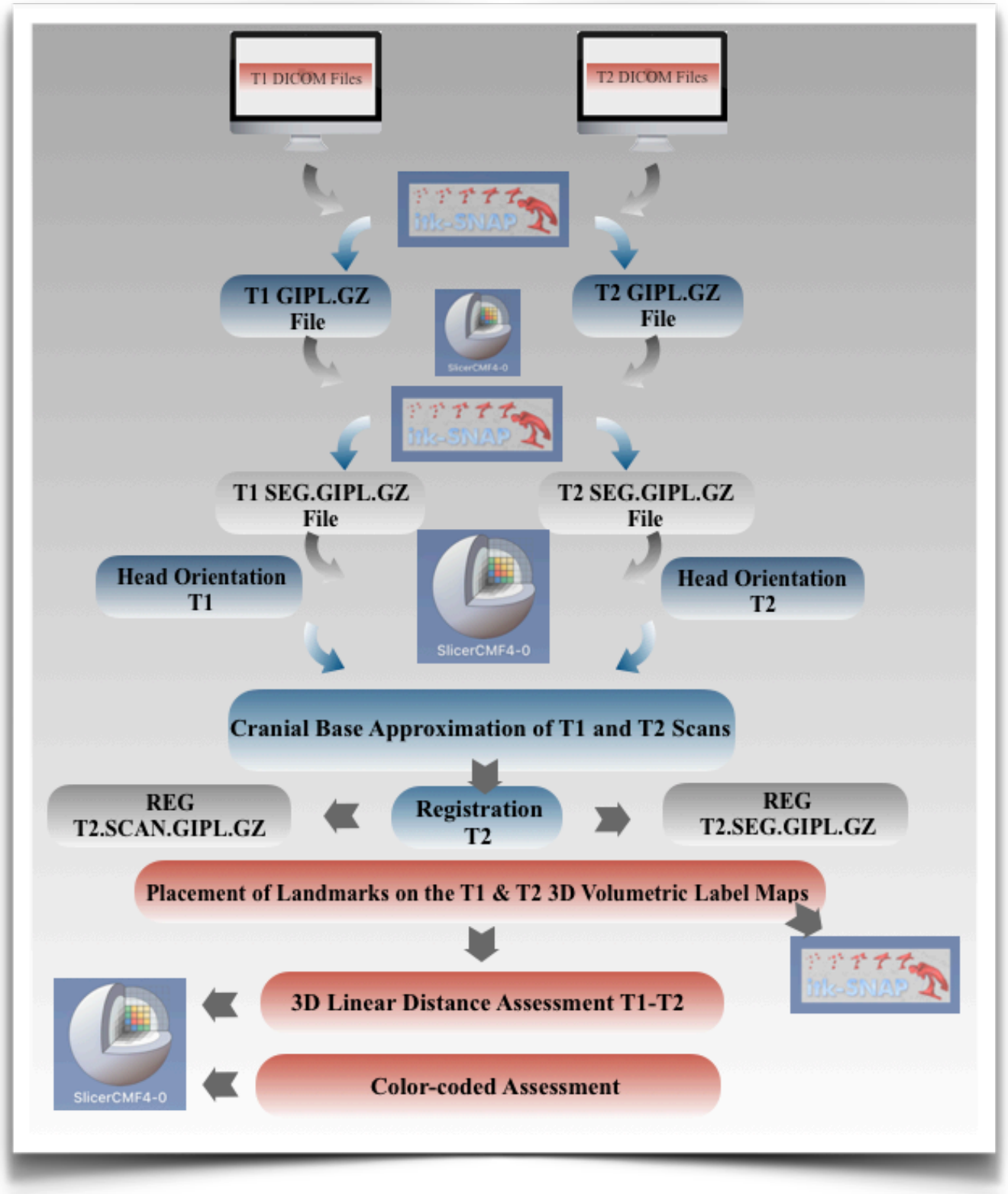


FIGURE 3.1 Flow Diagram CMFreg/slicer Method

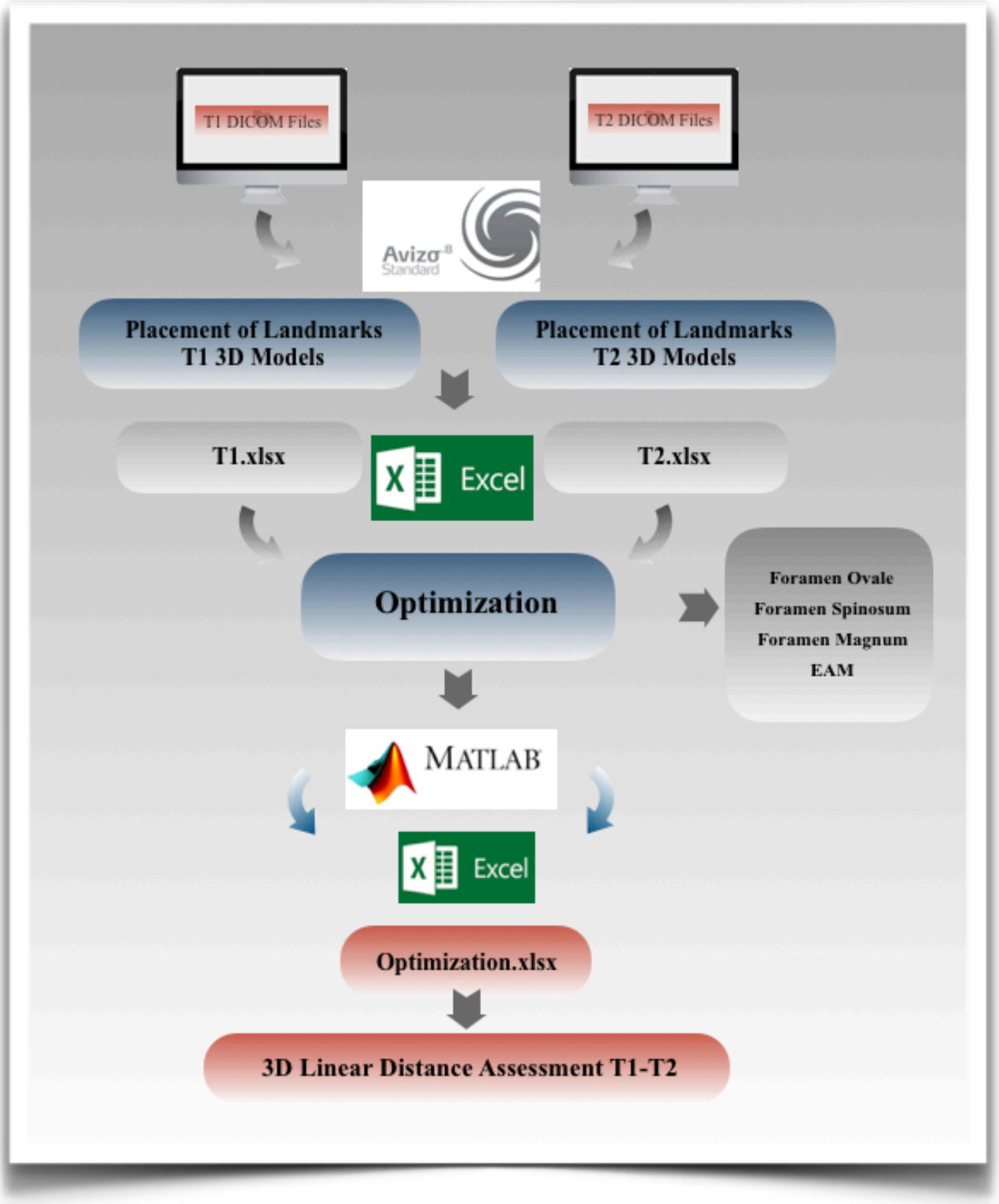


FIGURE 3.2 Flow Diagram Landmark-derived Method

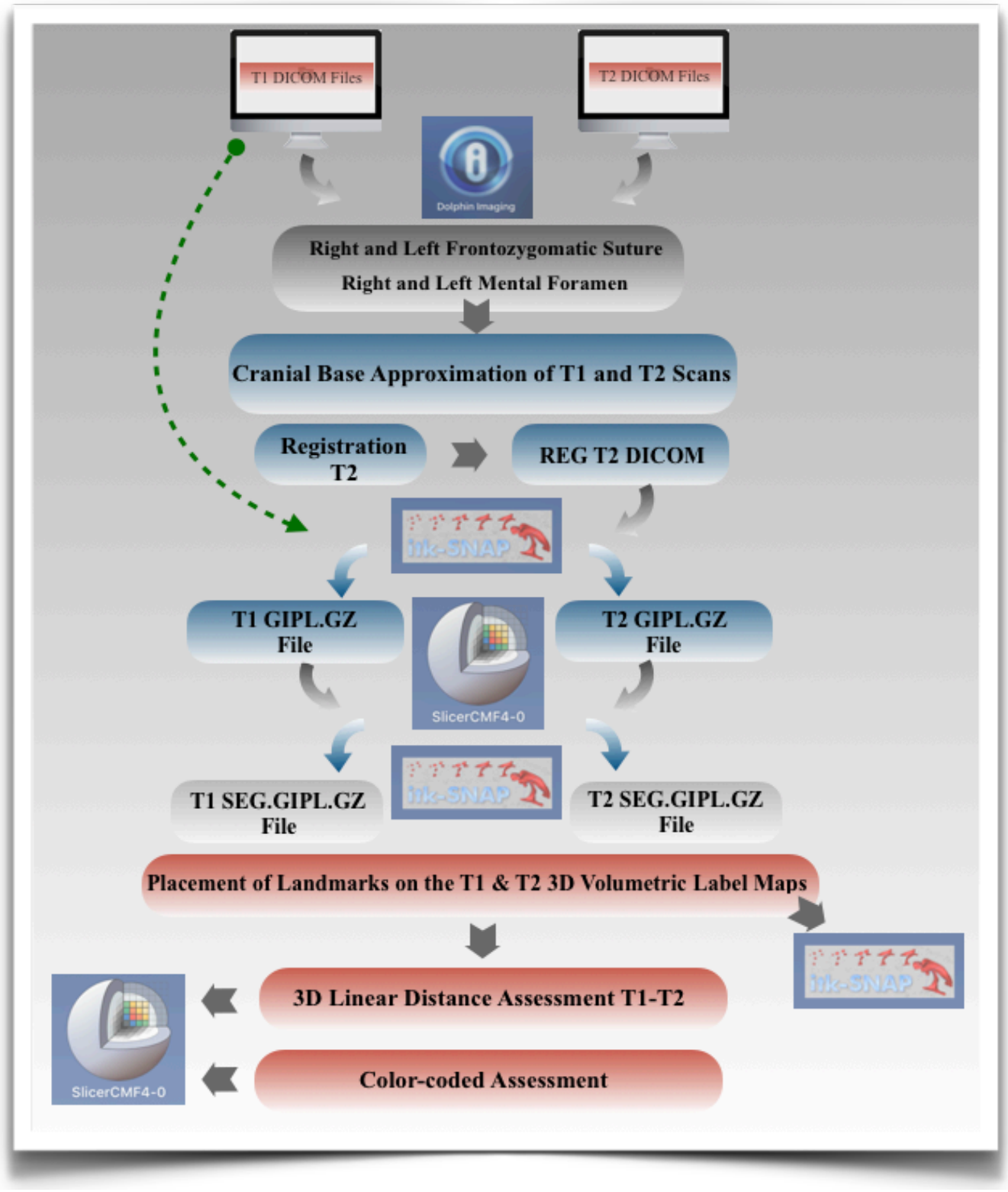


FIGURE 3.3 Flow Diagram Dolphin Method

3.3 Statistical Analysis

The volume data was manually entered into Microsoft Excel 2016 for MAC (Microsoft, Redmond, WA). The SPSS for MAC (Statistical Package for the Social Sciences, version 24) was used for statistical tests. For all tests, the statistical significance was set at *p*-value of 0.05.

3.3.1 Intra-examiner reliability of 3D superimposition per method

Intraclass Correlation Coefficient (ICC) was used to measure the level of agreement between the two repeated measurements within each method by the principal investigator for the continuous dependent variable (difference between T1-T2) taken on the two separate days for each assessed superimposition method. It has to be noted that for the voxel-based methods reliability was tested twice, ten cases each, one performing the complete superimposition with registration at the cranial base and one with landmark tracing only.

The chosen statistical model was a two-way mixed model, single measures with consistency, to ensure consistency in one rater's individual measurements while the subjects were chosen randomly.

The ICC values were interpreted by the general guidelines presented by Portney and Watkins:[45]

Table 3.3 Portney and Watkins' ICC guidelines[45]	
ICC > 0.90	Excellent agreement
0.75 < ICC < 0.89	Good agreement
0.51 < ICC < 0.74	Moderate agreement
ICC < 0.50	Poor agreement

Paired-sample T-test was performed to compare the means of corresponding measurements following both registrations on the anterior cranial base and the registration on the cranial base and the landmark tracing only for both voxel-based methods (CMFreg/Slicer and Dolphin). The significance level was set at 5%.

3.3.2 Intra-examiner reliability of 3D superimposition among methods

ICC was used to assess agreement between the measurements for the continuous dependent variable (3D linear distances - T2 – T1) among all the three methods. In this case, a single-measures with absolute agreement under two-way mixed model was chosen.

3.4 Results

To test reliability within each method, the same investigator repeated the superimposition of ten cases, with each trial being at least 1 week apart. To test the reliability among the three methods, the first trial of thirty-six cases of each method was used. A summary of results is presented in Tables 3.4 to 3.16.

3.4.1 Intra-examiner reliability of 3D superimposition per method

Voxel-based CMFreg/Slicer Method: Complete Superimposition

Using fourteen pre-determined linear distances, excellent agreement for intra-examiner reliability was found on seven skeletal landmarks in the 3D measurements as indicated by an ICC ≥ 0.904. Excellent agreement was also found for all four dental landmarks with ICC ≥ 0.961. Only three landmarks showed good intra-examiner reliability (APoint, PNS and OrR), ICC greater than 0.834. Menton and BPoint were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.968. All these ICC values are considered acceptable;

however, lower bound of CI of two landmarks (APoint and OrR) are below 0.50. (Table 3.4)

Table 3.4 Intra-Examiner Reliability of Linear Measurements - Complete Superimposition

Distances	ICC	Voxel-based CMFreg/Slicer Method 95% Confidence Interval	
		Lower Bound	Upper Bound
ANS T2-T1	0.933	0.731	0.983
APoint T2-T1	0.866	0.462	0.967
PNS T2-T1	0.895	0.579	0.974
OrR T2-T1	0.834	0.331	0.959
OrL T2-T1	0.915	0.656	0.979
Menton T2-T1	0.990	0.959	0.997
BPoint T2-T1	0.968	0.869	0.992
GoR T2-T1	0.967	0.866	0.992
GoL T2-T1	0.904	0.613	0.976
Incisal Edge #11 T2-T1	0.989	0.958	0.997
Apex #11 T2-T1	0.988	0.953	0.997
Incisal Edge #41 T2-T1	0.990	0.960	0.998
Apex #41 T2-T1	0.961	0.842	0.990
Pg T2-T1	0.926	0.702	0.982

Voxel-based CMFreg/Slicer Method: Landmark Only

Excellent agreement for intra-examiner reliability was found on twelve skeletal landmarks in the 3D measurements as indicated by an ICC ≥ 0.900, when only landmarks were retraced. Excellent agreement was also found for all four dental landmarks with ICC ≥ 0.962. Only two landmarks showed good intra-examiner reliability (APoint and OrR), ICC greater than 0.894. Menton and Pogonion were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.980. All lower bound of CI are above 0.50. (Table 3.5)

Table 3.5 Intra-Examiner Reliability of Linear Measurements - Landmarks Only			
Distances	**ICC**	**Voxel-based CMFreg/Slicer Method 95% Confidence Interval**	
		Lower Bound	**Upper Bound**
ANS T2-T1	0.940	0.757	0.985
APoint T2-T1	0.894	0.573	0.974
PNS T2-T1	0.900	0.599	0.975
OrR T2-T1	0.899	0.593	0.975
OrL T2-T1	0.930	0.719	0.983
Menton T2-T1	0.998	0.990	0.999
BPoint T2-T1	0.973	0.893	0.993
GoR T2-T1	0.973	0.889	0.993
GoL T2-T1	0.942	0.765	0.985
Incisal Edge #11 T2-T1	0.993	0.972	0.998
Apex #11 T2-T1	0.996	0.984	0.999
Incisal Edge #41 T2-T1	0.995	0.978	0.999
Apex #41 T2-T1	0.962	0.848	0.991
Pg T2-T1	0.980	0.920	0.995

Table 3.6 shows the differences between the first and second superimposition (registration) on the anterior cranial base. Mean differences between both complete superimpositions were less than 0.67mm. No statistically significant differences were found at any landmark (p-values > 0.05).

Table 3.6 Paired Sample T-test Voxel-based CMFreg/Slicer Method – Complete Superimposition				
Distances	**Voxel-based CMFreg/Slicer Method**			Sig.
	Mean	**95% Confidence Interval of the Difference**		
		Lower	**Upper**	
ANS T2-T1	0.17	-0.321	0.661	0.454
APoint T2-T1	0.33	-0.407	1.073	0.336
PNS T2-T1	0.34	-0.324	0.997	0.279
OrR T2-T1	-0.21	-0.579	0.154	0.222
OrL T2-T1	-0.01	-0.424	0.233	0.526
Menton T2-T1	0.18	-0.165	0.528	0.266
BPoint T2-T1	-0.24	-0.813	0.337	0.373
GoR T2-T1	0.44	-0.319	0.921	0.064
GoL T2-T1	0.67	-0.262	1.599	0.138
Incisal Edge #11 T2-T1	-0.14	-0.445	1.733	0.346
Apex #11 T2-T1	0.19	-0.215	0.591	0.318
Incisal Edge #41 T2-T1	0.12	-0.926	0.334	0.233
Apex #41 T2-T1	0.59	-0.517	0.636	0.822
Pg T2-T1	0.40	-0.441	1.246	0.308

Table 3.7 shows the differences between the first superimposition on the anterior cranial base and the landmark tracing. Mean differences between both trials were less than 0.74mm. No statistically significant differences were found at any landmark (p-values > 0.05).

Table 3.7 Paired Sample T-test Voxel-based CMFreg/Slicer Method Landmarks-only				
Distances	**Voxel-based CMFreg/Slicer Method**			Sig.
	Mean	**95% Confidence Interval of the Difference**		
		Lower	**Upper**	
ANS T3-T1	-0.03	-0.456	0.392	0.867
APoint T3-T1	0.47	-0.513	0.661	0.782
PNS T3-T1	0.41	-0.133	0.956	0.122
OrR T3-T1	-0.03	-0.351	0.288	0.828
OrL T3-T1	-0.04	-0.334	0.251	0.755
Menton T3-T1	-0.05	-0.226	0.126	0.534
BPoint T3-T1	0.02	-0.510	0.558	0.921
GoR T3-T1	0.24	-0.169	0.655	0.215
GoL T3-T1	0.48	-0.251	1.219	0.17
Incisal Edge #11 T3-T1	-0.15	-0.403	0.105	0.216
Apex #11 T3-T1	0.07	-0.165	0.310	0.508
Incisal Edge #41 T3-T1	0.01	-0.164	0.174	0.946
Apex #41 T3-T1	-0.05	-0.599	0.507	0.855
Pg T3-T1	0.20	-0.241	0.646	0.329

Landmark-derived Method:

Excellent agreement for intra-examiner reliability was found on eight skeletal landmarks in the 3D measurements as indicated by an ICC ≥ 0.913. Excellent agreement was also found for all four dental landmarks with ICC ≥ 0.946. OrL and PNS showed good and moderate intra-examiner reliability respectively, ICC ≥ 0.712. Menton and Pogonion were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.955. All these ICC values are considered acceptable; however, lower bound of CI of two landmarks (PNS and OrL) were below 0.50. (Table 3.8)

Mean differences between both complete superimpositions were as high as 1.168mm.

Statistically significant differences were found at five skeletal landmarks: PNS, OrL, Menton, BPoint, and GoL (p-values < 0.05). (Table 3.9)

Table 3.8 Intra-Examiner Reliability of Linear Measurements - Complete Superimposition

Distances	ICC	Landmark-derived Method 95% Confidence Interval	
		Lower Bound	Upper Bound
ANS T2-T1	0.952	0.808	0.988
APoint T2-T1	0.953	0.811	0.988
PNS T2-T1	0.712	-0.161	0.928
OrR T2-T1	0.913	0.650	0.978
OrL T2-T1	0.825	0.444	0.954
Menton T2-T1	0.955	0.831	0.989
BPoint T2-T1	0.926	0.732	0.981
GoR T2-T1	0.953	0.811	0.988
GoL T2-T1	0.948	0.805	0.987
Incisal Edge #11 T2-T1	0.968	0.873	0.992
Apex #11 T2-T1	0.959	0.834	0.99
Incisal Edge #41 T2-T1	0.946	0.784	0.987
Apex #41 T2-T1	0.968	0.873	0.992
Pg T2-T1	0.993	0.972	0.998

Table 3.9 Paired Sample T-test Landmark-derived Method – Complete Superimposition				
Distances	**Landmark-derived Method**			
	Mean	**95% Confidence Interval of the Difference**		**Sig.**
		Lower	**Upper**	
ANS T2-T1	0.74	-0.094	1.575	0.076
APoint T2-T1	-0.09	-0.985	0.804	0.824
PNS T2-T1	0.91	0.063	1.759	0.038
OrR T2-T1	0.71	-0.117	1.530	0.084
OrL T2-T1	0.89	0.061	1.710	0.038
Menton T2-T1	0.92	0.131	1.718	0.027
BPoint T2-T1	1.14	0.178	2.108	0.025
GoR T2-T1	0.74	-0.160	1.634	0.096
GoL T2-T1	1.17	0.595	1.741	0.001
Incisal Edge #11 T2-T1	0.49	-0.410	1.395	0.248
Apex #11 T2-T1	0.50	-0.360	1.361	0.221
Incisal Edge #41 T2-T1	0.97	0.272	1.672	0.012
Apex #41 T2-T1	0.44	-0.528	1.397	0.334
Pg T2-T1	-0.06	-0.537	0.425	0.797

Voxel-based Dolphin Method: Complete Superimposition

Using fourteen pre-determined linear distances, excellent agreement for intra-examiner reliability was found on all skeletal landmarks in the 3D measurements as indicated by an ICC ≥ 0.905. Excellent agreement was also found for all four dental landmarks with ICC ≥ 0.989. Menton and Pogonion were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.979. (Table 3.10)

Table 3.10 Intra-Examiner Reliability of Linear Measurements - Complete Superimposition			
Distances	**ICC**	**Voxel-based Dolphin Method 95% Confidence Interval**	
		Lower Bound	**Upper Bound**
ANS T2-T1	0.959	0.835	0.99
APoint T2-T1	0.911	0.641	0.978
PNS T2-T1	0.919	0.672	0.98
OrR T2-T1	0.905	0.616	0.976
OrL T2-T1	0.903	0.611	0.976
Menton T2-T1	0.993	0.970	0.998
BPoint T2-T1	0.949	0.793	0.987
GoR T2-T1	0.971	0.883	0.993
GoL T2-T1	0.943	0.772	0.986
Incisal Edge #11 T2-T1	0.996	0.984	0.999
Apex #11 T2-T1	0.996	0.983	0.999
Incisal Edge #41 T2-T1	0.989	0.956	0.997
Apex #41 T2-T1	0.989	0.955	0.977
Pg T2-T1	0.979	0.917	0.995

Voxel-based Dolphin Method: Landmark Only

Excellent agreement for the intra-examiner reliability was found on all skeletal landmarks in the 3D measurements as indicated by an ICC ≥ 0.916, when only landmarks were traced on the third trial. Excellent agreement was also found for all four dental landmarks with ICC ≥ 0.994. Menton and Pogonion were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.988. (Table 3.11)

Table 3.11 Intra-Examiner Reliability of Linear Measurements - Landmarks Only			
Distances	**ICC**	**Voxel-based Dolphin Method 95% Confidence Interval**	
		Lower Bound	**Upper Bound**
ANS T3-T1	0.968	0.870	0.992
APoint T3-T1	0.916	0.661	0.979
PNS T3-T1	0.920	0.678	0.980
OrR T3-T1	0.936	0.741	0.984
OrL T3-T1	0.920	0.679	0.980
Menton T3-T1	0.995	0.978	0.999
BPoint T3-T1	0.959	0.834	0.990
GoR T3-T1	0.974	0.896	0.994
GoL T3-T1	0.956	0.823	0.989
Incisal Edge #11 T3-T1	0.997	0.987	0.999
Apex #11 T3-T1	0.997	0.988	0.999
Incisal Edge #41 T3-T1	0.994	0.975	0.998
Apex #41 T3-T1	0.996	0.982	0.999
Pg T3-T1	0.988	0.952	0.997

Table 3.12 shows the differences between the first and second superimposition (registration) on the anterior cranial base. Mean differences between both complete superimpositions were less than 0.4mm. No statistically significant differences were found at any skeletal landmark (p-values > 0.05).

Table 3.12 Paired Sample T-test Voxel-based Dolphin Method – Complete Superimposition				
Distances	**Voxel-based Dolphin Method**			
	Mean	**95% Confidence Interval of the Difference**		Sig.
		Lower	**Upper**	
ANS T2-T1	0.13	-0.275	0.539	0.483
APoint T2-T1	0.19	-0.375	0.753	0.468
PNS T2-T1	-0.02	-0.294	0.254	0.871
OrR T2-T1	0.09	-0.175	0.360	0.454
OrL T2-T1	-0.03	-0.312	0.246	0.795
Menton T2-T1	-0.04	-0.350	0.272	0.781
BPoint T2-T1	0.29	-0.393	0.970	0.364
GoR T2-T1	-0.04	-0.448	0.374	0.842
GoL T2-T1	-0.03	-0.723	0.654	0.913
Incisal Edge #11 T2-T1	-0.02	-0.439	-0.048	0.020
Apex #11 T2-T1	0.16	-0.070	0.389	0.150
Incisal Edge #41 T2-T1	0.06	-0.183	0.300	0.597
Apex #41 T2-T1	0.14	-0.161	0.436	0.324
Pg T2-T1	0.40	-0.062	0.863	0.082

Table 3.13 shows the differences between the first superimposition on the anterior cranial base and the landmark tracing. Mean differences between both trials were less than 0.26mm. No statistically significant differences were found at any skeletal landmark (p-values > 0.05).

Table 3.13 Paired Sample T-test Voxel-based Dolphin Method – Landmark-only				
Distances	**Voxel-based Dolphin Method**			Sig.
	Mean	**95% Confidence Interval of the Difference**		
		Lower	**Upper**	
ANS T3-T1	0.09	-0.264	0.441	0.582
APoint T3-T1	-0.21	-0.690	0.262	0.335
PNS T3-T1	-0.03	-0.370	0.302	0.823
OrR T3-T1	-0.01	-0.249	0.237	0.955
OrL T3-T1	-0.02	-0.285	0.235	0.834
Menton T3-T1	-0.16	-0.431	0.119	0.232
BPoint T3-T1	0.09	-0.508	0.691	0.738
GoR T3-T1	0.26	-0.152	0.663	0.190
GoL T3-T1	-0.03	-0.634	0.564	0.898
Incisal Edge #11 T3-T1	-0.23	-0.408	-0.055	0.016
Apex #11 T3-T1	0.10	-0.095	0.289	0.282
Incisal Edge #41 T3-T1	0.08	-0.090	0.257	0.305
Apex #41 T3-T1	-0.04	-0.237	0.151	0.628
Pg T3-T1	0.11	-0.251	0.465	0.517

3.4.2 Intra-examiner reliability of 3D superimposition among methods

Good agreement for the intra-examiner reliability was found only at GoL, ICC = 0.759 when the three 3D superimposition methods were evaluated. Menton, BPoint and GoR showed moderate agreement as indicated by an ICC ≥ 0.549. Moderate agreement was also found for all four dental landmarks with ICC ≥ 0.518. ANS, APoint, PNS, OrR, OrL, and Pogonion were the skeletal landmarks with the lowest intra-examiner reliability showing poor agreement, ICC ≥ -0.071. (Table 3.14)

Table 3.14 Intra-Examiner Reliability of Linear Measurements Among the Three Superimposition Methods			
Distances	**ICC**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
ANS T2-T1	0.307	-0.059	0.589
APoint T2-T1	0.389	-0.023	0.662
PNS T2-T1	0.480	0.136	0.709
OrR T2-T1	-0.071	-0.337	0.23
OrL T2-T1	0.267	-0.074	0.551
Menton T2-T1	0.659	0.197	0.845
BPoint T2-T1	0.549	0.139	0.769
GoR T2-T1	0.646	0.374	0.809
GoL T2-T1	0.759	0.574	0.87
Incisal Edge #11 T2-T1	0.518	0.079	0.754
Apex #11 T2-T1	0.722	0.248	0.882
Incisal Edge #41 T2-T1	0.609	0.167	0.812
Apex #41 T2-T1	0.537	0.134	0.759
Pg T2-T1	0.402	0.029	0.659

When assessing both voxel-based methods (CMFreg/Slicer and Dolphin), excellent agreement for intra-examiner reliability was found on four skeletal landmarks (Me, BPoint, GoR and Pg) in the 3D measurements as indicated by an ICC $\geq$ 0.904. Excellent agreement was also found for all four dental landmarks with ICC $\geq$ 0.972. BPoint and GoR were the skeletal landmarks with the highest intra-examiner reliability showing ICC $\geq$ 0.943, and PNS and OrR the skeletal landmarks with the lowest intra-examiner reliability showing ICC = 0.545 and 0.596 respectively. (Table 3.15)

Table 3.15 Intra-Examiner Reliability of Linear Measurements – Voxel-based (CMFreg/Slicer and Dolphin) Superimposition Methods			
Distances	**ICC**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
ANS T2-T1	0.885	0.777	0.941
APoint T2-T1	0.863	0.733	0.93
PNS T2-T1	0.545	0.101	0.769
OrR T2-T1	0.596	0.211	0.793
OrL T2-T1	0.741	0.492	0.868
Menton T2-T1	0.904	0.901	0.974
BPoint T2-T1	0.943	0.889	0.971
GoR T2-T1	0.972	0.945	0.986
GoL T2-T1	0.787	0.582	0.892
Incisal Edge #11 T2-T1	0.977	0.955	0.988
Apex #11 T2-T1	0.975	0.952	0.987
Incisal Edge #41 T2-T1	0.983	0.967	0.991
Apex #41 T2-T1	0.972	0.749	0.935
Pg T2-T1	0.919	0.841	0.959

When assessing the voxel-based CMFreg/Slicer and the Landmark-derived methods, moderate agreement for the intra-examiner reliability was found only at GoL, ICC = 0.538. The rest of skeletal landmarks showed poor agreement as indicated by an ICC ≥ -0.137. Poor agreement for the intra-examiner reliability was also found for three dental landmarks with ICC ≥ 0.324. Menton and GoL were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.460, and ANS and OrR the skeletal landmarks with the lowest intra-examiner reliability showing ICC = 0.119 and -0.137 respectively. (Table 3.16)

Table 3.16 Intra-Examiner Reliability of Linear Measurements – Voxel-based CMFreg/Slicer and Landmark-derived Superimposition Methods

Distances	ICC	95% Confidence Interval	
		Lower Bound	Upper Bound
ANS T2-T1	0.119	-0.201	0.422
APoint T2-T1	0.231	-0.194	0.555
PNS T2-T1	0.289	-0.231	0.611
OrR T2-T1	-0.137	-0.539	0.254
OrL T2-T1	0.167	-0.177	0.475
Menton T2-T1	0.460	-0.176	0.751
BPoint T2-T1	0.348	-0.171	0.656
GoR T2-T1	0.394	-0.094	0.676
GoL T2-T1	0.538	0.125	0.76
Incisal Edge #11 T2-T1	0.335	-0.19	0.651
Apex #11 T2-T1	0.533	-0.193	0.806
Incisal Edge #41 T2-T1	0.426	-0.181	0.726
Apex #41 T2-T1	0.324	-0.177	0.635
Pg T2-T1	0.252	-0.214	0.574

When assessing the voxel-based Dolphin and the Landmark-derived methods, moderate agreement for the intra-examiner reliability was found only at GoL, ICC = 0.717. The rest of skeletal landmarks showed poor agreement as indicated by an ICC ≥ -0.081. Poor agreement for the intra-examiner reliability was also found for three dental landmarks with ICC ≥ 0.301. Menton and GoL were the skeletal landmarks with the highest intra-examiner reliability showing ICC ≥ 0.480, and OrR and OrL the skeletal landmarks with the lowest intra-examiner reliability showing ICC = -0.081 and 0.100 respectively. (Table 3.17)

Overall, voxel-based CMFreg/Slicer/Landmark-derived and voxel-based Dolphin/Landmark-derived generated the lowest agreement with ICC values as low as -0.081.

Table 3.17 Intra-Examiner Reliability of Linear Measurements - Voxel-based Dolphin and Landmark-derived Superimposition Methods

Distances	ICC	95% Confidence Interval	
		Lower Bound	Upper Bound
ANS T2-T1	0.143	-0.182	0.446
APoint T2-T1	0.210	-0.195	0.529
PNS T2-T1	0.353	-0.158	0.656
OrR T2-T1	-0.081	-0.377	0.241
OrL T2-T1	0.100	-0.154	0.372
Menton T2-T1	0.480	-0.201	0.772
BPoint T2-T1	0.335	-0.178	0.646
GoR T2-T1	0.406	-0.077	0.684
GoL T2-T1	0.717	0.313	0.871
Incisal Edge #11 T2-T1	0.301	-0.193	0.62
Apex #11 T2-T1	0.561	-0.192	0.824
Incisal Edge #41 T2-T1	0.397	-0.171	0.700
Apex #41 T2-T1	0.373	-0.169	0.678
Pg T2-T1	0.257	-0.209	0.577

3.5 Discussion

Cranial base superimposition of serial lateral cephalograms has provided clinicians with a visual assessment of overall hard and soft tissue changes resulting from treatment, either orthodontic, orthopedic or orthognathic surgery; and/or growth during a time frame. One of the major disadvantages of using a conventional cephalometric analysis is that a 3D information is depicted as 2D data and often limited to midline structures. Improvements in image registration algorithms have led to the development of new methods for CBCT volume superimposition to overcome the issues faced with generated 2D images.

The challenge of image registration is to superimpose CBCT volumes of patients with craniofacial changes due to the normal growth and/or treatment response at different time-points. In these situations, the different CBCT volumes may have dissimilar grey level intensity, field of view, and dental/skeletal components modified by growth and/or treatments, making the registration process more difficult and prone to failure. Therefore, this study aimed to assess three commonly used 3D superimposition methods and determine if they can adequately be used to superimpose pre- and post-treatment CBCT images of growing patients registered at the anterior cranial base and if there is any difference among them.

The voxel-based image registration is an automated registration technique whereby CBCT scans are superimposed by comparing the grey-level intensity of each voxel in a defined volume of interest in two scans to compute the rotation and translation required to align the two datasets.[31, 34] The basis of the method is to mask all skeletal structures so that only the cranial base is used for superimposition. The major strength of this method is that the registration does not depend on the precision of the 3D surface models. The cranial base models are only used to mask anatomic structures that change with growth and/or treatment. The registration procedures compare voxel by voxel of grey-level CBCT images, containing only the unmasked data, normally the cranial base, to calculate rotation and translation parameters between the two images.[37] The final output obtained by this method is shown in color-coded differences between surfaces.[46, 47]

The voxel-based CMFreg/Slicer method was introduced by Cevidanes et al.[46] and has been reported to be more efficient than the point landmark-derived method since it compares the non-changing reference structures in volumetric data voxel by voxel, is observer-independent,

and does not rely on specific landmarks, as in the case of the landmark-based method.[34] This technique allows a superimposition on the cranial base using approximately 300 thousand voxels and accurately matching structures using high gray scale levels.[46] Such method has the potential to be more reproducible and precise than the traditional 2D superimpositions.[48] Nevertheless, this registration process lacks a clinician-friendly user interface, it is too time-consuming and computing intensive to be applied in routine clinical use, making the process unworkable for clinicians and only suitable for researchers.

Lagravère[38] and DeCesare[41] reported the use of cranial base landmarks to determine the reference co-ordinate system for pre and post-treatment image superimposition. However, it has been stated that using cranial base landmarks to define the reference co-ordinate system is not simple since minor errors in landmark location can magnify the errors of landmarks of interest located far away from the origin of the co-ordinate system.[44, 49]

This method started with the development of a reference point for 3D cephalometric analysis with CBCT by Lagravere et al[19] in 2005. They used sagittal, axial and coronal volumetric slices as well as the 3D image reconstruction to determine landmark positions. AEML, AEMR, ELSA (chosen as the origin to the coordinate system), and DFM were used to define a 3D anatomical reference co-ordinate, as they are anatomical structures located in the cranial base area and in relative correct positions for determining orientation of planes.[50, 51] Lagravère et al[44] evaluated the potential errors associated with superimposition of serial CBCT images utilizing reference planes based on cranial base landmarks using a sensitivity analysis. Although individual cranial base reference points had an excellent level of reliability, they found that a positioning error of 0.25mm in the ELSA could produce up to 1.0mm error in other

cranial base landmark co-ordinates. These errors could be magnified to distant landmarks as they were shown in some cases where menton and infraorbital landmarks were displaced up to 6mm. Therefore, this sensitivity analysis demonstrated that a small envelope of error of 3D superimposition of serial CBCT images using four cranial base landmarks had a compound effect in establishing the 3D superimposition reference planes not being an appropriate approach for analyzing growth and treatment changes.

The last method tested in this study was a voxel-based method that uses Dolphin software. It also uses a mutual intensity information algorithm, as the CMFreg/Slicer method, and although it is only commercially available, the registration step is less time-consuming and more user-friendly, no segmentation is required for the registration, and only one software program is used to complete the superimposition process. However, because Dolphin does not take differences of scale into account for growing patients and uses a box of reference rather than only stable anatomic structures of interest, the registration often fails and requires repetition of approaches until adequate superimposition is obtained. Furthermore, 3D visualization of changes and 3D quantification of surface models was only possible after the construction of surface models was performed in other software, as the tools for segmentation in Dolphin currently utilize only thresholding filters.

The reliability of the three 3D superimposition methods was tested in this study by calculating the mean linear distances between the two models (T2-T1) at fourteen different anatomic regions. When the methods were analyzed individually, the ICC results showed good to excellent agreement for the intra-examiner reliability with CMFreg/Slicer and landmark-derived methods, and excellent intra-examiner reliability when CBCT images were

superimposed with Dolphin method. The slightly higher agreement observed with Dolphin method could be just a reflection of examiner's expertise since this was the last method assessed. Similar although less powerful results were reported by Nada et al.,[52] who tested the reproducibility of CBCT superimposition on the anterior cranial base and the zygomatic arches using voxel-based image registration of 3D CBCT scans from sixteen adult patients who underwent combined surgical orthodontic treatment. When the models were registered on the anterior cranial base, intra-observer reliability was reported to be moderate to good between the repeated superimpositions: the ICC ranged between 0.53 and 0.94 and the mean distances between the two models registered on the zygomatic arch remained within 0.5 mm. Likewise, Cevidanes et al.[37] studied the variability between observers in quantification of treatment outcome only using color-coded distance maps for different anatomic regions on 3D CBCT models registered on the anterior cranial base using a voxel-method method. They reported an inter-examiner range of measurements across anatomic regions equal or less than 0.5mm, which they considered to be clinically insignificant.

The reproducibility of the registration was also tested on both voxel-based (CMFreg/Slicer and Dolphin) methods. There were no evident differences found between the complete registration and the landmarking only, as demonstrated by an excellent agreement for the intra-examiner reliability. In addition, paired T-test showed no statistical significance with mean differences between complete superimposition and landmark only; even though the head orientation procedure was not performed with Dolphin method. Since differences $\leq$ 0.4mm are not clinically significant, the registration process of CMFreg/Slicer and Dolphin methods can be considered clinically reproducible. These results are in agreement with the reports from

Cevidanes et al,[37] who assessed cranial base superimposition in growing patients and Nguyen et al[53] and Ruellas et al[54] who tested regional superimpositions demonstrating similar findings.

On the other hand, when assessing reliability among the three methods, the ICC demonstrated less powerful agreement with a wide range of confidence interval. ICC values were the lowest when compared the landmark-derived method and the voxel-based (CMFreg/Slicer and Dolphin) methods. Moderate to excellent agreement; however, was observed for the intra-examiner reliability when compared the voxel-based methods against each other.

According to the present study, the landmark-derived method generated magnified errors since the 3D linear distances were higher when compared to the other two methods in all the defined landmarks. (Appendix 3.1) Although the method showed moderate to excellent agreement for the intra-examiner reliability when assessed individually, poor to moderate agreement was observed when all the methods were evaluated simultaneously. These results contradict the findings from DeCesare[41] study, who reported a reduced envelope of error using the 6-point correction algorithm optimized analysis instead of the 4-point when determining the co-ordinate system. Although, the landmark-derived registration method uses a number of landmarks as reference and they could be susceptible to landmark identification errors, reliability in landmark identification was determined to be adequate. Therefore, the most likely reason for the reduced reliability and increased measurement error may be the lack of stability of the reference areas. As the landmarks used to superimpose the pre and post-treatment images are located in the medial and posterior cranial base.

Since there is no gold standard, it is not possible yet to determine the accuracy of 3D superimposition methods. Therefore, it is unknown if the amount of change generated by the two voxel-based (CMFreg/Slicer and Dolphin) methods is closer to the real values, although it is a good start to know that two similar computing-based superimposition methods generate quite similar measurements.

3.6 Clinical Implications

- Three-dimensional cranial base superimpositions can be carried out with either an open-source software such as Slicer/ITK-Snap (voxel-based method) or with a commercially available software such as Dolphin (voxel-based method) or Avizo (landmark-derived method).
- CMFreg/Slicer has good to excellent reliability, uses an open-source software, but is time-consuming, up to 3 hours from the segmentation to color-coded maps.
- Dolphin 3D superimposition is highly reliable and considerably faster than the CMFreg/Slicer method but requires a license and uses a rectangular box instead of only stable areas as reference, which increases its failure rate requiring repetition.
- Landmark-derived method using Avizo software has good reliability but requires license and needs to be revisited to reduce measurement error.

3.7 Limitations

Although the 3D superimposition with the voxel-based methods was carried out using two different software programs, they use similar mutual intensity information open-source algorithm as that of pre-treatment, so the registered post-treatment images line up very closely

in both methods. However, the landmark-derived method uses landmarks and a mathematical algorithm to define a 3D anatomical reference co-ordinate system.

Quantification of the differences was performed by measuring the distance between the two surface models (T1 and T2) for the two voxel-based (CMFreg/Slicer and Dolphin) methods using Q3DC tool in 3D Slicer. By contrast, distances in the landmark-derived method were calculated manually through a formula using Excel after the optimization was completed, this could have led to an extra source of error.

Finally, the complexity of the voxel-based CMFreg/Slicer method increased substantially the processing time if a mistake was made even at one of the several steps during the procedure. Hence, decreasing the practicality and convenience for use in clinical practice. (Appendix 3.2) However, because Dolphin does not take differences of scale into account for growing patients and uses a box of reference rather than only stable anatomic structures of interest, the registration often fails and requires repetition of approaches until adequate superimposition is obtained. Furthermore, 3D visualization of changes and 3D quantification of surface models was only possible after the construction of surface models was performed in other software (ITK-Snap and CMFreg/Slicer), as the tools for segmentation in Dolphin currently utilize only thresholding filters.

3.8 Conclusions

The findings of the research indicate good to excellent intra-examiner reliability of the three 3D superimposition methods when assessed individually. However, when assessing reliability among the three methods, the ICC demonstrated less powerful agreement with a wide

range of confidence interval. ICC values were the lowest when compared the landmark-based method and the voxel-based (CMFreg/Slicer and Dolphin) methods. Moderate to excellent agreement was observed for the intra-examiner reliability when compared the voxel-based methods against each other. The landmark-based method generated the highest measurement error among the three methods.

3.9 Appendices

Appendix 3.1 Descriptives of Repeated Measures for All Distances				
			95% Confidence Interval	
Landmarks	**Superimposition Method**	**Mean**	**Lower Bound**	**Upper Bound**
ANS	Voxel-based CMFreg/Slicer Method	2.688	2.207	3.170
	Landmark-derived Method	6.000	5.083	6.917
	Voxel-based Dolphin Method	2.540	2.117	2.963
APoint	Voxel-based CMFreg/Slicer Method	2.541	2.098	2.985
	Landmark-derived Method	5.848	4.896	6.799
	Voxel-based Dolphin Method	2.671	2.220	3.122
PNS	Voxel-based CMFreg/Slicer Method	1.850	1.312	2.388
	Landmark-derived Method	2.871	2.410	3.333
	Voxel-based Dolphin Method	1.734	1.467	2.001
OrR	Voxel-based CMFreg/Slicer Method	1.583	1.204	1.963
	Landmark-derived Method	4.314	3.566	5.062
	Voxel-based Dolphin Method	1.420	1.184	1.656
OrL	Voxel-based CMFreg/Slicer Method	1.543	1.113	1.974
	Landmark-derived Method	4.471	3.768	5.173
	Voxel-based Dolphin Method	1.454	1.205	1.704
Me	Voxel-based CMFreg/Slicer Method	5.690	4.892	6.488
	Landmark-derived Method	9.210	7.912	10.509
	Voxel-based Dolphin Method	5.546	4.771	6.321
BPoint	Voxel-based CMFreg/Slicer Method	4.595	3.904	5.285
	Landmark-derived Method	7.966	6.663	9.268
	Voxel-based Dolphin Method	4.500	3.810	5.191
GoR	Voxel-based CMFreg/Slicer Method	4.623	4.075	5.172
	Landmark-derived Method	6.141	5.233	7.048
	Voxel-based Dolphin Method	4.692	4.129	5.255
GoL	Voxel-based CMFreg/Slicer Method	4.576	3.836	5.315
	Landmark-derived Method	5.569	4.833	6.305
	Voxel-based Dolphin Method	4.449	3.845	5.052
IE #11	Voxel-based CMFreg/Slicer Method	4.039	3.488	4.590
	Landmark-derived Method	7.280	6.137	8.424
	Voxel-based Dolphin Method	3.980	3.426	4.534
Apex #11	Voxel-based CMFreg/Slicer Method	4.303	3.595	5.010
	Landmark-derived Method	7.333	6.208	8.458

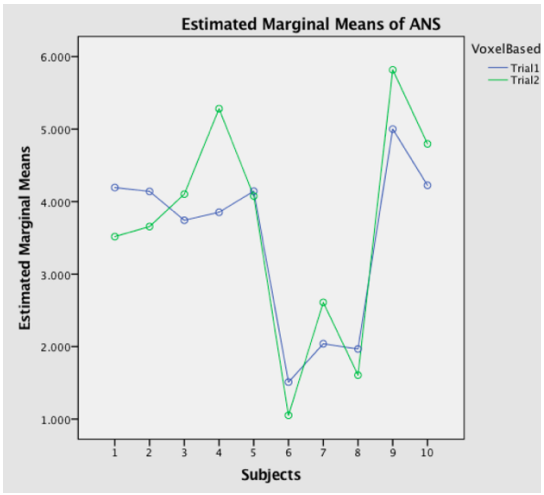
	Voxel-based Dolphin Method	4.383	3.681	5.084
IE #41	Voxel-based CMFreg/Slicer Method	5.121	4.492	5.749
	Landmark-derived Method	8.262	7.100	9.424
	Voxel-based Dolphin Method	5.237	4.656	5.819
Apex #41	Voxel-based CMFreg/Slicer Method	3.893	3.217	4.569
	Landmark-derived Method	7.015	5.831	8.200
	Voxel-based Dolphin Method	3.978	3.406	4.551
Pg	Voxel-based CMFreg/Slicer Method	5.086	4.340	5.832
	Landmark-derived Method	10.179	7.873	12.495
	Voxel-based Dolphin Method	5.076	4.434	5.718

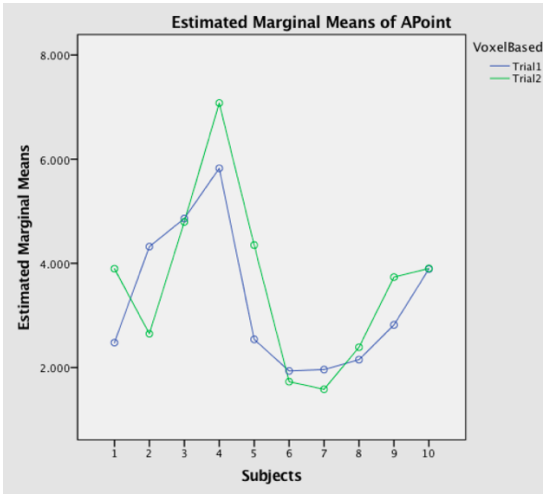
Appendix 3.2 Advantages and Disadvantages of 3D Superimposition Methods

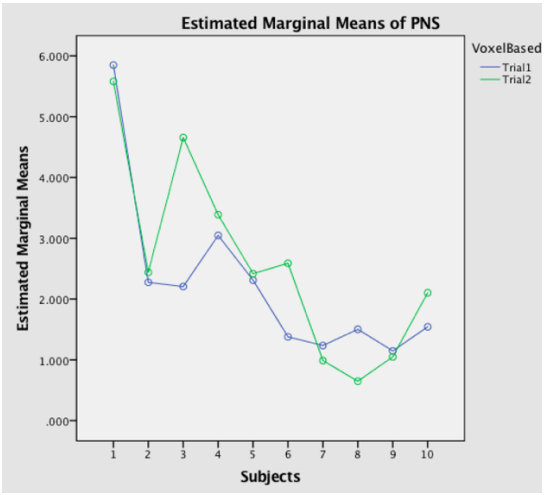
	Superimposition Methods		
Features	**CMFreg/Slicer**	**Dolphin**	**Landmark-derived**
Processing Time	Up to 3 hours from segmentation to color-coded maps. Cranial base registration itself takes from 5 - 60 minutes	Up to 50 minutes if numerical assessment and color–coded maps for visual evaluation are required. Cranial base registration itself takes about10 minutes, but it often fails, requiring repetition, which increasing time	Up to 35 minutes from landmarking to numerical assessment. Approximately 15 minutes for landmarking only. Numerical assessment requires another 15-20 minutes
Complexity	Highly complex for a naïve user. Multiple steps and two different software programs are used throughout the process	Easy to use but utilizes a volume cube for reference not only stable structures. 3D surfaces visual and quantitative evaluations after registration steps require similar steps to CMFreg/Slicer and need to be followed using Slicer and ITK-Snap software programs	Easy to use. Only landmark placement is required
Convenience/Access	Open-source software programs. 3D Slicer and ITK-Snap	License required	License required
Visual Assessment	Superimposition of surface models and color-coded maps for more thorough visual evaluation	Superimposition of surface models and color-coded maps using other software for more thorough visual evaluation	Not available

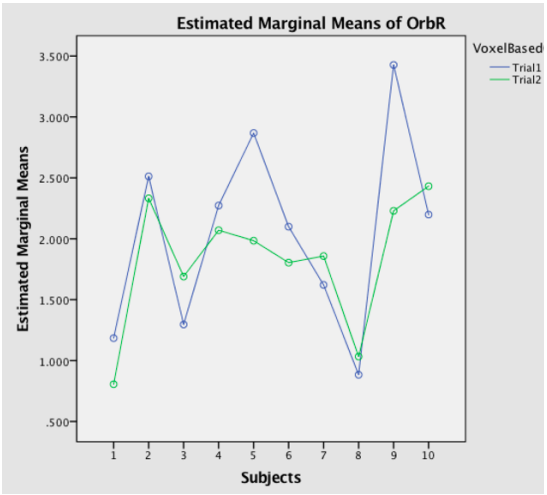
Reliability	Excellent reliability	Excellent reliability	Good reliability
Measurement Error	Less than 0.7mm	Less than 0.4mm	Less than 1.2mm when assessed reliability. However, there was a magnified error when compared to the voxel-based methods
Accuracy	No gold standard available	No gold standard available	No gold standard available

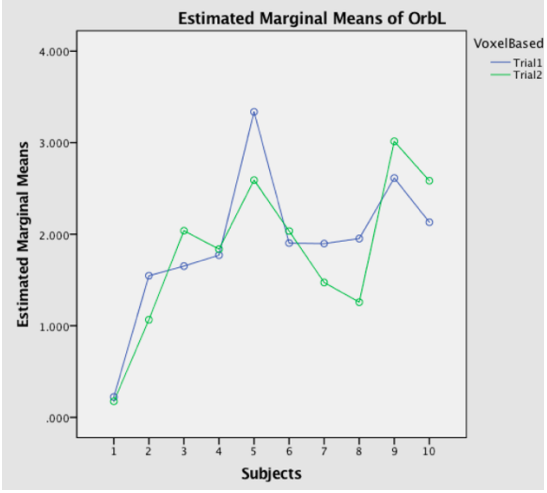
Appendix 3.3 Profile Plots of Intra-Examiner Reliability Voxel-based CMFreg/Slicer Method – Complete Superimposition

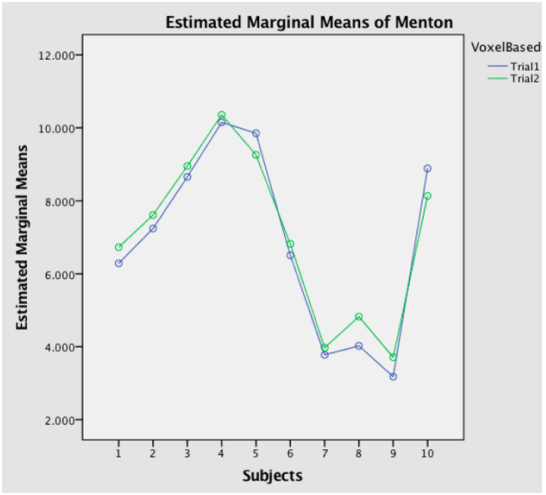


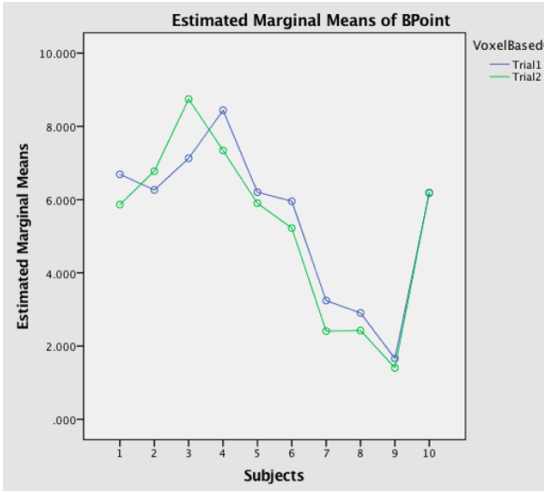


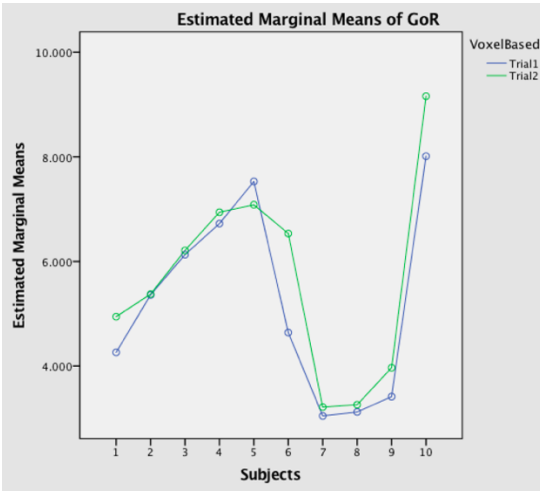


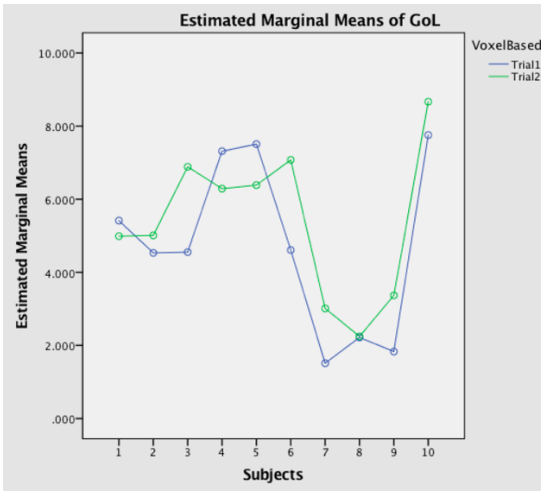


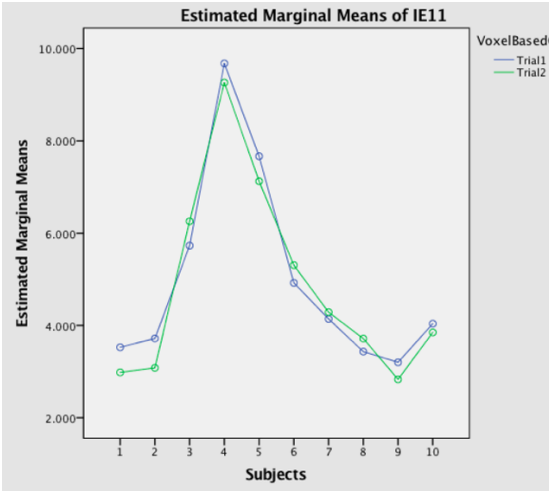

Estimated Marginal Means of OrbL
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2

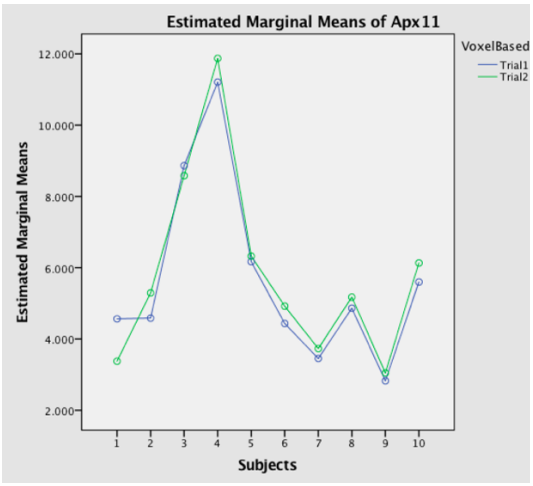

Estimated Marginal Means of Menton
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2

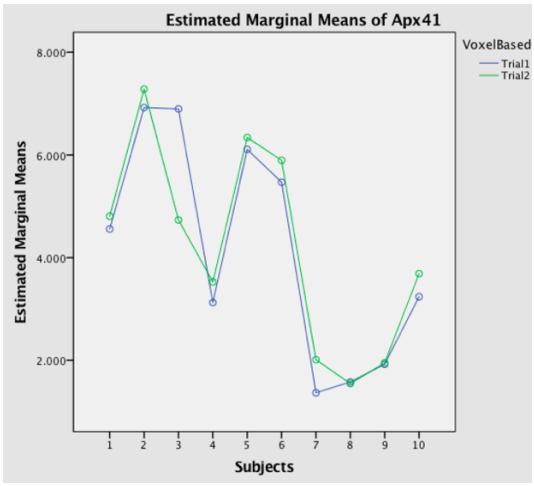

Estimated Marginal Means of BPoint
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2

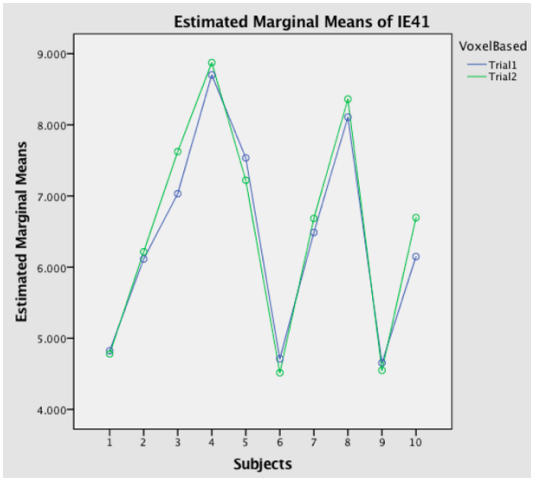

Estimated Marginal Means of GoR
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2

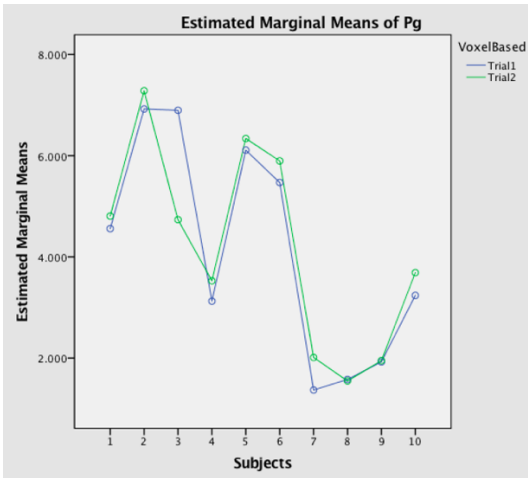

Estimated Marginal Means of GoL
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2


Estimated Marginal Means of IE11
Estimated Marginal Means
Subjects
VoxelBased
Trial1
Trial2

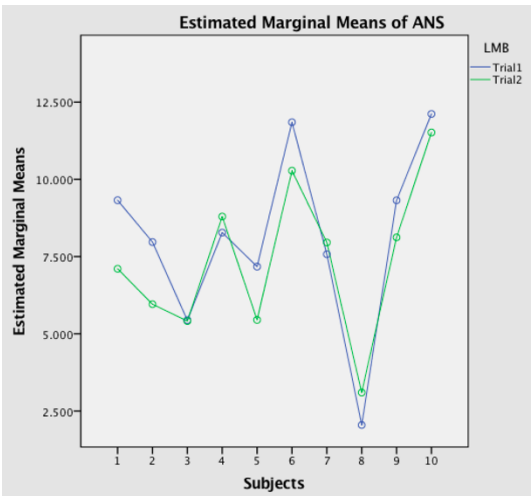


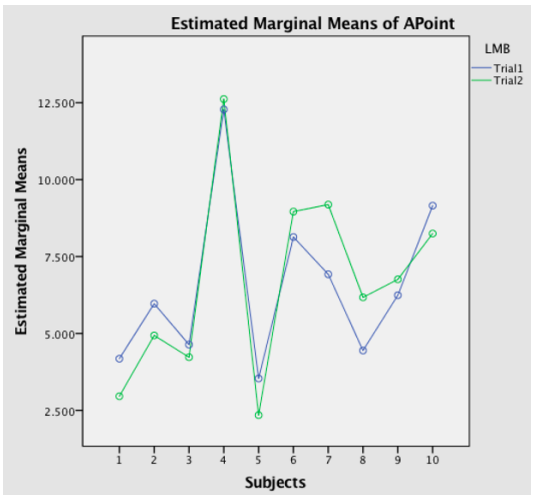


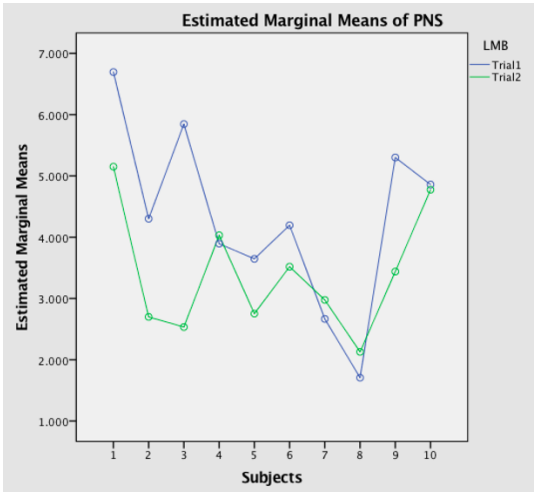


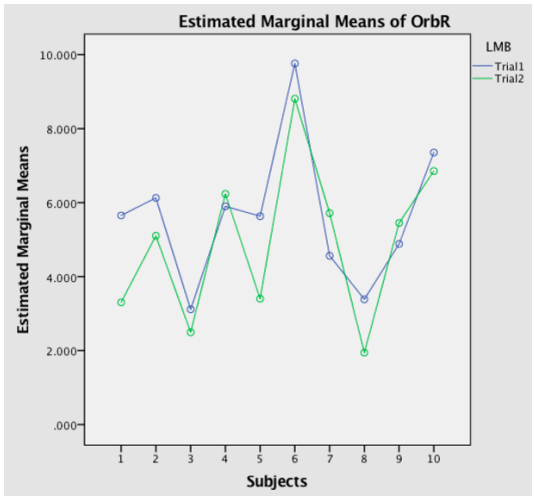


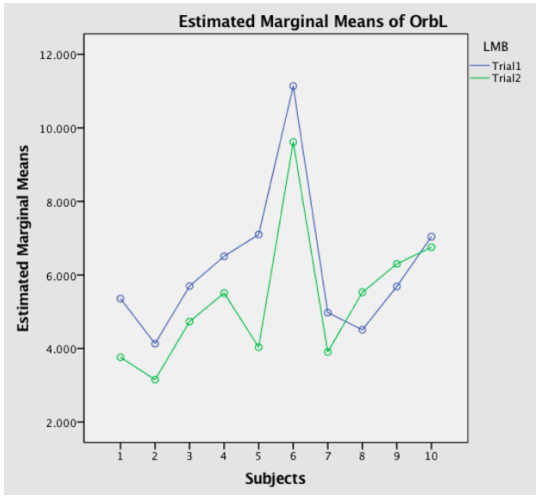
Appendix 3.4 Profile Plots of Intra-Examiner Reliability Landmark-derived Method

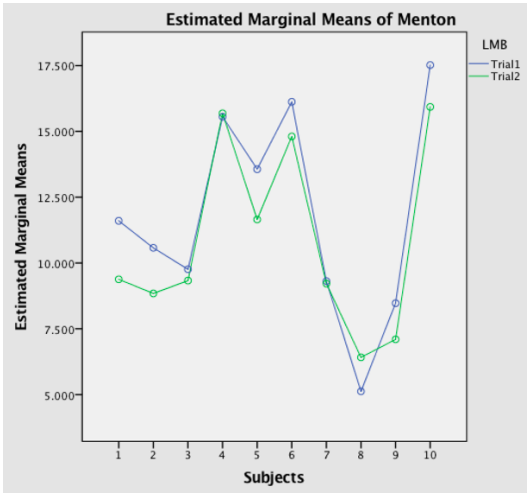


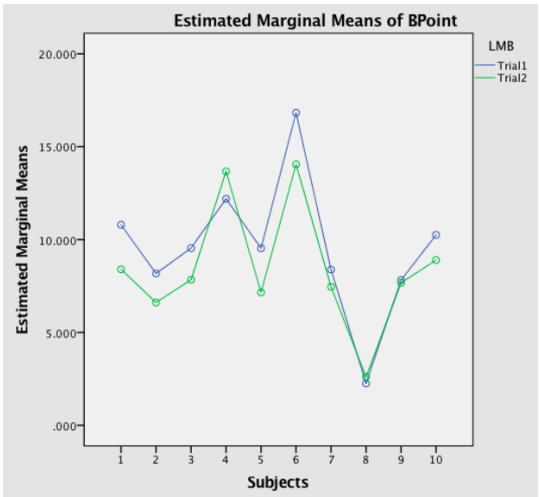


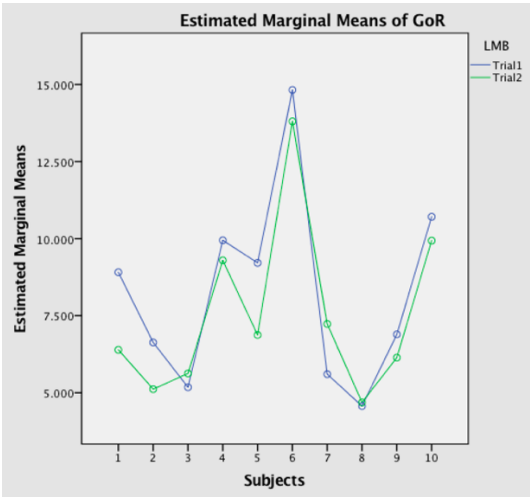

Estimated Marginal Means of PNS
LMB
Trial1
Trial2
Estimated Marginal Means
7.000
6.000
5.000
4.000
3.000
2.000
1.000
1 2 3 4 5 6 7 8 9 10
Subjects

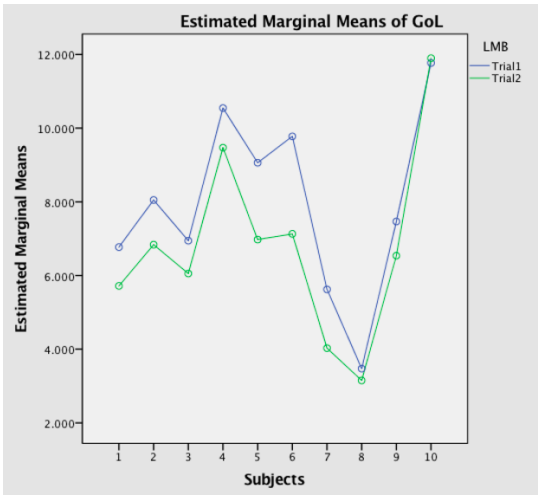

Estimated Marginal Means of OrbR
LMB
Trial1
Trial2
Estimated Marginal Means
10.000
8.000
6.000
4.000
2.000
.000
1 2 3 4 5 6 7 8 9 10
Subjects

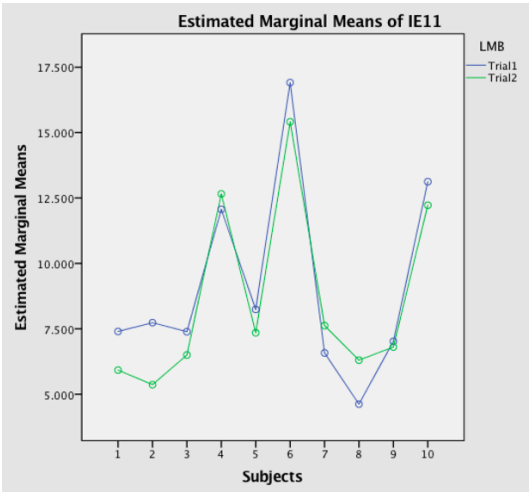

Estimated Marginal Means of OrbL
LMB
Trial1
Trial2
Estimated Marginal Means
12.000
10.000
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects

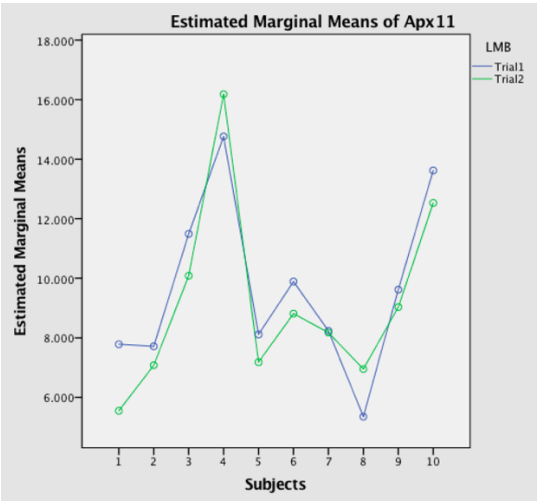

Estimated Marginal Means of Menton
LMB
Trial1
Trial2
Estimated Marginal Means
17.500
15.000
12.500
10.000
7.500
5.000
1 2 3 4 5 6 7 8 9 10
Subjects

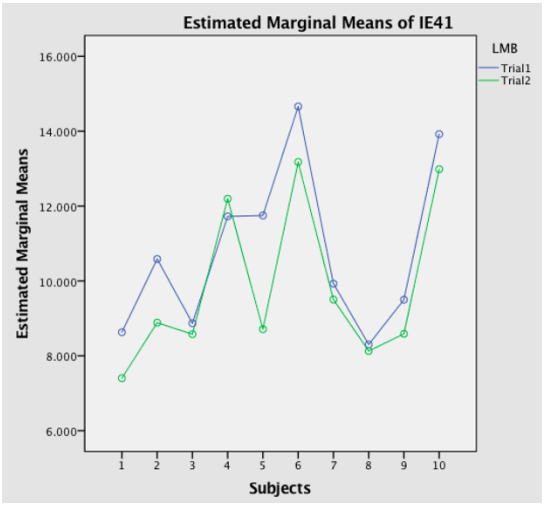

Estimated Marginal Means of BPoint
LMB
Trial1
Trial2
Estimated Marginal Means
20.000
15.000
10.000
5.000
.000
1 2 3 4 5 6 7 8 9 10
Subjects

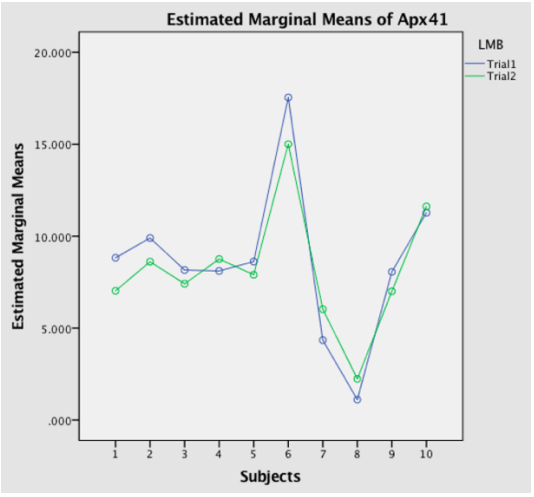

Estimated Marginal Means of GoR
LMB
Trial1
Trial2
Estimated Marginal Means
15.000
12.500
10.000
7.500
5.000
1 2 3 4 5 6 7 8 9 10
Subjects

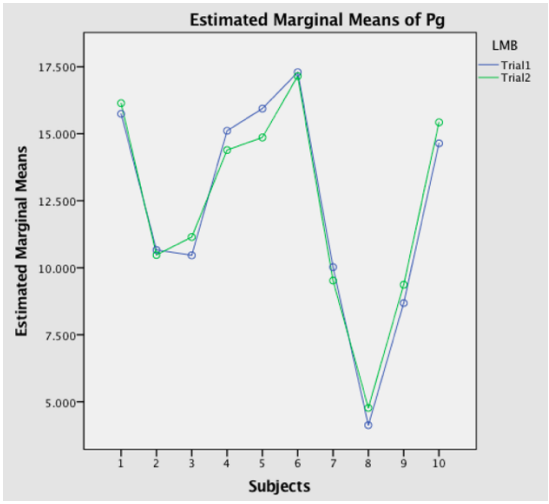

Estimated Marginal Means of GoL
LMB
Trial1
Trial2
Estimated Marginal Means
12.000
10.000
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects


Estimated Marginal Means of IE11
LMB
Trial1
Trial2
Estimated Marginal Means
17.500
15.000
12.500
10.000
7.500
5.000
1 2 3 4 5 6 7 8 9 10
Subjects

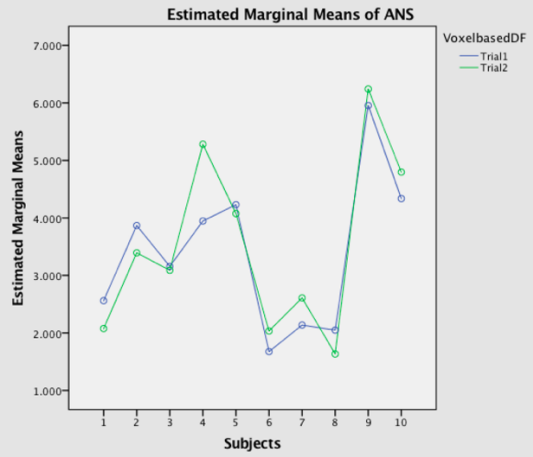

Estimated Marginal Means of Apx11
LMB
Trial1
Trial2
Estimated Marginal Means
18.000
16.000
14.000
12.000
10.000
8.000
6.000
1 2 3 4 5 6 7 8 9 10
Subjects

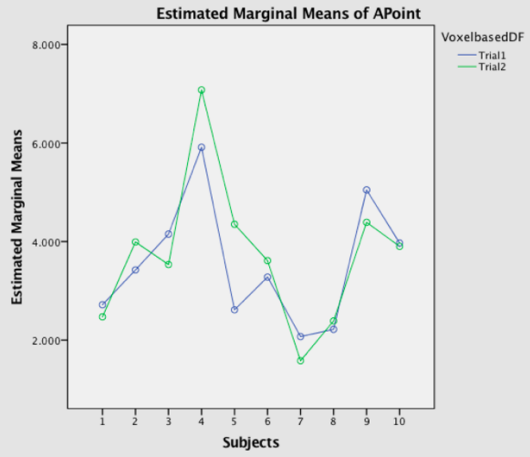

Estimated Marginal Means of IE41
LMB
Trial1
Trial2
Estimated Marginal Means
16.000
14.000
12.000
10.000
8.000
6.000
1 2 3 4 5 6 7 8 9 10
Subjects

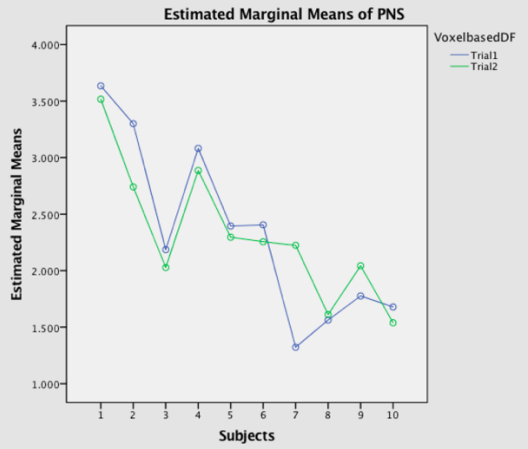

Estimated Marginal Means of Apx41
LMB
Trial1
Trial2
Estimated Marginal Means
20.000
15.000
10.000
5.000
.000
1 2 3 4 5 6 7 8 9 10
Subjects

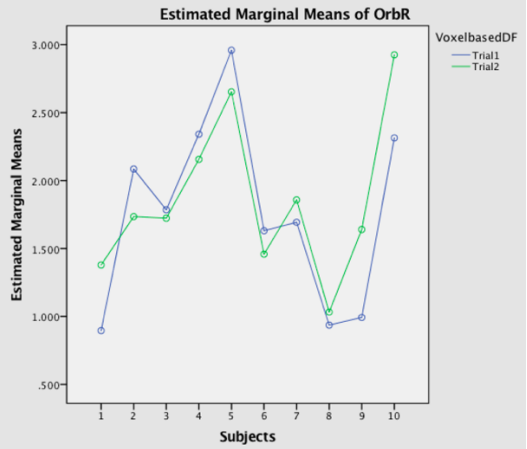

Estimated Marginal Means of Pg
LMB
Trial1
Trial2
Estimated Marginal Means
17.500
15.000
12.500
10.000
7.500
5.000
1 2 3 4 5 6 7 8 9 10
Subjects

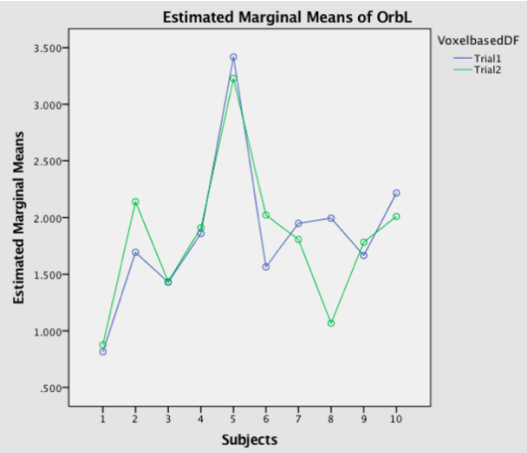
Appendix 3.5 Profile Plots of Intra-Examiner Reliability Voxel-based Dolphin Method – Complete Superimposition

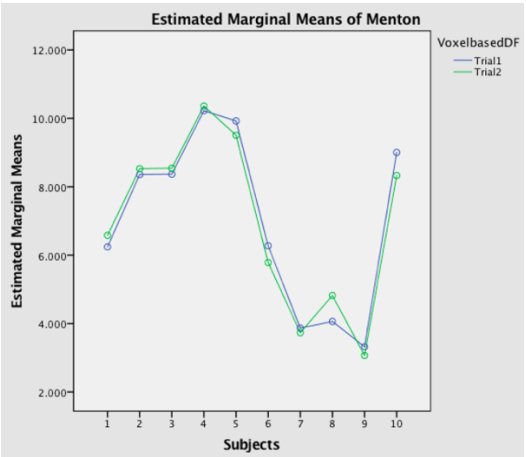


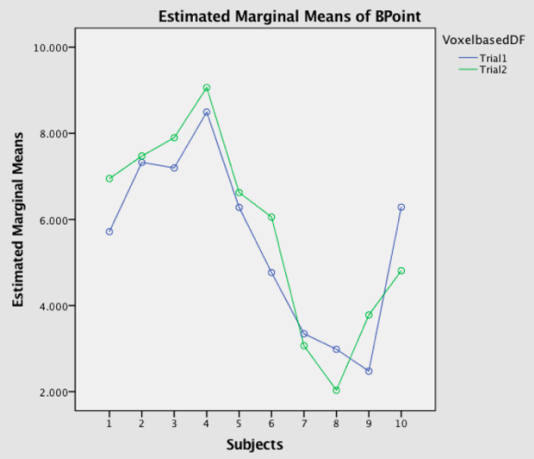


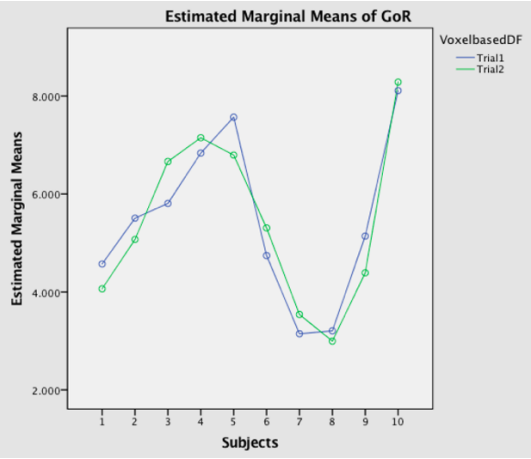


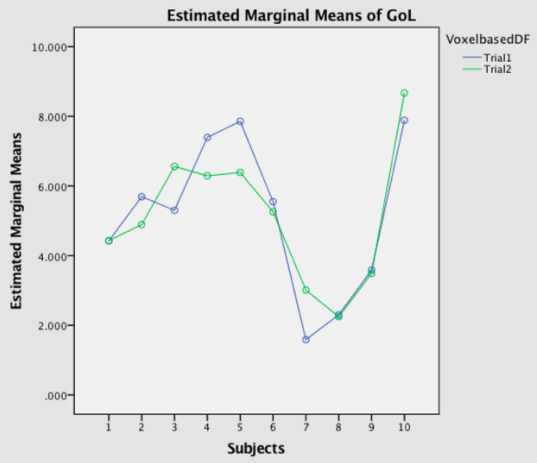


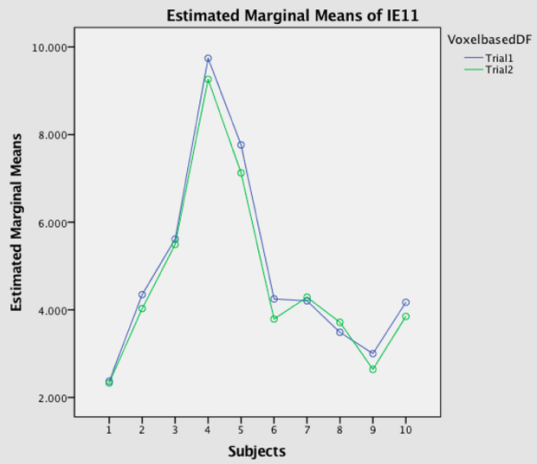


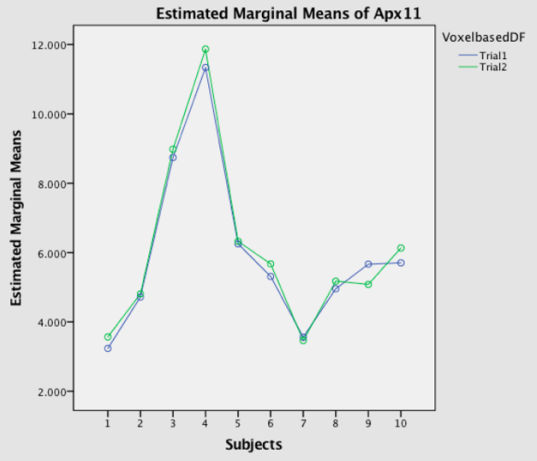


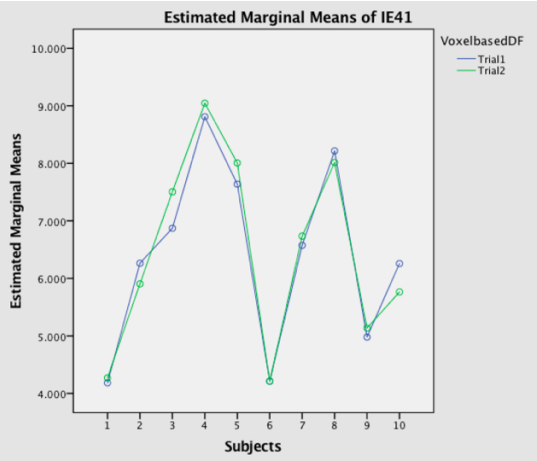

Estimated Marginal Means of BPoint
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
10.000
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects

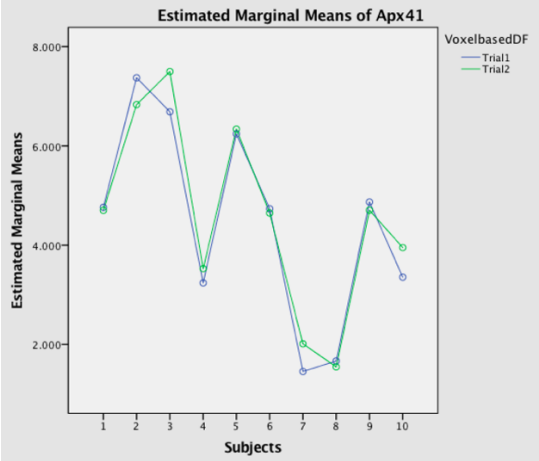

Estimated Marginal Means of GoR
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects

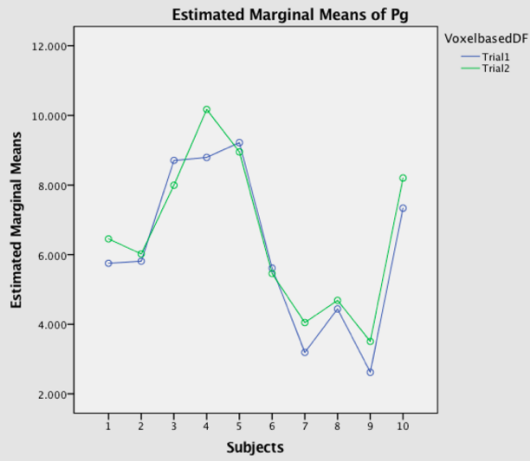

Estimated Marginal Means of GoL
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
10.000
8.000
6.000
4.000
2.000
.000
1 2 3 4 5 6 7 8 9 10
Subjects


Estimated Marginal Means of IE11
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
10.000
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects


Estimated Marginal Means of Apx11
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
12.000
10.000
8.000
6.000
4.000
2.000
1 2 3 4 5 6 7 8 9 10
Subjects

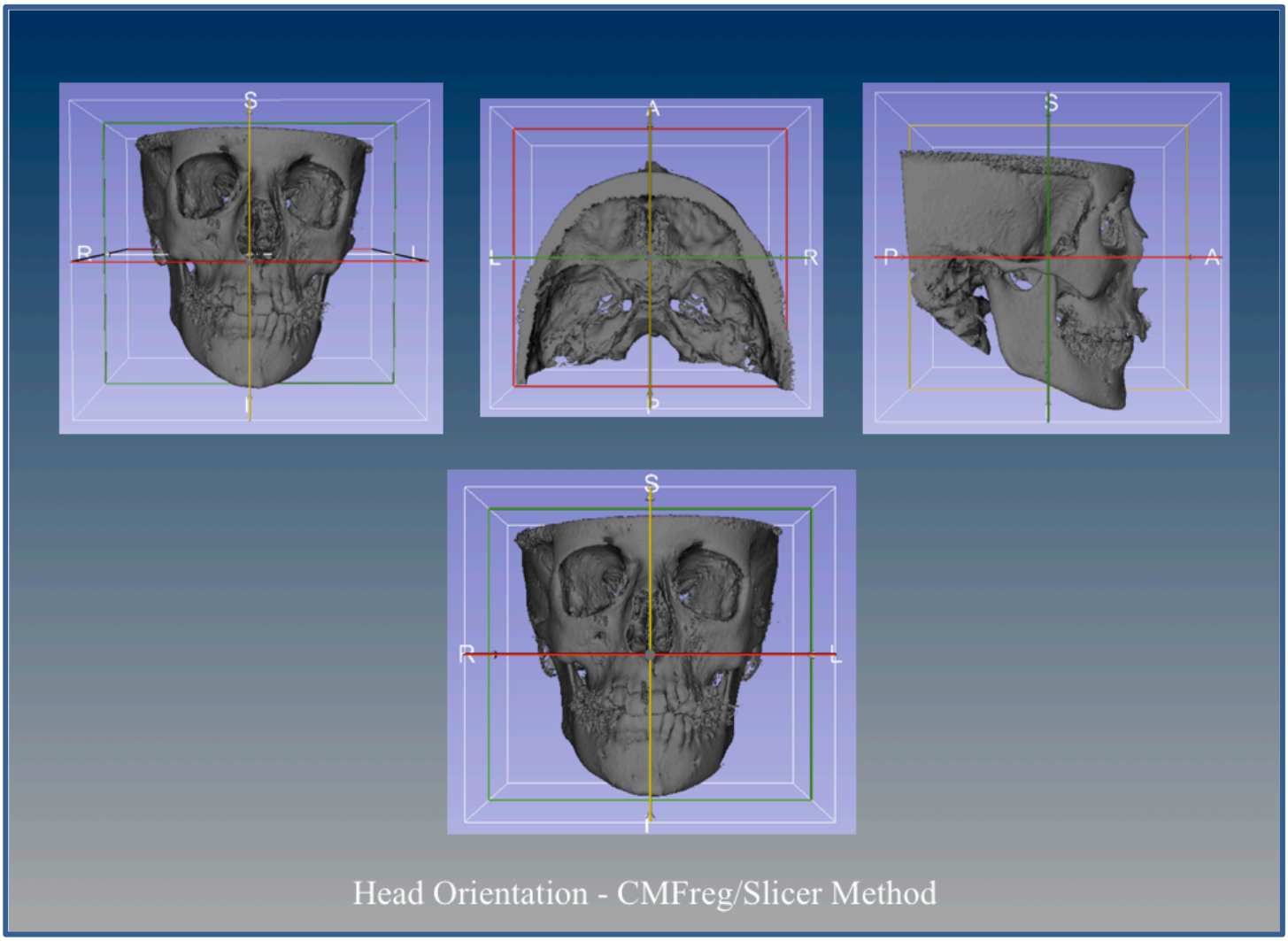

Estimated Marginal Means of IE41
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
10.000
9.000
8.000
7.000
6.000
5.000
4.000
1 2 3 4 5 6 7 8 9 10
Subjects


Estimated Marginal Means of Apx41
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
8.000
6.000
4.000
2.000
1
2
3
4
5
6
7
8
9
10
Subjects

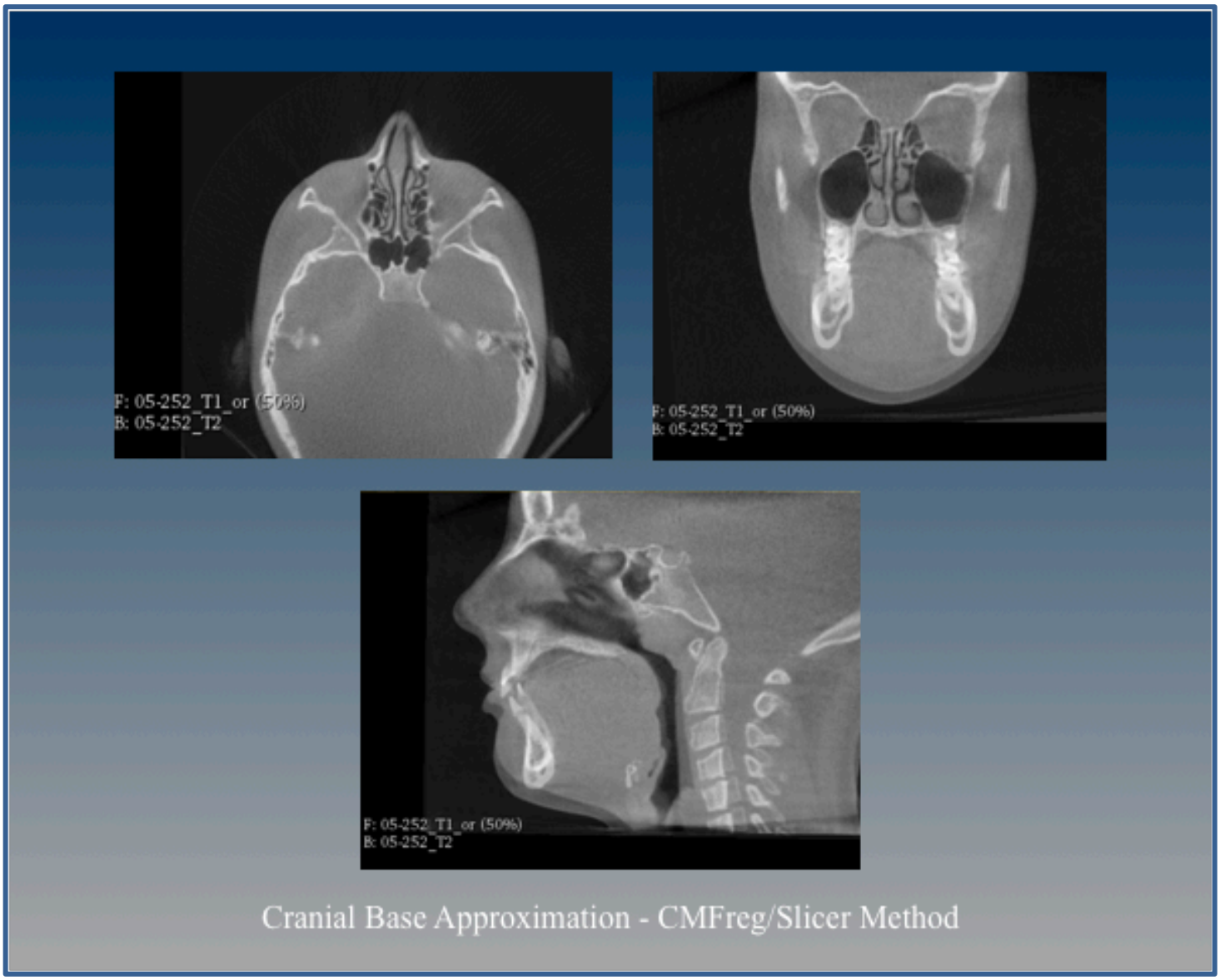

Estimated Marginal Means of Pg
VoxelbasedDF
Trial1
Trial2
Estimated Marginal Means
12.000
10.000
8.000
6.000
4.000
2.000
1
2
3
4
5
6
7
8
9
10
Subjects

Appendix 3.6 Steps for 3D Superimposition

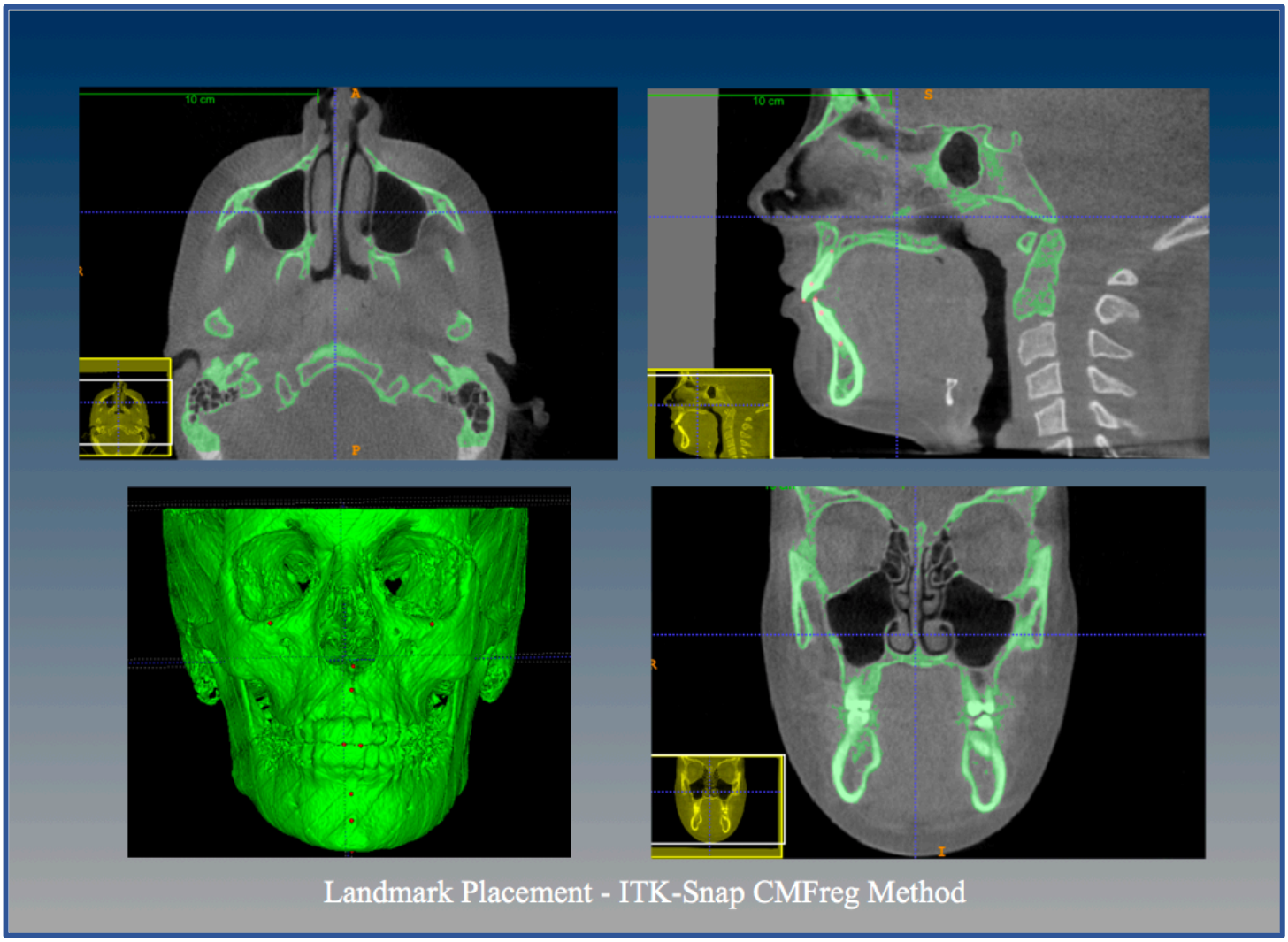
Appendix 3.6.1 Head orientation with CMFreg/Slicer method



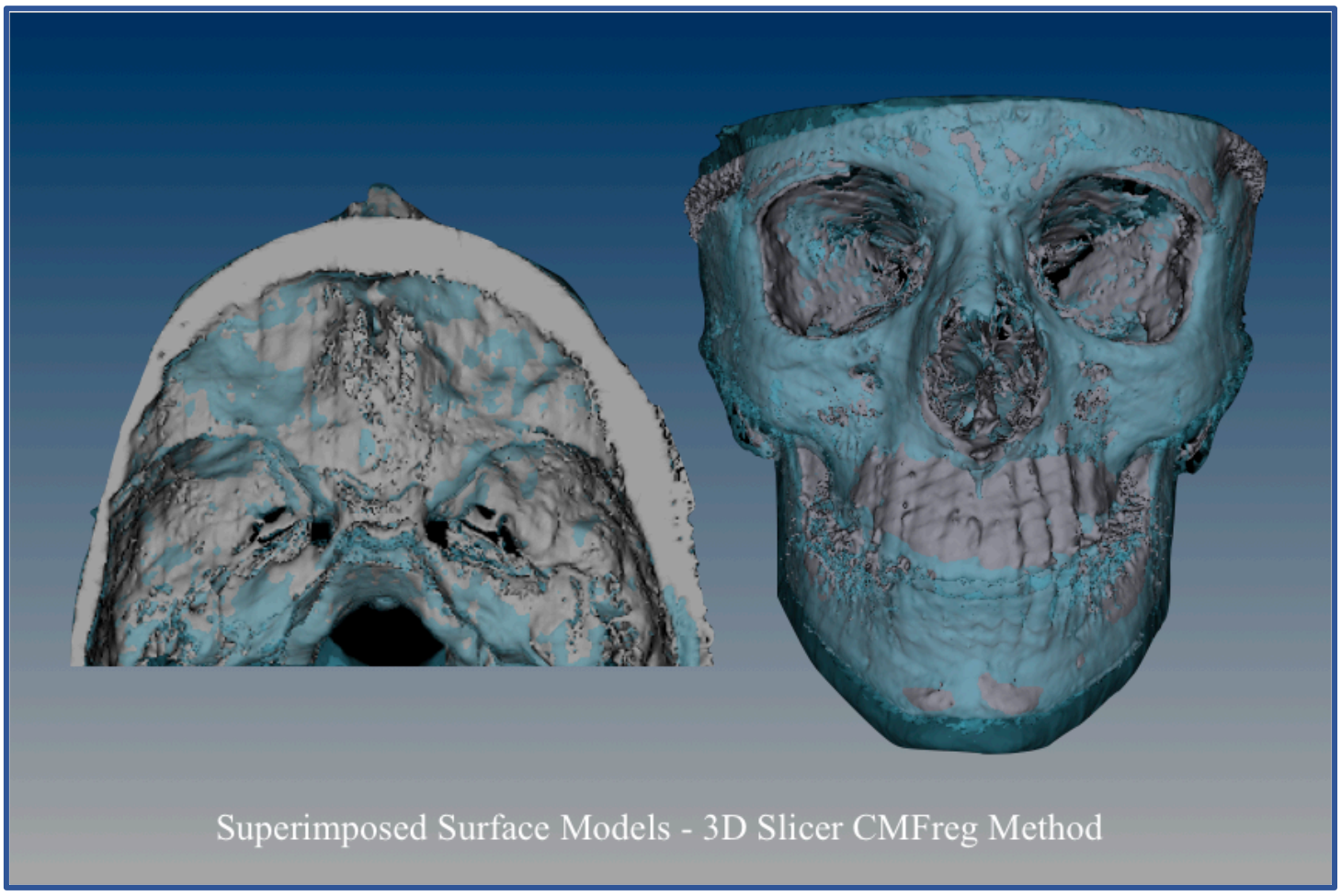
Appendix 3.6.2 Cranial base approximation with CMFreg/Slicer method



Appendix 3.6.3 Landmark placement step with CMFreg/Slicer method

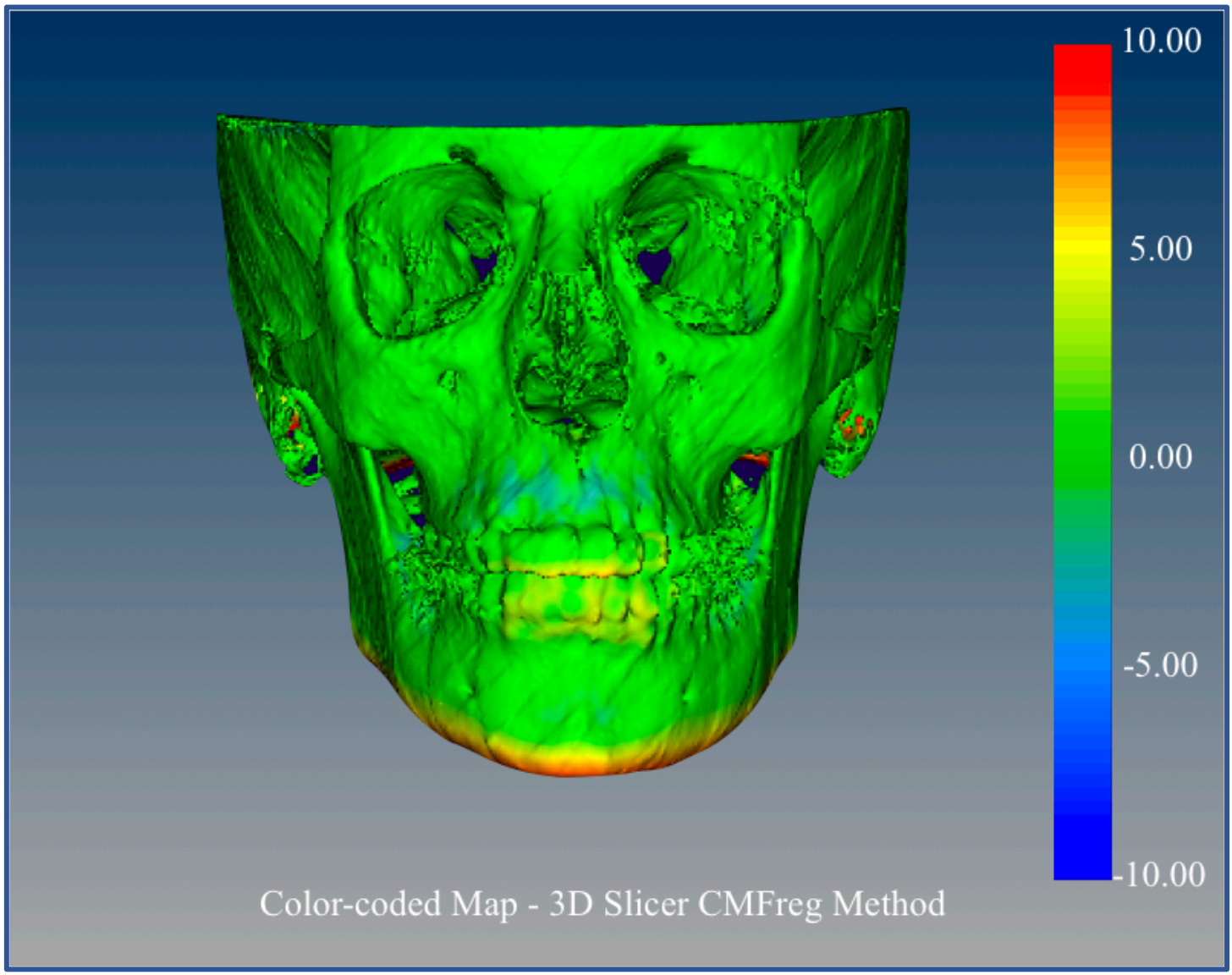


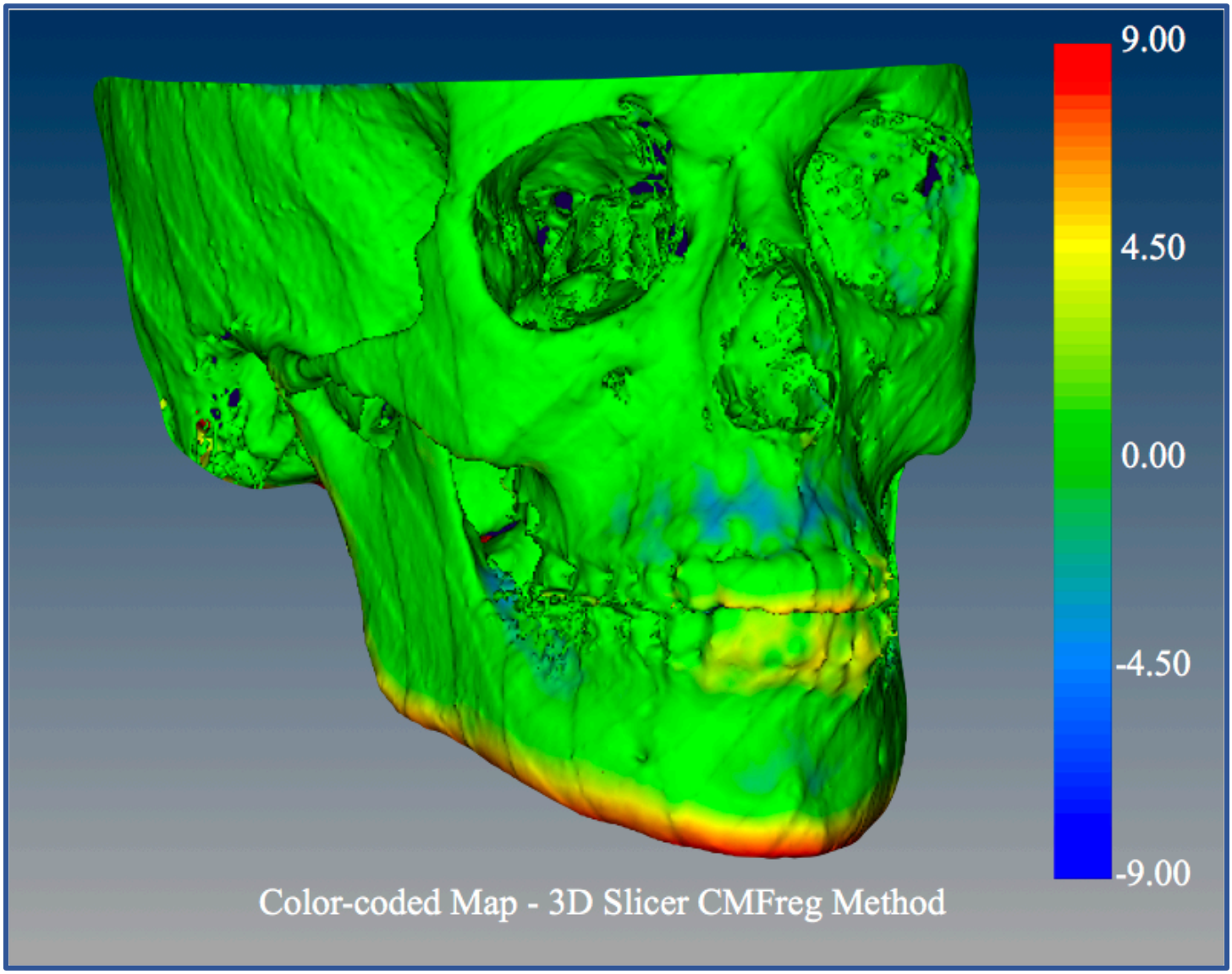
Appendix 3.6.4 Superimposed surface models with CMFreg/Slicer method

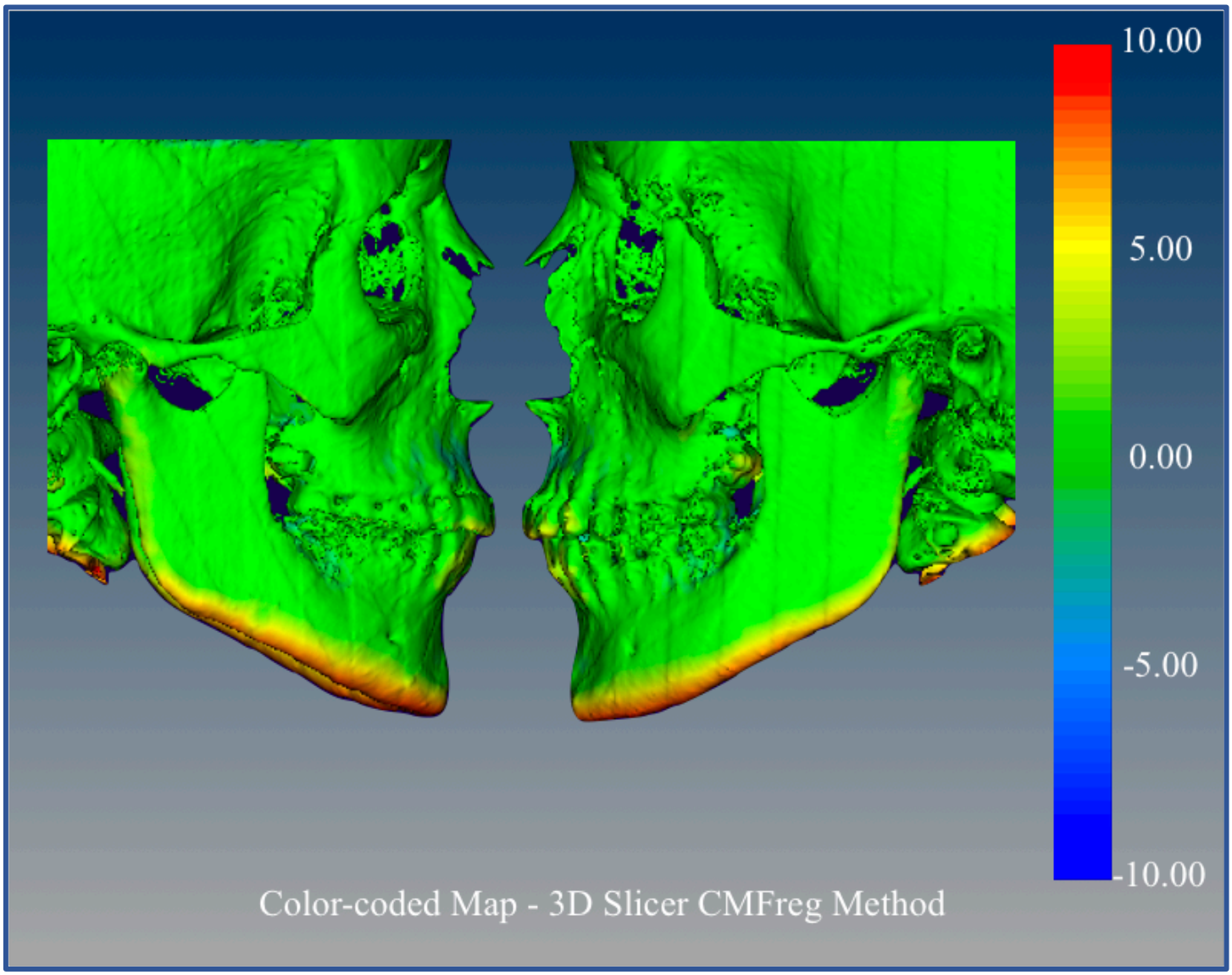


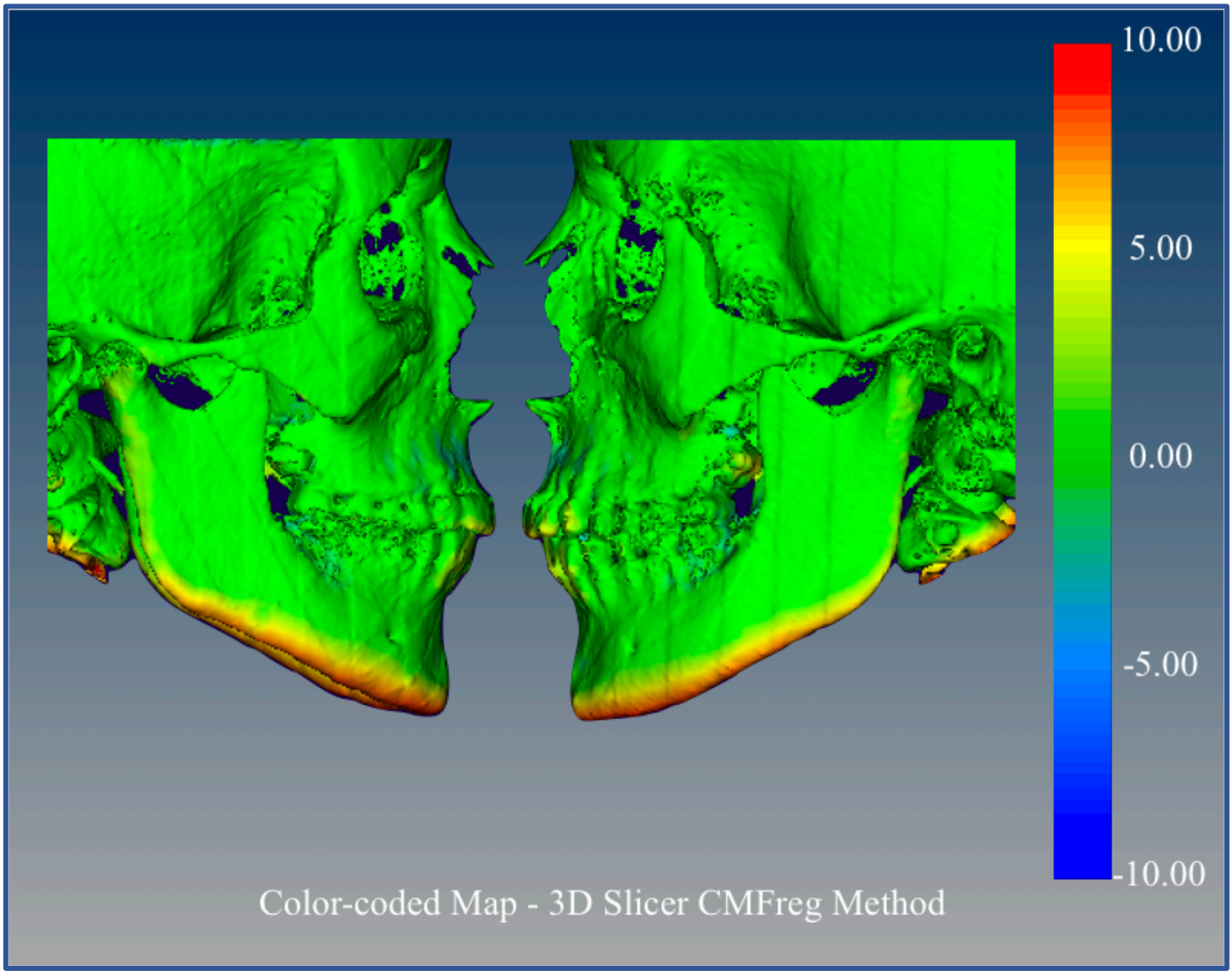
Appendix 3.6.5 Color-coded map with CMFreg/Slicer method

Case #1

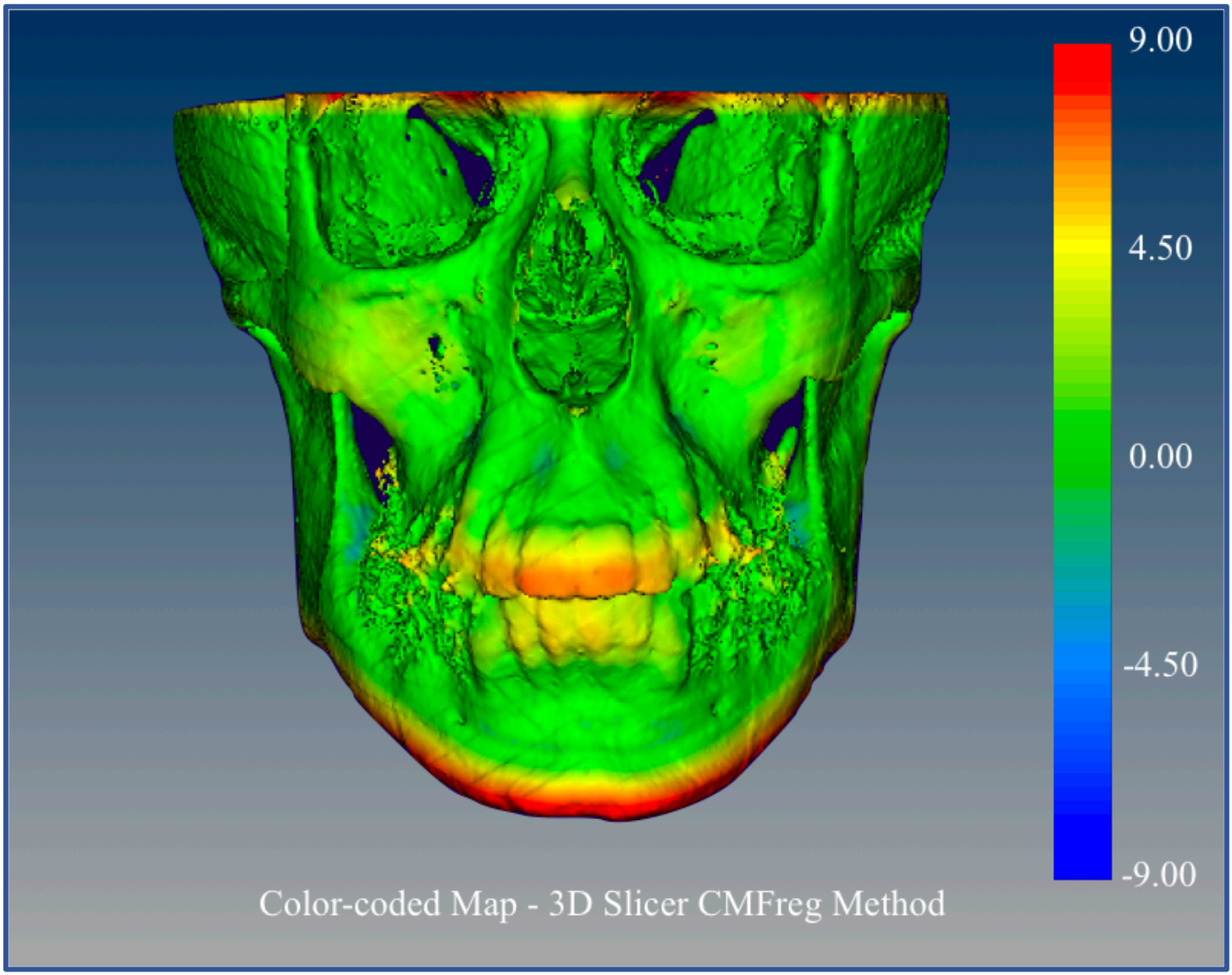


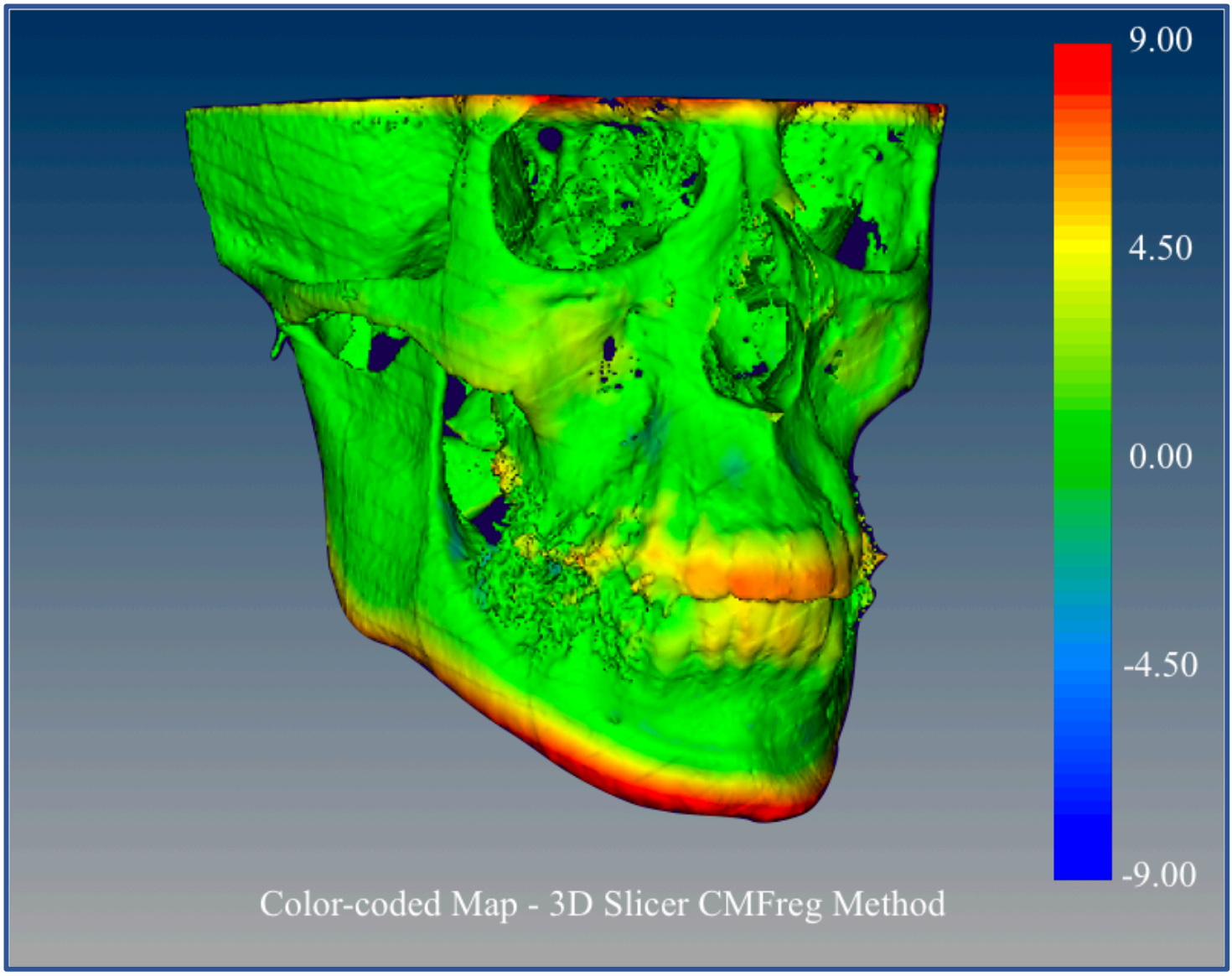

9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer CMFreg Method

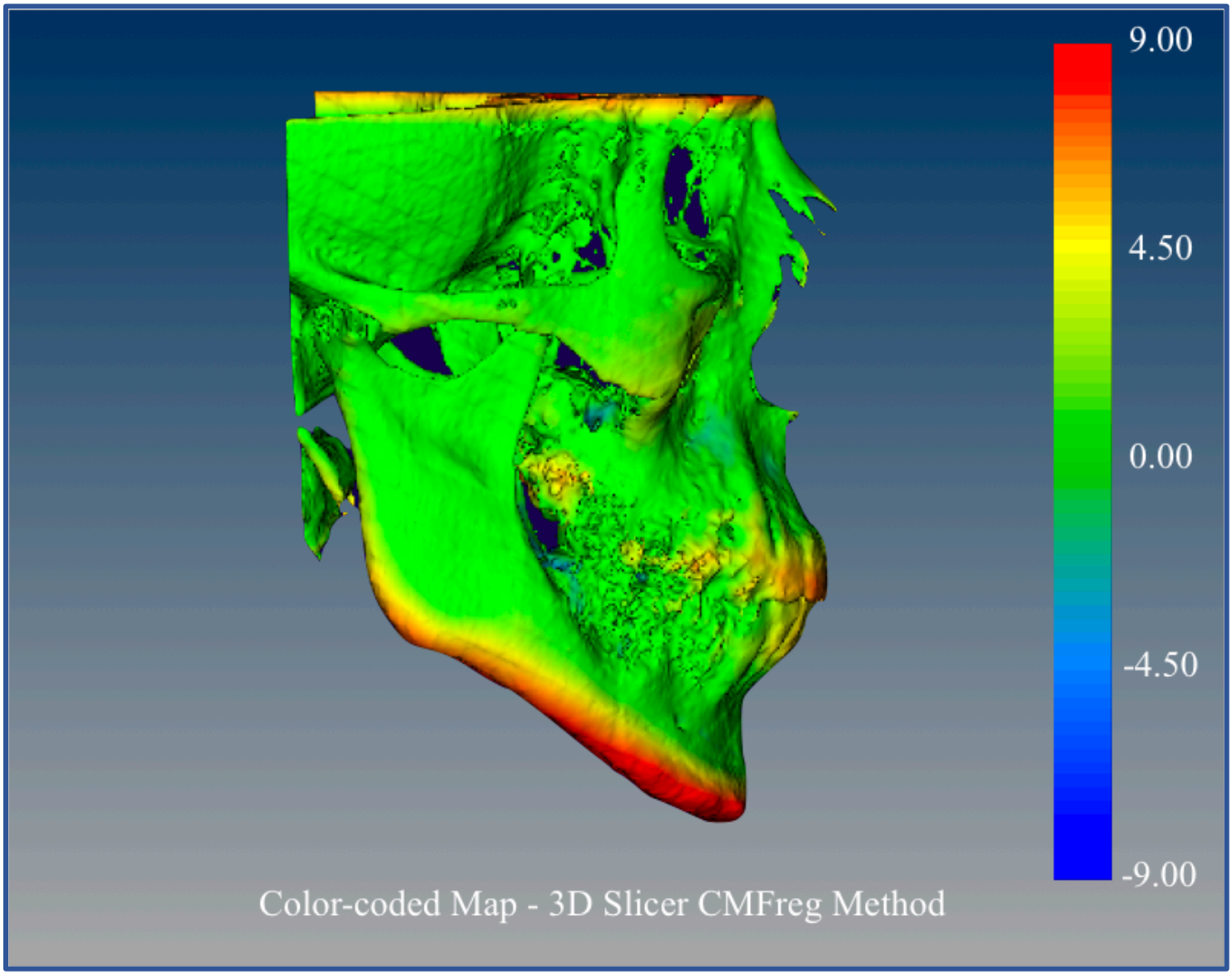

10.00
5.00
0.00
-5.00
-10.00
Color-coded Map - 3D Slicer CMFreg Method

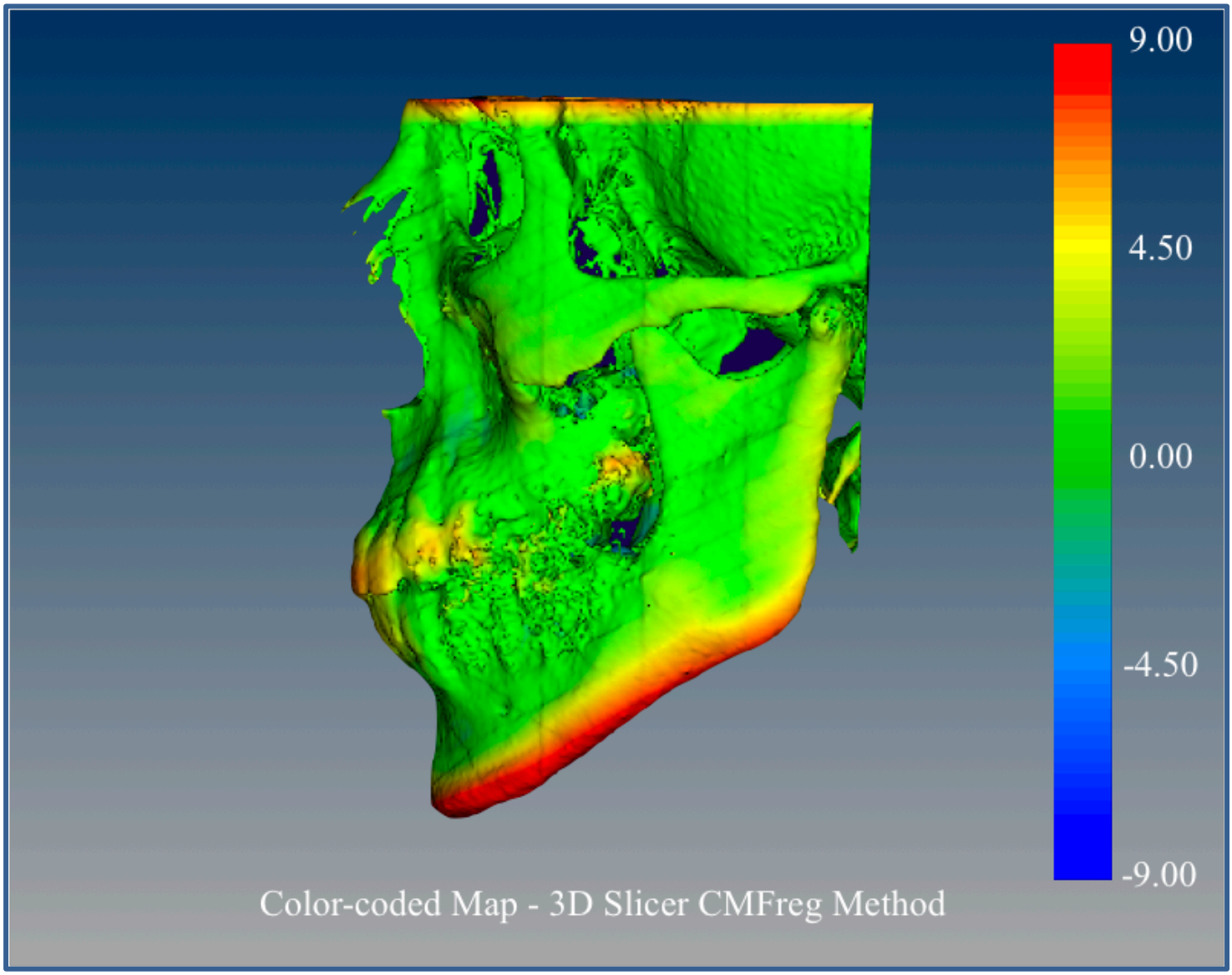

10.00
5.00
0.00
-5.00
-10.00
Color-coded Map - 3D Slicer CMFreg Method

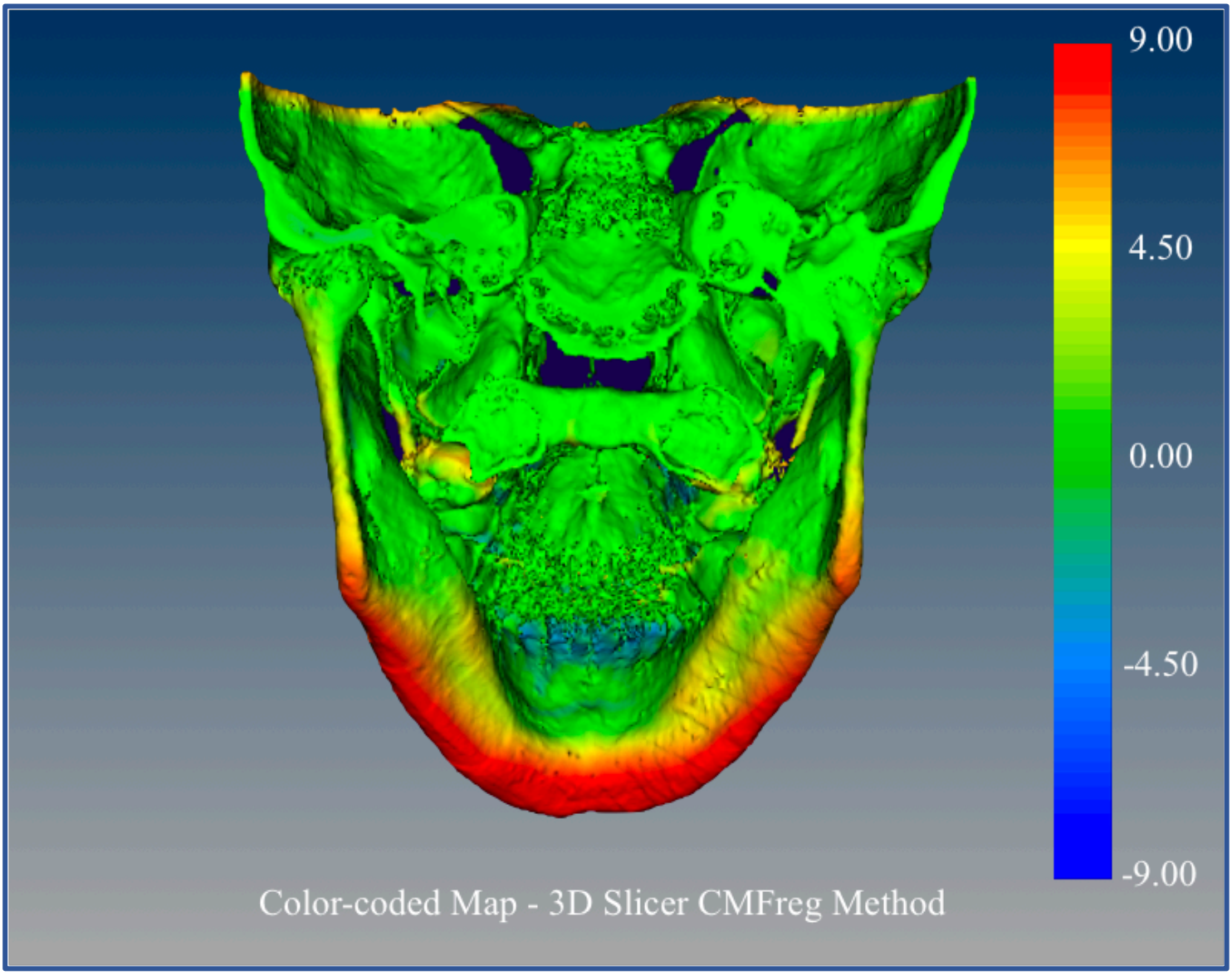
Case #2



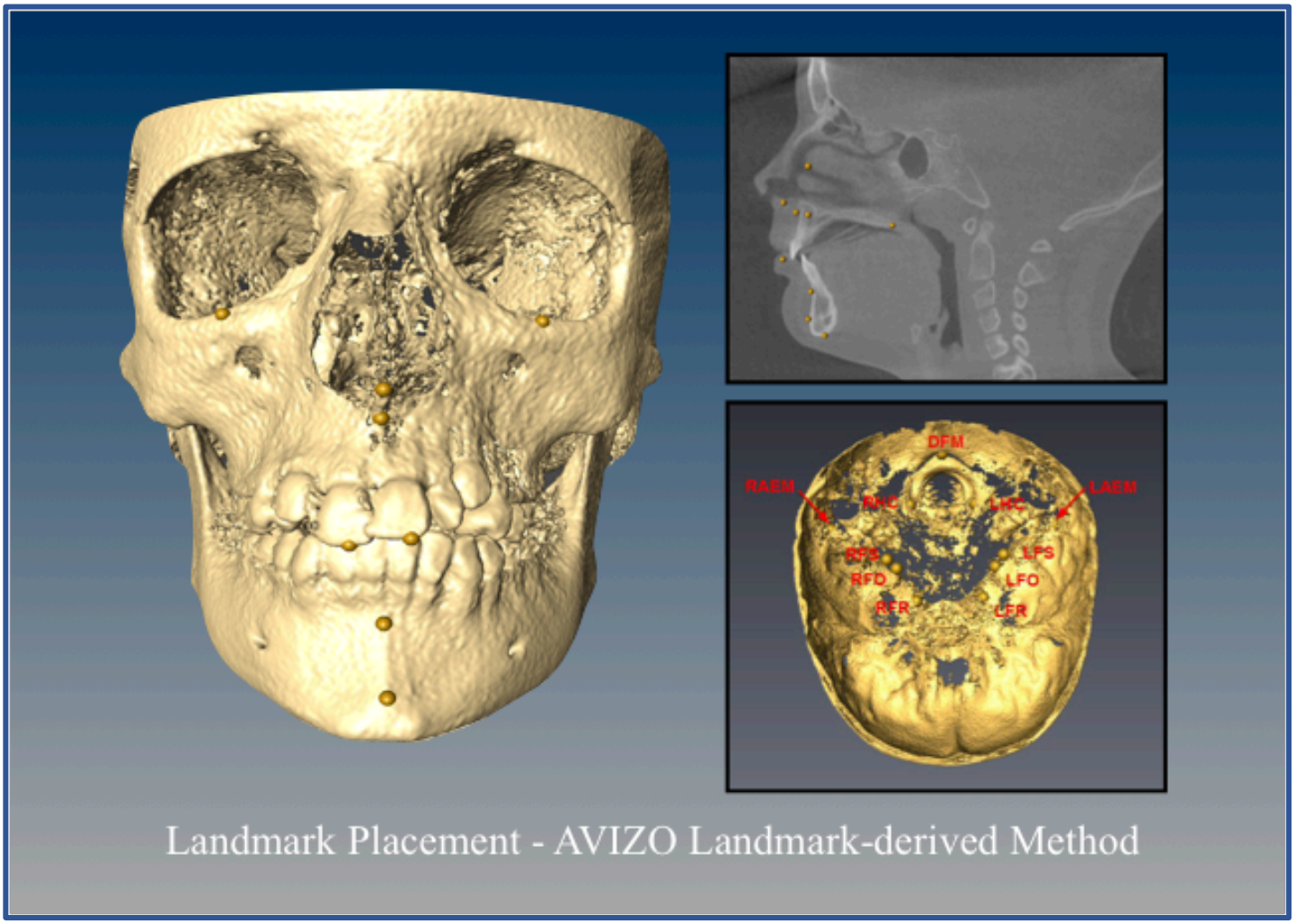

9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer CMFreg Method


9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer CMFreg Method

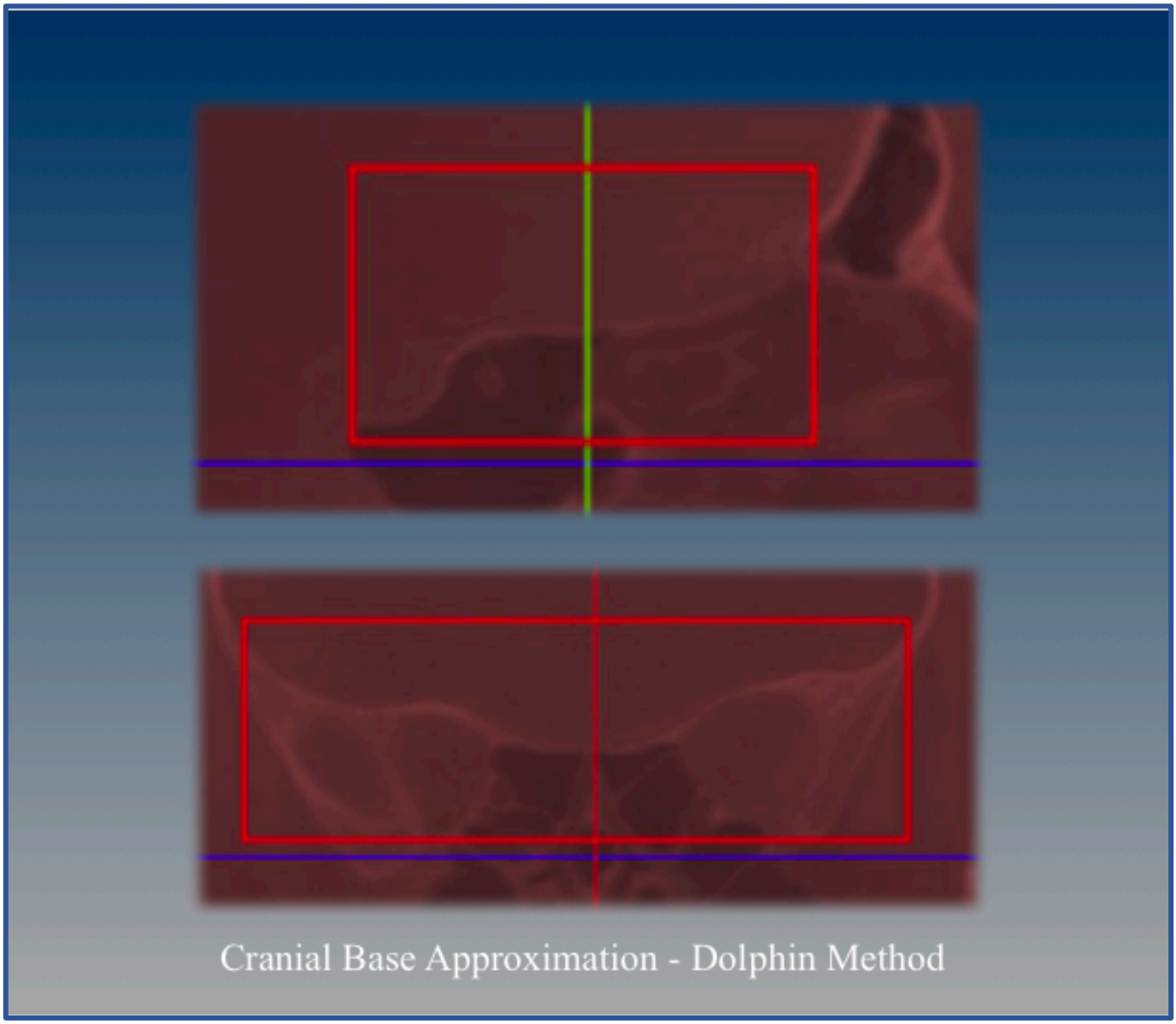

9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer CMFreg Method


9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer CMFreg Method

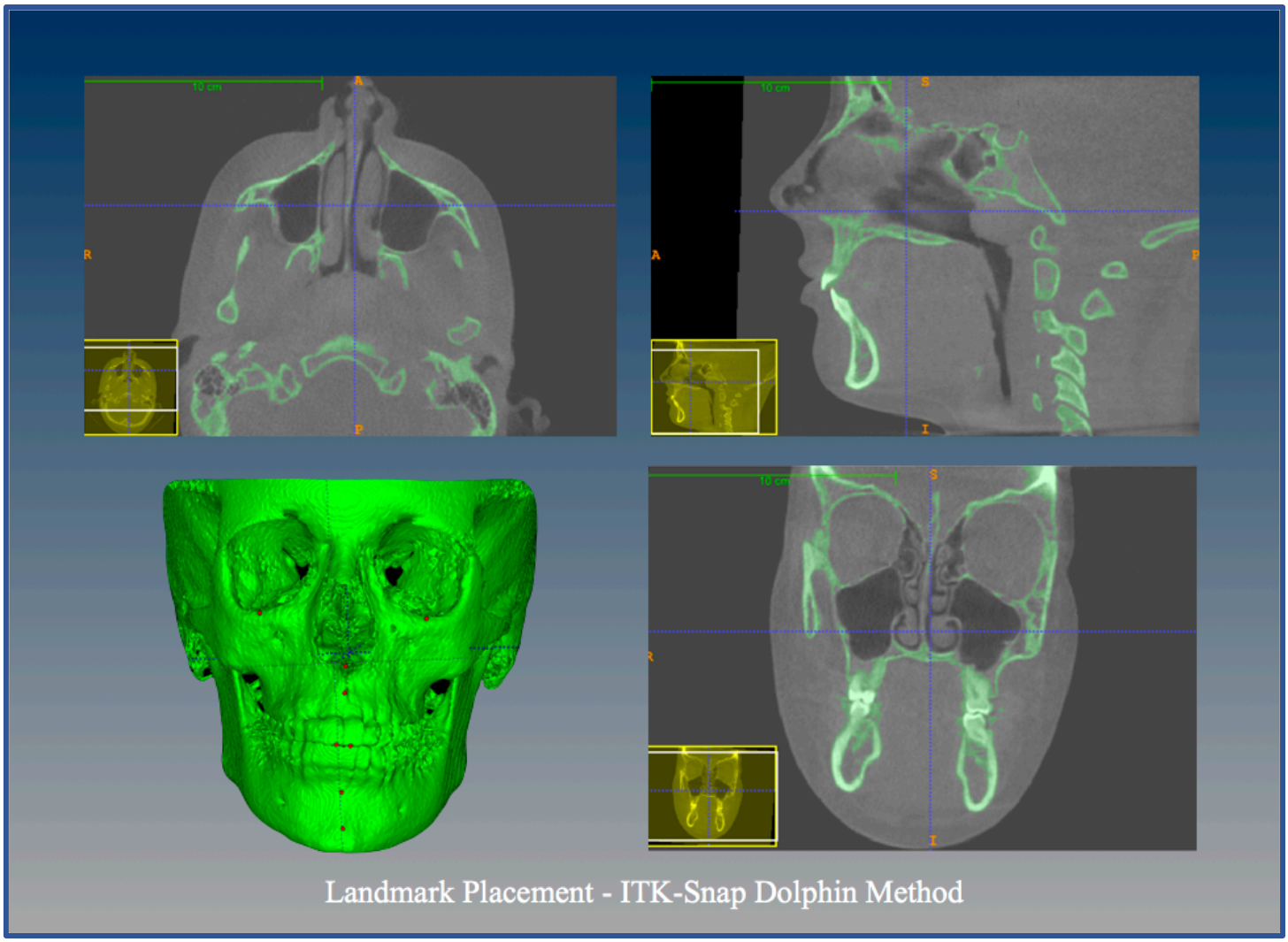
Appendix 3.6.6 Landmark placement with landmark-derived method



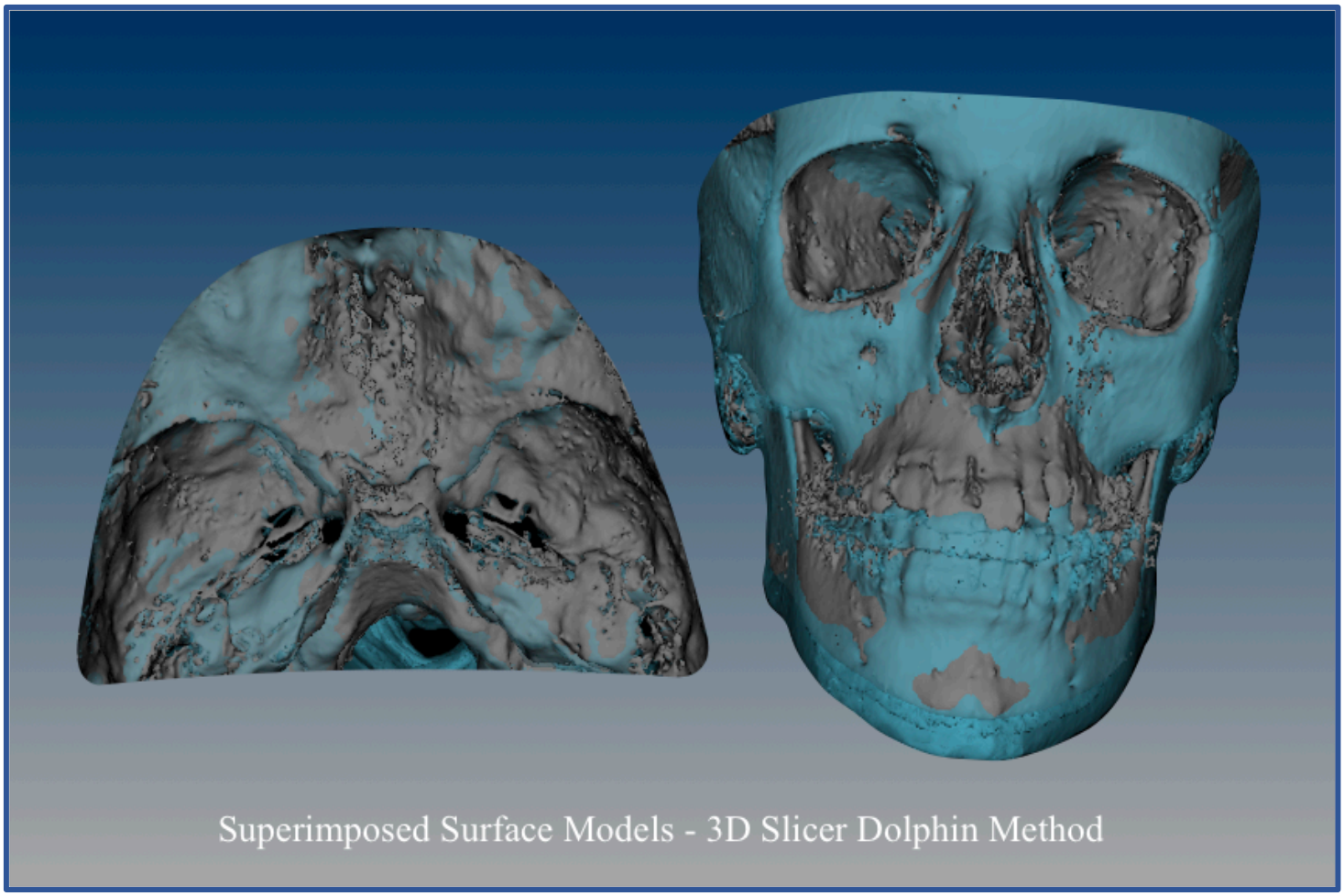
Appendix 3.6.7 Cranial base approximation with Dolphin method



Appendix 3.6.8 Landmark placement with Dolphin method

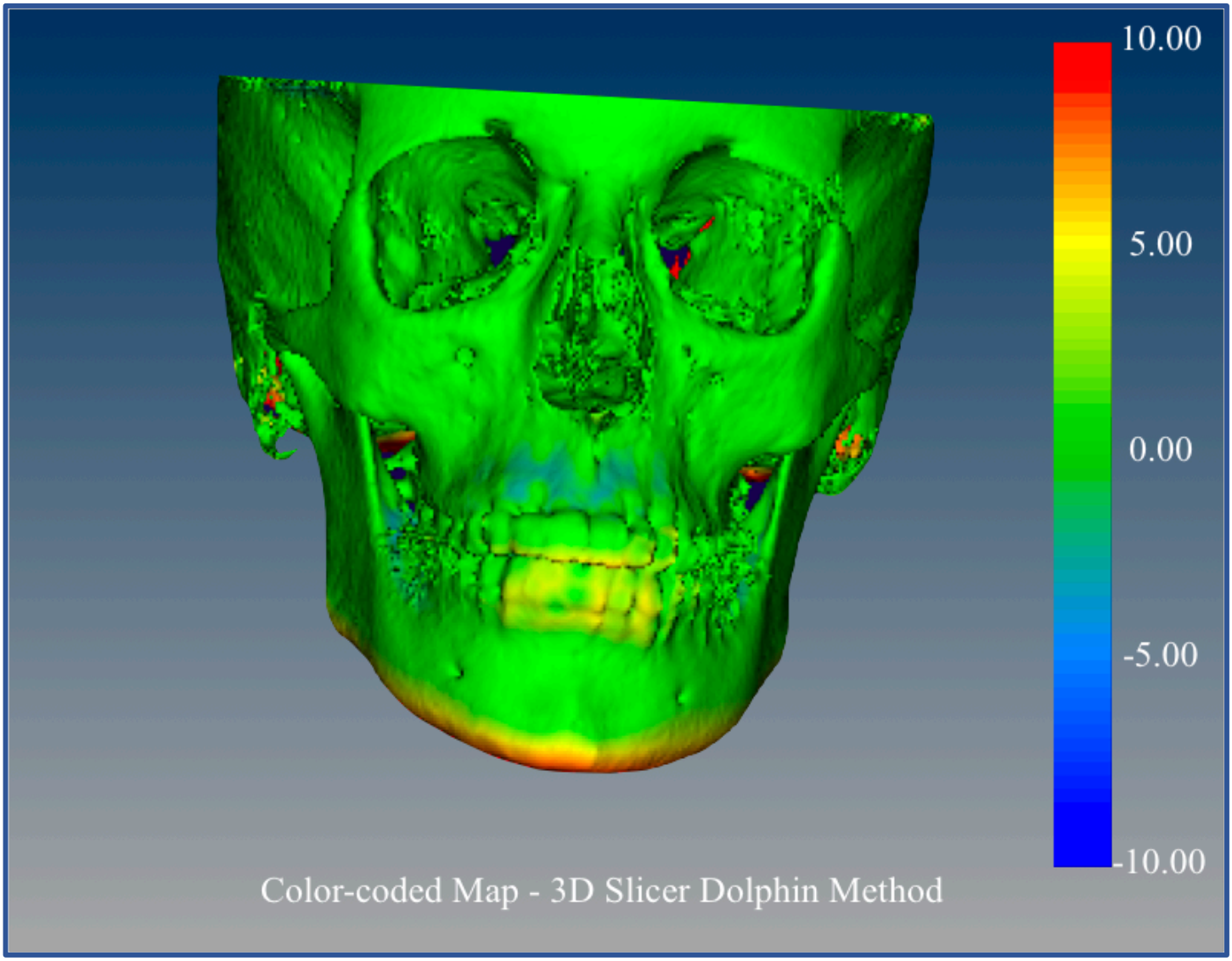


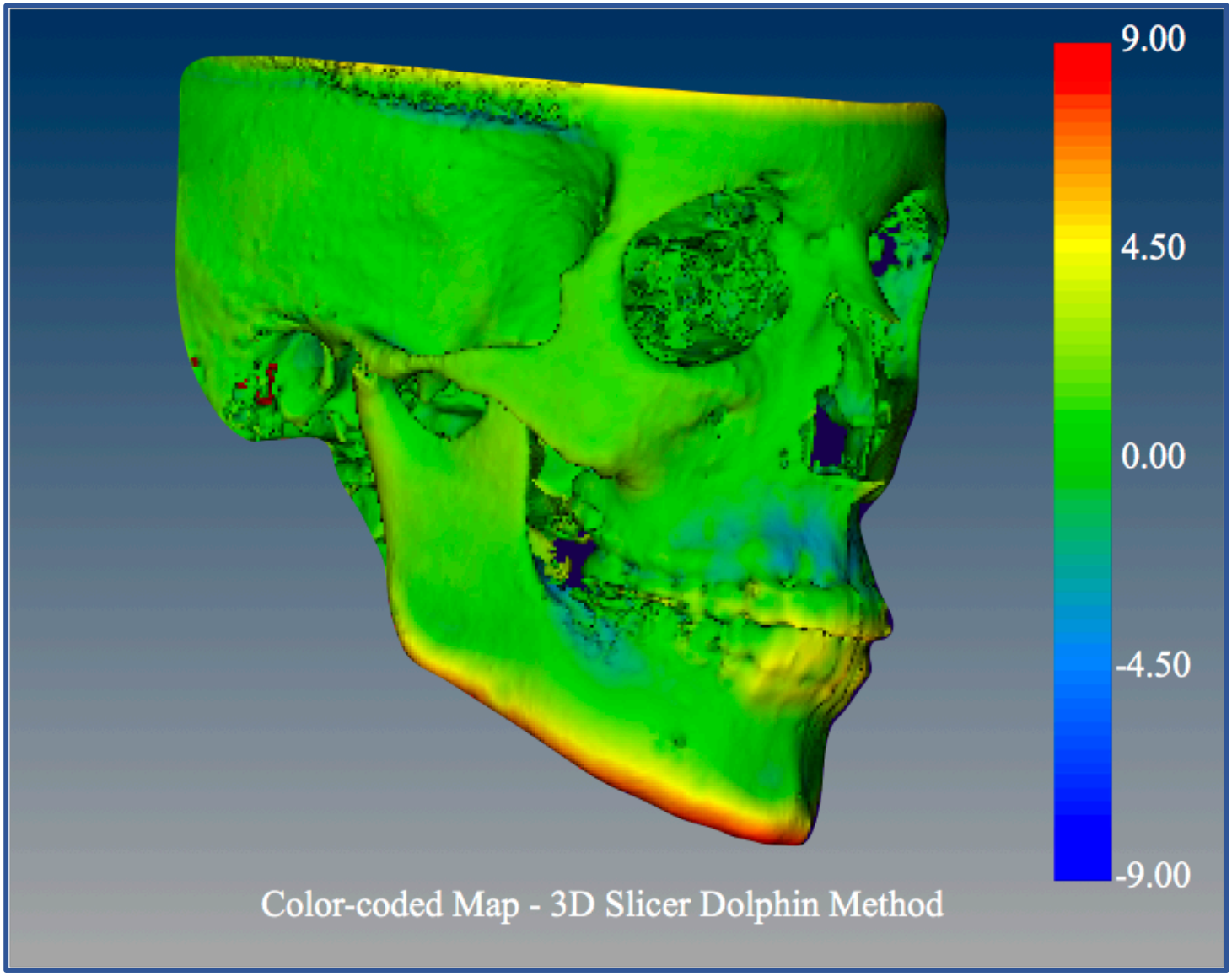
Appendix 3.6.9 Superimposed surface models with Dolphin method



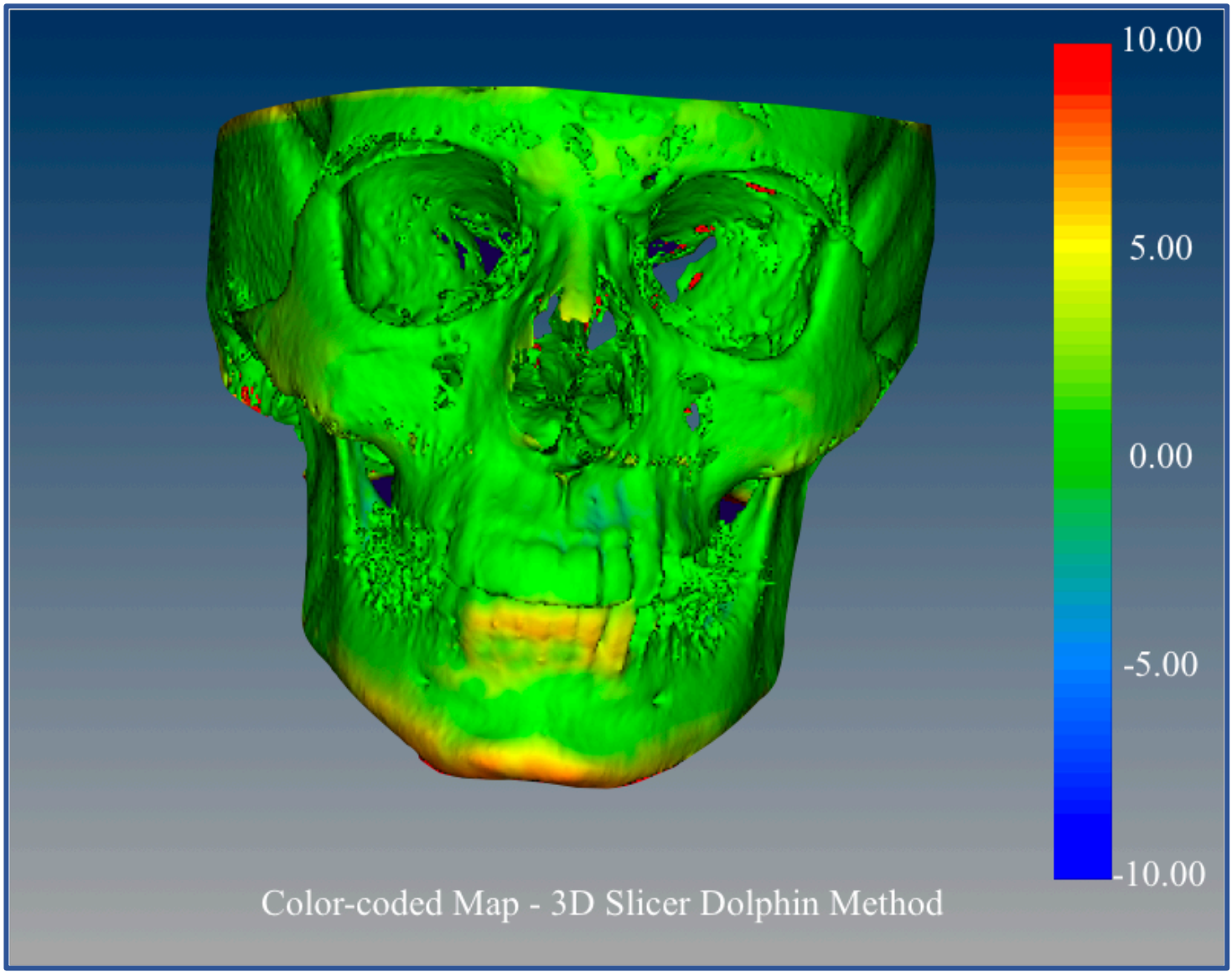
Appendix 3.6.10 Color-coded maps with Dolphin method

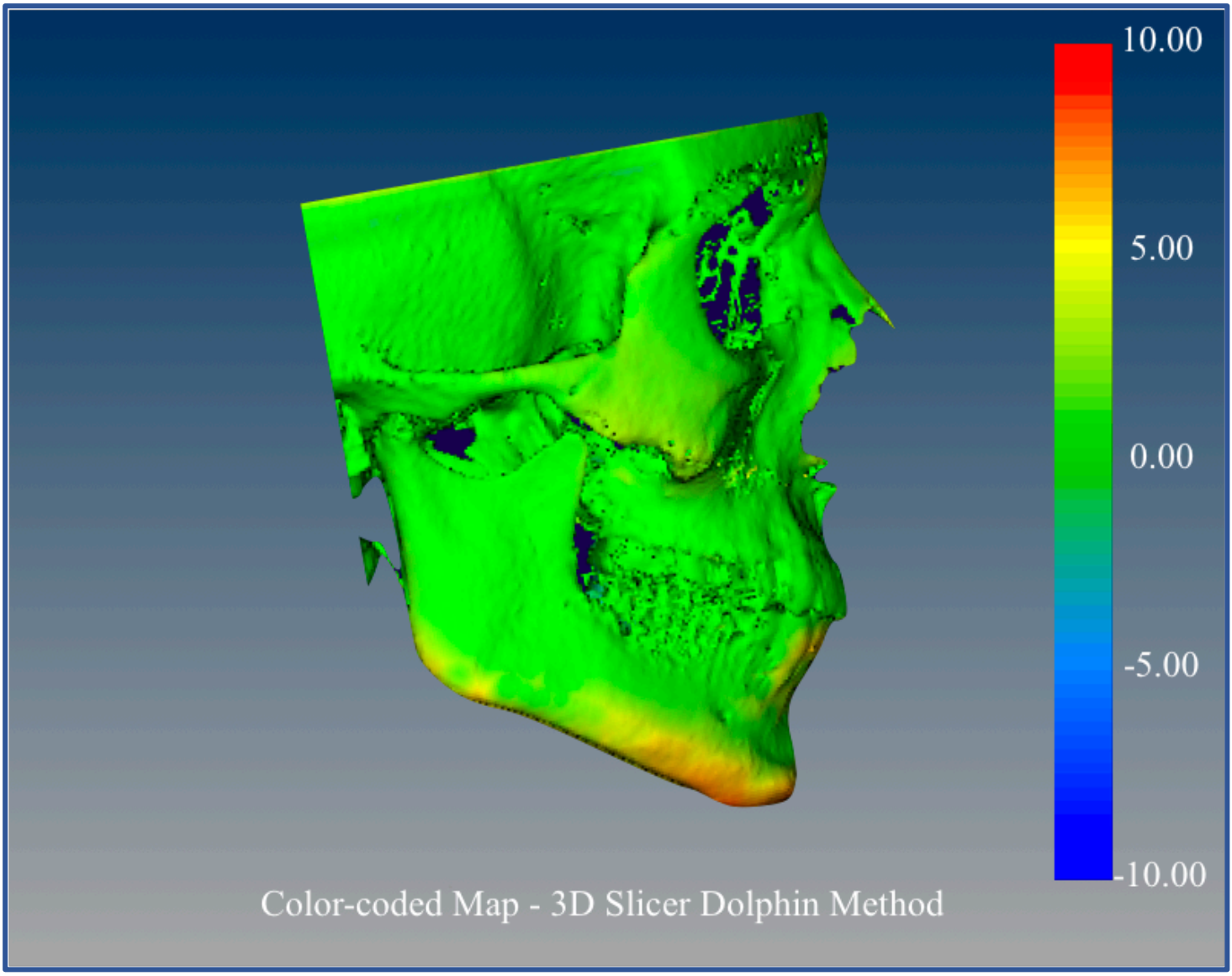
Case #1




9.00
4.50
0.00
-4.50
-9.00
Color-coded Map - 3D Slicer Dolphin Method

Case #2


10.00
5.00
0.00
-5.00
-10.00
Color-coded Map - 3D Slicer Dolphin Method


10.00
5.00
0.00
-5.00
-10.00
Color-coded Map - 3D Slicer Dolphin Method

3.10 References

1. Jacobson A, Jacobson R. Radiographic Cephalometry. Second Edition ed: Quintessence; 2006.
2. Duterloo H, Planché P. Handbook of cephalometric superimposition Hanover Park, IL : Quintessence Pub., c2011.; 2011.
3. Ghafari J, Engel FE, Laster LL. Cephalometric superimposition on the cranial base: a review and a comparison of four methods. Am J Orthod Dentofacial Orthop 1987;91(5):403-13.
4. Moore WJ, Lavelle CLB. Growth of the facial skeleton in the Hominoidea. New York: Academic Press 1974.
5. Bambha JK. Longitudinal cephalometric roentgenographic study of the face and cranium in relation to body height. J Am Dent Assoc 1961;63:776-99.
6. Björk A. The relationship of the jaws to the cranium. In: Lund-Strom A, ed. Introduction to orthodontics. New York: McGraw Hill Book Company; 1960:104-140.
7. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.
8. Björk A. The Use of Metallic Implants in the Study of Facial Growth in Children : Method and Application. Am. J. Phys. Anthrop 1968;29:243-54.
9. Bialocerkowski A, Klupp N, Bragge P. How to read and critically appraise a reliability article. International Journal of Therapy and Rehabilitation 2010;17(3):114-20.
10. Kimberlin C, Winterstein A. Validity and reliability of measurement intruments used in research. Am J Health-Syst Pharm 2008;65:2276-84.
11. Bartlett J, Frost C. Reliability, repeatibility and reproducibility: analysis of measuremement errors in continuous variables. Ultrasound Obstet Gynecol 2008;31:466-75.
12. Leonardi R, Giordano D, Maioranac F, Spampinato C. Automatic Cephalometric Analysis. Angle Orthodontist 2008;78(1):145-51.
13. Houston WJ. The analysis of errors in orthodontic measurements. Am J Orthod 1983;83:382-90.
14. Houston WJ, Maher RE, McElroy D, Sherriff M. Sources of error in measurements from cephalometric radiographs. Eur J Orthod 1986;8:149-51.
15. Kamoen A, Dermaut L, Verbeeck R. The clinical significance of error measurement in the interpretation of treatment results Eur J Orthod 2001;23:569-78.
16. Miller RL, Dijkman DJ, Riolo ML, Moyers RE. Graphic computerization of cephalometric data. J Dent Res 1971;50:1363-69.

17. Savage AW, Showfety KJ, Yancey J. Repeated measures analysis of geometrically constructed and directly determined cephalometric points. Am J Orthod Dentofacial Orthop 1987;91:295-99.

18. Cavalcanti MG, Yang J, Ruprecht A, Vannier MW. Accurate linear measurements in the anterior maxilla using orthoradially reformatted spiral computed tomography. Dentomaxillofac Radiol 1999;28:137-40.

19. Lagravère M, Major P. Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography. AJODO 2005;128:657-60.

20. Lagravère MO, Low C, Flores-Mir C, Carey J, et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. AJODO 2010;137(5):598-604.

21. Ruellas AE, Cevidanes LH, Phillips C, et al. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009;107(2):256-65.

22. Lagravère M, Gordon J, Flores-Mir C, et al. Cranial base foramen location accuracy and reliability in cone-beam computerized tomography. AJODO 2011;139:e203–e10.

23. Naji P, Alsufyani N, Lagravere M. Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist 2014;84:762-72.

24. Ludlow JB, Gubler M, Cevidanes L, et al. Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views. Am J Orthod Dentofacial Orthop 2009;136:e311-13.

25. Lou L, Lagravere MO, Compton S, Major PW, Flores-Mir C. Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007;104:402-11.

26. Berkowitz S. A multicenter retrospective 3D study of serial complete unilateral cleft lip and palate and complete bilateral cleft lip and palate casts to evaluate treatment: part 1—the participating institutions and research aims. Cleft Palate Craniofac J 1999;36:413-24.

27. Adams G, Gansky S, Miller A, et al. Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls. Am J Orthod Dentofacial Orthop 2004;126:397-409.

28. Halazonetis D. From 2-dimensional cephalograms to 3-dimensional computed tomography scans. Am J Orthod Dentofacial Orthop 2005;127:627-37.

29. Grauer D, Cevidanes L, Proffit W. Working with DICOM craniofacial images. AJODO 2009;136:460-70.

30. Terajima M, Yanagita N, Ozeki K, et al. Three dimensional analysis system for orthognathic surgery patients with jaw deformities. AJODO 2008;134:100-11.

31. Park JH, Tai K, Owtad P. 3-Dimensional Cone-Beam Computed Tomography Superimposition: A review. Seminars in Orthodontics 2015;21(4):263-73.

32. Steuer I. The cranial base for superimposition of lateral cephalometric radiographs. Am J Orthod 1972;61(5):493-500.

33. Afrand M. Anterior and middle cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2015.

34. Weissheimer A, Menezes L, Koerich L, Pham J, Cevidanes L. Fast three-dimensional superimposition of cone beam computed tomography for orthopaedics and orthognathic surgery evaluation. Int. J. Oral Maxillofac. Surg. 2015;44:1188-96.

35. Cevidanes LH, Motta A, Proffit WR, Ackerman JL, Stynere M. Cranial base superimposition for 3-dimensional evaluation of soft-tissue changes. AJODO 2010;137(4 Suppl):S120-S29.

36. Cevidanes LH, Bailey L'TJ, Tucker SF, Styner MA, et al. Three-dimensional cone-beam computed tomography for assessment of mandibular changes after orthognathic surgery. AJODO 2007;131:44-50.

37. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

38. Lagravère M, Secanell M, Major P, Carey J. Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;111:771-77.

39. Redmond R, Choi J, Mah J. A New Method for Superimposition of CBCT Scans. JCO 2010;XLIV(5):303-12.

40. Viola P, Wells W. Alignment by maximization of mutual information. Paper presented at: Fifth International Conference on Computer Vision, 1995.

41. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

42. Gkantidis N, Schauseil M, Pazera P, Zorkun B, et al. Evaluation of 3-Dimensional Superimposition Techniques on Various Skeletal Structures of the Head Using Surface Models. PLoS ONE 2015;10(2):1-20.

43. Ruellas AC, Tonello C, Gomes LR, et al. Common 3-dimensional coordinate system for assessment of directional changes. . Am J Orthod Dentofacial Orthop 2016;149(5):645-56.

44. Lagravère M, Major P, Carey J. Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images. Dentomaxillofacial Radiology 2010;39:400-08.

45. Portney L WM. Foundations of clinical research: applications to practice: Prentice Hall, Upper Saddle River; 2008.

46. Cevidanes LH, Bailey LJ, Tucker Jr. GR, Styner MA, et al. Superimposition of 3D cone-beam CT models of orthognathic surgery patients. Dentomaxillofacial Radiology 2005;34:369-75.

47. Cevidanes LH, Styner MA, Proffit WR. Image analysis and superimposition of 3-dimensional cone-beam computed tomography models. AJODO 2006;129:611-18.

48. Bazina M, Cevidanes L, Ruellas A, et al. Precision and reliability of Dolphin 3-dimensional voxel-based superimposition. Am J Orthod Dentofacial Orthop 2018;153:599-606.

49. Lagravère M, Hansen L, Harzer W, Major P. Plane orientation for standardization in 3-dimensional cephalometric analysis with computerized tomography imaging. AJODO 2006;129:601-04.

50. Danforth RA DI, Mah J. 3-D volume imaging for dentistry: a new dimension. J Calif Dent Assoc 2003;31:817-23.

51. Danforth RA PJ, Hall P. Cone beam volume tomography: an imaging option for diagnosis of complex mandibular third molar anatomical relationships. J Calif Dent Assoc 2003;31:847-52.

52. Nada R, Maal T, Breuning K, Berge S, et al. 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(2):e16520.

53. Nguyen T, Cevidanes L, George W. Validation of 3D mandibular regional superimposition methods for growing patients (Abstract). J Dent Res 2014;93 (Spec Iss B):784.

54. Ruellas A, Huanca L, Gomes MT, et al. Comparison and Reproducibility of Two Maxillary Regional Registration Methods. J Dent Res 2015;94(Spec Iss A):0915.

CHAPTER 4

Analysis of overall 3D maxillo-mandibular changes using three 3D superimposition methods

4.1 Introduction:

Cephalometric superimposition is known to be critically important when assessing growth and orthodontic-orthopedic treatment response.[1-3] Cephalometric tracings should be carefully superimposed in order to provide a reliable assessment of orthodontic/growth structural changes.[3] Longitudinal changes in craniofacial morphology caused by growth and treatment response can be measured by superimposing a series of lateral cephalograms, using relatively stable landmarks, such as the anterior cranial base, as reference. Studies have shown that about 90% of the anterior cranial base growth is completed by the age of 4-5 years and before other craniofacial structures.[4-6] Therefore, the anterior cranial base is considered a satisfactory reference for cephalometric superimpositions.[7]

The antero-posterior length of the pre-sphenoid region was reported to be stable after the age of 7 as assessed by cross-sectional and longitudinal studies when the spheno-ethmoidal synchondrosis ceases to grow.[4, 8-10] After that time a number of structures especially those associated with neural tissues remain stable and can be relied upon for superimposition. The most important structures for cranial base superimposition are the anterior wall of sella turcica below of the anterior clinoid process, which is stable after 5-6 years of age and the cribriform plate of the ethmoid bone, which is stable after 4-5 years of age. Other structures relative stable after 7 years of age are the ethmoidal crest that has shown to grow only until age of 6, the cerebral surfaces of the frontal bone associated with the orbits and the greater wings of the sphenoid bone.[11]

A 5-year longitudinal study assessing exclusively the growth of planum sphenoidale to dorsum sella reported its reasonable stability to be used for superimposition, but the same study

mentioned that the hypophyseal fossa deepened in a small sample of subjects who were observed for longer than 5 years.[5] It is also known that although the jugum sphenoidale shows minimal growth after 6 years of age bony apposition can occur in some cases up to 14 years of age. Bjork[12] also observed remodeling of sella turcica during growth, resulting in displacement of sella downward and backward. He recorded an elevation of the tuberculum sella in relation to other structures of the anterior cranial fossa.[4, 8]

The posterior cranial base is problematic due to growth that occurs during childhood and adolescence at the spheno-occipital synchondrosis and the remodeling that occurs at the surfaces of the occipital and posterior sphenoid bones including the posterior wall of sella. The anterior margin of the foramen magnum shows apposition until 16 years of age, the sphenoid-occipital synchondrosis initiates osseous connection at 12-15 years of age and the posterior sella goes under resorption until late teens.[11, 13]

The need to assess the behavior of craniofacial structures in response to orthodontic treatment and continuous growth and development has motivated the development of numerous superimposition techniques.[14-18] Different anatomical structures, cephalometric landmarks, lines and planes of reference have been used for this purpose, allowing quantitative analysis of growth and treatment based on changes of the facial skeleton of a particular individual over a period of time.[1, 16-18]

The Broadbent triangle (Na-S-Bo) and its registration point R were among the first anatomical areas used for superimpositions to assess overall changes. Thereafter, superimposition orienting the two tracings on the *Sella-Nasion* line with registration at *Sella* was proposed by the American Board of Orthodontics.[19] The major disadvantage of these

superimposition techniques is that they include regions of the cranial base that continue to remodel to a certain small degree during most of the growing years. Growth at the spheno-occipital synchondrosis as well as bone remodeling at *Nasion* and *Sella* are responsible for these changes. In 1955, Coben[20] presented the *Basion* Horizontal concept, with the *Basion* Horizontal being a plane constructed at the level of the anterior border of the foramen magnum parallel to Frankfort horizontal. Here, *Basion* was used as the point of reference for the analysis of craniofacial growth. The use of *Basion-Nasion* plane as an area of registration for overall evaluation of the dentofacial changes was suggested later by Ricketts et al.[21] However, as Melsen[9] observed the position of *Basion* is influenced by the remodeling processes on the surface of the clivus and on the anterior border of the foramen magnum, as well as by displacement of the occipital bone, which is associated with the growth in the spheno-occipital synchondrosis. Because *Nasion*, *Sella,* and *Basion* move during growth, the methods of overall superimposition on S-Na or Ba-Na lines have shown a low degree of validity.[15]

In addition to quantitative information, cephalometric superimpositions can provide relevant qualitative data. However, it is critical to have it obtained from serial lateral cephalograms taken under identical conditions of magnification, head posture, and radiological exposure; furthermore, the cephalometric tracings, as well as the superimpositions, must be accurate.[15]

In order to overcome some of the 2D imaging related-issues, there has been a dramatic increase in the use of CBCT in orthodontics and in dentistry over the last decade. CBCT allows assessment of three-dimensional dental, skeletal, and soft-tissue variations for both growing and non-growing patients.[22]

One of the key advantages of CBCT over 2D radiography is its ability to provide 3D volumetric, surface and sectional information about the craniofacial structures. This allowed orthodontists and researchers in the specialty to overcome the considerable limitations of 2D radiographs, including geometric distortion, alignment of the imaging device, magnification, overlapping structures and inadequate head position.[23] In the initial stages of introducing CBCT for orthodontic purposes, there was a tendency to convert the 3D data set to a 2D image since analyses of the images in this format were the only methods known for evaluating relationships of the structures in the maxillofacial complex.[24, 25] As it was not the optimal approach, researchers took up the challenge of developing new methods to measure 3D angles, distances, and volumes and to superimpose CBCT images.[26-28]

Hence, superimposition techniques using 3D images have been evolving since the first introduction in the field. Their application has been tested and proved extensively in orthognathic surgery to show surgical outcomes. This possibility opens an extraordinary clinical field in growing patients for a 3D follow-up of craniofacial changes and assessment of treatment results providing a complete analysis in all three planes of space.

To date, only one study[29] has tested a method to superimpose 3D images to determine craniofacial changes in growing patients. Therefore, a thorough assessment of superimposition methods of 3D images in growing patients is required. Our study will test three methods for cranial base superimposition using CBCT scans, two voxel-based (CMFreg/Slicer and Dolphin) and one landmark-derived method. The results of the current study will assess the overall 3D changes generated by each method and determine which method produces less measurement error.

4.2 Material and Methods

4.2.1 Study population:

A retrospective, observational longitudinal study was carried out on individuals receiving comprehensive orthodontic treatment at the University of Alberta. Thirty-six patients were selected from a population of teenagers from 11 to 14 years.

A total of eighty-two patients who had initial (pre-orthodontic treatment) and final (post-orthodontic treatment) full FOV CBCT images generated between 2010 and 2015 were selected. Thirty-four individuals with a CVM stage above 5 at pre-treatment and twelve patients with either pre or post-treatment CBCT images with lack of complete cranial base visibility were excluded. Thus, a final sample of thirty-six patients was included.

The mean age of patients at the time of the initial CBCT was 12.4 ± 0.9 years (Cervical Vertebrae Maturation index [CVM] stage 3-4). The mean age at final CBCT was 14.3 ± 0.8 years. The sample included seventeen males and nineteen females.

The interval between T1 and T2 ranged from 22 - 25 months apart. Fourteen patients presented Class I malocclusion, eight patients presented mild Class II malocclusion and fourteen patients presented mild Class III malocclusion. All patients received a non-extraction treatment and included rapid maxillary expansion, full fixed appliances, and intermaxillary elastics.

Patients with any noted syndromic anomalies were excluded as well as CBCT imaging taken with machines other than iCAT Volumetric Scanner.

This study only analyzed gathered data from patients that participated in clinical trials. No additional imaging was requested. Ethics approval was previously obtained by the Institutional Health Research Ethics Board at the University of Alberta.

4.2.2 Data Collection:

CBCT volumetric data were taken using the iCAT Volumetric Scanner at 120 kV, 5mA and 8.9sec. Images were obtained and converted to Digital Imaging and Communications in Medicine (DICOM) format using the iCAT software with a voxel size of 0.3mm. Utilizing AVIZO (Visualization Sciences Group, Burlington, MA), ITK-Snap (Penn Image Computing and Science Laboratory at the University of Pennsylvania), 3D SLICER (Open Platform for the Medical Image Computing Community) and DOLPHIN software, the DICOM format images were rendered into a volumetric image. Sagittal, axial and coronal volumetric slices, as well as the 3D image reconstructions, were used to determine the landmark positions, required for all three methods.

Two images, pre-treatment and post-treatment separated by 22-25 months interval were analyzed for each patient. Analysis of the images was carried out by one researcher using the respective superimposition techniques (CMFreg/Slicer, Dolphin and landmark-derived). Extensive training was required prior to superimposing with each method. CMFreg/slicer demanded more training time due to multiple steps, followed by Dolphin and the landmark-derived method. Fourteen landmarks, used in previous studies, were marked on three-dimensional images at T1 and T2 with each of the three methods to assess reliability. (Tables 3.1 and 3.2)

Voxel-based CMFreg/Slicer Method:

This method uses two different open-source programs ITK-Snap (http://www.itksnap.org) and 3D Slicer (http://www.slicer.org). Using ITK-Snap software program (version 2.0.0) pre-treatment and post-treatment DICOM files were opened and converted to GIPL (Guys Imaging Processing Lab) format for easy processing. Segmentations then were created using the GIPL.GZ files for both pre and post treatment scans using the intensity segmenter function on 3D Slicer software program (version 4.7.0) to construct 3D volumetric label maps. The same range intensity level was adequate for the whole sample, mainly due to the same machine and machine settings used for all scans and similar bone density among the subjects. Once the segmentations were completed and saved, ITK-Snap was used to segment the area of the cranial base to be used as a reference for the superimposition using semi-automatic segmentation. Both scan and segmentation were downloaded so the area of the cranial base to be segmented could be manually "painted" using the software. The software then automatically removed the lower density, and the non-painted areas, leaving only the cranial base. At this point we have a complete T1 scan, a T1 segmented cranial base, a complete T2 scan, and a T2 segmented cranial base. The software then combined each image with its respective cranial base.

Then, surface models were created using the pre-treatment scan and segmentation in 3D Slicer to re-orient the head. Using the Transform function pre and post-treatment images were reoriented on the sagittal plane utilizing Foramen Magnum, Crista Galli and Glabella, on the vertical plane using Frankfort horizontal (Porion-Orbitale) and on the transverse plane using Porion to Porion as references.

Once the head orientation step was completed,[30] the post-treatment image was manually approximated in relation to pre-treatment image using 3D Slicer, a matrix for cranial base approximation and post-treatment cranial base approximation scan and segmentation files were created. Then using ITK-Snap, the post-treatment mask was cropped before the cranial base registration.

The registration (superimposition) of the post-treatment image upon the pre-treatment image was carried out on the segmented cranial base, using the craniomaxillofacial (CMF) tool and the setting growing rigid automatic registration in 3D Slicer. During the superimposition, T2 was reoriented guided by the best fit of the outlines of the anterior cranial base and automatically superimposed on a static T1, creating a registered T2 surface model.

Once the superimposition was completed, the pre-treatment scan and segmentation, as well as the registered post-treatment scan and segmentation, were landmarked using ITK-Snap. 3D landmarks were identified using the three views (axial, sagittal and coronal) for consistency of landmark location.[30]

After placing the defined landmarks to pre and post-images, 3D surface models were created using 3D Slicer for all the levels used in ITK-Snap. These models were utilized to measure the absolute differences between the pre and post-treatment images by utilizing the Q3DC module of Slicer CMF (Quantification of directional changes in each plane of the three planes of space).[30] 3D linear distances between T1 and T2 of corresponding landmarks were quantified in the transversal (x-axis), antero-posterior (y-axis) and vertical (z-axis) direction. (Fig. 3.1 and Appendix 3.5)

Landmark-derived Method

Using AVIZO software, the DICOM files were rendered into a volumetric image using 512 x 512 matrices giving a range of 400–420 DICOM slices. Sagittal, axial and coronal multiplanar slices, as well as the 3D image reconstructions, were used to determine the landmark positions. The principal investigator marked the landmarks using the virtual spherical marker of 0.5mm diameter in the x, y and z-axes. The center of the spherical marker is considered as the position of the landmark by the software, therefore the size of the spherical marker does not affect the position of the landmark.

Given the coordinates of three reference landmarks for a plane, 3D visualization software can compute the plane, however, entering the three-point coordinates usually is a time-consuming repetitive manual process. A similar argument applies to determine the perpendicular distance. In order to resolve this issue, this study reproduced the mathematic procedure in Microsoft Excel. This allowed the reference planes and perpendicular distances to be automatically calculated whenever the landmark coordinates were updated.

Four landmarks were required to define a 3D anatomical reference co-ordinate system. The left and right auditory external meatus (AEML and AEMR, respectively) and the dorsum foramen magnum (DFM) were selected as suggested by previous research. The fourth point, ELSA, defined as the midpoint between the left and right foramen spinosum[31] was selected as the origin of the new Cartesian co-ordinate system. From the origin, 3D positional co-ordinates for the AEML, AEMR and DFM were determined.[32]

The optimization formulation used in this study was the 6-point algorithm, that not only optimizes the location of the same three points (i.e. AEML, AEMR and DFM) as used in the 4-point algorithm but also includes both foramen ovale (right and left) in each image.[33, 34] The addition of two extra landmarks (FOR and FOL) in the optimization analysis was shown to reduce the envelope of error when determining the co-ordinate system.[32]

Once all final coordinates were obtained in AVIZO software, the dataset was exported into a spreadsheet database (Excel 2007, Microsoft, Redmond, Wash). Matlab (Matrix Laboratory) software was then used to generate the optimization.

Once data was optimized, linear distances between the 3-D coordinates were calculated using the following Euclidean distance formula:

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2}$$

With this formula, *d* represented the distance (mm) between two anatomic landmarks, while (X1, Y1, Z1) and (X2, Y2, Z2) are the respective coordinates of any two given landmarks of 3D interest. Each landmark was included in multiple linear measurements of different orientations to be able to assess all dimensions (superior-inferior, anterior-posterior, right-left). (Fig 3.2 and Appendix 3.5)

Voxel-based – Dolphin Method

For each patient, T1 and T2 CBCT images were approximated using 4 landmarks located at the right and left frontozygomatic sutures and the right and left mental foramen and

superimposed on the cranial base using voxel-based superimposition tool in Dolphin 3D (Chatsworth, CA -version 11.8.06.15 premium). The area of the cranial base used for superimposition was defined by a red box in the three different slice views (axial, sagittal and coronal). The superimposition was achieved by moving the T2 image in relation to the T1 image creating a registered T2 image.

Then the slice views (axial, sagittal and coronal) were used to confirm the precision of Dolphin 3D superimposition. Once this step was completed, the registered post-treatment scans were exported as DICOM files and opened in ITK-snap software to convert them into GIPL format similar to the procedure done with the CMFreg/Slicer method. 3D slicer was then used to segment the whole skull using Intensity Segmenter tool, with the same intensity level for all cases to remove any potential error due to the segmentation process. Thus, a surface model of post-treatment segmentation was created for each particular patient. Then pre and post-treatment images were ready for landmarking using ITK-Snap.

After placing the defined landmarks to pre and post-treatment images, 3D surface models were created using 3D Slicer for all the levels used in ITK-Snap. These models were utilized to measure the absolute differences between the pre and post-treatment images by applying the Q3DC function (Quantification of directional changes in each plane of the three planes of space). 3D linear distances between T1 and T2 of corresponding landmarks were quantified in the transversal (x-axis), antero-posterior (y-axis) and vertical (z-axis) direction. (Fig 3.3 and Appendix 3.5)

4.3 Statistical Analysis

The volume data was manually entered into Microsoft Excel 2016 for MAC (Microsoft, Redmond, WA). The SPSS for MAC (Statistical Package for the Social Sciences) version 24 was used to run all statistical tests for this study. For all tests, statistical significance was set at *p*-value of 0.05.

A sample size of 36 individuals was selected without a power analysis due to lack of preliminary study at the time of data collection. 3D changes in the craniofacial complex were assessed by one-way repeated measures analysis of variance (ANOVA) followed by post-hoc analysis. Descriptive statistics were generated for each variable (Appendix 4.1).

Prior to performing the one-way repeated measures ANOVA statistical analysis, the data was checked for model assumptions. All data was checked for normality via Q-Q plots and box plots of the distance of the difference between T1 and T2 of each dependent variable with each method (Appendix 4.3). The assumption of normality was met for most of the linear distances as assessed visually. Although, few linear distances violated this assumption, the one-way repeated measures ANOVA is considered robust to some violations of normality.

The assumption of sphericity was not met as assessed by Mauchly's test of sphericity. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

The Bonferroni post hoc test was applied last for testing all possible pairwise combinations of the superimposition methods. This test provided the statistical significance level (p-value) for each pairwise comparison, as well as the confidence intervals for the mean difference for each comparison. A *p*-value less than 0.05 was considered significant.

4.4 Results

The one-way repeated measures ANOVA revealed evidence of a statistically significant difference between the mean of distances T2-T1 when comparing CMFreg/Slicer method to Landmark-derived method and when comparing Dolphin method to Landmark-derived method in the overall 3D at all dependent variables. Pairwise comparisons/Post hoc analysis with a Bonferroni adjustment results are presented in Table 4.1.

Maxilla

ANS had an overall 3D change of approximately 2.5mm when analyzed by the both voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 6mm. When assessed per components, all the three methods had the same tendency; the vertical change was the most evident, showing an inferior direction, followed by the antero-posterior dimension, with anterior drift. Although very minimum, there was also a change in the right-left dimension. Table 4.2

A-Point had an overall 3D change similar to ANS, 2.5mm on average when analyzed by the two voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 5.8mm. Similar to ANS, when assessed per components, all the three methods had the same tendency; the vertical change was the most evident, showing an inferior direction, followed by the antero-posterior dimension, which was mainly anterior, but it varied based on upper incisor inclination. Although very minimum, there was also a change in the right-left dimension. Table 4.2

PNS showed less overall 3D change than ANS and A-Point, it moved approximately 1.75mm when analyzed by both voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-

derived method, by contrast, showed an overall 3D change of about 2.8mm. When assessed per components, all the three methods had the same tendency; vertical variation was the most evident, showing an inferior direction, followed by the antero-posterior dimension, with posterior drift. Although very minimum, there was also a change in the right-left dimension. Table 4.2

Orbitale right and left had even less change than ANS, A-Point and PNS, the voxel-based (CMFreg/Slicer and Dolphin) methods showed an average 3D change of 1.45mm, the landmark-derived method, by contrast, showed an overall 3D change of 2.8mm. When assessed per components, all the three methods had the same tendency; more antero-posterior variation than in any other dimension, mainly in an anterior direction, followed by vertical changes, with a superior direction. Although very minimum, there was also a change in the right-left dimension. Table 4.2

Mandible

Menton is one of the areas that demonstrated the greatest amount of change. It had an overall 3D variation of approximately 5.5mm when analyzed by the both voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 9.2mm. When assessed per components, all the three methods had the same tendency; the vertical change was the most evident, showing an inferior direction, followed by the antero-posterior dimension, with anterior drift. Although very minimum, there was also a change in the right-left dimension. Table 4.2

BPoint changed approximately 4.5mm on average when analyzed by the two voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 7.9mm. Similar to menton, when assessed per components, all the three methods had the same tendency; the vertical change was more evident in an inferior direction, followed by the antero-posterior dimension, which was mainly anterior but it varied based on lower incisor inclination. Although very minimum, there was also a variation in the right-left dimension. Table 4.2

Gonion right and left changed approximately 4.5mm when analyzed by both voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 6mm. When assessed per components, all the three methods had the same tendency; more vertical variation than in any other dimension, with an inferior direction. There were also antero-posterior changes with posterior drift not anterior as it occurred at menton and Bpoint, as well as in the right-left dimension, mainly towards the external surface. Table 4.2

Pogonion showed similar variations to Menton, it changed approximately 5mm when analyzed by both voxel-based (CMFreg/Slicer and Dolphin) methods, the landmark-derived method, by contrast, showed an overall 3D change of about 10.18mm. When assessed per components, all the three methods had the same tendency; the vertical change was the most evident, showing an inferior direction, followed by the antero-posterior dimension, with anterior drift. Although very minimum, there was also a variation in the right-left dimension. Table 4.2

Incisal edge and apex of the upper right central maxillary incisor (#11) and the lower right central mandibular incisor (#41) were also registered but only with the aim to evaluate changes

at APoint and BPoint. Assessing the axial inclination and anterior-posterior changes would not be appropriate since superimpositions were at the cranial base and not regional superimpositions.

Insignificant differences were found in the overall 3D at all dependent variables when comparing CMFreg/Slicer to Dolphin method (Table 4.1). The least difference was found at Pg with a mean of 0.01mm (CI 95% -0.47, 0.49) and the largest at OrR with a mean of 0.16mm (CI 95% -0.26, 0.58). All mean differences were less than the voxel size (0.4mm), and they were considered of no statistical and clinical significance.

By contrast, significant differences were found between CMFreg/Slicer method and landmark-derived method (Table 4.1). The least difference was found at GoL with a mean of 0.99mm (CI 95% -0.01, 1.99) and the largest at pogonion with a mean of 5.09mm (CI 95% 2.43, 7.75). The same trend was observed when compared Dolphin and landmark-derived method (Table 4.1). The least difference was found at GoL with a mean of 1.12mm (CI 95% 0.42, 1.83) and the largest at Pg with a mean of 5.10mm (CI 95% 2.49, 7.72).

Table 4.1 One-way Repeated Measures ANOVA - Pairwise Comparisons

			95% Confidence Interval		
Landmarks	**Superimposition Method**	**Mean**	**Lower Bound**	**Upper Bound**	**p-value**
ANS	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.312	2.119	4.505	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.148	-0.211	0.508	0.920
	Voxel-Based Dolphin Method - Landmark-derived Method	3.460	2.327	4.594	0.000

A-Point	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.306	2.201	4.412	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.129	-0.256	0.515	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.177	2.038	4.316	0.000
PNS	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	1.021	0.239	1.804	0.007
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.116	-0.474	0.707	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	1.138	0.594	1.681	0.000
Orbitale Right	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	2.731	1.622	3.840	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.163	-0.257	0.583	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	2.894	1.876	3.912	0.000
Orbitale Left	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	2.927	2.026	3.829	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.089	-0.308	0.486	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.016	2.164	3.868	0.000
Menton	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.521	2.155	4.886	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.144	-0.286	0.573	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.664	2.372	4.957	0.000
B-Point	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.371	1.880	4.862	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.094	-0.305	0.494	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.466	1.965	4.966	0.000
Gonion Right	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	1.517	0.416	2.618	0.004

	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.068	-0.160	0.297	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	1.449	0.345	2.552	0.007
Gonion Left	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	0.993	-0.008	1.995	0.052
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.127	-0.578	0.832	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	1.120	0.415	1.825	0.001
IE #11	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.242	1.972	4.512	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.059	-0.148	0.266	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.301	1.994	4.608	0.000
Apex #11	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.030	1.957	4.103	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.080	-0.193	0.353	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	2.950	1.920	3.981	0.000
IE #41	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.141	1.918	4.364	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.117	-0.073	0.307	0.394
	Voxel-Based Dolphin Method - Landmark-derived Method	3.025	1.776	4.274	0.000
Apex #41	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	3.122	1.714	4.531	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.085	-0.442	0.612	1.000
	Voxel-Based Dolphin Method - Landmark-derived Method	3.037	1.739	4.336	0.000
Pogonion	Voxel-Based CMFreg/Slicer Method - Landmark-derived Method	5.093	2.432	7.754	0.000
	Voxel-Based CMFreg/Slicer Method - Voxel-based Dolphin Method	0.010	-0.467	0.487	1.000

	Voxel-Based Dolphin Method - Landmark-derived Method	5.103	2.486	7.719	0.000

Table 4.2 Descriptive Statistics - Three Methods per Components

Landmarks	Superimposition Method	Mean	95% Confidence Interval	
			Lower Bound	Upper Bound
ANS R-L	Voxel-Based CMFreg/Slicer Method	0.057	-0.140	0.254
	Landmark-derived Method	0.185	-0.622	0.992
	Voxel-based Dolphin Method	0.096	-0.146	0.338
ANS A-P	Voxel-Based CMFreg/Slicer Method	1.103	0.582	1.625
	Landmark-derived Method	1.146	0.435	1.856
	Voxel-based Dolphin Method	1.210	0.730	1.690
ANS S-I	Voxel-Based CMFreg/Slicer Method	-1.717	-2.130	-1.305
	Landmark-derived Method	-4.373	-5.604	-3.142
	Voxel-based Dolphin Method	-1.680	-2.060	-1.301
A-Point R-L	Voxel-Based CMFreg/Slicer Method	0.036	-0.165	0.237
	Landmark-derived Method	0.223	-0.546	0.992
	Voxel-based Dolphin Method	0.127	-0.079	0.333
A-Point A-P	Voxel-Based CMFreg/Slicer Method	0.628	0.279	0.978
	Landmark-derived Method	0.534	-0.081	1.150
	Voxel-based Dolphin Method	0.786	0.430	1.141
A-Point S-I	Voxel-Based CMFreg/Slicer Method	-1.900	-2.410	-1.390
	Landmark-derived Method	-4.503	-5.742	-3.265
	Voxel-based Dolphin Method	-2.076	-2.617	-1.536
PNS R-L	Voxel-Based CMFreg/Slicer Method	-0.064	-0.238	0.110
	Landmark-derived Method	-0.079	-0.398	0.240
	Voxel-based Dolphin Method	0.092	-0.108	0.292
PNS A-P	Voxel-Based CMFreg/Slicer Method	-0.151	-0.515	0.214
	Landmark-derived Method	-0.650	-1.156	-0.144
	Voxel-based Dolphin Method	-0.124	-0.522	0.274
PNS S-I	Voxel-Based CMFreg/Slicer Method	-1.217	-1.528	-0.905
	Landmark-derived Method	-1.974	-2.537	-1.411
	Voxel-based Dolphin Method	-1.115	-1.394	-0.836
Orbitale Right R-L	Voxel-Based CMFreg/Slicer Method	0.192	-0.045	0.428
	Landmark-derived Method	0.353	-0.455	1.162

	Voxel-based Dolphin Method	0.018	-0.083	0.120
Orbitale Right A-P	Voxel-Based CMFreg/Slicer Method	0.964	0.667	1.262
	Landmark-derived Method	1.164	0.644	1.685
	Voxel-based Dolphin Method	1.076	0.788	1.363
Orbitale Right S-I	Voxel-Based CMFreg/Slicer Method	0.126	0.108	0.360
	Landmark-derived Method	2.054	0.980	3.129
	Voxel-based Dolphin Method	0.105	0.145	0.354
Orbitale Left R-L	Voxel-Based CMFreg/Slicer Method	-0.074	-0.213	0.064
	Landmark-derived Method	-0.404	-1.268	0.460
	Voxel-based Dolphin Method	-0.084	-0.217	0.049
Orbitale Left A-P	Voxel-Based CMFreg/Slicer Method	1.055	0.735	1.376
	Landmark-derived Method	1.358	0.938	1.778
	Voxel-based Dolphin Method	1.217	0.942	1.492
Orbitale Left S-I	Voxel-Based CMFreg/Slicer Method	0.198	0.017	0.413
	Landmark-derived Method	2.057	0.967	3.148
	Voxel-based Dolphin Method	0.242	0.053	0.431
Menton R-L	Voxel-Based CMFreg/Slicer Method	-0.117	-0.445	0.211
	Landmark-derived Method	0.247	-0.691	1.185
	Voxel-based Dolphin Method	0.013	-0.323	0.349
Menton A-P	Voxel-Based CMFreg/Slicer Method	0.900	0.159	1.641
	Landmark-derived Method	1.122	0.521	2.765
	Voxel-based Dolphin Method	0.994	0.272	1.716
Menton S-I	Voxel-Based CMFreg/Slicer Method	-4.922	-5.680	-4.163
	Landmark-derived Method	-7.144	-8.534	-5.755
	Voxel-based Dolphin Method	-4.964	-5.719	-4.208
B-Point R-L	Voxel-Based CMFreg/Slicer Method	-0.042	-0.402	0.318
	Landmark-derived Method	0.132	-0.832	1.097
	Voxel-based Dolphin Method	0.000	-0.353	0.353
B-Point A-P	Voxel-Based CMFreg/Slicer Method	0.711	0.064	1.358
	Landmark-derived Method	0.794	0.559	2.148
	Voxel-based Dolphin Method	0.786	0.144	1.428
B-Point S-I	Voxel-Based CMFreg/Slicer Method	-3.812	-4.535	-3.090
	Landmark-derived Method	-5.854	-7.369	-4.339
	Voxel-based Dolphin Method	-3.832	-4.555	-3.109
Gonion Right R-L	Voxel-Based CMFreg/Slicer Method	1.296	0.996	1.596
	Landmark-derived Method	1.496	1.078	1.915
	Voxel-based Dolphin Method	1.404	1.109	1.699
Gonion Right A-P	Voxel-Based CMFreg/Slicer Method	-1.205	-1.764	-0.646
	Landmark-derived Method	-2.651	-3.994	-1.308

	Voxel-based Dolphin Method	-0.977	-1.504	-0.451
Gonion Right S-I	Voxel-Based CMFreg/Slicer Method	-3.536	-4.252	-2.820
	Landmark-derived Method	-3.638	-4.435	-2.840
	Voxel-based Dolphin Method	-3.841	-4.522	-3.160
Gonion Left R-L	Voxel-Based CMFreg/Slicer Method	-1.230	-1.519	-0.941
	Landmark-derived Method	-1.130	-1.462	-0.797
	Voxel-based Dolphin Method	-1.069	-1.399	-0.740
Gonion Left A-P	Voxel-Based CMFreg/Slicer Method	-1.396	-2.042	-0.750
	Landmark-derived Method	-2.135	-3.083	-1.186
	Voxel-based Dolphin Method	-0.989	-1.522	-0.455
Gonion Left S-I	Voxel-Based CMFreg/Slicer Method	-3.138	-3.857	-2.419
	Landmark-derived Method	-3.929	-4.753	-3.105
	Voxel-based Dolphin Method	-3.696	-4.362	-3.031
Pogonion R-L	Voxel-Based CMFreg/Slicer Method	-0.037	-0.381	0.306
	Landmark-derived Method	0.063	-1.068	1.193
	Voxel-based Dolphin Method	-0.001	-0.329	0.327
Pogonion A-P	Voxel-Based CMFreg/Slicer Method	1.025	0.322	1.729
	Landmark-derived Method	0.115	-1.753	1.983
	Voxel-based Dolphin Method	1.095	0.422	1.769
Pogonion S-I	Voxel-Based CMFreg/Slicer Method	-4.216	-4.949	-3.483
	Landmark-derived Method	-5.746	-8.723	-2.770
	Voxel-based Dolphin Method	-4.425	-5.087	-3.762

4.5 Discussion

An important part in the field of orthodontics is the assessment of changes due to growth and/or treatment effects. Thus, when evaluating the dentofacial changes that occur prior, during and after orthodontic treatment, orthodontists are interested in observing specific areas of variations.[35] As a result, superimposition of serial images is performed on stable structures to observe the impact of these changes in the facial structures. These cephalometric superimpositions involve the evaluation of changes in the overall face, in the maxilla and its dentition, as well as in the mandible and its dentition.[15, 36]

Cranial base superimpositions provide an overall assessment of the growth and treatment changes of the facial structures, including the amount and direction of maxillary and mandibular growth or displacement, changes in maxillary-mandibular relationships, and the relative changes in the soft tissues such as the nose, lips, and chin.[15, 36]

The maxilla develops post-natally entirely by intramembranous ossification and growth occurs by apposition of bone at the sutures that attach the maxilla to the cranium and cranial base and by surface remodeling. The sutures connecting the maxilla posteriorly and superiorly are ideally located to allow its downward and forward repositioning. Bone is added at the tuberosity, creating additional space into which the primary and then the permanent posterior dentition successively erupts. As the maxilla grows downward and forward, its front surfaces are remodeled, and bone is removed from most of the anterior surface, almost the entire anterior surface of the maxilla is an area of resorption, not apposition. The overall growth changes are the result of both a downward and forward translation of the nasomaxillary complex and a simultaneous surface remodeling. The palatal area is carried downward and forward along with the rest of the maxilla, but at the same time, bone is removed on the nasal side and added on the oral side, creating an additional downward and forward movement of the palate.[37-39]

From the sagittal perspective, the area of the ANS drifts inferiorly; and the APoint also moves inferiorly and slightly posteriorly. These findings are similar to our results, except APoint drifted slightly anteriorly instead, which could have been the effect of a change in the axial inclination of upper incisors moving APoint forward. According to Graber[40] for every 4mm that the posterior nasal spine drifts posteriorly, it drifts approximately 3mm inferiorly. Our results confirm these findings; although, PNS in the sample assessed drifted inferiorly slightly

more than it did posteriorly. One of the reasons behind this change could have been the effect of maxillary expansion performed as part of the orthodontic treatment in all the individuals included in the study.

Associated with inferior displacement of the midfacial complex bony resorption occurs along the floor of the nasal cavity, whereas apposition occurs on the palate and orbital floor.[41] Supporting this finding, all the three methods used in the present study observed superior and anterior movement of orbitale right and left. Although, the landmark-derived method showed means as high as twice the means observed with the voxel-based methods in the overall 3D.

Growth of the nasomaxillary complex continues throughout childhood and adolescence, with substantially greater vertical than anteroposterior growth potential, as it was observed with all the three 3D superimposition methods in this study. Although vertical maxillary growth rates peak during adolescence, at approximately the same time as stature, anteroposterior maxillary growth remains more or less constant, with no distinct adolescent spurt.[40] As the individuals included in this study were at CVM stage 3-4 at the initial records, the peak of growth was expressed during treatment, as observed in all the anatomical areas assessed.

In contrast to the maxilla, both endochondral and periosteal activity are important in the growth of the mandible. As a growth site, the chin is almost inactive. It is translated downward and forward, as the actual growth occurs at the mandibular condyle and along the posterior surface of the ramus. In essence, the body of the mandible grows longer as the ramus moves away from the chin, and this occurs by removal of bone from the anterior surface of the ramus and deposition of bone on the posterior surface.[37, 39, 42]

Viewed in its lateral projection, the postero-inferior and superior border of the ramus, including the condyle, and the posterosuperior aspect of the coronoid process are depository throughout the period of active growth.[42] The results of the assessment with the three 3D superimposition methods in this present study show a posterior movement of gonion right and left confirming the findings from previous growth studies. The symphysis, especially the superior aspect, becomes wider due to superior and posterior drift of its posterior aspect.[43] Our findings show pogonion as one of the areas of greatest variation observed throughout the two-year period. Although the landmark-derived method showed a mean difference as high as twice the mean difference measured with the voxel-based methods in the overall 3D, all the 3D superimposition methods used in this study observed a substantial change of Pogonion between pre and post treatment images. It mainly drifted inferiorly, although some antero-posterior component was also noted.

Literature shows that there is resorption on the anterior aspect of the symphysis above the bony chin.[40, 42] BPoint was assessed in this study and it was found that it moved inferiorly and anteriorly, when assessed with the three 3D superimposition methods, although the antero-posterior dimension varied considerably based on the lower incisor inclination.

Widening of the body of the mandible occurs through deposition of bone along the buccal surface and transverse rotation of the right and left corpii. The mandible also widens due to bony deposition along its posterior surface, which produces a longer and wider body. Between 7 and 15 years of age, biantegonial and bigonial widths increase approximately 10mm and 12mm, respectively.[44, 45]Although it was not the most important dimension that varied at gonion right and left, transverse change did happen, and it was in an outward direction. The

greatest change at gonion right and left was found in the vertical direction, followed by the antero-posterior dimension as assessed with the three 3D superimposition methods. They changed inferiorly and posteriorly, supporting the known principle that the vertical and the antero-posterior aspects of craniofacial growth are less mature and have greater postnatal growth potential than the transverse dimension.

Orthodontic treatment mechanics can also play a role in the changes observed in the craniofacial complex in growing patients. This study included patients that underwent orthodontic therapy during the 24 months of assessment. Although the sample used in this study included individuals with different malocclusions, the entire sample received maxillary expansion, and almost three quarters had either Class II or Class III intermaxillary elastics as part of the treatment. This mechanics could have moved mandible in a backward and downward direction, explaining some of the changes observed in the mandibular landmarks. They could have also moved upper and lower incisors as a result of the dentoalveolar compensation required to treat mild Class II and Class III malocclusions, changing the position of APoint and BPoint.

From the results of this study, it could be observed that the three 3D superimposition methods demonstrated an overall 3D change in the craniofacial complex during an average of 24 months of evaluation (mean age of 12.4 years - CVM 3-4 at initial records). Both voxel-based methods (CMFreg/Slicer and Dolphin) showed similar mean differences between pre and post images with no statistical significance. On the other hand, the landmark-derived method exhibited mean differences as high as twice as the mean differences obtained with any of the voxel-based methods in the overall 3D assessment. When the methods assessed the changes at each landmark per components, eight skeletal landmarks (ANS, APoint, PNS, Menton, Bpoint,

GoR, GoL and Pg) showed the highest variation in the superior-inferior component, with inferior direction, and two skeletal landmarks (OrR and OrL) in the antero-posterior component, with anterior drift. Similar to the overall 3D evaluation, the landmark-derived method exhibited the highest mean differences when assessed per components, being the superior-inferior component, which demonstrated the most substantial variation.

The magnitude of variation obtained with both voxel-based methods (CMFreg/Slicer and Dolphin) appears to be within the range of change observed by previous studies. However, as the included individuals had orthodontic treatment, it is not possible to determine if the amount of change seen at the specific landmarks in the maxilla and mandible was due to growth only, or it was a combination of growth and treatment effects. In addition, as none of the methods used to assess changes in the maxillofacial complex are considered the gold standard for 3D superimposition; the accuracy of the results cannot be determined. Therefore, even with the availability of 3D imaging, quantification of growth is still an area for research.

4.6 Limitations

Some important limitations exist in our study. First is the lack of a gold standard for 3D superimposition. Although two out of the three methods tested in this study showed very minor differences between them and the mean differences were not statistically significant, it is not possible to determine the accuracy of the results.

Second is the stability of the cranial base structures used for the landmark-derived method. The CBCTs included in our study were on average 24 months apart during adolescent

years, with a CVM stage of 3-4 at the initial scan. Ideally, the method should have use structures on the anterior cranial base that has demonstrated to be stable after 7 years of age.

Third is the possible effect of the segmentation process and the different software programs used for the full superimposition as well as the landmark identification as a source of measurement error in 3D radiographic imaging.

Fourth, the surface model construction in CBCT is based on the voxel-based data. A threshold value specifies each structure whether it is bone or soft tissue. The threshold value and gray value entered by the operator in to the CBCT machine determines the image accuracy. Also, the CBCT imaging lacks beam homogeneity which means that the gray value of the voxels of the CBCT of the same individual at different time points differ.[46, 47]

Fifth is the small sample size, which limited the assessment of treatment effects on each malocclusion (Class I, Class II and Class III) in order to obtain meaningful information and analyze in depth the magnitude, direction, and etiology of the variations.

Finally, due to the lack of a control group differentiation between the treatment and normal growth changes was not possible.

4.7 Conclusions

Overall 3D findings are in concordance with previous studies that assessed changes in the craniofacial area in growing patients. Two of the three methods (CMFreg/Slicer and Dolphin) used in this study showed similar mean differences; however, the accuracy of the results could not be determined since none of them have been considered the gold standard for

3D superimposition in growing patients. In addition, due to the lack of a control group differentiation between the treatment and normal growth changes was not possible.

4.8 Appendices

Appendix 4.1 Descriptives of Repeated Measures for All Distances				
			95% Confidence Interval	
Landmarks	**Superimposition Method**	**Mean**	**Lower Bound**	**Upper Bound**
ANS	Voxel-based CMFreg/Slicer Method	2.688	2.207	3.170
	Landmark-derived Method	6.000	5.083	6.917
	Voxel-based Dolphin's Method	2.540	2.117	2.963
APoint	Voxel-based CMFreg/Slicer Method	2.541	2.098	2.985
	Landmark-derived Method	5.848	4.896	6.799
	Voxel-based Dolphin's Method	2.671	2.220	3.122
PNS	Voxel-based CMFreg/Slicer Method	1.850	1.312	2.388
	Landmark-derived Method	2.871	2.410	3.333
	Voxel-based Dolphin's Method	1.734	1.467	2.001
Orbitale Right	Voxel-based CMFreg/Slicer Method	1.583	1.204	1.963
	Landmark-derived Method	4.314	3.566	5.062
	Voxel-based Dolphin's Method	1.420	1.184	1.656
Orbitale Left	Voxel-based CMFreg/Slicer Method	1.543	1.113	1.974
	Landmark-derived Method	4.471	3.768	5.173
	Voxel-based Dolphin's Method	1.454	1.205	1.704
Menton	Voxel-based CMFreg/Slicer Method	5.690	4.892	6.488
	Landmark-derived Method	9.210	7.912	10.509
	Voxel-based Dolphin's Method	5.546	4.771	6.321
B-Point	Voxel-based CMFreg/Slicer Method	4.595	3.904	5.285
	Landmark-derived Method	7.966	6.663	9.268
	Voxel-based Dolphin's Method	4.500	3.810	5.191
Gonion Right	Voxel-based CMFreg/Slicer Method	4.623	4.075	5.172
	Landmark-derived Method	6.141	5.233	7.048
	Voxel-based Dolphin's Method	4.692	4.129	5.255
Gonion Left	Voxel-based CMFreg/Slicer Method	4.576	3.836	5.315
	Landmark-derived Method	5.569	4.833	6.305
	Voxel-based Dolphin's Method	4.449	3.845	5.052
IE #11	Voxel-based CMFreg/Slicer Method	4.039	3.488	4.590
	Landmark-derived Method	7.280	6.137	8.424
	Voxel-based Dolphin's Method	3.980	3.426	4.534
Apex #11	Voxel-based CMFreg/Slicer Method	4.303	3.595	5.010
	Landmark-derived Method	7.333	6.208	8.458

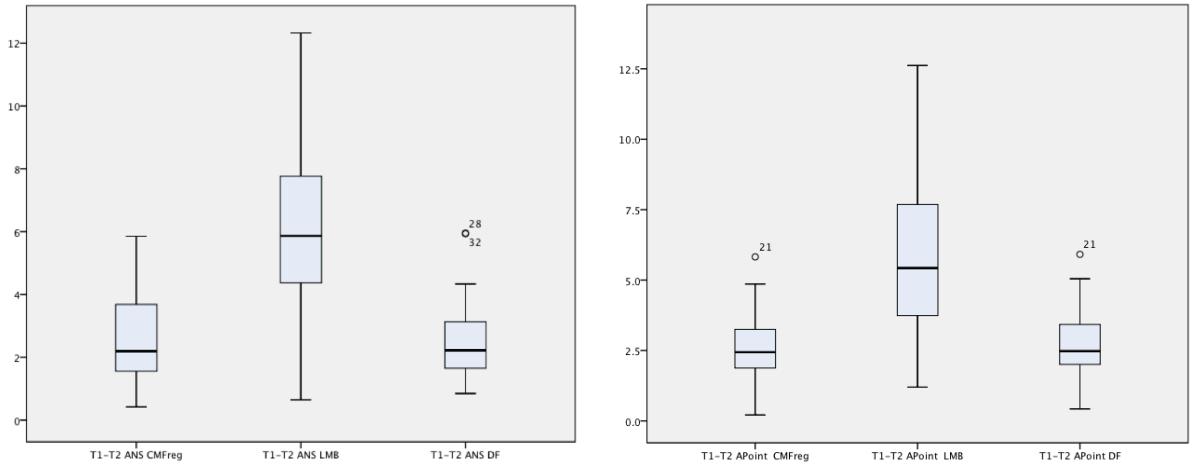
	Voxel-based Dolphin's Method	4.383	3.681	5.084
IE #41	Voxel-based CMFreg/Slicer Method	5.121	4.492	5.749
	Landmark-derived Method	8.262	7.100	9.424
	Voxel-based Dolphin's Method	5.237	4.656	5.819
Apex #41	Voxel-based CMFreg/Slicer Method	3.893	3.217	4.569
	Landmark-derived Method	7.015	5.831	8.200
	Voxel-based Dolphin's Method	3.978	3.406	4.551
Pogonion	Voxel-based CMFreg/Slicer Method	5.086	4.340	5.832
	Landmark-derived Method	10.179	7.873	12.495
	Voxel-based Dolphin's Method	5.076	4.434	5.718

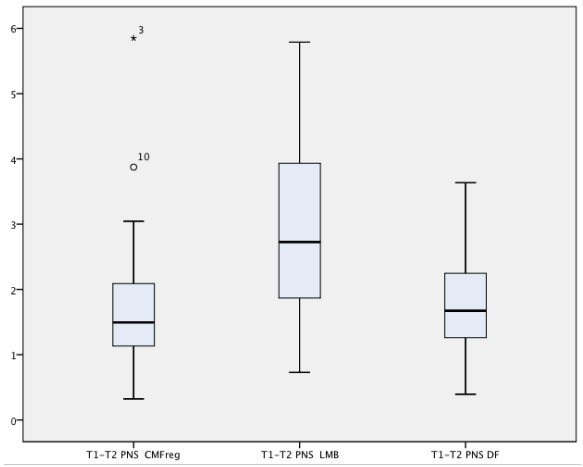
Appendix 4.2 Hypothesis Tests for Repeated Measures Anova Statistics

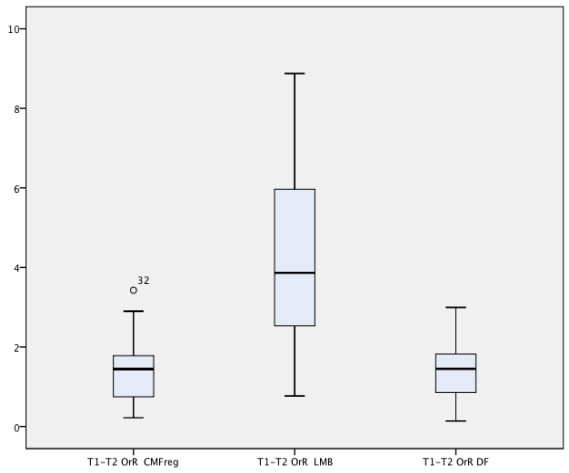
H0: The mean difference between T1 and T2 of the fourteen different linear measurements are equal on the three methods.

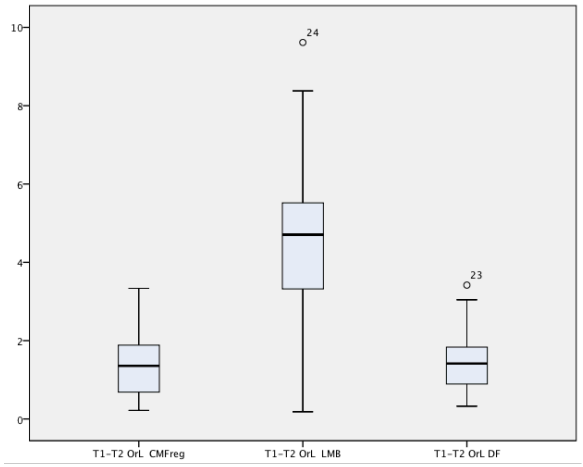
Ha: The mean difference between T1 and T2 of the fourteen different linear measurements are not equal on the three methods.

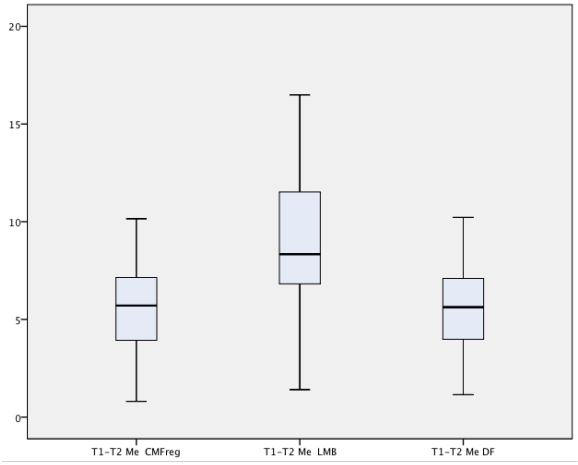
Appendix 4.3 Box Plots of Estimated Marginal Means of Distances

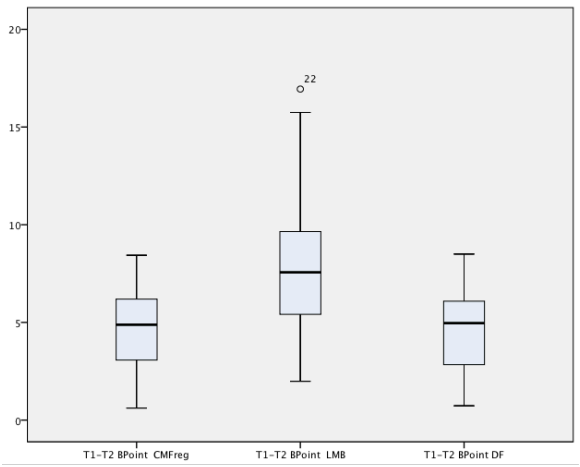


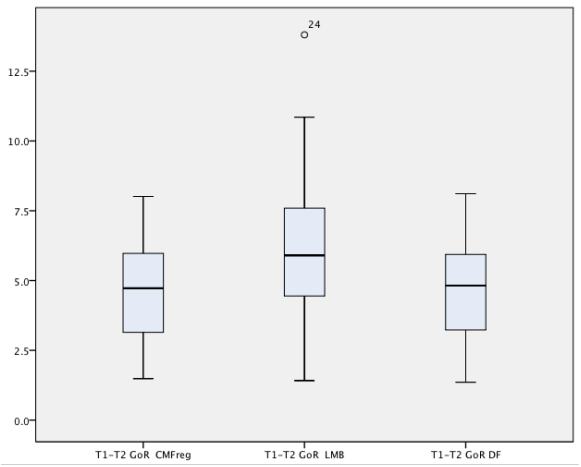

6
5
4
3
2
1
0
3
10
T1-T2 PNS CMFreg
T1-T2 PNS LMB
T1-T2 PNS DF

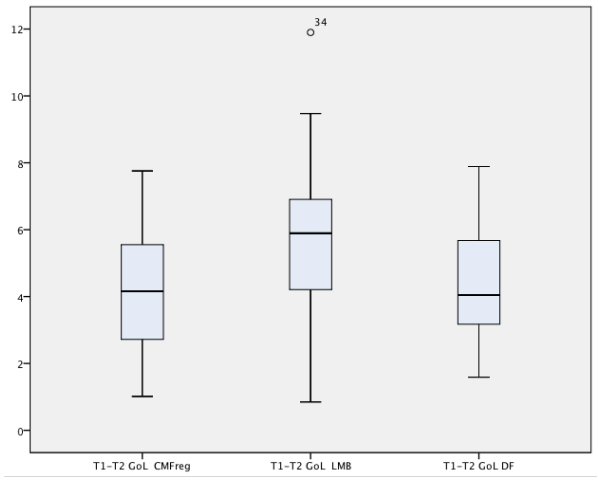

10
8
6
4
2
0
32
T1-T2 OrR CMFreg
T1-T2 OrR LMB
T1-T2 OrR DF

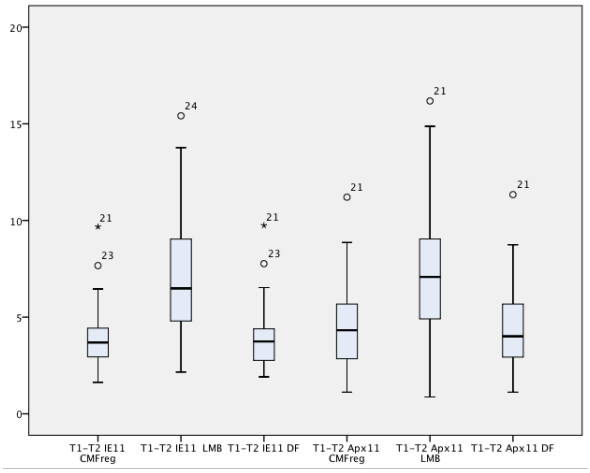

10
8
6
4
2
0
24
23
T1-T2 OrL CMFreg
T1-T2 OrL LMB
T1-T2 OrL DF

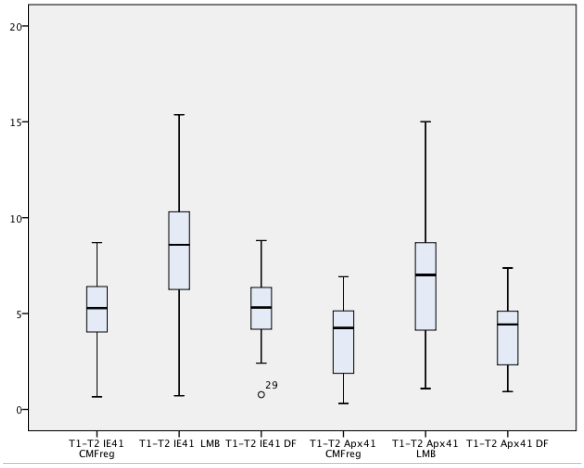

20
15
10
5
0
T1-T2 Me CMFreg
T1-T2 Me LMB
T1-T2 Me DF

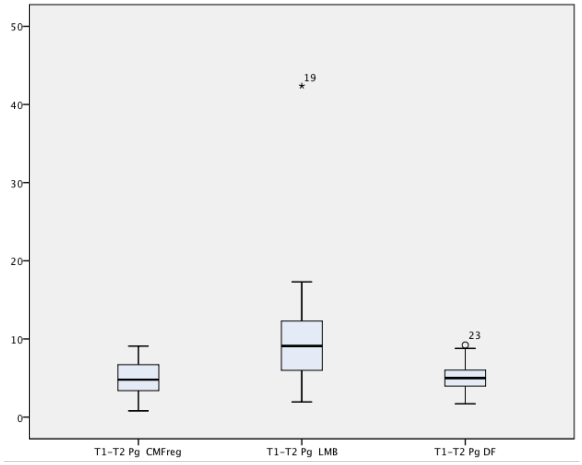

20
15
10
5
0
22
T1-T2 BPoint CMFreg
T1-T2 BPoint LMB
T1-T2 BPoint DF

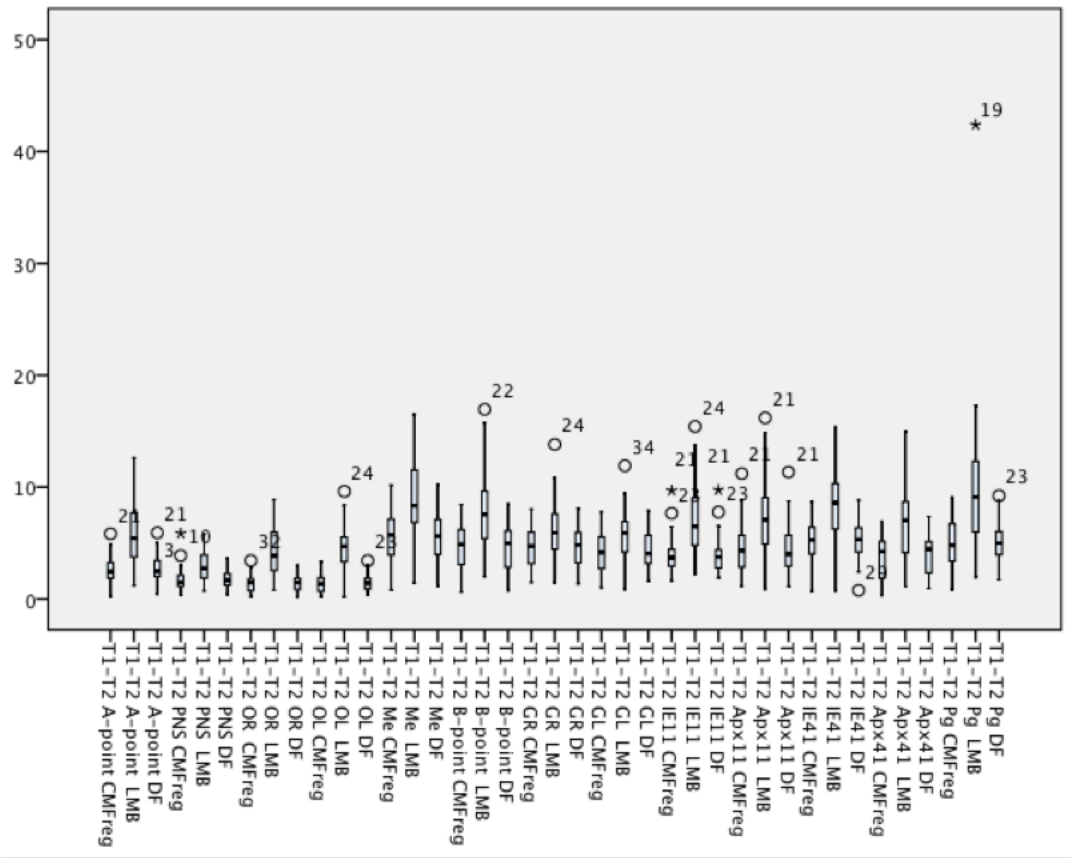

12.5
10.0
7.5
5.0
2.5
0.0
24
T1-T2 GoR CMFreg
T1-T2 GoR LMB
T1-T2 GoR DF


34
T1-T2 GoL CMFreg
T1-T2 GoL LMB
T1-T2 GoL DF

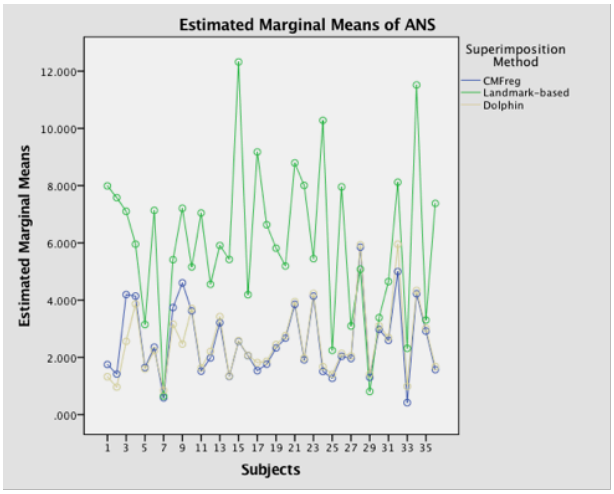

21
24
21
21
23
23
21
21
T1-T2 IE11 CMFreg
T1-T2 IE11 LMB
T1-T2 IE11 DF
T1-T2 Apx11 CMFreg
T1-T2 Apx11 LMB
T1-T2 Apx11 DF

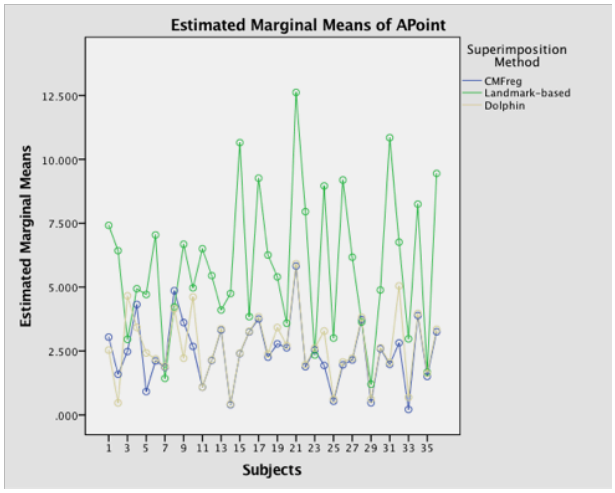

29
T1-T2 IE41 CMFreg
T1-T2 IE41 LMB
T1-T2 IE41 DF
T1-T2 Apx41 CMFreg
T1-T2 Apx41 LMB
T1-T2 Apx41 DF

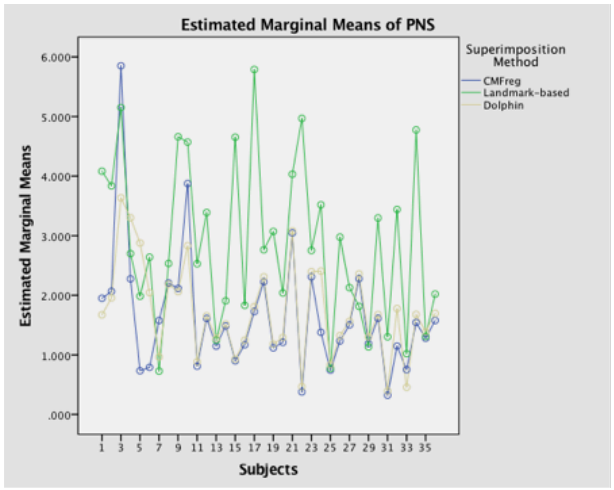

19
23
T1-T2 Pg CMFreg
T1-T2 Pg LMB
T1-T2 Pg DF

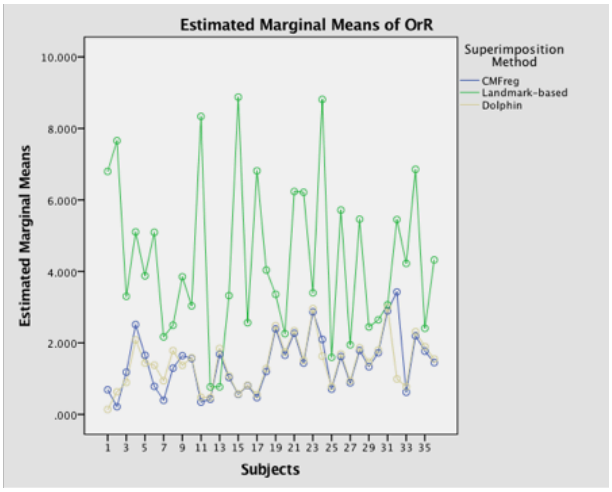


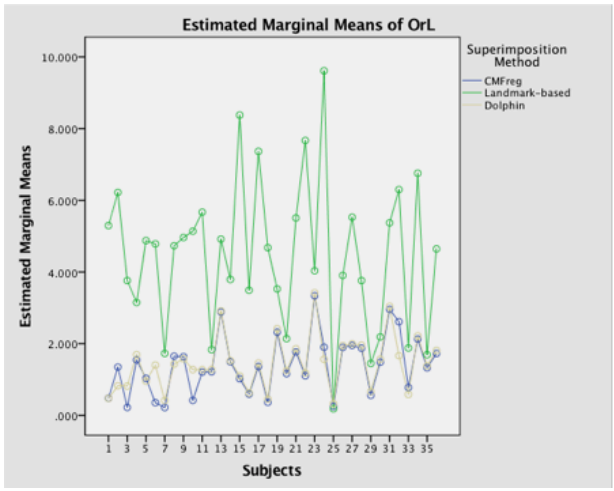
Appendix 4.4 Profile Plots Comparing Estimated Marginal Means of Distances Among Methods

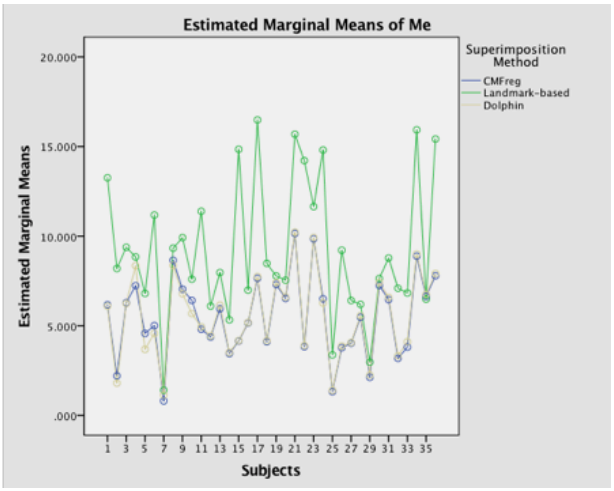


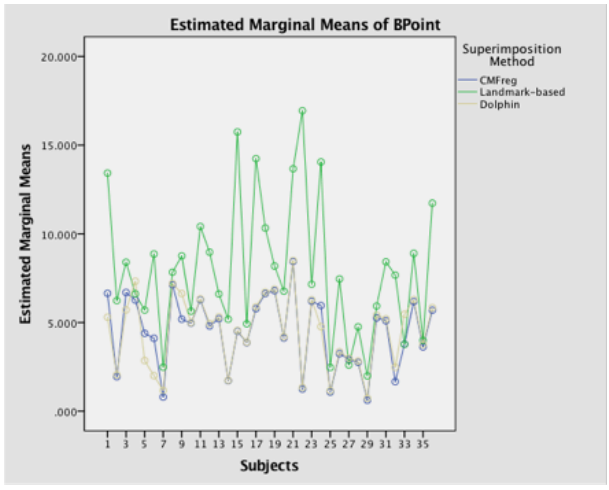


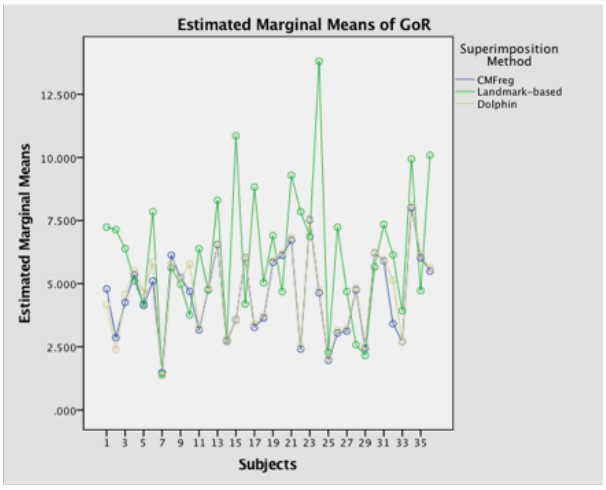

Estimated Marginal Means of PNS
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

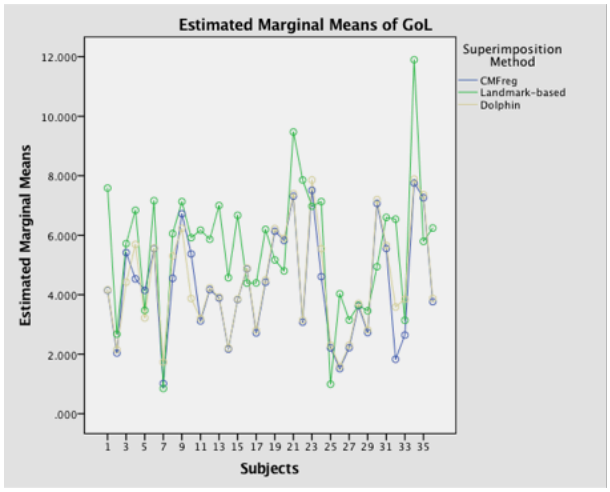

Estimated Marginal Means of OrR
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

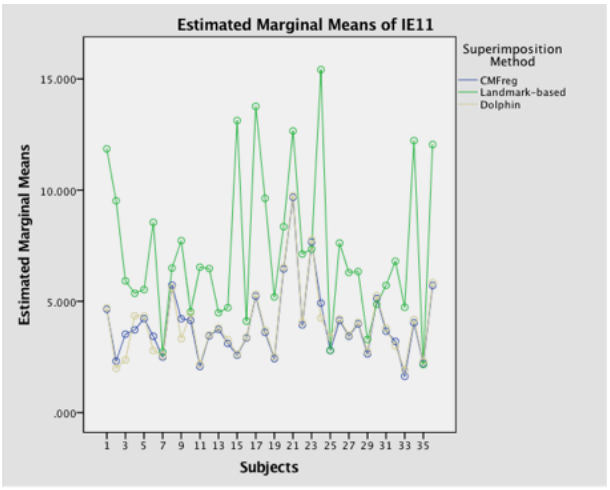

Estimated Marginal Means of OrL
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

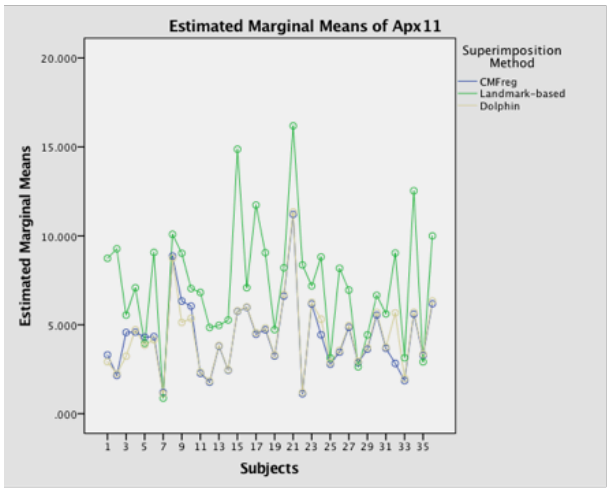

Estimated Marginal Means of Me
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

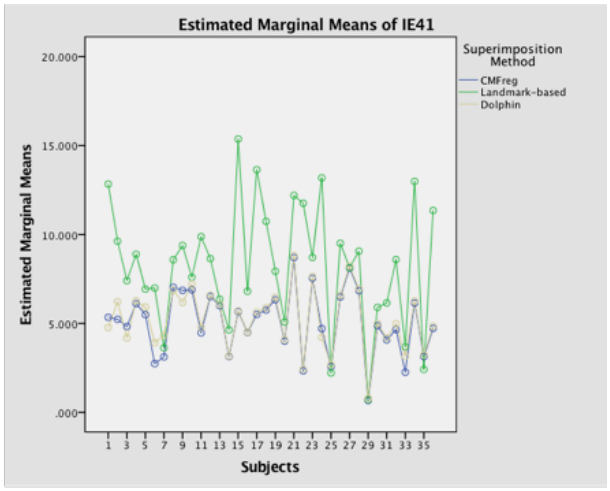

Estimated Marginal Means of BPoint
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

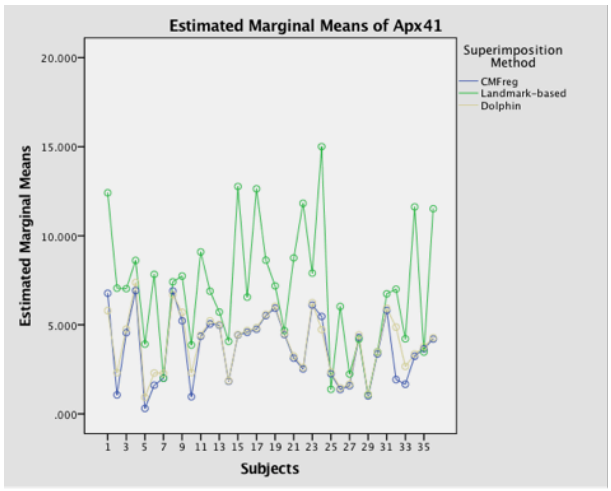

Estimated Marginal Means of GoR
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

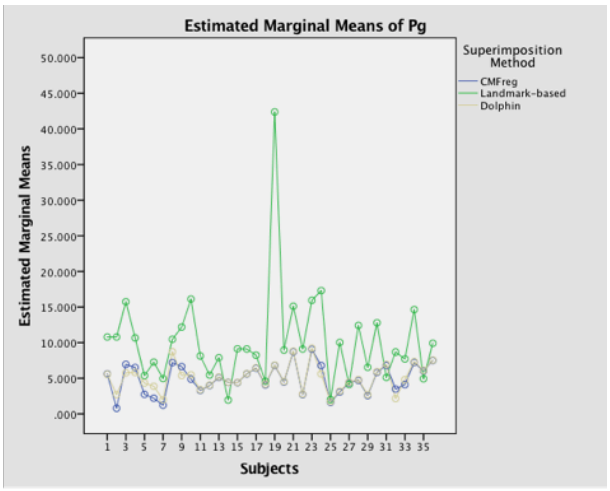

Estimated Marginal Means of GoL
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin


Estimated Marginal Means of IE11
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin


Estimated Marginal Means of Apx11
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin


Estimated Marginal Means of IE41
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin


Estimated Marginal Means of Apx41
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin


Estimated Marginal Means of Pg
Estimated Marginal Means
Subjects
Superimposition Method
CMFreg
Landmark-based
Dolphin

4.9 References

1. Ghafari J, Engel FE, Laster LL. Cephalometric superimposition on the cranial base: a review and a comparison of four methods. Am J Orthod Dentofacial Orthop 1987;91(5):403-13.

2. Bergersen EO. A comparative study of cephalometric superimposition. Angle Orthod 1961;31(4):216-29.

3. Arat ZM, Turkkahraman H, English JD, Gallerano RL, Boley JC. Longitudinal growth changes of the cranial base from puberty to adulthood. A comparison of different superimposition methods. Angle Orthod 2010;80(4):537-44.

4. Ford E. Growth of the human cranial base. Am J Orthod 1958;44(7):498-506.

5. Steuer I. The cranial base for superimposition of lateral cephalometric radiographs. Am J Orthod 1972;61(5):493-500.

6. Bambha JK. Longitudinal cephalometric roentgenographic study of the face and cranium in relation to body height. J Am Dent Assoc 1961;63:776-99.

7. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.

8. Knott V. Change in cranial base measures of human males and females from age 6 years to early adulthood. Growth 1971;35(2):145-58.

9. Melsen B. The cranial base. Acta Odontol Scand 1974;32 (Suppl 62):86-101.

10. Afrand M. Anterior and middle cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2015.

11. American Board of Orthodontics 2D Cranial Base Superimposition. https://www.americanboardortho.com. Accessed on March 10, 2018.

12. Björk A. Cranial base development: a follow-up x-ray study of the individual variation in growth occurring between the ages of 12 and 20 years and its relation to brain case and face development. Am J Orthod 1955;41(3):198-225.

13. Currie K. Posterior cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2017.

14. Ricketts RM. A four-step method to distinguish orthodontic changes from natural growth. J Clin Orthod 1975;9:208-15.

15. Athanasios E Athanasiou. Orthodontic Cephalometry. Denmark: Mosby-Wolfe; 1995.

16. Baumrind S, Frantz RC. The reliability of head film measurements: 1. Landmark identification. Am J Orthod 1971;60(2):111-27.

17. Baumrind S, Miller D, Molthen R. The reliability of head film mensurements: 2. Dentitional angular and linear measures. Am J Orthod 1971;60(5):505-17.

18. Baumrind S, Miller D, Molthen R. The reliability of head film measurements: 3. Tracing superimposition. Am J Orthod 1976;70(6):617-44.

19. American Board of Orthodontics. Examination Information Manual. . St. Louis: American Board of Orthodontics; 1990.

20. Coben SE. The integration of facial skeletal variants. Am J Orthod 1955;41:407-334.

21. Ricketts RM, Bench RW, Gugino CF, Hilgers JJ, Schulhof RJ. Bioprogressive Therapy. Denver, Colorado: Rocky Mountain Orthodontics; 1979.

22. Kaygısız E, Tortop T. Cone Beam Computed Tomography in Orthodontics. Intechopen 2017;DOI: 10.5772/intechopen.68555.

23. Kapila S, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. Dentomaxillofacial Radiology 2015;44(1):20140282.

24. Kumar V, Ludlow JB, Mol A, Cevidanes L. Comparison of conventional and cone beam CT synthesized cephalograms. Dentomaxillofac Radiol 2007;36:263-69.

25. Moshiri M, Scarfe WC, Hilgers ML, et al. Accuracy of linear measurements from imaging plate and lateral cephalometric images derived from cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2007;132:550-60.

26. Stratemann SA, Huang JC, Maki K, Hatcher DC, Miller AJ. Evaluating the mandible with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2010;137:S58-S70.

27. Lee M, Kanavakis G, Miner RM. Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review. Angle Orthod 2015;85(1):3-10.

28. Pittayapat P, Limchaichana-Bolstad N, Willems G, Jacobs R. Three-dimensional cephalometric analysis in orthodontics: a systematic review. Orthod Craniofac Res 2014;17:69-91.

29. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

30. Ruellas AC, Tonello C, Gomes LR, et al. Common 3-dimensional coordinate system for assessment of directional changes. . Am J Orthod Dentofacial Orthop 2016;149(5):645-56.

31. Lagravère M, Major P. Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography. AJODO 2005;128:657-60.

32. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

33. Lagravère M, Gordon J, Flores-Mir C, et al. Cranial base foramen location accuracy and reliability in cone-beam computerized tomography. AJODO 2011;139:e203–e10.

34. Lagravère M, Major P, Carey J. Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images. Dentomaxillofacial Radiology 2010;39:400-08.

35. Kristensen B. Cephalometric Superimposition: Growth and Treatment Evaluation. The Royal Dental College: Aarhus 1989.

36. Ruellas A, Huanca L, Gomes MT, et al. Comparison and Reproducibility of Two Maxillary Regional Registration Methods. J Dent Res 2015;94(Spec Iss A):0915.

37. Proffit W, Fields HW, Sarver DM. Chapter 2: Concepts of Growth and Development in Contemporary Orthodontics: Mosby; 2012.

38. Van der Linden F. Facial Growth and Facial Orthopedics: Quintessence; 1989.

39. Enlow DH. Essentials of Facial Growth. 2nd ed: Needham Press; 2008.

40. Carlson DS, Buschang PH. Chapter 1: Craniofacial Growth and Development: Developing a Perspective in Orthodontics: Current Principles and Techniques. Graber, Lee W.; Vanarsdall, Robert L.; Vig, Katherine W. L.; Huang, Greg J. Sixth Edition ed. St. Louis, Missouri: Elsevier Health Sciences. Kindle Edition; 2017.

41. Björk A, Skieller V. Growth of the maxilla in three dimensions as revealed radiographically by the implant method. Br J Orthod 1977;4:53-64.

42. Enlow D.H, HDB. A study of the postnatal growth of the human mandible. Am J Orthod 1964;50(1):25-50.

43. Buschang P.H, Julien K, Sachdeva R, et al. Childhood and pubertal growth changes of the human symphysis. Angle Orthod 1992;62:203-10.

44. Hersby RM, Marshall SD, Dawson DV, et al. Transverse skeletal and dentoalveolar changes during growth. Am J Orthod Dentofac Orthop 2006;130:721-31.

45. Lux CJ, Conradt C, Burden D, et al. Transverse development of the craniofacial skeleton and dentition between 7– 15 years of age: a longitudinal postero-anterior cephalometric study. Eur J Orthod 2004;26:31-42.

46. Loubele M, Jacobs R, Maes F, et al. Image quality vs radiation dose of four cone beam computed tomography scanners. Dentomaxillofac Radiol 2008;37(6):309-18.

47. Damstra J, Fourie Z, Huddleston Slater JJR, Ren Y. Reliability and the smallest detectable difference of measurements on 3-dimensional cone-beam computed tomography images. Am J Orthod Dentofacial Orthop 2011;140(3):e107-e14.

Chapter 5
General Discussion

5.1 Introduction

Superimposition of cephalometric headfilms taken at defined intervals is used by researchers and clinicians to help in orthodontic diagnosis, treatment planning and obtaining a general view of growth changes and treatment outcomes in the dentofacial complex.[1-3] A reference is required for a superimposition to be able to determine what changes occurred. Such references must be consistently visible and relatively easy to identify in the cephalograms of the individual, and they must be stable within the time frame of the observation period.[1] In addition, references must satisfy two conditions: validity and reliability. Validity, the predominant prerequisite, requires that the reference truly represent an anatomical entity and that the superimposition technique depict as close as possible the actual biological changes. Reliability relates to precision and reproducibility.[1, 4-6]

Many types of superimposing methods have been used for 2D images; for example, lateral cephalograms. However, 2D imaging does not fully represent a 3D structure, because much of the information is lost when 3D structures are converted to 2D images.[7-9] Thus, while 2D cephalometric superimposition is the conventional method used to evaluate growth and treatment outcomes, superimposition of CBCT scans, nowadays, allows a 3D visualization of these effects. Similar to cephalometric tracings, 3D models constructed from CBCT scans could be superimposed manually by registering common stable landmarks or by best fit of stable anatomical regions.[10-12] Improvements in image registration algorithms have made the superimposition of CBCT volumes the state-of-the-art technique for 3D outcomes assessment.[12, 13]

The main methods of 3D cephalometric superimposition that are used for clinical diagnosis and assessment of orthodontic treatment outcomes are: (1) voxel-based,[12-17] (2) landmark-based,[18, 19] and (3) surface-based.[12, 20] Most of the limitations of these 3D superimposition techniques are related to imaging and landmark identification flaws and software/hardware related errors.

This research analyzed two voxel-based [CMFreg (Craniomaxillofacial registration) and Dolphin] and one point-based (LMD) superimposition methods. Surface-method was not included due to lack of software and training availability.

The main objectives of this thesis were to:

1. Determine and compare the intra-rater reliability generated by three 3D cephalometric superimposition methods over an average period of 24 months using CBCT images acquired from a specific scan, the iCAT machine.
2. Determine and compare the changes observed by three 3D cephalometric superimposition methods over an average period of 24 months using CBCT images.

The research questions of this thesis were:

1. Which 3D superimposition method demonstrates the highest repeatability (intra-rater reliability) in CBCT images, and can be used to assess changes in growing patients?
    a. Which landmarks are the most repeatable in three-dimensional craniofacial images?
2. Are the overall 3D changes equal among the 3D superimposition methods?
    a. Which dimension shows the most noticeable variation?

5.2 Summary of Findings

The research started by determining the intra-rater reliability of fourteen anatomical landmarks in CBCT images. Most of the landmarks are commonly used in 2D cephalometric analysis, however it is slightly different to identify on a 3D dimensional image.

In general, all the proposed 3D landmarks showed good to excellent agreement for intra-examiner reliability when assessed on each method separately and therefore, could be used for future growth and treatment effects assessment in three-dimensional superimposition. The level of agreement among the three methods when aligning pre and post-treatment images was moderate to poor at thirteen landmarks and only good to moderate at one landmark (GoL). The highest agreement was between the two voxel-based (CMFreg/Slicer and Dolphin) methods when comparing methods in pairs.

The voxel-based (CMFreg/Slicer and Dolphin) methods used to superimpose 3D CBCT images in this study showed similar 3D changes. The landmark-derived method generated higher measurements compared to the voxel-based methods. From the results of this study, it could also be observed that the overall 3D variation of the maxillofacial complex during 24 months of evaluation (mean age of 12.4 years - CVM 3-4 at initial records) was mainly vertical with some antero-posterior component.

5.3 Clinical Implications

- Three-dimensional cranial base superimpositions can be carried out with either an open-source software such as Slicer/ITK-Snap (voxel-based method), or with a commercially

available software such as Dolphin (voxel-based method) or Avizo (landmark-derived method).

- CMFreg/Slicer has good to excellent reliability, uses an open-source software, but is time-consuming, up to 3 hours from the segmentation to color-coded maps.
- Dolphin 3D superimposition is highly reliable and considerably faster than the CMFreg/Slicer method but requires license and uses a rectangular box instead of only stable areas as reference, which increases its failure rate requiring repetition.
- Landmark-derived method using Avizo software has good reliability but requires license and needs to be revisited to reduce measurement error.

5.4 Study Limitations

All of the three 3D superimposition methods used in this study required a significant learning curve. CMFreg/Slicer method had the highest level of complexity among all the three methods and used two different software programs (3D Slicer and ITK-Snap) throughout the process. Although it includes systematic steps to obtain a high level of precision it is highly time-consuming. Depending on the level of expertise and type of computer used the whole superimposition from beginning to end (including obtaining measurements and color-coded maps) it could take up to 3 hours. Dolphin method, on the other hand, is faster and user-friendlier. A whole superimposition could take approximately 10 minutes; however, if the clinician wants to assess numbers, scans are required to be loaded in ITK-Snap for landmark placement and then measure using Q3DC tool in 3D Slicer. These additional steps increase the working time up to 30 minutes. The landmark-derived method appears to be simpler, since it only requires landmark placement similar as in a 2D cephalometric analysis, although in a 3D

image. Avizo software was used in the landmark-derived method, although the software is not difficult to use, it does not allow viewing the landmarks in all three planes at the same time. The researcher requires changing the plane continuously to check the landmark position in different planes.

Although the 3D superimposition with the voxel-based methods was carried out using two different software programs (Slicer and Dolphin), they use similar mutual intensity information open-source algorithm as that of pre-treatment, so the registered post-treatment images line up very closely in both methods. However, the landmark-derived method uses landmarks and a mathematical algorithm to define a 3D anatomical reference co-ordinate system. In addition, the areas used to superimpose upon in the cranial base with the landmark-derived method seemed to be still remodeling, therefore unstable and poorly reliable to be utilized as a reference in a growing patient. Ideally, the method should use structures in the anterior cranial base that has demonstrated to be stable after 7 years of age.

Quantification of the differences was performed by measuring the distance between the two surface models (T1 and T2) for the two voxel-based (CMFreg/Slicer and Dolphin) methods using Q3DC tool in 3D Slicer. By contrast, distances in the landmark-derived method were calculated manually after the optimization was completed, which could have led to an extra source of error.

The possible effect of the segmentation process and the different software used for the full superimposition as well as the landmark identification as a source of measurement error in three-dimensional radiographic imaging. The surface model construction in CBCT is based on the voxel-based data. A threshold value specifies each structure whether it is bone or soft tissue.

The threshold value and gray value entered by the operator in to the CBCT machine determines the image accuracy. Also, the CBCT imaging lacks beam homogeneity which means that the gray value of the voxels of the CBCT of the same individual at different time points differ.[21, 22]

Last but not least is the lack of gold standard for 3D superimposition. Although two out of the three methods tested in this study showed very minor differences between them and the mean differences were not statistically significant, it is not possible to determine the accuracy of the results.

5.5 Future Research

1. The sample included in our study had two CBSTs taken at 24 months apart. Growth assessment studies are usually done longitudinally and monitor patients for long-term. In order to get a clear image of growth changes in the maxillofacial complex assessed by 3D superimposition techniques, repeating the study on CBCTs taken farther apart or assessing growth on serial CBCTs is recommended.

2. Type of malocclusion and direction of growth as well as orthodontic treatment type should be assessed thoroughly in future studies when evaluating growth changes. It would help determine more specifically growth changes from treatment effects. Although, both factors were superficially analyzed in this study, sample size was small to obtain meaningful information.

3. Landmark-derived method should be revisited to reduce the magnification errors in maxilla and mandible in growing patients. The method could be tested by moving the reference landmarks used for superimposition more towards the anterior cranial base.

4. Future research should focus on determining the gold standard method to accurately assess growth changes or treatment effects using CBCTs images.

5.6 References

1. Duterloo H, Planché P. Handbook of cephalometric superimposition Hanover Park, IL : Quintessence Pub., c2011.; 2011.

2. De Clerck HJ, Nguyen T, Koerich L, Cevidanes L. Three-dimensional assessment of mandibular and glenoid fossa changes after bone-anchored Class III intermaxillary traction. Am J Orthod Dentofacial Orthop 2012;142:25-31.

3. Gu Yan, Jr MJ. Cephalometric Superimpositions. A Comparison of Anatomical and Metallic Implant Methods. Angle Orthodontist 2008;78(6):967-76.

4. Bialocerkowski A, Klupp N, Bragge P. How to read and critically appraise a reliability article. International Journal of Therapy and Rehabilitation 2010;17(3):114-20.

5. Kimberlin C, Winterstein A. Validity and reliability of measurement intruments used in research. Am J Health-Syst Pharm 2008;65:2276-84.

6. Bartlett J, Frost C. Reliability, repeatibility and reproducibility: analysis of measuremement errors in continuous variables. Ultrasound Obstet Gynecol 2008;31:466-75.

7. Berkowitz S. A multicenter retrospective 3D study of serial complete unilateral cleft lip and palate and complete bilateral cleft lip and palate casts to evaluate treatment: part 1—the participating institutions and research aims. Cleft Palate Craniofac J 1999;36:413-24.

8. Adams G, Gansky S, Miller A, et al. Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls. Am J Orthod Dentofacial Orthop 2004;126:397-409.

9. Halazonetis D. From 2-dimensional cephalograms to 3-dimensional computed tomography scans. Am J Orthod Dentofacial Orthop 2005;127:627-37.

10. Grauer D, Cevidanes L, Proffit W. Working with DICOM craniofacial images. AJODO 2009;136:460-70.

11. Terajima M, Yanagita N, Ozeki K, et al. Three dimensional analysis system for orthognathic surgery patients with jaw deformities. AJODO 2008;134:100-11.

12. Park JH, Tai K, Owtad P. 3-Dimensional Cone-Beam Computed Tomography Superimposition: A review. Seminars in Orthodontics 2015;21(4):263-73.

13. Weissheimer A, Menezes L, Koerich L, Pham J, Cevidanes L. Fast three-dimensional superimposition of cone beam computed tomography for orthopaedics and orthognathic surgery evaluation. Int. J. Oral Maxillofac. Surg. 2015;44:1188-96.

14. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

15. Cevidanes LH, Motta A, Proffit WR, Ackerman JL, Stynere M. Cranial base superimposition for 3-dimensional evaluation of soft-tissue changes. AJODO 2010;137(4 Suppl):S120-S29.

16. Cevidanes LH, Bailey L'TJ, Tucker SF, Styner MA, et al. Three-dimensional cone-beam computed tomography for assessment of mandibular changes after orthognathic surgery. AJODO 2007;131:44-50.

17. Viola P, Wells W. Alignment by maximization of mutual information. Paper presented at: Fifth International Conference on Computer Vision, 1995.

18. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

19. Lagravère M, Secanell M, Major P, Carey J. Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;111:771-77.

20. Gkantidis N, Schauseil M, Pazera P, Zorkun B, et al. Evaluation of 3-Dimensional Superimposition Techniques on Various Skeletal Structures of the Head Using Surface Models. PLoS ONE 2015;10(2):1-20.

21. Loubele M, Jacobs R, Maes F, et al. Image quality vs radiation dose of four cone beam computed tomography scanners. Dentomaxillofac Radiol 2008;37(6):309-18.

22. Damstra J, Fourie Z, Huddleston Slater JJR, Ren Y. Reliability and the smallest detectable difference of measurements on 3-dimensional cone-beam computed tomography images. Am J Orthod Dentofacial Orthop 2011;140(3):e107-e14.

BIBLIOGRAPHY

1. Adams G, Gansky S, Miller A, et al. Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls. Am J Orthod Dentofacial Orthop 2004;126:397-409.

2. Afrand M. Anterior and middle cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2015.

3. Almeida MA, Cevidanes L, Phillips C, Motta A, et al. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009;107(2):256-65.

4. Almukhtar A, Ju X, Khambay B, McDonald J, Ayoub A. Comparison of the accuracy of voxel based registration and surface based registration for 3D assessment of surgical change following orthognathic surgery. PLoS ONE 2014;9(4):e93402.

5. American Board of Orthodontics. Examination Information Manual. . St. Louis: American Board of Orthodontics; 1990.

6. American Board of Orthodontics 2D Cranial Base Superimposition. https://www.americanboardortho.com. Accessed on March 10, 2018.

7. Angelopoulos C. Cone Beam Tomography Imaging Anatomy of the Maxillofacial Region. Dent Clin North Am 2008;52:731-52.

8. Arat ZM, Turkkahraman H, English JD, Gallerano RL, Boley JC. Longitudinal growth changes of the cranial base from puberty to adulthood. A comparison of different superimposition methods. Angle Orthod 2010;80(4):537-44.

9. Athanasios E Athanasiou. Orthodontic Cephalometry. Denmark: Mosby-Wolfe; 1995.

10. Bambha JK. Longitudinal cephalometric roentgenographic study of the face and cranium in relation to body height. J Am Dent Assoc 1961;63:776-99.

11. Bartlett J, Frost C. Reliability, repeatibility and reproducibility: analysis of measuremement errors in continuous variables. Ultrasound Obstet Gynecol 2008;31:466-75.

12. Baumrind S, Frantz RC. The reliability of head film measurements: 1. Landmark identification. Am J Orthod 1971;60(2):111-27.

13. Baumrind S, Miller D, Molthen R. The reliability of head film mensurements: 2. Dentitional angular and linear measures. Am J Orthod 1971;60(5):505-17.

14. Baumrind S, Miller D, Molthen R. The reliability of head film measurements: 3. Tracing superimposition. Am J Orthod 1976;70(6):617-44.

15. Bazina M, Cevidanes L, Ruellas A, et al. Precision and reliability of Dolphin 3-dimensional voxel-based superimposition. Am J Orthod Dentofacial Orthop 2018;153:599-606.

16. Beebe D, Rausch J, Byars K, Lanphear B, Yolton K. Persistent Snoring in Preschool Children: Predictors and Behavioral and Developmental Correlates. Pediatrics 2012;130:382-89.

17. Bergersen EO. A comparative study of cephalometric superimposition. Angle Orthod 1961;31(4):216-29.

18. Berkowitz S. A multicenter retrospective 3D study of serial complete unilateral cleft lip and palate and complete bilateral cleft lip and palate casts to evaluate treatment: part 1—the participating institutions and research aims. Cleft Palate Craniofac J 1999;36:413-24.

19. Bialocerkowski A, Klupp N, Bragge P. How to read and critically appraise a reliability article. International Journal of Therapy and Rehabilitation 2010;17(3):114-20.

20. Björk A. The relationship of the jaws to the cranium. In: Lund-Strom A, ed. Introduction to orthodontics. New York: McGraw Hill Book Company; 1960:104-140.

21. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.

22. Björk A. Cranial base development: a follow-up x-ray study of the individual variation in growth occurring between the ages of 12 and 20 years and its relation to brain case and face development. Am J Orthod 1955;41(3):198-225.

23. Björk A, Skieller V. Growth of the maxilla in three dimensions as revealed radiographically by the implant method. Br J Orthod 1977;4:53-64.

24. Buschang P.H, Julien K, Sachdeva R, et al. Childhood and pubertal growth changes of the human symphysis. Angle Orthod 1992;62:203-10.

25. Carlson DS, Buschang PH. Chapter 1: Craniofacial Growth and Development: Developing a Perspective in Orthodontics: Current Principles and Techniques. Graber, Lee W.; Vanarsdall, Robert L.; Vig, Katherine W. L.; Huang, Greg J. Sixth Edition ed. St. Louis, Missouri: Elsevier Health Sciences. Kindle Edition; 2017.

26. Cavalcanti MG, Yang J, Ruprecht A, Vannier MW. Accurate linear measurements in the anterior maxilla using orthoradially reformatted spiral computed tomography. Dentomaxillofac Radiol 1999;28:137-40.

27. Cevidanes LH, Motta A, Proffit WR, Ackerman JL, Stynere M. Cranial base superimposition for 3-dimensional evaluation of soft-tissue changes. AJODO 2010;137(4 Suppl):S120-S29.

28. Cevidanes L, Benavides Erika, Ludlow J, Ruellas A. Orthodontic diagnosis and treatment planning with cone-beam computed tomography imaging. In: Graber, editor. Orthodontics Current Principles and Techniques. USA: Elsevier; 2017.

29. Cevidanes LH, Bailey L'TJ, Tucker SF, Styner MA, et al. Three-dimensional cone-beam computed tomography for assessment of mandibular changes after orthognathic surgery. AJODO 2007;131:44-50.

30. Cevidanes LH, Heymann A, Cornelis M, DeClerck HJ, Tulloch JFC. Superimposition of 3-dimensional cone-beam computed tomography models of growing patients. AJODO 2009;136:94-99.

31. Cevidanes LH, Styner MA, Proffit WR. Image analysis and superimposition of 3-dimensional cone-beam computed tomography models. AJODO 2006;129:611-18.

32. Cevidanes LH, Bailey LJ, Tucker Jr. GR, Styner MA, et al. Superimposition of 3D cone-beam CT models of orthognathic surgery patients. Dentomaxillofacial Radiology 2005;34:369-75.

33. Cevidanes LH, Styner MA, Proffit WR. Image analysis and superimposition of 3-dimensional cone-beam computed tomography models. AJODO 2006;129:611-18.

34. Coben SE. The integration of facial skeletal variants. Am J Orthod 1955;41:407-334.

35. Cochrane Consumers and Communication Group resources for authors. Data template extraction. Available from: https://cccrg.cochrane.org/author-resources; 2013. Accessed on May 7, 2016.

36. Currie K. Posterior cranial base growth and development changes as assessed through CBCT imaging in adolescents [University of Alberta]; 2017.

37. Damstra J, Fourie Z, Huddleston Slater JJR, Ren Y. Reliability and the smallest detectable difference of measurements on 3-dimensional cone-beam computed tomography images. Am J Orthod Dentofacial Orthop 2011;140(3):e107-e14.

38. Danforth RA DI, Mah J. 3-D volume imaging for dentistry: a new dimension. J Calif Dent Assoc 2003;31:817-23.

39. Danforth RA PJ, Hall P. Cone beam volume tomography: an imaging option for diagnosis of complex mandibular third molar anatomical relationships. J Calif Dent Assoc 2003;31:847-52.

40. DeCesare A, Secanell M, Lagravère M, Carey J. Multiobjective optimization framework for landmark measurement error correction in three-dimensional cephalometric tomography. Dentomaxillofacial Radiology 2013;42:1-10.

41. De Clerck HJ, Nguyen T, Koerich L, Cevidanes L. Three-dimensional assessment of mandibular and glenoid fossa changes after bone-anchored Class III intermaxillary traction. Am J Orthod Dentofacial Orthop 2012;142:25-31.

42. De Oliveira AE, Cevidanes LH, Phillips C, et al. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009;107(2):256-65.

43. Duterloo H, Planché P. Handbook of cephalometric superimposition Hanover Park, IL : Quintessence Pub., c2011.; 2011.

44. Enlow DH. Essentials of Facial Growth. 2nd ed: Needham Press; 2008.

45. Enlow D.H, HDB. A study of the postnatal growth of the human mandible. Am J Orthod 1964;50(1):25-50.

46. Ford E. Growth of the human cranial base. Am J Orthod 1958;44(7):498-506.

47. Ghafari J, Engel FE, Laster LL. Cephalometric superimposition on the cranial base: a review and a comparison of four methods. Am J Orthod Dentofacial Orthop 1987;91(5):403-13.

48. Gkantidis N, Schauseil M, Pazera P, Zorkun B, et al. Evaluation of 3-Dimensional Superimposition Techniques on Various Skeletal Structures of the Head Using Surface Models. PLoS ONE 2015;10(2):1-20.

49. Grauer D, Cevidanes L, Proffit W. Working with DICOM craniofacial images. AJODO 2009;136:460-70.

50. Gu Yan, Jr MJ. Cephalometric Superimpositions. A Comparison of Anatomical and Metallic Implant Methods. Angle Orthodontist 2008;78(6):967-76.

51. Halazonetis D. From 2-dimensional cephalograms to 3-dimensional computed tomography scans. Am J Orthod Dentofacial Orthop 2005;127:627-37.

52. Hatcher D. Maxillofacial imaging. In McNeill C, ed: Science and practice of occlusion. Chicago: Quintessence Pub. Co.; 1997.

53. Hersby RM, Marshall SD, Dawson DV, et al. Transverse skeletal and dentoalveolar changes during growth. Am J Orthod Dentofac Orthop 2006;130:721-31.

54. Heymann G, Cevidanes L, Cornelis M, De Clerck H, et al. Three-dimensional analysis of maxillary protraction with intermaxillary elastics to miniplates. AJODO 2010;137:274-84.

55. Houston WJ. The analysis of errors in orthodontic measurements. Am J Orthod 1983;83:382-90.

56. Houston WJ, Maher RE, McElroy D, Sherriff M. Sources of error in measurements from cephalometric radiographs. Eur J Orthod 1986;8:149-51.

57. Jacobson A, Jacobson R. Radiographic Cephalometry. Second Edition ed: Quintessence; 2006.

58. Kamoen A, Dermaut L, Verbeeck R. The clinical significance of error measurement in the interpretation of treatment results Eur J Orthod 2001;23:569-78.

59. Kang AH, Kim MK, Kim HJ, Zhengguo P, Lee SH. Accuracy assessment of image-based surface meshing for volumetric computed tomography images in the craniofacial region. Journal of Craniofacial Surgery 2014;25:2051-55.

60. Kapila S, Conley R, Harrell Jr W. The current status of cone beam computed tomography imaging in orthodontics. Dentomaxillofacial Radiology 2011;40:24-34.

61. Kapila S, Nervina JM. CBCT in orthodontics: assessment of treatment outcomes and indications for its use. Dentomaxillofacial Radiology 2015;44(1):20140282.

62. Kaygısız E, Tortop T. Cone Beam Computed Tomography in Orthodontics. Intechopen 2017;DOI: 10.5772/intechopen.68555.

63. Kimberlin C, Winterstein A. Validity and reliability of measurement intruments used in research. Am J Health-Syst Pharm 2008;65:2276-84.

64. Knott V. Change in cranial base measures of human males and females from age 6 years to early adulthood. Growth 1971;35(2):145-58.

65. Kristensen B. Cephalometric Superimposition: Growth and Treatment Evaluation. The Royal Dental College: Aarhus 1989.

66. Kumar V, Ludlow JB, Mol A, Cevidanes L. Comparison of conventional and cone beam CT synthesized cephalograms. Dentomaxillofac Radiol 2007;36:263-69.

67. Lagravère M, Secanell M, Major P, Carey J. Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;111:771-77.

68. Lagravère MO, Low C, Flores-Mir C, Carey J, et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-

dimensional cone-beam computerized tomography images. AJODO 2010;137(5):598-604.

69. Lagravère M, Major P. Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography. AJODO 2005;128:657-60.

70. Lagravère M, Hansen L, Harzer W, Major P. Plane orientation for standardization in 3-dimensional cephalometric analysis with computerized tomography imaging. AJODO 2006;129:601-04.

71. Lagravère M, Major P, Carey J. Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images. Dentomaxillofacial Radiology 2010;39:400-08.

72. Lagravère M, Gordon J, Flores-Mir C, et al. Cranial base foramen location accuracy and reliability in cone-beam computerized tomography. AJODO 2011;139:e203–e10.

73. Lenza MA, Carvalho AA, Lenza EB, et al. Radiographic evaluation of orthodontic treatment by means of four different cephalometric superimposition methods. Dental Press J. Orthod 2015;20(3):29-36.

74. Leonardi R, Giordano D, Maioranac F, Spampinato C. Automatic Cephalometric Analysis. Angle Orthodontist 2008;78(1):145-51.

75. Lee M, Kanavakis G, Miner RM. Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review. Angle Orthod 2015;85(1):3-10.

76. Lou L, Lagravere MO, Compton S, Major PW, Flores-Mir C. Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007;104:402-11.

77. Loubele M, Jacobs R, Maes F, et al. Image quality vs radiation dose of four cone beam computed tomography scanners. Dentomaxillofac Radiol 2008;37(6):309-18.

78. Ludlow JB, Gubler M, Cevidanes L, et al. Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views. Am J Orthod Dentofacial Orthop 2009;136:e311-13.

79. Lux CJ, Conradt C, Burden D, et al. Transverse development of the craniofacial skeleton and dentition between 7– 15 years of age: a longitudinal postero-anterior cephalometric study. Eur J Orthod 2004;26:31-42.

80. Melsen B. The cranial base. Acta Odontol Scand 1974;32 (Suppl 62):86-101.

81. Melsen B, Melsen F. The postnatal development of the palatomaxillary region studied on human autopsy material. Am J Orthod 1982;82:329-42.

82. Miller RL, Dijkman DJ, Riolo ML, Moyers RE. Graphic computerization of cephalometric data. J Dent Res 1971;50:1363-69.

83. Moher D, Shamseer L, Clarke M, Ghersi D, et al. Preferred Reporting Items for systematic Reviews and Meta-Analyses: The PRISMA Statement Systematic Reviews 2015;4(1):3-9.

84. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual. Available from: http://www.cosmin.nl/cosmin_checklist.html; 2012. Accessed on March 3, 2017.

85. Moore WJ, Lavelle CLB. Growth of the facial skeleton in the Hominoidea. New York: Academic Press 1974.

86. Moshiri M, Scarfe WC, Hilgers ML, et al. Accuracy of linear measurements from imaging plate and lateral cephalometric images derived from cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2007;132:550-60.

87. Nada R, Maal T, Breuning K, Berge S, et al. 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(2):e16520.

88. Naji P, Alsufyani N, Lagravere M. Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist 2014;84:762-72.

89. Nguyen T, Cevidanes L, George W. Validation of 3D mandibular regional superimposition methods for growing patients (Abstract). J Dent Res 2014;93 (Spec Iss B):784.

90. Park JH, Tai K, Owtad P. 3-Dimensional Cone-Beam Computed Tomography Superimposition: A review. Seminars in Orthodontics 2015;21(4):263-73.

91. Pittayapat P, Limchaichana-Bolstad N, Willems G, Jacobs R. Three-dimensional cephalometric analysis in orthodontics: a systematic review. Orthod Craniofac Res 2014;17:69-91.

92. Portney L WM. Foundations of clinical research: applications to practice: Prentice Hall, Upper Saddle River; 2008.

93. Proffit W, Fields HW, Sarver DM. Chapter 2: Concepts of Growth and Development in Contemporary Orthodontics: Mosby; 2012.

94. Redmond R, Huang J, Bumann A, Mah J. Three-Dimensional Radiographic Analysis in Orthodontics. JCO 2005;XXXIX(7):421-28.

95. Redmond R, Choi J, Mah J. A New Method for Superimposition of CBCT Scans. JCO 2010;XLIV(5):303-12.

96. Ribeiro Carvalho F, Cevidanes LH, Motta AT, Oliveira Almeida MA, Phillips C. Three-dimensional assessment of mandibular advancment 1 year after surgery. AJODO 2010;137(4):S53.e1-S53.e12.

97. Ricketts RM. A four-step method to distinguish orthodontic changes from natural growth. J Clin Orthod 1975;9:208-15.

98. Ricketts RM, Bench RW, Gugino CF, Hilgers JJ, Schulhof RJ. Bioprogressive Therapy. Denver, Colorado: Rocky Mountain Orthodontics; 1979.

99. Ruellas AC, Tonello C, Gomes LR, et al. Common 3-dimensional coordinate system for assessment of directional changes. Am J Orthod Dentofacial Orthop 2016;149(5):645-56.

100. Ruellas A, Huanca L, Gomes MT, et al. Comparison and Reproducibility of Two Maxillary Regional Registration Methods. J Dent Res 2015;94(Spec Iss A):0915.

101. Ruellas AC, Tonello C, Alonso N, et al. Author's response. Am J Orthod Dentofacial Orthop 2016;150(3):398-400.

102. Savage AW, Showfety KJ, Yancey J. Repeated measures analysis of geometrically constructed and directly determined cephalometric points. Am J Orthod Dentofacial Orthop 1987;91:295-99.

103. Steuer I. The cranial base for superimposition of lateral cephalometric radiographs. Am J Orthod 1972;61(5):493-500.

104. Stratemann SA, Huang JC, Maki K, Hatcher DC, Miller AJ. Evaluating the mandible with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2010;137:S58-S70.

105. Terajima M, Yanagita N, Ozeki K, et al. Three dimensional analysis system for orthognathic surgery patients with jaw deformities. AJODO 2008;134:100-11.

106. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality Life Research 2012;21:651-57.

107. Toyama Hino C, Cevidanes L, Nguyen T, De Clerck H, et al. Three-dimensional analysis of maxiallry changes associated with facemask and rapid maxillary expansion compared with bone anchored maxillary protraction. Am J Orthod Dentofacial Orthop 2013;144:705-14.

108. Van der Linden F. Facial Growth and Facial Orthopedics: Quintessence; 1989.

109. Viola P, Wells W. Alignment by maximization of mutual information. Paper presented at: Fifth International Conference on Computer Vision, 1995.

110. Weissheimer A, Menezes L, Koerich L, Pham J, Cevidanes L. Fast three-dimensional superimposition of cone beam computed tomography for orthopaedics and orthognathic surgery evaluation. Int. J. Oral Maxillofac. Surg. 2015;44:1188-96.