Three-Dimensional Skeletal and Dental Relations in Orthodontic Patients with Class I and Class II Division 1 Malocclusions

by

Alycia Jessica Sam

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

Master of Science

Medical Sciences –Orthodontics

University of Alberta

Abstract

Introduction: The fundamental purpose of this thesis was to develop a novel three-dimensional (3D) cephalometric analysis that utilizes a 3D Cartesian coordinate system to classify orthodontic patients into different malocclusion groups, *per* skeletal and dental relationships they share. A review of scientific literature identified reliable 3D landmarks that can potentially describe such relationships, and a reliability study using a pilot sample was conducted to ensure that any landmarks used in subsequent studies were indeed reliable. This process also provided the opportunity to compare two-dimensional (2D) normative values with those found in 3D and measurements taken from right and left sides. The two most common overall malocclusion types and of interest to this project were: Class I and Class II Division 1 (Class II-1).

Methods: Sixty pre-orthodontic cone beam computed tomography (CBCT) patient scans (Class I: n=30, Class II-1: n=30) were used as the sample population. Forty landmarks were identified on each, and appropriate measurements between them made. A multivariate statistical analysis, discriminant analysis (DA), testing 8 arbitrarily chosen measurements was run to determine which, if any, 3D linear and angular measurements discriminate 'best' between these two orthodontic classifications.

Results: The DA produced a non-significant overall result (p>0.05). Although not totally suggestive for, a univariate ANOVA model may lead to believe that 5 of these linear measurements be effectively useful in predicting malocclusion classification of

orthodontic patients: *infraorbital (right) – mental foramen (right), infraorbital (left) –*
mental foramen (left), 2.6 root apex – infraorbital (left) – 2.6 buccal, infraorbital (right) –
infraorbital (left) – 2.6 buccal, 1.6 root apex – infraorbital (right) – 1.6 buccal). Use of
various predictor measurements or combination of them will likely produce different
results. Interestingly, linear distance *infraorbital – menton* was found to be greater in the
Class I than Class II-1 malocclusion group bilaterally. Further research is required to
refine the findings of this research.

Conclusion: Continued testing of different 3D measurements is needed, to reveal
those that can 'best' discriminate between different malocclusion groups. The
complexity of any malocclusion and fundamental overlaps between distinct traits of
these different types makes the allocation of patients into clear-cut definitions
challenging. It may be prudent to reconsider the way in which patients are classified into
malocclusion groups and move away from the traditional Class I, Class II, and Class III
definitions.

Preface

This thesis is an original work by Alycia Jessica Sam. Research ethics approval from the University of Alberta Research Ethics Board was granted, with the project name "Three-dimensional skeletal and dental relations in orthodontic patients with Class I and Class II malocclusions".

Chapter 2 of this thesis has been accepted for publication as A. Sam, K. Currie, M. Lagravere-Vich, C. Flores-Mir, and H. Oh, "Reliability of different 3D cephalometric landmarks in CBCT: A systematic review", *Angle Orthodontics*, 2018.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Manuel Lagravere-Vich for his guidance and mentorship throughout this entire effort. You have taught me more than you will ever know, and for that I am sincerely grateful for. I would also like to express my gratitude to Dr. Carlos Flores-Mir and Dr. Heesoo Oh for being a part of my thesis supervising committee. This would not have been possible without every one of you.

Lastly I would like to thank my family and friends for always believing in me, and providing the support and encouragement to accomplish my dreams. There was never anything too far out of reach in your eyes. To my family –I simply could not have asked for a better one. To my friends –you are truly an extension of my family.

Table of Contents

Lists of Tables

Lists of Figures

Lists of Appendices

Chapter 1 : Introduction

Alycia Sam, Manuel Lagravere-Vich, Carlos Flores-Mir, and Heesoo Oh

Statement of the problem

Clinical examinations, medical history and imaging supports are essential components of comprehensive orthodontic diagnosis and treatment planning. Together, they provide a detailed description of anatomic and physiologic facts of an individual and offers useful information regarding skeletal base, dental relations, and overlying soft-tissue drape. A patient's craniofacial complex must be fully understood in terms of its three-dimensional (3D) anatomy, in a way that its 3D integrity is accurately displayed. Preservation of this integrity is kept through ensuring that structural identifiers are described in all three planes of space: transversely, vertically and antero-posteriorly. This is akin to the Cartesian coordinate system along the x-, y- and z-coordinates. Factors like the influence of head orientation adds additional layers to the complexity of such a feat.[1]

Radiographic imaging is one of the most important adjunctive supports available to clinicians. Two-dimensional (2D) cephalometric radiography has been the standardized technique used by clinicians to capture these relationships, providing insight when classifying orthodontic patients into a malocclusion type. Three separate 2D cephalometric images would be required to describe all three-planes of space (posterior-anterior, basilar and lateral) though challenges still exist. Within the realm of orthodontics, lateral cephalograms are considered customary and a routine part of diagnostic records. They are used in decision-making for orthodontic and/or

orthognathic surgical treatment approaches, growth predictions, and illustration of therapeutic changes or depiction of craniofacial growth over time through superimpositions at different time points. This traditional idea of it being part of the minimum set of diagnostic records required is debated nowadays.[2] [3]

Lateral cephalometric radiography presents issues or disadvantages in its use. These include but are not limited to magnification, distortion, and superimposition of bilateral structures.[4] Proper identification of structural landmarks is the basis for analysis of traditional 2D cephalometric images and as such are subjected to these issues. Cone beam computed tomography (CBCT) can overcome these challenges because for all intents and purposes is a 3D representation of a 3D object. And so there is also the absence of magnification or distortion while providing the ease of bilateral structure identification. These improvements are associated with increased accuracy of landmark identification and corresponding measurements generated from them.[5]

Novel comparisons between 2D and 3D cephalometric measurements has proven the existence of significant differences for some parameters.[6] Increased complexity for orthodontic case diagnosis and treatment planning lends the need for evaluation and possibly reconsideration of the simple acceptance of traditional modalities in daily orthodontic practices.

Research objectives

Three research objectives were set:

Objective #1: To evaluate the reliability, intra-examiner, of landmark identification in 3D cephalometric imaging.

Objective #2:

    a) To examine the potential differences between 2D normative values and 3D novel skeletal and dental measurements, for those linear and angular measurements that can be obtained with both methods of imaging.

    b) To examine the potential differences between 3D measurements taken from right and left sides.

Objective #3: To investigate the existence of alternative 3D skeletal and dental relationships to adequately categorize orthodontic malocclusions, specifically Class I and Class II-1.

References

1. Ruellas, A.C., et al., *Common 3-dimensional coordinate system for assessment of directional changes.* (1097-6752 (Electronic)).
2. Rischen, R.J., et al., *Records needed for orthodontic diagnosis and treatment planning: a systematic review.* (1932-6203 (Electronic)).
3. Durao, A.R., et al., *Validity of 2D lateral cephalometry in orthodontics: a systematic review.* (2196-1042 (Electronic)).
4. Karatas, O.H. and E. Toy, *Three-dimensional imaging techniques: A literature review.* (1305-7456 (Print)).
5. Machado, G.L., *CBCT imaging - A boon to orthodontics.* (1013-9052 (Print)).
6. Nalcaci, R., O. Ozturk F Fau - Sokucu, and O. Sokucu, *A comparison of two-dimensional radiography and three-dimensional computed tomography in angular cephalometric measurements.* (0250-832X (Print)).

Chapter 2 : Reliability of different three-dimensional landmarks obtained from CBCT reconstructions: A Systematic Review

Alycia Sam, Kris Currie, Carlos Flores-Mir, Manuel Lagravere-Vich, and Heesoo Oh

Introduction

Cephalometric radiography is a standardized radiographic technique employed to provide a better understanding of an individual's craniofacial structures in different planes of space: antero-posteriorly, vertically, and transversely. The relationship between the dentition and underlying skeletal base can also be understood in terms of these spatial relationships, in addition to the soft tissue drape that overlies it. They are particularly useful when considering surgical-orthodontic treatment approaches, when growth predictions are needed, and for cephalometric superimpositions of pre-treatment, progress, and post-treatment images to better illustrate therapeutic changes or monitoring craniofacial growth.[1, 2] However, each type of cephalometric radiographs (lateral, posterior-anterior, and basilar) depicts only two of these three dimensions and consequently, the choice in landmarks chosen to make linear and angular cephalometric measurements can have a profound effect on the values obtained. Thus, their combined use in cephalometric analyses for classification of "normal" relationships may not necessarily portray the entire picture at hand.

The landmarks routinely used in the analysis of conventional 2D lateral cephalometric radiographs are chosen based on their ability to be reliably identified.[2] The distances/angles between these landmarks are measured, and then compared to one or various sets of standardized norms that provide an indication of antero-posterior,

transverse and vertical relationships within the craniofacial complex of an individual at a given time. Radiographic findings are then compared with clinical findings. Even with rigid assessment criteria, issues with image distortion and superimposition of bilateral structures may pose significant limitations to the interpretation of this data.[3, 4] Often times a 2D antero-posterior (PA) cephalogram will be a valuable adjunct to routine orthodontic diagnosis and treatment planning as it provides valuable information, especially in the transverse direction, eliminating superimposition of certain bilateral structures. This makes it easier to interpret potential facial asymmetries. Even though chosen landmarks may be easily identified and reproducible, it is imperative to question true meaningfulness as this transverse dimension is often unaccounted for without additional imaging.

With the recent emergence of CBCT and its escalating availability, it is by no surprise that it has several appreciable applications in orthodontics. Nevertheless, there is a significant improved potential for craniofacial relationships to be visualized and understood in 3D rather than the traditional 2D. The limitations once imposed by 2D lateral and frontal cephalometric interpretation may now be overcome –volumetric data contained within the voxels of a single 360-degree CBCT scan is invaluable in reconstructing and understanding skeletal, dental, and soft tissue drape relationships in all three-planes of space.[4] Accuracy of landmark identification and placement is now enhanced as each occupies a specific location along a coordinate system of x-y-z axes. As such, it is possible that discrepancies existing between conventional lateral cephalometric analysis and 3D cephalometric analyses are attributed to the fact that measurements are made between two lines in the prior, whereas CBCT imaging

alternatively affords the possibility for measurements to be made between two planes.[5] This in itself may be enough to suggest that practitioners must strive to question whether cephalometric norms routinely used in 2D evaluation can be directly applied to 3D imaging.

Naji *et al.* found that although most "old" landmarks are reliable for use in 3D cephalometric analysis, specific nerve foramina in the maxilla and mandible provide better landmarks in 3D imaging that are more reliable and reproducible than others. It has been argued that the obliquity of foramen such as the infraorbital foramina and oral incisive foramina tend to pose different challenges as locating their centre point can be difficult.[6] Others have also introduced cranial foramina including ovale, spinosum, rotundum, and hypoglossal canals.[7]

A prior systematic review (SR) examining the reliability and reproducibility of 3D cephalometric landmarks using CBCT was published by Lisboa *et al.* in 2014.[8] Their period of search ended in October 2014, thus the decision to further explore this area of interest based on the increasing popularity of CBCT imaging (and its corresponding significance demonstrated by the abundance of scientific studies increasingly available every day) and ever evolving applications it has on clinical practice that changes rapidly. As such, the search period for our review was vaster, inclusive from 1998 (first introduction of CBCT into dentistry) to October 2017. The databases we aim to search were also more widespread than those previously considered. At least 17 additional articles reviewed in our second selection phase, as they were published between October 2014 and October 2017.

Therefore, the purpose of this SR is to investigate the available scientific literature to evaluate the accuracy and reliability of different 3D cephalometric landmarks in CBCT imaging, and to see if they depict any variability with those conventionally used and accepted when evaluating 2D cephalometric films.

Methods

Protocol and registration

This SR follows as closely as possible the methodology as detailed by the PRISMA guidelines[9] for the transparent reporting of systematic reviews and meta-analyses. A review protocol or protocol registration is available through the International prospective register of systematic reviews (PROSPERO), under the registration number: CRD42018098764.

Eligibility criteria

An extensive search of available scientific literature was carried out electronically, with only those studies that examined the accuracy and reliability of 3Dl cephalometric anatomic landmarks using CBCT considered for review. No language or study restrictions were placed. Unpublished materials were not excluded.

Information sources

Databases searched included: PubMed, MEDLINE via OvidSP, EBMR and EMBASE via OvidSP, Scopus, and Web of Science. To ensure that a wide range of academic literature was well represented, Google Scholar was used as an adjunctive search tool to discover other scholarly sources that may exist in this area of interest. The first 100 relevant hits were evaluated from this 'grey literature' and considered for inclusion.

<u>Search</u>

The strategic design was developed through consultation with a health sciences research librarian using appropriate keywords and their combinations. The full electronic search strategy for each database is illustrated in *Appendix 1.1*.

<u>Study selection</u>

Evaluation of the selected articles was staged in a two-step process to determine their eligibility. First, each individual article title and abstract was screened by each of the 2 reviewers (AS, KC) independently. This step was to make sure each article pertained to the following topics: three-dimensional imaging, anatomic landmarks, cephalometric analysis, and accuracy and/or reliability of findings. Next, decisions for final eligibility of the article was made based on full-text assessments by the same reviewers. They were not blinded to the authors and results of the studies. Any disagreements in article evaluations among reviewers were resolved by discussion and consensus, or by the introduction of a third reviewer (ML) to mediate when deemed necessary.

<u>Data collection</u>

Data collection was done in duplicate. The key features of eligible articles were documented by each reviewer including study design, sample size, observers involved, number of repetitions, type of CBCT scanner used, method of viewing 3D imaging (multiplanar rendering view *(i.e. axial, coronal, and sagittal)* only, 3D volumetric view only, or both), and utilized landmarks. The statistical results, including intra- and/or inter-examiner measurement error with standard deviation of important landmarks, and conclusions for every study were also retrieved.

Risk of bias among included studies

Individual articles then underwent a methodological quality scoring, adapted from a process described in previous related studies with modifications based on a research methodology series for reliability articles.[8, 10, 11] Each criterion for judgment was open to discussion among reviewers to aim to limit risk of bias and to serve as a baseline for assessments. An article was awarded one point when the criterion was fulfilled completely. When two requirements were required for a single criterion, half of a point was awarded when it was partially fulfilled and zero points when neither requirement was met. Each article received a grading score dependent on these specified criteria, and then categorized per its overall quality of evidence and strength of its recommendations. Articles were categorized into groupings based on the methodological quality/magnitude of scoring: excellent or high (76% or more), good or moderate (51-75%), and poor or low (50% or less). It must be noted that this is a non-validated assessment tool.

Synthesis of results

A meta-analysis was not justifiable for this topic as studies were very diverse, both in their study designs and report of relevant findings.

Results

Study selection

After the identification process and removal of duplicates, 140 records were screened for further consideration. The titles and associated abstracts of these records were then assessed for eligibility. After the first selection phase, only 76 articles were further considered and read in full-text for a second selection phase. The reasoning

behind excluded full-text articles at this stage can be found in *Appendix 2.2*. Excluded

articles also included those using CBCT scans of dry human skulls –often mounted with

water to simulate soft tissue attenuation of facial features. As such, a final total of 13

articles satisfied the selection criteria and were included in this review. A detailed outline

of our selection process, from identification through to articles included, is illustrated

Figure 2.1.
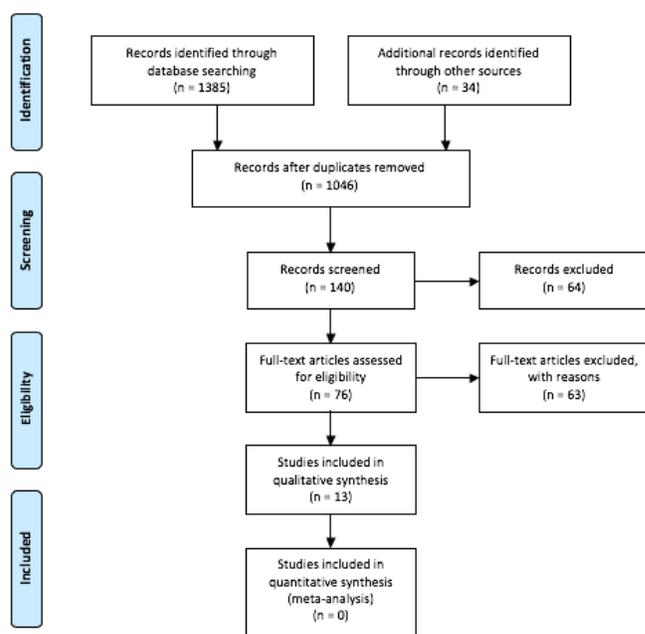
Figure 2.1: PRISMA 2009 flow diagram

In comparison to the previously published SR[8], our follow-up gained three

additional articles[12-14] and exclusion of four others considered in the previous SR.

The reason for these findings is that the prior review's search criteria ended much

earlier in October 2014, and it opted to combine CBCT landmark reliability studies using

both human patient scans and dry human skulls. Thus, any discrepancy between the

prior and latter reviews reflect these differences. It should also be noted that one of the

three additional articles[12] retrieved by our review yielded excellent or high

methodological quality scoring. Considering that overall it was only one of four included

review articles to obtain this scoring, it has the potential to offer invaluable solid insight

into this area of study. The second added article[13] introduces new landmarks and

measurements to shift the traditional 2D cephalometric analysis paradigm towards a

novel 3D one. And lastly, the third added article[13] offers insights into the use of

landmark-based superimposition in 3D.

Study characteristics

Selected articles were published between 2008 and 2017 in several diverse

medical/dental journals. All selected articles were written in English apart from two; one

was written in French and another written in Korean. The articles were obtained,

although English versions were not accessible at the time and thus the decision was

made for them to be excluded. All articles were retrospective and cross-sectional in

nature (CBCT imaging collected before the research project). Sample sizes ranged from

10-100 for subjects and 2-11 for assessors. CBCT scans of human subjects from pre-

existing data sets were chosen. The types of scanner, computer software, and methods

in viewing the 3D imaging varied greatly between studies and is described in detail in

Table 2.1.

As expected the aim of all studies was to investigate firstly the reliability (intra-

and/or inter-rater measurements) of anatomic landmarks in 3D cephalometric analysis,

which were reported statistically with one or more of the following: intra-class correlation

coefficients (ICC), Bland Altman testing, mean error and standard deviations, 3D

scatterplots, and Pearson's correlation coefficient. The choice of landmarks to quantify

identification errors was highly variable, with 'skeletal only' landmarks reported in two

articles versus the combination of 'skeletal and dental' landmarks reported in the remainder eleven articles. Some studies had the additional benefit of reporting on accuracy between true and observed measurements of 3D landmarks, comparison with conventional lateral cephalometric measurements, or effects of voxel size on accuracy and reliability.

The methodological assessment tool used, Methodological Scoring for Reliability Articles, is outlined in *Appendix 2.3*. Summary of the scores imparted to reliability articles is found in *Table 2.2*. In general, weaknesses included inadequate description of sample characteristics of subjects (e.g., sex, age, inclusion criteria, exclusion criteria, specific database used), no justification or calculation for sample sizes, and lack of explanation regarding dealing with cofounders such as exclusion criteria and employment of randomization.

Table 2.1: Summary characteristics of included articles

	Article	Sample size, type of data, sex, and age	Observers (number, experience)	Repetitions and intervals	Type of CBCT (scanner/software)	Method of viewing 3D imaging	Landmark identification	Reliability Statistical analysis	Results	Methodological quality score
1	Ghoneima et al. (2017)[14]	N = 20 F = 11 M = 9 Ages: 8-15 Mean age 11 years Pre- and post-Herbst treatment CBCT's, T1 and T2	Number not specified; Experience not specified	Each pre- and post-treatment CBCT: T_1, T_2 Interval not specified	i-CAT 3D (Xoran Technologies, Ann Harbor, Michigan) 0.3mm voxel size; Dolphin Imaging 11.8 Premium (Dolphin, Chatsworth, California, USA)	MPRV and 3D-VRV simultaneously	Skeletal and dental landmarks Landmarks = 7	Reliability of 3D landmark-based superimposition methods ICC	- All landmarks had an ICC >0.90, except ACP-x and PNS-y - Ba, Na, A point, ANS, B point, Pg, Me, U1, L1 Most reliable landmarks in x- and y-coordinates; - Landmark-based superimposition method reliable, although less than surface-based and voxel based	Moderate
2	Da Neiva et al.[12] (2015)	N = 12 Human CBCT scans	N = 3; 1: Student with undergraduate degree in dentistry	1-week intervals Each type of visualization: T_1, T_2, T_3	i-CAT 3D (Xoran Technologies, Ann Harbor, Michigan) 0.4mm voxel size;	MPRV individually 3D-VRV individually	Skeletal and dental MPRV = (30-VRV);	3D landmark identification in CBCT, using two different visualization techniques (MPRV vs. 3D-VRV);	- More highly reliable values in intra- than inter-observer assessment - MPRV more highly reliable values in landmark identification than 3D-VRV	High

	3
Study	Lee et al.[13] (2015)
Sample	N = 100
Observers	N = 2;
Calibration	Primary: Each type of visualization: T_1, T_2; 1-week interval. Other: T_1; Not repeated
Sample type	Human CBCT scans
	Not specified
Observer qualification	2: Certification in orthodontics; 3: Master's degree in orthodontics
Age / Sex	Ages: 20-43; F = 8; M = 4
Software	InVivo Dental 5.1 (Anatomage Inc., San Jose, Calif); ILUMA Ultra (Imtec 3M, Ardmore, Okla)
Voxel size	0.3mm voxel size;
3D-VRV = 30	Not specified whether 3D-VRV seen simultaneously with MPRV
Notes	*Zygomatic-maxillary suture right and left could not be correctly displayed in this view
Landmark type	Skeletal: estimation of maxillary and mandibular basal bone; Non-traditional cephalometric landmarks = 2
Statistics	ICC; Mean and standard deviation
Outcome	Reliability of two novel 3D cephalometric landmarks;
Results	- Landmarks on midsagittal plane demonstrated higher reliability - Landmarks on condyle demonstrated lower reliability <u>Clinically reliable in both techniques (ICC >0.75):</u> - B, Pg, Me, ANS, rGo, lMCo, and lUM1 <u>Poor clinical reliability in both techniques (ICC <0.45):</u> - rCo, IRP, rZS, lZS - Overall, favourable intra- and inter-examiner reliability for both maxillary and mandibular centroid landmark
Quality	Moderate

4	Naji et al.[6] (2014)	N = 30 N = 2; Sex not specified	Mixed dentition with all permanent molars erupted in occlusion	Randomly selected scans (N = 30)	Human CBCT scans Not specified	Primary: T_1, T_2, T_3 1-week intervals Others: T_1 Not repeated	InVivo Dental 5.1 (Anatomage Inc., San Jose, Calif) Next Generation i-CAT (Imaging Science International, Hatfield, Pa) Avizo 7.0 (Visualization Sciences Group, Burlington, Mass)	0.3mm voxel size; Not specified whether 3D-VRV seen simultaneously with MPRV	Skeletal and dental Non-traditional cephalometric landmarks = 42	3D scatterplots Reliability of several anatomic 3D cephalometric landmarks; Mean and standard deviation ICC	- Intra- and inter-examiner reliability for x, y, and z coordinates for all landmarks ICC >0.95 with CI of 0.88-0.99 - Intraexaminer, mostly <0.5mm - Interexaminer, all <1.4mm Mean differences of measurements for landmarks: Landmarks reliable to be used in 3D analysis: - Mental foramina, infraorbital foramina, inferior hamulus, dens axis, foramina transversium of atlas, medial	Moderate

and lateral condyles of the mandible, superior clinoid processes, and mid-clinoid

5	Zamora et al.[15] (2012)	N = 15	N = 2;	Each type of visualization: T_1, T_2	CBCT i-CAT (Imaging Sciences International, Hatfield, PA)	MPRV and 3D-VRV simultaneously	Skeletal only Landmarks = 41	Intra- and inter-observer reliability for landmark identification; all landmarks ICC>0.95	- Intra- and inter-examiner reliability for x, y, and z coordinates for all landmarks ICC>0.95	Moderate
	Ages: 12-17									
		Human CBCT scans	1, 2: Six years' experience or background in orthodontics	1-week intervals	0.4mm voxel size;	Explanation of standardized protocol for landmark identification		Mean and standard deviation	- Highest values in z-axis (ICC >0.996)	
		F = 73.4% M = 26.6%	Previously trained in locating cephalometric landmarks		Beta NemoStudio (Software Nemotec SL, Madrid, Spain)			ICC	Landmarks with no errors in determination: - Nasion, Sella, left porion, point A, anterior nasal spine, pogonion, gnathion, menton, frontozygomatic sutures, first lower molars, upper and lower incisors	
		Ages: 8-27							Landmarks with more than 6 errors in determination: - Supraorbital, right zygion, posterior nasal spine	

#	Author (year)	Sample		Subjects	Intervals / Timing	Software	Device / Voxel	Visualization	Landmark type	Outcome / Reliability measures	Results	Quality
6	Frongia et al.[16] (2012)	Pre-surgical orthodontic and impacted maxillary canine patients	N = 10	N = 2; Human CBCT scans	1,2: Experts in orthodontics	1-week intervals	Each type of visualization: T_1, T_2, T_3	Not specified	0.133mm voxel size;	MPRV and 3D-VRV simultaneously	Skeletal and dental	Hard tissue landmarks = 21	Reliability and repeatability of landmark identification in 3D; Mean and standard deviation	Pearson's correlation coefficient	- Pearson's correlation coefficient demonstrated a strong correlation (>0.7) for both intra- and inter-observer repeatability of landmark identification. <u>Most reliable landmarks:</u> - Na, S, Ba, PoL, A, Ans, FzR, FzL, Pg, Me, Gn, B36, B46, UIR, LIR - Highest SD values in z-axis (0.20-0.24mm)	Moderate
7	Schilicher et al.[17] (2012)	Orthognathic patients Ages: 18.9±1.2 F = 5 M = 5	N = 19	N = 9; Human CBCT scans	1-9: Second- or third-year orthodontic residents	No intervals	Each type of visualization: T_1	Simplant OMS Software 13.0 (Materialise Dental NV, Leuven, Belgium)	Hitachi CB MercuRay (Hitachi Medico Technology, Tokyo, Japan) 0.2-0.376mm voxel size;	MPRV and 3D-VRV simultaneously	Skeletal and dental	Landmarks = 32	Interexaminer consistency and precision of landmark identification; Standard deviation	- Average consistency across all landmarks was 1.64mm <u>Landmarks with greatest consistency:</u>	Moderate

8	Hassan et al.[18] (2011)	N = 10	N = 11;	Each type of visualization: T_1, T_2	NewTom 3G CBCT (QR- SRL, Verona, Italy)	First: 3D-VRV only	Skeletal and dental	Landmarks = 22	Intra- and interexaminer precision of landmark identification;	Range for precision of landmark measurements:
	Human CBCT scans	1-11: Orthodontic residents	At least 1 day between intervals	0.3mm voxel size;	Second: MPRV and 3D-VRV simultaneously (starting from landmark coordinates determined				Mean and standard deviation	- 0.29 +/- 0.17mm (upper incisor right) and 2.82+/- 7.53mm (portion right)

- Adding MPRV to 3D-VRV improved precision of identifying cephalometric landmarks; exceptions are:

Moderate

	F = 13 M = 6	Calibrated: definition of landmarks	Six months to complete	Dolphin Imaging 10.1 (Dolphin, Chatsworth, California, USA)					Pearson's correlation coefficient	Landmarks with poorest consistency:

- Left and right maxillary cant point, left and right orbitale, right porion

- Landmarks located along curves do not have clear anatomic boundaries, and are more erroneous

- Sella, left and right maxillary incisor crown tip, basion, right mandibular incisor crown tip

Ages: 18-35; 21.2±7.9

#	Study	Sample	Radiation exposure	Software	Imaging method	Landmarks	Statistics	Results	Quality
9	Lagravere et al.[4] (2010)	N = 10 F = 6 M = 4 Ages: 18-23 Human CBCT scans	Maximum 2 hours per individual session Not specified	Dolphin-3D v. 11 (Dolphin Imaging and Management Systems, Chatworth, CA)			Cronbach's α	- Intraobserver total reliability, mean and standard deviation, ranged from 0.48 (0.43) to 1.92 (4.69) mm - Interobserver reliability, Cronbach's α, of all landmarks below 0.7	
		N = 3; Primary: T₁, T₂, T₃ Intervals not specified Others: T₁ Not repeated		NewTom 3G (AFP Imaging, Elmsford, NY) Voxel size not specified;	No specification if 3D-VRV MPRV and 3D-VRV simultaneously viewed Conventional 2D cephalograms	Skeletal and dental Landmarks = 18	Intra- and inter-observer reproducibility of all landmark identification (ICC > 0.9) in x, y and z coordinates Mean and standard deviation ICC	- High reliability of landmark identification using 3D imaging and 2D digital cephalometrics; Intraobserver variability in CBCT: - Generally <1.0mm - x-axis: orbitale left, Sella, basion, anterior nasal spine, posterior nasal spine, condylion right (1.0-2.0mm) - x axis: porion right and left	Moderate

from 3D-VRV only)

- Nasion, menton, orbitale left, and Sella turcica

		(2.62 and 3.37mm, highest)
		- y-axis: gonion right and left, porion left and posterior nasal spine (1.0-2.0mm)
		- z-axis: B-point, mandibular incisor root apex left (1.0-2.0mm)
Sex not specified	AMIRA (Mercury Computer Systems, Berlin, Germany)	Interobserver variability in CBCT:
		- Generally >1.0mm
		- x-axis: orbitale right and left, porion right and left, condylion right and left (>2.0mm, highest)
		- y-axis: gonion right and left, anterior nasal spine (>2.0mm, highest)
Adolescents		- z-axis: gonion right and left, mandibular incisor root apex (>2.0mm, highest)

10	Ludlow et al.[2] (2009)	N = 20	N = 5;	Each type of visualization: T_1, T_2, T_3, T_4	NewTom 3G (QR-NIM s.r.l., Verona, Italy)	MPRV and 3D-VRV simultaneously	Skeletal and dental	Comparison of landmark identification between MPRV derived from CBCT and conventional 2D cephalograms;	Generally, landmark identification more precise with MPRV than conventional 2D cephalograms (13 of 22 subjects)	Moderate
		Human CBCT scans	1,2: Experienced oral and maxillofacial radiologists	No intervals specified	0.4mm voxel size;	Conventional 2D cephalograms	Landmarks = 22	Difference of the mean (ODM)	- Sella (0.7mm, lowest)	
									- Soft-tissue pogonion (2.6mm, highest)	
		Sex not specified	3: Third-year radiology resident	Observation sessions spread over 2-week period	Dolphin Imaging 10.1 (Dolphin, Chatsworth, California, USA)			Difference of each observer from each other (DEO) ANOVA	Average DEO variability in MPRV:	
									- Anteroposterior (ANS)	
									- Caudal-cranial (A-point, pogonion, porion, soft-tissue A-point, and soft-tissue pogonion)	
									- Mediolateral (condylion, mandibular incisor tip, maxillary incisor tip, orbitale, and porion)	
		Ages not specified	4: Experienced orthodontist					Paired t tests	DEO variability by 3 directional axes in MPRV:	

11	Chien et al.[19] (2009)	Pre-surgical orthodontic patients	5: Second-year orthodontic resident	Each type of visualization: T_1, T_2	i-CAT (Imaging Sciences International, Sacramento, CA)	MPRV and 3D-VRV simultaneously	Skeletal and dental	Intra- and inter-observer reliability of landmark identification using 3D CBCT imaging and 2D digital cephalogram;	Intraobserver reliability, standard error >1mm in 3D:	High
		N = 10	N = 6;						- x-direction (L1 lingual gingival border, L1 root, orbitale, porion, sigmoid notch)	
									- y-direction (supramentale, gonion, L1 lingual gingival border, L1 root, midramus, ramus point, U1 root)	
		Human CBCT scans	1-6: Second-year orthodontic residents	At least 1-week intervals	Voxel size not specified;	Conventional 2D cephalograms	Landmarks = 27	Mean and standard deviation	Intraobserver reliability, more than 15% variation for one observer:	
									- x-direction (orbitale, porion)	
									- y-direction (supramentale, L1 lingual gingival border, L1 root, ramus point)	
		Sex not specified			Dolphin Imaging 10.0 (Chatsworth, CA)		ICC		Interobserver reliability, standard error >1mm in 3D:	
									- x-direction (condylion, orbitale)	

#	Study	Sample	Operators	Ages	Image type	Settings	Time points	Machine	Voxel	Viewing	Landmark type	Landmarks	Purpose	Statistics	Results	Evidence
12	Lagravere et al.[3] (2009)	N = 24 (12 baseline, 12 six-months in treatment)	N = 5	Ages not specified	Human CBCT scans	Not specified	Primary: T_1, T_2, T_3 At least 1-week intervals	NewTom 3G (AFP Imaging, Elmsford, NY)	Voxel size not specified;	MPRV and 3D-VRV simultaneously	Skeletal and dental	Landmarks = 44	Intra- and inter-examiner reliability of landmark identification using 3D CBCT imaging; those landmarks previously used in traditional 2D imaging;	Mean and standard deviation	- High reproducibility of all measurements (ICC > 0.8) in x-, y- and z-coordinates - y-direction (gonion, midramus, ramus point) Interobserver reliability, more than 15% variation in 3D: - y-direction (gonion, midramus, ramus point) Intraexaminer reliability: - ICC >0.97 for all landmarks - Variability generally <1.5mm; exception is in z-axis: piriform right (1.53mm, highest) - Variability >1.0mm in x-axis: auditory external meatus right and left, zygomaxillary left	Moderate

Sex not specified	Others: T$_1$ Not repeated	AMIRA (Mercury Computer Systems, Berlin, Germany)	ICC	Interexaminer reliability:
				- Variability in >1.0mm in y-axis: none
				- Variability in >1.0mm in z-axis: A point, B point, piriform right and left, ectomolare right and left (>1.0mm)
				- ICC >0.92 for most landmarks; exceptions are in x-axis: auditory external meatus right and left, orbit right and left (0.8-0.9, lowest)
				- Variability greater than in intraexaminer; orbit left (3.61mm, highest)
				- Variability in x-axis: orbit right and left (>2.5mm), zygomaxillary right and left (>1.5mm)
				- Variability in y-axis: auditory external meatus left, piriform left, orbit right and left, MB 36

apex, MB 46 apex, anterior nasal spine (>1.5mm), none >2.5mm

- Variability in z-axis: piriform right and left (>2.5mm), ectomolare right and left (>1.5mm)

Ages not specified

ANOVA

Landmarks presenting with statistical differences with other landmarks in same region:

-Auditory external meatus right and left (x-axis and y-axis)

- Most landmarks in skeletal facial region

- Upper first molar (26B) and mesio-buccal apex (26A), mesio-buccal apex (36A and 46A)

Maxillary expansion treatments

13	De Oliveira et al.[20] (2008)	N = 12	N = 3;	Human CBCT scans	1: Orthodontist	Each type of visualization: T₁, T₂, T₃	NewTom 3G (AFP Imaging, Elmsford, NY)	MPRV (slices) and 3D-VRV simultaneously	Skeletal and dental	Intra- and inter-observer reliability of landmark identification using 3D imaging;	Reliability for x-, y- and z-coordinates with ICC ≥0.90, intra- and inter-observer respectively:	Moderate

Each type of visualization: T_1, T_2, T_3

3-day intervals

0.4mm voxel size;

Landmarks = 30

Intra- and inter-observer reliability of landmark identification using 3D imaging;

Reliability for x-, y- and z-coordinates with ICC ≥0.90, intra- and inter-observer respectively:

- x (80%, 66.6%)

- y (83.33%, 50%)

- z (93.33%, 80%)

Least reliable landmarks:

- x-coordinate (right and left condylion, ICC = 0.46-0.66)

- y-coordinate (right and left ramus point, ICC = 0.29- 0.68 and right and left tuberosity, ICC = 0.48-0.77)

- z-coordinate (right and left condylion, ICC = 0.28-0.5)

ICC

2: Dental radiologist — Dolphin 3D pre-release version 1 (Dolphin Imaging & Management Systems, Chatsworth, CA) — ANOVA

3: Third-year dental student

Sex not specified

Ages not specified; inclusion

criteria ages 13-50

Presurgical patients; 6 skeletal Class II and 6 skeletal Class III

Table 2.2: Methodological scores for reliability articles

Article	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	Total	%
Ghoneima et al.[13] (2017)	✓	✓	✓	✗	✗	✗	▷	✗	✗	✓	✓	✗	▷	✓	✗	✓	✗	✓	✗	✓	▷	✓	✓	12.5	54.3
Da Neiva et al.[11] (2016)	✓	✓	✓	✓	✗	✓	▷	✓	✓	✗	✓	✓	▷	✓	▷	✓	✗	✓	✓	✓	▷	✓	✓	18.5	80.4
Lee et al.[12] (2015)	✓	✓	▷	▷	✓	▷	▷	✗	✗	✗	✓	▷	▷	✓	✓	✓	✗	✓	✓	✓	▷	✓	✓	15.0	65.2
Naji et al.[5] (2014)	✓	✓	▷	✗	✗	▷	✓	✗	✗	▷	✓	▷	✗	✓	✓	✓	✗	✓	✓	✓	✗	✓	✓	15.0	65.2
Zamora et al.[14] (2012)	✓	✓	✓	✗	✗	▷	▷	✓	✓	▷	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	▷	18.0	78.2
Frongia et al.[15] (2014)	✓	✓	✓	▷	✗	▷	▷	✓	▷	✓	▷	✗	▷	✓	✓	✗	✗	✓	✓	✓	✗	▷	▷	13.5	58.7
Schilcher et al.[16] (2012)	✓	✓	✓	✓	✗	✓	▷	✓	✓	✓	✓	✓	✓	✓	▷	✓	✓	✓	✓	✓	▷	✓	✓	20.5	89.1
Hassan et al.[17] (2011)	✓	✓	✓	▷	✗	✓	▷	✓	✓	✓	▷	✓	✓	✓	✗	✓	✗	▷	✗	✓	▷	▷	▷	15.5	67.4
Lagravere et al.[3] (2010)	✓	▷	▷	✗	✓	✓	✓	✗	✗	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	15.0	65.0
Ludlow et al.[1] (2009)	✓	✓	▷	✓	✗	✓	▷	✓	✓	✗	▷	✓	✓	✓	▷	▷	✗	✓	✓	✓	✗	▷	✓	16.0	69.6
Chien et al.[18] (2009)	✓	✓	▷	✓	✗	✓	▷	✓	✓	✗	✓	✗	▷	✓	✓	✓	▷	✓	✓	✓	▷	✓	✓	17.5	76.1

Lagravere et al.[2] (2009)	✓	△	△	✗	✗	✓	△	✗	✗	✗	✓	✓	✓	✗	✓	✓	✓	✗	✓	13.5	58.7
De Oliveira et al.[19] (2008)	△	✓	✓	✗	✓	△	✓	✓	△	△	✗	✓	✓	△	✓	✗	✓	△	✓	17.0	73.9

Legend:

✓   Fulfilled satisfactorily the methodological criteria (1.0 check point)

△   Fulfilled partially the methodological criteria (0.5 check point)

✗   Fulfillment of no methodological criteria (0.0 check point)

Risk of bias within studies

A possible source of bias within each article was based on timing of the records. As many of these studies were retrospective and it is unethically sound to expose patients to radiation solely for reliability quantification purposes, investigators are reliant on the use of pre-existing data sets for subject populations. Ideally repeated CBCTs would be obtained to assess reliability of imaging *per se*, but would be unethical to do. As this review is interested in reliability, it becomes problematic if a study utilizes a set not representative of the spectrum of individuals to generalize findings in a research or clinical context. Additionally, although a few studies mentioned the use of randomization, few described how and none of them reported using sequence generation within their data set to ensure randomization was somewhat reflected when extracting their subject sample.

The more observers involved in the measurements of a single study, the greater potential for measurement error due to individual expertise. Also, as some authors are involved in more than one study of this sort, it is possible for the use of the same pre-existing database for patient CBCT scans across multiple studies. This could pose a significant problem as it will artificially inflate the reliability values if this were the case.

Results of individual studies

Due to the heterogeneity of these studies, the specific characteristics of each and key data is reported in *Table 1 and Table 4,* respectively. Notable statistical results, as detailed in *Table 4*, encompasses a summary of pertinent

statistical reliability measures for various landmarks listed by included studies. For instance, this includes such measures as intra-class coefficients (ICCs) and standard mean error (SME) for x-, y-, and z-coordinates of cephalometric landmarks. It also depicts differences between the reliability of landmarking between intra- and inter-examiners. Typically, intra- was higher than inter-examiner reliability in landmark identification. In terms of landmark characteristics themselves, skeletal landmarks present similar reliabilities compared to dental landmarks; variability was dependent on the challenges a specific location poses.

Table 2.3: Summary of notable statistical results for included articles

Article	Landmark	Landmark ICC (95% CI)	
		X	Y
Ghoneima et al.[14] (2017)	Anterior clinoid process of Sella (ACP)	0.83 (0.35, 0.99)	0.99 (0.95, 0.99)
	Basion (Ba)	0.99 (0.99, 1.00)	1.00 (0.99, 1.00)
	Nasion (Na)	0.99 (0.98, 1.00)	0.96 (0.86, 0.99)
	A point	0.99 (0.99, 1.00)	0.98 (0.94, 0.99)
	Anterior nasal spine (ANS)	0.98 (0.96, 0.99)	0.99 (0.98, 1.00)
	Posterior nasal spine (PNS)	0.94 (0.74, 0.99)	0.78 (0.27, 0.99)
	B point	0.99 (0.97, 1.00)	0.98 (0.94, 0.99)
	Pogonion (Pg)	1.00 (0.98, 1.00)	0.98 (0.94, 0.99)
	Mention (Me)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)
	Incisal tip of the upper central incisor (U1)	0.94 (0.81, 1.00)	0.99 (0.98, 1.00)
	Incisal tip of the lower central incisor (L1)	0.99 (0.98, 1.00)	1.00 (0.98, 1.00)

Article	Landmark	3D intraobserver (ICC)			3D interobserver (ICC)			MPR intraobserver (ICC)			MPR interobserver (ICC)			Clinical reliable?
		X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z	
Da Neiva et al.[12] (2016)	B point (B)	0.96	0.95	0.86	0.96	0.84	0.79	0.98	0.99	0.92	0.98	0.99	0.91	Both
	Pogonion (Pog)	0.95	0.94	0.90	0.93	0.84	0.88	0.98	0.99	0.97	0.98	0.99	0.86	Both
	Menton (Me)	0.95	0.94	0.96	0.94	0.85	0.91	0.98	1.0	0.98	0.98	0.99	0.99	Both
	Anterior nasal spine (ANS)	0.95	0.94	0.97	0.92	0.92	0.93	1.0	0.98	0.99	1.0	0.96	0.99	Both
	Left mandibular gonion (IGo)	0.98	0.95	0.92	0.96	0.92	0.93	0.99	0.95	0.93	0.99	0.95	0.90	Both
	Left medial mandibular condyle (IMco)	0.93	0.95	0.96	0.92	0.93	0.96	0.99	0.99	1.0	0.98	0.99	0.99	Both
	Left upper molar point (IUM1)	0.89	0.95	0.97	0.89	0.90	0.94	0.99	1.0	0.98	0.97	1.0	0.97	Both

Legend:

3D = 3D-virtual reconstruction view

MPR = multiplanar reconstruction view

Article	Landmark	Intraexaminer 3D distance (mm)		Interexaminer 3D distance (mm)	
		Mean	Standard deviation	Mean	Standard deviation
Lee et al.[13] (2015)	Average maxillary centroid	0.76	0.32	0.89	0.49
	Average mandibular centroid	0.57	0.39	0.82	0.43

Article	Landmark	X-axis (mm)		Y-axis (mm)		Z-axis (mm)	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Naji et al.[6] (2014)	Mental foramen right	0.48	0.32	0.39	0.26	0.31	0.27
	Mental foramen left	0.40	0.20	0.44	0.28	0.40	0.36
	Dens axis	0.46	0.28	0.42	0.20	0.13	0.18
	Right transversium atlas	0.34	0.19	0.47	0.27	0.50	0.37
	Left transversium atlas	0.41	0.20	0.38	0.26	0.39	0.34
	Inferior right hamulus	0.47	0.35	0.50	0.28	0.37	0.41
	Inferior left hamulus	0.54	0.37	0.45	0.28	0.36	0.35
	Right infraorbital	0.51	0.47	0.51	0.20	0.57	0.54
	Left infraorbital	0.51	0.31	0.54	0.37	0.66	0.43
	Superior right clinoid process	0.60	0.51	0.33	0.18	0.17	0.23
	Superior left clinoid process	0.64	0.68	0.41	0.20	0.16	0.23
	Mid-clinoid	0.56	0.45	0.43	0.25	0.56	0.92
	Lateral right condyle	0.22	0.22	0.53	0.27	0.59	0.51
	Medial right condyle	0.25	0.24	0.52	0.30	0.51	0.36
	Medial left condyle	0.32	0.19	0.55	0.29	0.71	0.35
	Lateral left condyle	0.21	0.17	0.54	0.32	0.54	0.35

Article	Region	Landmark	SD_X (mm)	SD_Y (mm)	SD_X (mm)	% of error
Zamora et al.[15] (2012)	Cranial	Nasion (Na)	0.36	0.32	0.49	0.00
		Sella turcica (S)	0.89	0.41	0.61	0.00
		Basion (Ba)	0.79	0.51	0.49	0.00
		Left porion (PoL)	0.90	0.46	0.37	0.00
	Maxilar	Point A (A)	0.86	0.41	0.93	0.00
		Anterior nasal spine (Ans)	0.96	0.66	0.30	0.00
		Incisal edge upper right central incisor (UIR)	0.86	0.31	0.28	0.00
	Orbital-zygomatic	Right frontozygomatic suture (FzR)	0.28	0.41	0.37	0.00
		Left frontozygomatic suture (FzL)	0.32	0.57	0.38	0.00
	Mandibular	Pogonion (Pg)	0.16	0.23	0.67	0.00
		Menton (Me)	0.50	0.63	0.24	0.00
		Gnathion (Gn)	0.15	0.44	0.45	0.0
		First lower left molar (B36)	0.55	0.52	0.83	0.0
		First lower right molar (B46)	0.54	0.37	0.94	0.0
		Incisal edge lower right central incisor (LIR)	0.75	0.31	0.31	0.0

Article	Operator	Intra- and inter-observer SD of 3D measurements (mm)			Intra- and inter-observer Pearson's correlation coefficient of 3D cephalometric measurements		
		X-axis	Y-axis	Z-axis	T1-T2	T2-T3	T1-T3
Frongia et al.[16] (2014)	A	0.17	0.18	0.20	0.9997	0.9998	0.9998
	B	0.18	0.24	0.24	0.9997	0.9997	0.9997
					T1-T1	T2-T2	T3-T3
	A-B	0.18	0.22	0.23	0.7908	0.7913	0.7897

Article	Landmark	Overall consistency (mm)	X-axis consistency (mm)	Y-axis consistency (mm)	Z-axis consistency (mm)
Schlicher et al.[17] (2012) NB: The number in parentheses is that landmark's rank for each column.	Sella	0.50 ± 0.24 (1)	0.14 ± 0.13 (1)	0.31 ± 0.23 (4)	0.23 ± 0.16 (1)
	Left maxillary incisor crown tip	0.58 ± 0.28 (2)	0.39 ± 0.32 (3)	0.23 ± 0.16 (2)	0.24 ± 0.15 (2)
	Right maxillary incisor crown tip	0.59 ± 0.25 (3)	0.39 ± 0.31 (4)	0.17 ± 0.14 (1)	0.31 ± 0.21 (3)
	Basion	0.85 ± 0.32 (4)	0.33 ± 0.25 (2)	0.35 ± 0.26 (5)	0.32 ± 0.28 (4)
	Right mandibular incisor crown tip	0.91 ± 0.60 (5)	0.54 ± 0.43 (10)	0.37 ± 0.26 (8)	0.37 ± 0.26 (8)
	Nasion	1.02 ± 0.50 (6)	0.48 ± 0.35 (8)	0.62 ± 0.49 (16)	0.33 ± 0.20 (5)
	Right maxillary incisor root apex	1.05 ± 0.46 (7)	0.56 ± 0.44 (11)	0.63 ± 0.38 (18)	0.36 ± 0.31 (7)
	Left mandibular incisor crown tip	1.13 ± 0.69 (8)	0.50 ± 0.41 (9)	0.46 ± 0.35 (11)	0.53 ± 0.43 (11)
	ANS	1.15 ± 0.49 (9)	0.47 ± 0.35 (7)	0.36 ± 0.25 (6)	0.76 ± 0.52 (17)
	Point A	1.20 ± 0.59 (10)	0.47 ± 0.36 (5)	1.07 ± 0.60 (27)	0.34 ± 0.24 (6)
	Left maxillary incisor root apex	1.20 ± 0.53 (11)	0.47 ± 0.36 (6)	0.71 ± 0.40 (19)	0.38 ± 0.30 (9)
	Gnathion	1.35 ± 0.61 (12)	0.67 ± 0.44 (15)	0.72 ± 0.38 (20)	0.78 ± 0.59 (19)
	Left mandibular incisor root apex	1.50 ± 0.73 (13)	0.58 ± 0.46 (12)	0.76 ± 0.61 (21)	0.70 ± 0.58 (16)
	Point B	1.50 ± 0.72 (14)	0.65 ± 0.43 (14)	0.87 ± 0.63 (22)	0.55 ± 0.39 (13)

Article	Landmark	Total precision: 3D, 3D + MPR (mm)	P-value: statistically significant differences between 3D and 3D + MPR	Interobserver reliability (Cronbach's alpha)
Hassan et al.[18] (2011) NB: Statistically significant P-values are bolded.	Orbitale right	1.00 (1.14)	0.11	0.06
	Orbitale left	1.00 (1.07)	**0.04**	0.37
	Nasion	0.64 (0.70)	**0.001**	0.66
	Anterior nasal spine	0.93 (2.21)	0.06	0.55
	Posterior nasal spine	0.96 (1.54)	**0.04**	0.58
	A-point	0.71 (0.80)	0.19	0.28
	Upper incisor right	0.29 (0.17)	0.07	0.58
	Upper incisor left	0.30 (0.20)	0.48	0.40
	Lower incisor right	0.48 (0.50)	0.87	0.69
	Lower incisor left	0.62 (1.35)	0.19	0.67
	B-point	0.74 (0.50)	0.3	0.27
	Pogonion	0.77 (0.81)	0.42	0.57
	Menton	1.00 (1.58)	**0.03**	0.49
	Gonion right	0.88 (0.62)	0.77	0.21
	Upper right molar	0.98 (3.39)	0.95	-0.33
	Upper left molar	0.62 (0.90)	0.58	0.25

Article	Landmark	Intraexaminer mean differences of coordinates (mm)					
		X		Y		Z	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Lagravere et al.[4] (2010)	Nasion (N)	0.68	0.48	0.86	0.72	1.78	1.15
	A-point (A)	0.92	0.24	0.80	0.35	0.77	0.60
	B-point (B)	1.51	1.03	0.54	0.32	1.81	1.69
	Pogonion (Pg)	1.44	1.03	0.71	0.33	1.22	0.74
	Gnathion (Gn)	1.42	1.05	0.93	0.75	0.73	0.84
	Menton (Me)	1.51	0.94	1.21	1.10	0.55	0.46
	Sella (S)	1.21	0.80	0.41	0.31	0.57	0.25
	Basion (Ba)	1.23	0.78	0.97	0.60	1.03	0.33
	Posterior nasal spine (PNS)	1.56	1.11	1.03	0.84	0.47	0.21
	Upper central incisor tip right (U1T right)	0.61	0.29	0.53	0.30	0.53	0.35
	Upper central incisor root apex right (U1R right)	0.52	0.29	0.98	0.87	1.24	1.16
	Lower central incisor tip right (L1T right)	1.53	1.06	0.72	0.45	0.65	0.58
	Lower central incisor root apex right (L1R right)	1.30	0.95	1.30	0.90	1.38	0.64
	Upper central incisor tip left (U1T left)	0.78	0.60	0.44	0.12	0.58	0.34
	Upper central incisor root apex left (U1R left)	1.11	1.07	0.79	0.72	1.21	0.97
	Lower central incisor tip left (L1T left)	1.11	0.72	0.43	0.25	0.49	0.26

Article	Landmark	Difference for every observer (DEO) variability by directions of MPR views (mm)		
		Anteroposterior (AP)	Caudal-cranial (CC)	Mediolateral (ML)
Ludlow et al.[2] (2009)	Sella	0.65	0.66	1.05
	ANS	1.43	0.73	0.66
	A-point	0.74	2.01	0.68
	Pogonion	0.69	1.91	1.35
	Soft-tissue A-point	0.78	1.93	0.79
	Soft-tissue pogonion	1.31	3.98	1.44
	Condylion	1.82	1.01	2.55
	Mandibular incisor tip	0.63	0.67	2.06
	Maxillary incisor tip	0.62	0.76	1.99
	Orbitale	2.80	0.80	5.76
	Porion	1.46	3.46	7.14

Article	Landmark	Intraobserver, mean 3D error (SD) (mm)		Interobserver, mean 3D error (SD) (mm)		Interobserver, ICC	
		X-coordinate	Y-coordinate	X-coordinate	Y-coordinate	X-coordinate	Y-coordinate
Chien et al.[19] (2009)	Subspinale	0.71 (0.79)	1.16 (0.78)	0.53 (0.56)	0.98 (0.79)	0.98	0.97
	Condylion	1.08 (0.84)	0.83 (0.62)	1.23 (0.73)	0.91 (0.70)	0.96	0.97
	Gonion	1.03 (0.87)	1.34 (1.12)	0.97 (0.83)	1.27 (1.20)	0.96	0.81
	Midramus	0.38 (0.38)	1.64 (1.28)	0.24 (0.25)	1.54 (1.02)	0.99	0.75
	Orbitale	1.13 (1.59)	0.50 (0.46)	1.06 (1.19)	0.40 (0.33)	0.95	0.99
	Ramus point	0.49 (0.39)	1.80 (1.84)	0.48 (0.45)	2.71 (2.11)	0.99	0.64
	Supramentale	0.42 (0.40)	1.15 (1.06)	0.28 (0.26)	0.97 (0.83)	1.00	0.86
	L1 root	0.93 (1.21)	1.23 (1.50)	0.52 (0.49)	0.89 (0.60)	0.98	0.90
	Porion	1.10 (2.13)	0.83 (0.83)	0.88 (0.71)	0.66 (0.69)	0.94	0.99

Article	Landmark	Intraexaminer absolute mean measurement difference (SD) (mm)			Interexaminer absolute mean measurement difference (mm)		
		X-coordinate	Y-coordinate	Z-coordinate	X-coordinate	Y-coordinate	Z-coordinate
Lagravere et al.[3] (2009)	Ectomolare left (EkmL)	0.55 (0.27)	0.68 (0.31)	1.45 (0.46)	0.99 (0.56)	1.18 (0.53)	2.44 (0.92)
	Ectomolare right (EkmR)	0.60 (0.36)	0.70 (0.38)	1.46 (0.59)	0.92 (0.60)	1.36 (0.59)	2.18 (0.72)
	Upper first molar, right (16B)	0.29 (0.39)	0.53 (0.23)	0.46 (0.23)	0.36 (0.38)	0.63 (0.21)	0.58 (0.23)
	Mesio-buccal apex (16A)	0.46 (0.19)	0.43 (0.14)	0.55 (0.42)	0.73 (0.31)	0.67 (0.27)	0.95 (0.52)
	Buccal apex (14B)	0.43 (0.42)	0.44 (0.33)	0.57 (0.24)	0.58 (0.46)	0.41 (0.25)	0.70 (0.34)
	Buccal apex (14A)	0.51 (0.19)	0.47 (0.16)	0.80 (0.41)	0.62 (0.31)	0.51 (0.20)	0.94 (0.51)
	Upper canine, right (13B)	0.37 (0.19)	0.37 (0.20)	0.57 (0.19)	0.47 (0.21)	0.42 (0.16)	0.98 (0.29)
	Canine apex (13A)	0.51 (0.18)	0.45 (0.18)	0.67 (0.24)	0.63 (0.33)	0.59 (0.24)	0.84 (0.30)
	Upper canine, left (23B)	0.36 (0.17)	0.30 (0.14)	0.59 (0.26)	0.56 (0.21)	0.44 (0.18)	1.03 (0.24)
	Canine apex (23A)	0.43 (0.18)	0.47 (0.19)	0.69 (0.32)	0.74 (0.48)	0.72 (0.49)	0.98 (0.61)
	Upper first premolar, left (24B)	0.41 (0.24)	0.43 (0.36)	0.66 (0.36)	0.52 (0.35)	0.51 (0.55)	0.65 (0.28)
	Buccal apex (24A)	0.40 (0.18)	0.46 (0.19)	0.76 (0.58)	0.63 (0.33)	0.50 (0.28)	0.86 (0.50)
	Upper first molar, left (26B)	0.22 (0.16)	0.47 (0.28)	0.54 (0.29)	0.35 (0.27)	0.57 (0.33)	0.69 (0.33)
	Mesio-buccal apex (26A)	0.56 (0.21)	0.53 (0.45)	0.86 (0.51)	0.70 (0.37)	0.76 (0.34)	1.34 (0.76)
	Lower first molar, left (36B)	0.42 (0.24)	0.37 (0.11)	0.41 (0.21)	0.55 (0.29)	0.53 (0.22)	0.69 (0.24)
	Lower canine, left (33B)	0.35 (0.15)	0.30 (0.20)	0.67 (0.23)	0.64 (0.44)	0.53 (0.39)	0.85 (0.22)
	Lower first molar, right (46B)	0.39 (0.18)	0.41 (0.14)	0.51 (0.14)	0.47 (0.26)	0.63 (0.27)	0.59 (0.27)
	Lower canine, right (43B)	0.37 (0.23)	0.38 (0.21)	0.62 (0.26)	0.41 (0.21)	0.43 (0.23)	1.04 (0.28)
	Foramen spinosum, left (FSL)	0.39 (0.31)	0.48 (0.36)	0.67 (0.37)	0.74 (0.49)	0.42 (0.10)	0.46 (0.21)
	Foramen spinosum, right (FSR)	0.38 (0.29)	0.37 (0.15)	0.40 (0.33)	0.70 (0.47)	0.37 (0.13)	0.52 (0.00)
	Center coordinate point (ELSA)	0.48 (0.17)	0.55 (0.25)	0.52 (0.27)	1.04 (0.53)	0.39 (0.26)	0.42 (0.32)
	Auditory external meatus, left (AEML)	1.46 (0.60)	0.83 (0.47)	0.40 (0.30)	3.40 (1.30)	0.45 (0.59)	0.48 (0.12)
	Auditory external meatus, right (AEMR)	1.22 (0.88)	0.76 (0.29)	0.42 (0.33)	3.09 (1.08)	0.59 (0.40)	0.33 (0.22)
	Dorsum foramen	0.70 (0.39)	0.66 (0.48)	0.88 (1.28)	0.87 (0.49)	0.82 (0.46)	0.38 (0.21)

		magnum (DFM)					

Article	Landmark	Intraobserver reliability, ICC			Interobserver reliability, ICC		
		X	Y	Z	X	Y	Z
De Oliveira et al.[20] (2008)	Left condylion	0.66	1.00	0.50	0.65	0.97	0.49
	Right condylion	0.46	0.99	0.29	0.46	0.98	0.28
	Left ramus point	0.95	0.68	1.00	0.94	0.44	0.98
	Right ramus point	0.98	0.51	0.99	0.95	0.29	0.99
	Left tuberosity	0.73	0.75	0.96	0.71	0.57	0.89
	Right tuberosity	0.73	0.77	0.94	0.71	0.48	0.90

In general, midsagittal plane landmarks tend to demonstrate better consistency in identification compared to bilateral landmarks. The ease of locating these landmarks along midlines may come naturally to most clinicians, as manipulation and interpretation of the CBCT sagittal views are quite like customary 2D lateral cephalograms. Midsagittal plane landmarks recommended for use in 3D include *Sella*, *basion*, *nasion*, *anterior nasal spine*, *A-point*, *B-point*, *pogonion*, *gnathion*, and *menton*. Bilateral landmarks demonstrating variable consistency in identification include those on the *condyles*, *orbitale*, *porion*, and *lingula*. Moreover, this is further complicated by the fact that some located along broad curvatures or have indistinct boundaries are more difficult to locate, and thus more erroneous in identification. This is pertaining predominately to CBCT reconstructions. Dental landmarks demonstrating greatest consistency are incisor crown tips, tooth root apices, and defined points on teeth. Some non-traditional landmarks are also recommended for use, some of which are: *infraorbital foramina*, *mental foramina*, and possibly even *fronto-zygomatic sutur*es. Novel 3D landmarks such as *maxillary and mandibular centroid* landmarks have also shown favourable reliability.

Synthesis of results

A meta-analysis was not possible. The methodologies of the selected

studies were highly heterogeneous, posing a challenge when considering

combining their results together. Specifically, in terms of landmarks, not all

studies evaluated the same landmarks making the comparison more challenging.

Some of these were traditionally-used cephalometric landmarks, whereas others

were non-traditional in nature.

<u>Risk of bias across studies</u>

At the current time, the Grading of Recommendations, Assessment,

Development and Evaluation (GRADE) approach to evaluating the quality of

evidence has not been adapted to reliability in a SR.

<u>Additional analysis</u>

No additional analyses were performed.

Discussion

<u>Summary of Evidence</u>

The review of available scientific literature evaluating the reliability of

various 3D landmarks revealed distinctive advantages and challenges than those

conventionally used in 2D cephalometrics; thus, reliability of 3D landmarks

should be examined independently. For example, bilateral landmarks including

the *midramus*, *orbitale, ramus point*, and *sigmoid notch*, demonstrate more

consistent identification in 3D than 2D. This is likely explained by the 2D

limitation of structural superimposition being overcome. Each left and right sides

of a landmark can be evaluated independently, in a location in all three planes of

space, without any other structures impeding its interpretation. It should be noted

however, that due to the unfamiliarity of routinely locating these landmarks along a transverse axis (as done in 2D cephalometric radiography), bilateral landmarks tend to show more variability than those located in the midline. De Oliveira *et al.* found that two bilateral landmarks demonstrated poor reliability in one of the three axes; *right/left ramus* in the y-coordinate and *right/left condylion* in the z-coordinate.[20] Many bilateral landmarks are located along broad curvatures; it may also pose a challenge for the eye to detect the most prominent point or depression of the structure at hand. As stated by Lagravere *et al.*[3] differences in landmark identification error in the axes may differ and as such, certain landmarks may be useful in detecting changes in one axis but not another. Landmarks that demonstrate considerable variability in the x-coordinate are not suitable for use in width (transverse) measurements of the craniofacial complex. For example, *condylion*, *orbitale* and *porion* has demonstrated statistically greater variability in the medio-lateral (M-L) direction, x-axis, in MPR views and may not be suitable to use in taking width measurements

One of the main limitations of 2D imaging is that a 3D object is reduced to two-planes of space. The findings of Ludlow *et al.* are only one of several to support this thought, as more precise landmark identification was obtained with most multiplanar reconstruction view (MPRV) in 3D than 2D cephalograms. Of these landmarks *Sella* demonstrated the lowest variability of 0.7mm, whereas *soft-tissue pogonion* showed the highest variability of 2.6mm.[2] This is in agreement with another study that found a high reproducibility (ICC>0.9) of all measurements traditionally employed in 2D cephalometric analyses.[4] Chien *et*

al. found that for intraobserver identification of some landmarks was greater in 3D over 2D, while others showed the opposite pattern. For instance, landmarks with greater variations in 2D than 3D included *orbitales* and *sigmoid notch* in y-locations whilst the opposite was true for *U1 labial gingival borde*r in x-location and *L1 tip*, *L6 occlusal*, *menton* and *Sella* in y-location.[19] This could be explained by the fact that this study looked at 27 landmarks in total, which was more numerous than the other two. In fact, the higher number of landmarks analyzed affords a benefit as more likely meaningful errors will be noted. With that said it is important to realize that ease of identifiability of points does not necessarily translate to clinically meaningful implications for those points. This may artificially give a sense of reliability.

After reviewing selected articles, it became apparent that there is sometimes a discrepancy between the accuracy and reliability of identifying left versus right structures. The manifestation may be attributed to purely individual examiner's systematic error. Another hypothetical and plausible explanation for this could be the neuropsychological linkage between left- and right-handedness, and its effect on preferences of the human brain. Right-handed artists have been shown to prefer their subjects on the right with light sources from the left. Left-handed artists tend to demonstrate the opposite trend.[21] Extrapolating, this may imply some influence of handedness in an evaluator's spatial orientation of CBCT scans and their identification of 3D landmarks.

There is a recognizable trend that midline landmarks such as *Sella* and *A-point* shows the same consistency, if not greater, in landmark identification as in

2D. In contrast to bilateral structures, this may be facilitated by the familiarity of observers with radiographic interpretation in the sagittal plane, used in lateral head films. MPRV display in 3D software provides an avenue to limit the magnitude of superimposition of multiple structures, as slices can be set to a particular thickness when investigating an area of interest.

Three-dimensional objects occupy a specific location on an x-y-z coordinate system. Although the maximum mean difference was minimal, one study noted that the y-coordinate was more reliable than the x- and z-coordinates among observers. Per De Oliveira *et al.*, the least reliable landmark identification in these axis's are as follows: *right and left condylion* in the x-coordinate or mediolateral direction, *right and left ramus point* closely followed by *right and left tuberosity* in the y-coordinate or antero-posterior direction, and *right and left condylion* in the z-coordinate or caudal-cranial direction.[20] In contrast, *Chien et al.* highlighted that some difficulty arose determining best estimates in y-locations for *gonion*, *L1 tip*, *Sella*, and *U1 tip*. They suggest that most have a difficult time locating the y-location of structures like *gonion*, *midramus*, and *ramus point* since a precise vertical position must be established along these broadly curved structures in 2D and 3D. Most of these inaccuracies are linked to a line parallel to its curvature. Specifically for 3D, erroneous measures may be attributed to the inappropriate use of surface display shading utilized by the operator.[19]

Most traditionally used cephalometric landmarks are reproducible both in 2D and 3D imaging modalities. Since 3D has the enhanced ability to fulfill the precision of a third dimension, it makes one wonder if there are also non-

traditional cephalometric landmarks that are more reproducible and clinically meaningful in CBCT analyses. Using 42 newly defined anatomic landmarks Naji *et al.* concluded that the mean differences of all intra- and inter-examiner measurements was less than 1.4mm. Moreover, if a centre coordinate point is chosen using x-, y-, and z-coordinates to locate a specific landmark, the analysis of its reliability among evaluators is maximized and the more impactful these differences will be clinically. In fact, bilateral *mental foramina*, *dens axis*, bilateral *transversium atlas*, bilateral *inferior hamulus*, *right infraorbital foramen*, *medial right condyle*, and *lateral left condyle* show 0.5mm or even less of a difference. However, one should be mindful in its application, as not all non-traditional landmarks should be routinely used. These authors also noted that bilateral *posterior-inferior concha* and *oral incisive foramen* among other landmarks demonstrated a difference greater than 1mm.[6]

Since 3D cephalometric landmarking is still a new concept, labelling landmarks with a variability of clinical significance is not necessarily correct *per se*. Clinical significance of cephalometric landmarks with a variability of less than 0.5mm is unlikely, whereas between 0.5mm and 1.0mm may be likely. This differs from 2D, as cephalometric landmarks less than 1.0mm are unlikely to have clinical significance.[22] Thus, if linear and angular measurements are taken using these landmarks their clinical implications may be considered even less.

To date, three methods for visualization of data exported from CBCTs have been described in literature. These are MPVR (MPRV), 3D-virtual

reconstruction view (3D-VRV), and volume rendering or segmentation. The latter involves using computer software to separate 3D image features of interest using image voxel labels to generate models with enhanced quality potential.[23] After performing the PRISMA search, none of the selected articles satisfying our selection criteria involved the reliability of landmark identification after volume rendering or segmentation, and thus this method was not included in this review.

When evaluating the effects of software MPRV versus 3D-VRV for anatomic landmark identification, two of the included studies offer some valuable insight in this area. MPRV has been shown to be more highly reliable than 3D-VRV when considering these two types of visualizations independently.[12] However, most software used to import and view DICOM image formats of CBCT scans have the capability for both of these modalities to be viewed simultaneously. In fact, one study found that this simultaneous viewing actually improved the precision of most landmarks being identified, 15 of 22 to be exact, six of which demonstrating statistical significant with a p>0.05 (*nasion, menton, orbitale left, and Sella turcia*). Only 3 of these 22 landmarks did not show improvement when adding MPRV to 3D-VRV (*gonion left*, *upper right* and *left central incisors*).[18] This is supported by another study, which noted a strong correlation (p>0.7) of this relationship in both intra- and inter-observer repeatability.

It is an interesting finding that viewing MPRV and 3D-VRV concurrently did not improve the precision of identifying the upper right and left central incisors, and that MPR alone demonstrates consistency in accurate landmark

identification. A reasonable explanation for this is that the root apex of mandibular incisors is typically difficult to identify in the sagittal view due to superimposition of the root apices of canines, lateral incisors, or even central incisors.

Limitations

Since CBCT technology is a recent development (circa 1998) and its integration into routine practice in dentistry is relatively recent, all our selected studies are from no earlier than 2009. Although there have been prior attempts to synthesize a single document conveying all the research that has been done in understanding applications of landmark identification in 3D techniques, studies cover a broad range of topics and should not be automatically unified. As such, we opted to place more rigid exclusion criteria than those prior to narrow in on our area of interest.

One of the exclusion criteria we chose to include after the synthesis of selected articles was studies using human dry skulls. This was because the soft tissue attenuation of facial structures could not be accounted for in absolute, despite researchers' best attempts with fluid-filled units.

All selected studies underwent a methodological quality assessment carried out by a single examiner. There is not a gold standard methodological quality assessments tool used in reliability studies presently. This poses difficulties when trying to emphasize the relative weight of certain studies on the overall conclusions.

Conclusions

- Midsagittal plane, followed by bilateral structures demonstrate highest reliability.

- Landmarks with lowest reliability include those marked on the condyle and other anatomic structures with prominent curvatures without definitive boundaries.

- Newly introduced anatomical landmarks for the 3D imaging, such as maxillary and mandibular centroid, have demonstrated adequate intra- and inter-examiner reliabilities.

- A minimum number of dental landmarks were reported on with many demonstrating good to excellent reliability.

References

1. Devereux, L., et al., How important are lateral cephalometric radiographs in orthodontic treatment planning? (1097-6752 (Electronic)).
2. Ludlow, J.B., et al., Precision of cephalometric landmark identification: Cone-beam computed tomography vs conventional cephalometric views. American Journal of Orthodontics and Dentofacial Orthopedics, 2009. **136**(3): p. 312.e1-312.e10.
3. Lagravere, M.O., et al., Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments. (0003-3219 (Print)).
4. Lagravère, M.O., et al., Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. American Journal of Orthodontics and Dentofacial Orthopedics, 2010. **137**(5): p. 598-604.
5. van Vlijmen, O.J.C., et al., A comparison between 2D and 3D cephalometry on CBCT scans of human skulls. International journal of oral and maxillofacial surgery, 2010. **39**(2): p. 156-160.
6. Naji, P., N.A. Alsufyani, and M.O. Lagravere, Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist, 2014. **84**(5): p. 762-772.

7. Lagravere M.O., G.J.M., Flores-Mir C., Carey J., Heo G., Major P.W., Cranial base foramen location accuracy and reliability in cone-beam computerized tomography. Am J Orthod Dentofacial Orthop, 2011. **139**(3): p. 203-10.

8. Lisboa, C.d.O., et al., Reliability and reproducibility of three-dimensional cephalometric landmarks using CBCT: a systematic review. Journal of Applied Oral Science, 2015. **23**(2): p. 112-119.

9. Knobloch, K., P.M. Yoon U Fau - Vogt, and P.M. Vogt, Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. (1878-4119 (Electronic)).

10. van Vlijmen, O.J., et al., Evidence supporting the use of cone-beam computed tomography in orthodontics. (1943-4723 (Electronic)).

11. Bialocerkowski, A., N. Klupp, and P. Bragge, How to read and critically appraise a reliability article. International journal of therapy and rehabilitation 2010. **17**(3): p. 112-114.

12. Da Neiva, M.B., et al., Evaluation of cephalometric landmark identification on CBCT multiplanar and 3D reconstructions. Angle Orthodontist, 2015. **85**(1): p. 11-17.

13. Lee, M., G. Kanavakis, and R.M. Miner, Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review. Angle Orthodontist, 2015. **85**(1): p. 3-10.

14. Ghoneima, A., et al., Accuracy and reliability of landmark-based, surface-based and voxel-based 3D cone-beam computed tompography superimposition methods Orthod Craniofac Res., 2017. **20**(4): p. 227-236.

15. Zamora, N., et al., A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis. Medicina Oral, Patologia Oral y Cirugia Bucal, 2012. **17**(4): p. e678-e688.

16. Frongia, G., et al., Assessment of the reliability and repeatability of landmarks using 3-D cephalometric software. Cranio, 2012. **30**(4): p. 255-263.

17. Schlicher, W., et al., Consistency and precision of landmark identification in three-dimensional cone beam computed tomography scans. European journal of orthodontics, 2012. **34**(3): p. 263-275.

18. Hassan, B., et al., Precision of identifying cephalometric landmarks with cone beam computed tomography in vivo. European journal of orthodontics, 2013. **35**(1): p. 38-44.

19. Chien, P.C., et al., Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. Dentomaxillofacial Radiology, 2009. **38**(5): p. 262-273.

20. de Oliveira, A.E.F., et al., Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2009. **107**(2): p. 256-265.

21. Strachan, A., In the brain of the beholder: the neuropsychological basis of aesthetic preferences. Harvard Brain 2000: p. 7.
22. Mah, J.K., J.C. Huang, and H. Choo, Practical applications of cone-beam computed tomography in orthodontics. The Journal of the American Dental Association, 2010. **141**: p. 7S-13S.
23. Imaging, V., Amira 5: Amira User's Guide. 1999-2009. p. 40-42.

Appendices

Appendix 2.1: Search strategies for various electronic databases

Database	Keywords	Results
PubMed	(((anatomic landmarks[MeSH Terms]) AND cephalometry[MeSH Terms]) AND (cone beam computed tomography OR imaging, three-dimensional OR anatomy, cross-sectional[MeSH Terms])) AND (dimensional measurement accuracy OR reproducibility of results[MeSH Terms]) AND (("1998/01/01"[PDat] : "2017/10/01"[PDat]) AND Humans[Mesh])	68
Medline (via OvidSP)	Search group 1: anatomic landmarks.mp. OR exp Anatomic Landmarks/ Search group 2: cephalometry.mp. OR exp Cephalometry/ OR craniometry.mp. Search group 3: cone beam computed tomography.mp. OR exp Cone-Beam Computed Tomography/ or cone-beam CAT scan.mp. OR cone beam computerized tomography.mp. OR volumetric computed tomography.mp. OR CBCT.mp. or digital volume tomography.mp. OR DVT.mp. or imaging, three-dimensional.mp. OR exp Imaging, Three-Dimensional/ or three dimensional image.mp. OR 3D imaging.mp. OR three-dimensional computer assisted.mp.	67

	Search group 4: dimensional measurement accuracy.mp. OR exp Dimensional Measurement Accuracy/ or reproducibility of results.mp. OR exp "Reproducibility of Results"/ or Bland-Altman.mp. OR reliability.mp. OR validity.mp. OR precision.mp. OR reproducibility of findings.mp OR intraclass correlation coefficient.mp. Search group 5: 1 AND 2 AND 3 AND 4	
EBMR (via OvidSP)	Same as for Medline (via OvidSP)	0
EMBASE (via OvidSP)	Same as for Medline (via OvidSP)	54
Scopus	(("imaging, three-dimensional" OR "three dimensional image" OR "3D imaging" OR "three-dimensional computer-assisted" OR three-dimensional) AND PUBYEAR > 1997) AND (("cone-beam computed tomography" OR "cone-beam computerized tomography" OR cbct OR "volumetric computed tomography" OR "digital volume tomography" OR dvt OR cone-beam OR cone) AND PUBYEAR > 1997) AND ((cephalometr* OR craniometr*) AND PUBYEAR > 1997) AND (("anatomic* landmark" OR landmark OR structure) AND PUBYEAR > 1997)	976
Web of Science	Set #1: TS=(cephalometr*) OR TS=(craniometr*) *DocType=All document types; Language=All languages;* Set #2: TS=(anatomic* landmark) *DocType=All document types; Language=All languages;* Set #3:	221

TS=(reproducibil*) OR TS=(reliabil*) OR
TS=(precision) OR TS=(valid*) OR TS=(accura*)
OR TS=(intraclass correlation coefficient) OR
TS=(Bland-Altman)
*DocType=All document types; Language=All
languages;*

Set #4:

TS=(cone-beam computed tomography) OR
TS= (cone beam computed tomography) OR
TS=(cone-beam CAT scan) OR TS= (volumetric
computed tomography) OR TS=(CBCT) OR
TS=(digital volume tomography) OR TS=(DVT)
OR TS=(imaging, three-dimension*) OR
TS=(imaging) OR TS=(three-dimension*) OR
TS=(three dimension* image) OR TS=(3D
imaging) OR TS=(three-dimension*)
*DocType=All document types; Language=All
languages;*

Set #5:

#4 AND #3 AND #2 AND #1
*DocType=All document types; Language=All
languages;*

Appendix 2.2: Articles excluded in phase 2

Author	Reason
Adams et al.[23] (2004)	• No CBCT taken • Skulls manually measured with calipers
Lagravere & Major[24] (2005)	• Explores specific landmark for use as a reference point in 3D cephalometric analysis
Olzewski et al.[25] (2006)	• Uses medical CT
Subramanya et al.[26] (2006)	• Outlines series of steps to create 3D surface standards
Lou et al.[27] (2007)	• Systematic review
Periagno et al.[28] (2008)	• CBCT's of human dry skulls • Evaluates accuracy of linear measurements
Schlueter et al.[29] (2008)	• Assesses volumetric measurements of condyles at different window levels and widths

Brown et al. [30] (2009)	• CBCT's of human dry skulls • Evaluates accuracy of linear measurements
Hassan et al.[31] (2009)	• CBCT's of human dry skulls • Evaluates influence of patient head position on linear measurements
Moreira et al. [32] (2009)	• CBCT's of human dry skulls • Evaluates accuracy of linear and angular measurements
Couceiro & de Vasconcellos[33] (2010)	• Mentions that "3D images" were printed on photo paper • Unclear if 3D images refer to CBCT-generated cephalograms
Kim et al.[34] (2010)	• Abstract available in English • Full-text available in Korean
Kim et al.[35] (2010)	• Human hemi-mandibles were used • Not true representation of crania-facial complex
Ogawa et al.[36] (2010)	• Outlines a new cephalometric analysis method from different aspects
Chen et al.[37] (2011)	• Compares systems for cephalometric landmark annotation
Cheung et al.[38](2011)	• Generation of norms based on CBCT and 3D photogrammetry in Chinese population • Creation of a database
Damstra et al.[39] (2011)	• Evaluates reliability of linear and angular measurements
Gribel et al.[40] (2011)	• Uses algebraic calculations for measurement translation from 2D into 3D
Lagravere et al.[41] (2011)	• Assesses plane orientation
Medelnik et al.[42] (2011)	• Uses human frozen cadaver heads for landmark identification
Tomasi et al.[43] (2011)	• Human mandibles were used • Not true representation of craniofacial complex
van Vlijmen et al.[44] (2011)	• Compares different CBCT scanners for cephalometric analysis
Wong et al.[45] (2011)	• Generation of norms based on CBCT in Chinese population • Creation of a database
Yildirim et al.[46] (2011)	• Uses medical CT
Reychler[47] (2011)	• Compares cone-beam and low-dose computed tomography
Frongia et al.[48] (2012)	• Compares scanning positions

Frongia et al.[16] (2012)	• CBCT's of human dry skulls • Evaluates influence of patient head position on cephalometric measurements
Kim et al.[49] (2012)	• Compares multi-detector CT and cone-beam CT
Kim et al.[50] (2012)	• Explores 3D mandibular shape models
Patcas et al.[51] (2012)	• Compares CBCT and multi-detector CT
Santos et. al[52] (2012)	• Evaluates reliability and reproducibility of linear measurements
Shibata et al.[53] (2012)	• Compares four 3D coordinate systems
Titiz et al.[54] (2012)	• Uses medical CT
Frongia et al.[55] (2013)	• Explores new method to identify reference system of skull
Gaia et al.[56] (2013)	• Landmarks selected specifically for study • Le Fort 1 osteotomy
Katkar et al. [57] (2013)	• Evaluates reliability of 3D cephalometric landmarks between different CBCT machines
Mendonca et al.[58] (2013)	• Compares cranial anthropometry to 3D photogrammetry and CT • Does not specify that CBCT was used • Patients with cranial asymmetry
Olszewski et al.[59] (2013)	• Compares cone-beam and low-dose computed tomography
Farronato et al.[60] (2014)	• No numerical reliability data included for landmarks
Fuyamada et al. [61] (2014)	• Examines influence of dental experience level in landmark reproducibility in 3D using different methods
Gaia et al.[62] (2013)	• Used landmarks of interest to Le Fort I osteotomy • Compares MSCT and CBCT scans
Gaia et al.[63] (2014)	• Used landmarks of interest to Le Fort I osteotomy
Hwang et al.[64] (2014)	• Investigates plane orientation method
Jacquet et al.[65] (2014)	• Investigates thickness and width of mandible
Jung et al.[66] (2014)	• Uses medical CT
Kim et al.[67] (2014)	• Uses medical CT
Lee et al.[68] (2014)	• Evaluates maxillary and mandibular alveolar and basal bone widths
Lee et al.[69]	• Uses medical CT

(2014)	
Liang et al.[70] (2014)	• Does not specify what type of CT scan obtained • Scans were taken in hospital setting
Fernandes et al.[71] (2014)	• Uses human dry skulls for landmark identification
Metzler et al.[72] (2014)	• Uses a mannequin head for measurements
Jung et al.[73] (2015)	• Evaluates three re-orientation methods
Lisboa et al.[8] (2015)	• Systematic review
Pittayapat et al.[74] (2015)	• Evaluates only Sella turcia-specific reference system
Gupta et al.[75] (2015)	• Evaluates algorithm for automatic detection of landmarks in 3D
Dias et al.[76] (2015)	• Uses 3D skull models to investigate craniometrics analyses
Nam & Hong[77] (2016)	• Investigates 3D soft-tissue computer software
Lemieux et al.[78] (2016)	• Uses human dry skulls for landmark identification
Wikner et al.[79] (2016)	• Uses partially edentulous human dry skulls
Almuzian et al.[80] (2016)	• Investigates effects of Le Fort I osteotomy on airway using CBCT
Ruellas et al.[81] (2016)	• Uses 3D skull models built based on CBCT images
Gupta et al.[82] (2016)	• Comparison of 3D measurements using algorithms and manual identification
Zhang et al.[83] (2016)	• Investigates proposed landmark digitization method

1. Devereux, L., et al., *How important are lateral cephalometric radiographs in orthodontic treatment planning?* (1097-6752 (Electronic)).
2. Ludlow, J.B., et al., *Precision of cephalometric landmark identification: Cone-beam computed tomography vs conventional cephalometric views.* American Journal of Orthodontics and Dentofacial Orthopedics, 2009. **136**(3): p. 312.e1-312.e10.
3. Lagravere, M.O., et al., *Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments.* (0003-3219 (Print)).
4. Lagravère, M.O., et al., *Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images.* American Journal of Orthodontics and Dentofacial Orthopedics, 2010. **137**(5): p. 598-604.

5. van Vlijmen, O.J.C., et al., *A comparison between 2D and 3D cephalometry on CBCT scans of human skulls.* International journal of oral and maxillofacial surgery, 2010. **39**(2): p. 156-160.

6. Naji, P., N.A. Alsufyani, and M.O. Lagravere, *Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT.* Angle Orthodontist, 2014. **84**(5): p. 762-772.

7. Lagravere M.O., G.J.M., Flores-Mir C., Carey J., Heo G., Major P.W., *Cranial base foramen location accuracy and reliability in cone-beam computerized tomography.* Am J Orthod Dentofacial Orthop, 2011. **139**(3): p. 203-10.

8. Lisboa, C.d.O., et al., *Reliability and reproducibility of three-dimensional cephalometric landmarks using CBCT: a systematic review.* Journal of Applied Oral Science, 2015. **23**(2): p. 112-119.

9. Knobloch, K., P.M. Yoon U Fau - Vogt, and P.M. Vogt, *Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias.* (1878-4119 (Electronic)).

10. van Vlijmen, O.J., et al., *Evidence supporting the use of cone-beam computed tomography in orthodontics.* (1943-4723 (Electronic)).

11. Bialocerkowski, A., N. Klupp, and P. Bragge, *How to read and critically appraise a reliability article.* International journal of therapy and rehabilitation 2010. **17**(3): p. 112-114.

12. Da Neiva, M.B., et al., *Evaluation of cephalometric landmark identification on CBCT multiplanar and 3D reconstructions.* Angle Orthodontist, 2015. **85**(1): p. 11-17.

13. Lee, M., G. Kanavakis, and R.M. Miner, *Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review.* Angle Orthodontist, 2015. **85**(1): p. 3-10.

14. Ghoneima, A., et al., *Accuracy and reliability of landmark-based, surface-based and voxel-based 3D cone-beam computed tompography superimposition methods* Orthod Craniofac Res., 2017. **20**(4): p. 227-236.

15. Zamora, N., et al., *A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis.* Medicina Oral, Patologia Oral y Cirugia Bucal, 2012. **17**(4): p. e678-e688.

16. Frongia, G., et al., *Assessment of the reliability and repeatability of landmarks using 3-D cephalometric software.* Cranio, 2012. **30**(4): p. 255-263.

17. Schlicher, W., et al., *Consistency and precision of landmark identification in three-dimensional cone beam computed tomography scans.* European journal of orthodontics, 2012. **34**(3): p. 263-275.

18. Hassan, B., et al., *Precision of identifying cephalometric landmarks with cone beam computed tomography in vivo.* European journal of orthodontics, 2013. **35**(1): p. 38-44.

19. Chien, P.C., et al., *Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-*

dimensional cone beam computed tomography in vivo. Dentomaxillofacial Radiology, 2009. **38**(5): p. 262-273.

20. de Oliveira, A.E.F., et al., *Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2009. **107**(2): p. 256-265.

21. Strachan, A., *In the brain of the beholder: the neuropsychological basis of aesthetic preferences.* Harvard Brain 2000: p. 7.

22. Mah, J.K., J.C. Huang, and H. Choo, *Practical applications of cone-beam computed tomography in orthodontics.* The Journal of the American Dental Association, 2010. **141**: p. 7S-13S.

23. Adams, G.L., et al., *Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls.* American Journal of Orthodontics and Dentofacial Orthopedics, 2004. **126**(4): p. 397-409.

24. Lagravère, M.O. and P.W. Major, *Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography.* American Journal of Orthodontics and Dentofacial Orthopedics, 2005. **128**(5): p. 657-660.

25. Olszewski, R., et al., *3D CT-based cephalometric analysis: 3D cephalometric theoretical concept and software.* Neuroradiology, 2006. **48**(11): p. 853-862.

26. Subramanyan, K., M. Palomo, and M. Hans. *Creation of three dimensional craniofacial standards from CBCT images.*

27. Lou, L., et al., *Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2007. **104**(3): p. 402-411.

28. Periago, D.R., et al., *Linear accuracy and reliability of cone beam CT derived 3-dimensional images constructed using an orthodontic volumetric rendering program.* Angle Orthodontist, 2008. **78**(3): p. 387-395.

29. Schlueter, B., et al., *Cone beam computed tomography 3D reconstruction of the mandibular condyle.* Angle Orthodontist, 2008. **78**(5): p. 880-888.

30. Brown, A.A., et al., *Linear accuracy of cone beam CT derived 3D images.* Angle Orthodontist, 2009. **79**(1): p. 150-157.

31. Hassan, B., P. Van Der Stelt, and G. Sanderink, *Accuracy of three-dimensional measurements obtained from cone beam computed tomography surface-rendered images for cephalometric analysis: Influence of patient scanning position.* European journal of orthodontics, 2009. **31**(2): p. 129-134.

32. Moreira, C.R., et al., *Assessment of linear and angular measurements on three-dimensional cone-beam computed tomographic images.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2009. **108**(3): p. 430-436.

33. Couceiro, C.P. and O. de Vasconcellos Vilella, *2D / 3D Cone-Beam CT images or conventional radiography: Which is more reliable?* Dental Press Journal of Orthodontics, 2010. **15**(5): p. 40-41.

34. Kim, J., D. Lee, and S. Lee, *Comparison of the observer reliability of cranial anatomic landmarks based on cephalometric radiograph and three-dimensional computed tomography scans.* Journal of the Korean Association of Oral and Maxillofacial Surgeons, 2010. **36**(4): p. 262-269.

35. Kim, T.S., et al., *A comparison of cone-beam computed tomography and direct measurement in the examination of the mandibular canal and adjacent structures.* Journal of endodontics, 2010. **36**(7): p. 1191-1194.

36. Ogawa, N., et al., *Application of cone beam CT 3D images to cephalometric analysis.* Orthodontic Waves, 2010. **69**(4): p. 138-150.

37. Chen, J., et al. *A new annotation method for 3D cephalometric landmark in CBCT*.

38. Cheung, L.K., et al., *Three-dimensional cephalometric norms of Chinese adults in Hong Kong with balanced facial profile.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2011. **112**(2): p. e56-e73.

39. Damstra, J., et al., *Reliability and the smallest detectable difference of measurements on 3-dimensional cone-beam computed tomography images.* American Journal of Orthodontics and Dentofacial Orthopedics, 2011. **140**(3): p. e107-e114.

40. Gribel, B.F., et al., *From 2D to 3D: An algorithm to derive normal values for 3-dimensional computerized assessment.* Angle Orthodontist, 2011. **81**(1): p. 5-12.

41. Lagravre, M.O., et al., *Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2011. **111**(6): p. 771-777.

42. Medelnik, J., et al., *Accuracy of anatomical landmark identification using different CBCT- and MSCT-based 3D images : An in vitro study.* Journal of Orofacial Orthopedics, 2011. **72**(4): p. 261-278.

43. Tomasi, C., et al., *Reliability and reproducibility of linear mandible measurements with the use of a cone-beam computed tomography and two object inclinations.* Dentomaxillofacial Radiology, 2011. **40**(4): p. 244-250.

44. van Vlijmen, O.J.C., et al., *Measurements on 3D models of human skulls derived from two different cone beam CT scanners.* Clinical oral investigations, 2011. **15**(5): p. 721-727.

45. Wong, R.W.K., A.C.M. Chau, and U. Hägg, *3D CBCT McNamara's cephalometric analysis in an adult southern Chinese population.* International journal of oral and maxillofacial surgery, 2011. **40**(9): p. 920-925.

46. Yildirim, E., et al., *Evaluation of differences between two and three dimensional cephalometric measurements.* Gulhane Medical Journal, 2011. **53**(1): p. 43-49.

47.	Reychler, H.P., *Reproducibility of 3D cephalometric landmarks on cone-beam and low-dose computed tomography.* 20th International Conference on Oral and Maxillofacial Surgery, ICOMS 2011 Santiago Chile, 2011. **40**(10): p. 1160.

48.	Frongia, G., M.G. Piancino, and P. Bracco, *Cone-beam computed tomography: Accuracy of three-dimensional cephalometry analysis and influence of patient scanning position.* Journal of Craniofacial Surgery, 2012. **23**(4): p. 1038-1043.

49.	Kim, M., et al., *Evaluation of accuracy of 3D reconstruction images using multi-detector CT and cone-beam CT.* Imaging Science in Dentistry, 2012. **42**(1): p. 25-33.

50.	Kim, S.G., et al., *Development of 3D statistical mandible models for cephalometric measurements.* Imaging Science in Dentistry, 2012. **42**(3): p. 175-182.

51.	Patcas, R., et al., *Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: A tale of two CTs.* Dentomaxillofacial Radiology, 2012. **41**(8): p. 637-644.

52.	Santos, T.d.S., et al., *Evaluation of reliability and reproducibility of linear measurements of cone-beam-computed tomography.* Indian Journal of Dental Research, 2012. **23**(4): p. 473-478.

53.	Shibata, M., et al., *Reproducibility of three-dimensional coordinate systems based on craniofacial landmarks: a tentative evaluation of four systems created on images obtained by cone-beam computed tomography with a large field of view.* Angle Orthodontist, 2012. **82**(5): p. 776-784.

54.	Titiz, I., et al., *Repeatability and reproducibility of landmarks--a three-dimensional computed tomography study.* European journal of orthodontics, 2012. **34**(3): p. 276-286.

55.	Frongia, G., P. Bracco, and M.G. Piancino, *Three-dimensional cephalometry: a method for the identification and for the orientation of the skull after cone-bean computed tomographic scan.* Journal of Craniofacial Surgery, 2013. **24**(3): p. e308-11.

56.	Gaia, B.F., et al., *Comparison of precision and accuracy of linear measurements performed by two different imaging software programs and obtained from 3D-CBCT images for le Fort i osteotomy.* Dentomaxillofacial Radiology, 2013. **42**(5).

57.	Katkar, R.A., et al., *Comparison of observer reliability of three-dimensional cephalometric landmark identification on subject images from Galileos and i-CAT cone beam CT.* Dento-Maxillo-Facial Radiology, 2013. **42**(9): p. 20130059.

58.	Mendonca, D.A., et al., *Comparative study of cranial anthropometric measurement by traditional calipers to computed tomography and three-dimensional photogrammetry.* Journal of Craniofacial Surgery, 2013. **24**(4): p. 1106-1110.

59. Olszewski, R., et al., *Reproducibility of three-dimensional cephalometric landmarks in cone-beam and low-dose computed tomography.* Clinical oral investigations, 2013. **17**(1): p. 285-292.

60. Farronato, G., et al., *Assessment of inter- and intra-operator cephalometric tracings on cone beam CT radiographs: comparison of the precision of the cone beam CT versus the latero-lateral radiograph tracing.* Progress in Orthodontics, 2014. **15**: p. 1.

61. Fuyamada, M., et al., *Reproducibility of maxillofacial landmark identification on three-dimensional cone-beam computed tomography images of patients with mandibular prognathism: Comparative study of a tentative method and traditional cephalometric analysis.* Angle Orthodontist, 2014. **84**(6): p. 966-973.

62. Gaia, B.F., et al., *Validity of three-dimensional computed tomography measurements for le Fort i osteotomy.* International journal of oral and maxillofacial surgery, 2014. **43**(2): p. 197-203.

63. Gaia, B.F., et al., *Accuracy and reliability of linear measurements using 3-dimensional computed tomographic imaging software for Le Fort I Osteotomy.* British Journal of Oral & Maxillofacial Surgery, 2014. **52**(3): p. 258-263.

64. Hwang, J.J., et al., *Factors influencing superimposition error of 3D cephalometric landmarks by plane orientation method using 4 reference points: 4 point superimposition error regression model.* PLoS ONE [Electronic Resource], 2014. **9**(11): p. e110665.

65. Jacquet, W., et al., *The influence of slice orientation in 3D CBCT images on measurements of anatomical structures.* IFMBE Proceedings, 2014. **41**: p. 229-232.

66. Jung, S.Y., et al., *Analysis of mandibular structure using 3D facial computed tomography.* Otolaryngology - Head and Neck Surgery (United States), 2014. **151**(5): p. 760-764.

67. Kim, H.J., et al., *Construction and validation of the midsagittal reference plane based on the skull base symmetry for three-dimensional cephalometric craniofacial analysis.* Journal of Craniofacial Surgery, 2014. **25**(2): p. 338-342.

68. Lee, K.M., H.S. Hwang, and J.H. Cho, *Comparison of transverse analysis between posteroanterior cephalogram and cone-beam computed tomography.* Angle Orthodontist, 2014. **84**(4): p. 715-719.

69. Lee, S.H., et al., *Three-dimensional architectural and structural analysis - A transition in concept and design from Delaire's cephalometric analysis.* International journal of oral and maxillofacial surgery, 2014. **43**(9): p. 1154-1160.

70. Liang, C., et al., *Norms of mcnamara's cephalometric analysis on lateral view of 3D CT imaging in adults from Northeast China.* Journal of Hard Tissue Biology, 2014. **23**(2): p. 249-254.

71. Fernandes, T.M.F., et al., *Comparison between 3D volumetric rendering and multiplanar slices on the reliability of linear measurements on CBCT*

images: an in vitro study. Journal of Applied Oral Science, 2015. **23**(1): p. 56-63.

72. Metzler, P., et al., *Validity of the 3D VECTRA photogrammetric surface imaging system for cranio-maxillofacial anthropometric measurements.* (1865-1569 (Electronic)).

73. Jung, P.K., G.C. Lee, and C.H. Moon, *Comparison of cone-beam computed tomography cephalometric measurements using a midsagittal projection and conventional two-dimensional cephalometric measurements.* Korean Journal of Orthodontics, 2015. **45**(6): p. 282-288.

74. Pittayapat, P., et al., *Reproducibility of the sella turcica landmark in three dimensions using a sella turcicaspecific reference system.* Imaging Science in Dentistry, 2015. **45**(1): p. 15-22.

75. Gupta, A., et al., *A knowledge-based algorithm for automatic detection of cephalometric landmarks on CBCT images.* (1861-6429 (Electronic)).

76. Dias P Fau - Neves, L., et al., *CraMs: Craniometric Analysis Application Using 3D Skull Models.* (1558-1756 (Electronic)).

77. Nam, K.U. and J. Hong, *Is Three-Dimensional Soft Tissue Prediction by Software Accurate?* (1536-3732 (Electronic)).

78. Lemieux, G., et al., *Precision and accuracy of suggested maxillary and mandibular landmarks with cone-beam computed tomography for regional superimpositions: An in vitro study.* American Journal of Orthodontics & Dentofacial Orthopedics, 2016. **149**(1): p. 67-75.

79. Wikner, J., et al., *Linear accuracy and reliability of volume data sets acquired by two CBCT-devices and an MSCT using virtual models: a comparative in-vitro study.* (1502-3850 (Electronic)).

80. Almuzian, M., et al., *Effects of Le Fort I Osteotomy on the Nasopharyngeal Airway-6-Month Follow-Up.* (1531-5053 (Electronic)).

81. Ruellas, A.C., et al., *Common 3-dimensional coordinate system for assessment of directional changes.* (1097-6752 (Electronic)).

82. Gupta, A., et al., *Accuracy of 3D cephalometric measurements based on an automatic knowledge-based landmark detection algorithm.* (1861-6429 (Electronic)).

83. Zhang J Fau - Gao, Y., et al., *Automatic Craniomaxillofacial Landmark Digitization via Segmentation-Guided Partially-Joint Regression Forest Model and Multiscale Statistical Features.* (1558-2531 (Electronic)).

I. <u>Study design (15✓)</u>
  A. Objective –description of measurement or procedure under investigation (✓)
  B. Objective –outline of what is known from previous studies (✓)
  C. Sample characteristics –subjects described (✓)
  D. Sample characteristics –assessors described (✓)
  E. Sample size –subjects adequate (✓)
  F. Sample size –assessors adequate (✓)
  G. Sample representation –subjects representative of population (✓)
  H. Sample representation –assessors representative of population (✓)
  I. Sample qualifications/experience –all assessors with necessary experience (✓)
  J. Sample subject variability –heterogeneous subjects (✓)
  K. Minimization of random error –equipment described (✓)
  L. Minimization of random error –subjects described (✓)
  M. Minimization of random error –assessors described (✓)
  N. Clinically stable subjects –yes (✓)
  O. Period of time between measurements –adequate (✓)

II. <u>Study measurements (4✓)</u>
  P. Measurement method –appropriate to the objective (✓)
  Q. Blind measurement –blinding (✓)
  R. Reliability –adequate level of intra-observer agreement (✓)
  S. Reliability –adequate level of inter-observer agreement (✓)

III. <u>Statistical analysis (2✓)</u>
  T. Statistical analysis –appropriate for data (✓)
  U. Confounders –confounders included in analysis (✓)

IV. <u>Results (2✓)</u>
  V. Meaningfulness (e.g. ICC, SEM, CI, kappa) –provided (✓)
  W. Generalized to clinical/research context –yes (✓)

MAXIMUM NUMBER OF ✓s =

Chapter 3 : Reliability of three-dimensional skeletal and dental landmarks

Alycia Sam, Manuel Lagravere-Vich, Heesoo Oh, Carlos Flores-Mir, Hamey Lee, and Andrew Hoang

Introduction

The success of orthodontic treatment is dependent on a clinician's ability to precisely generate a diagnosis and treatment plan tailored to a patient's individual needs. A clinician's experienced judgement is arguably essential to this process, although it is also just as important for them to understand a patient's treatment motivations and to order any appropriate records that will help address these needs and ultimate goals of treatment.

2D cephalometric radiographs are routinely taken as part of orthodontic records. Craniofacial structures are located by choosing defined points, and then analyzed to determine the specific relationships they share with others. Comparison with standardized "norms" often dictates the presence or absence of skeletal, dental and/or soft-tissue disharmonies. The magnitude of these discrepancies can further infer the severity of the problem. With that said, it is essential to recognize the limitations of landmark identification in 2D imaging as it is subjected to magnification errors, distortion, and superimposition of bilateral structures among other issues.[1]

CBCT imaging has afforded the ability to overcome limitations of 2D, and now opens the possibility for depiction of the same relationships but now in all three planes of space (x-, y- and z- axes). With more information comes the need

for improved understanding of the additional information to enable its full

potential use. Landmarks utilized in the analysis of craniofacial structures should

not be only identifiable, but also meaningful and reproducible. Only through this a

thorough understanding of CBCT advancements and its integration into patient

evaluation will the benefits of this technology be truly felt. The first step is to

decipher which 3D landmarks are reliable for placement within observers that is

the objective of this study, and then to move forward with our understanding from

there.

There has been lots of interest in landmark reliability studies in the past

and present. Even back in 2007, Lou *et al.* reviewed available published

information on this topic to gain a better understanding of its developments. All

but one of eight included articles used a sample of human dry skulls that may not

consider other factors like soft-tissue attenuation of the human face, nevertheless

findings were conclusive –each landmark demonstrates a distinctive pattern of

error that ultimately leads to some degree of measurement inaccuracy. It is

possible for reliability error in 2D imaging between time-points minimized to a

variation of 0.5mm or less given practice.[2] Can this be extrapolated to 3D

imaging also? Other findings suggested that traditional landmarks have similar

identification reliabilities in both 2D and 3D spiral CT, though challenged by the

postulations that only 2 of 3 coordinates were used (with omission of the third)

and that conventional 2D cephalometric landmarks may not be optimal when all

three coordinates are used as in 3D.[3]

The objective of this study was to evaluate both intra-examiner reliability of those landmarks in 3D that identify those optimal for 3D quantitative imaging.

Methods

Pre-orthodontic CBCT patient scans (iCAT, Imaging Science International, Hatfield, PA, USA) were obtained from an existing University of Alberta graduate orthodontic program database. An initial review was conducted to acquire full field of view (FOV) CBCTs taken between January 2013 and December 2016. For intra-reliability studies, a total of 10 patients were taken from this database (Class I = 5, Class II-1 = 5). Subjects from this sample population were between the ages of 11.1 and 19.1 years (mean age = 14.1 years), and made up of 5 females and 5 males. A standardized protocol for the iCAT machine was utilized: large FOV 16 X 13.3 cm, voxel size 0.30 mm, 120 kVp, 18.54 mAS, and 8.9 seconds was used for imaging procedures. Raw images were exported as DICOM files, and then uploaded into both Avizo Version 8.1 software (Visualization Sciences Group, Burlington, MA, USA) and Dolphin Imaging 11.7 Premium (Dolphin Imaging & Management Solutions, Chatsworth, CA, USA) for analysis. CBCT scans were identified by codes to maintain anonymity of patients and use in randomization for examiner blinding. Ethics approval was granted from the University of Alberta Research Ethics Board.

The process of selecting which individual landmarks to explore began with the conduction of a systematic review (SR), to thoroughly investigate the available scientific literature that evaluates the reliability of various 3D cephalometric landmarks in 3D imaging. Chapter 2, titled *Reliability of different*

three-dimensional landmarks obtained from CBCT reconstructions: A Systematic

Review, contains the comprehensive compilation of the findings. All this

considered, we designed this study to include and explore the reliability of 3D

cephalometric landmarks that exemplified vital findings of this SR: traditional and

non-traditional landmarks, skeletal and dental landmarks, and differences

between and among observers for identification consistency. Landmarks of

interest were those believed to be useful in representing structures able to

express unique craniofacial relationships that help explain variations in growth

patterns amongst individuals. The one caveat of 3D landmark selection was that

the landmark needed to demonstrate clinically relevant reliability. Hence, an

intra-reliability study evaluating all 3D landmarks of interest to authors to begin

with. An inclusive list and their descriptions is found in *Appendix 3.1*. From these

46 landmarks, those demonstrating clinical relevance to a set of standards

(briefed in the discussion section), were used in all other successive studies in

this thesis.

Statistical Analysis

Standard statistical software package (SPSS Statistics 24 for Mac, IBM)

was used to generate intraclass correlation coefficient (ICC) for the assessment

of intra-examiner reliability. The ICC was chosen to depict the consistency and

reproducibility of landmarking taken at different time-points by the primary

investigator, and thereafter by an additional two examiners. A calculation of

sample size was not performed as a total of 10 subjects (Class I = 5, Class II-1 =

5) for intra-reliability was deemed an adequate sample in inferring significance from the results.

Statistical intra-reliability of 46 landmarks was examined, including 30 skeletal and 16 dental described in *Table 2.1*, using a three-coordinate (x, y and z) system to locate each point spatially.

The mean error and standard deviation for each landmark was assessed. Guidelines for clinically significant examiner variations were followed as outlined in Lagravere et al[4]; that is 1mm for intra-reliability. This cut-off point differs from one study to the next, thus they should not be taken as absolute. With that said, these guidelines were followed in a general sense and a few exceptions were indeed made. ICC's were also generated for each landmark in each axis, following recommendations from Portney *et al.*[5] for interpretation of these values.

Results

Avizo volumetric landmarking for intra-examiner reliability

Mean error for thirteen landmarks were found to have clinical significance (>1mm) in one or more coordinate. In the x-coordinate, this includes *orbitale right* and both *porion*. Several landmarks show variation in the y-coordinate including bilateral *jugal point*, bilateral *antegonion*, and *gonion right*. Even a greater number of landmarks demonstrate this variability in the z-coordinate such as *nasion*, *A-point*, *B-point*, *pogonion*, bilateral *jugal point*, *gonion right*, and bilateral *ramus point*. *Table 2.2* provides a comprehensive compilation of mean error and

standard deviation for each landmark in each coordinate. Mean errors with clinical significance are bolded.

All 46 landmarks demonstrated excellent intra-examiner reliability in each x-, y-, and z-coordinates. With respect to the landmarks themselves, both x- and y-coordinates had ICC's of 0.997 or greater. More variability was seen in landmarks in the z-coordinate, although miniscule, with ICC's ranging between 0.962 and 1.000. The ICC's for the Avizo volumetric landmarking in each coordinate is detailed in *Table 2.3*.

Discussion

The common distinction made between the ease of 2D versus 3D structural landmarking is that they are similar for those in the mid-sagittal plane, although superior for bilateral ones in the latter as demonstrated in Chapter 2. In accordance with that thought, our reliability analysis of landmark placement in 3D revealed excellent intra-examiner ICC>0.96 irrespective of their sagittal plane position. Zamora *et al.* had similar findings with coordinates of all landmarks demonstrating excellent reliability with an ICC>0.95. Furthermore, many landmarks with no errors in determination were actually those in the midsagittal including but not limited to *Sella*, *nasion*, *pogonion*, *menton*, *first molars*, and *upper and lower incisors*.[6] It has been suggested locating bilateral landmarks may be more erroneous due to their location on broad curvatures that makes it more challenging to detect a precise prominence or depression.[7]

It is interesting to note that some authors have reported variability in examiner landmark identification between left and right sides of a bilateral 3D

structures. For example, only *condylion right* in x-axis, *porion left* in y-axis, and *mandibular incisor root apex* left was found to have an intra-observer variability of 1.0-2.0mm whereas others generally had a variation of <1.0mm. Some have attributed this discrepancy due to systematic error by the examiner, whereas others suggest that there may in fact be a neuropsychological linkage between left- and right-handedness that manifests as a preference to a side.[8] No such pattern was seen in this study, with the exception of *orbitale right* and *left*.

The use of a Cartesian coordinate system enabled each landmark to be precisely located in three planes of space; plotted along three planes to comprise a unique coordinate in x-, y-, and z-axes. Since the uniqueness of the coordinate is dependent its precise identification along three different axes, it is possible for landmark identification error to be greater in one than another. This variation was seen in this study as well, with most landmarks z-axis showing the most and x-axis showing the least amount of variation. All things considered, the axis reported with most and least variability fluctuates between studies since it depends on the type of landmarks analyzed. Understanding this fact will aid clinicians in choosing suitable landmarks to use in the calculation of cephalometric measurements and interpretation.[9] For instance, *orbitale* has exhibited inconsistency in x-axis identification and as such should be avoided for making transverse width measurements.[10]

A major goal of this study was to investigate the reliability of not only traditional cephalometric landmarks, but also those non-traditionally used. We opted to include a considerable number of dental and novel skeletal landmarks in

our selection and analysis. This was not based on specific criteria, as the available scientific studies on 3D landmark reliability are heterogeneous in nature and cannot be easily combined. Consequently, the choice in landmark inclusion in this study had to be subjective to a degree. Literature on this topic was reviewed for a collection of 3D landmarks tested for reliability in the past, taking note of those that are meaningful in describing craniofacial pattern differences that have displayed excellent reliabilities. For example, if *infraorbital* has shown superior reliability for landmark identification than its traditional counterpart *orbitale*, this "new" reference landmark may be even more useful in construction of reference planes utilized in cephalometric analyses. One should take note that given a specific landmark name in 3D, explicit definitions are not necessarily the same between studies and should be examined when interpreting meaning. The definition of the landmark itself should be described in all three planes of space.

Novel skeletal landmarks in 3D may offer advantages to those traditionally used such as infraorbital being used in place of orbitale in construction of planes and measurements. This progressiveness could offer a benefit since this study found that when assessing intra-reliability *orbitale right* that a greater mean error than *orbitale left* and bilateral *infraorbita*l. Some authors have even based their entire study on newly defined landmarks such as Naji et al.[11], whereas others created unique landmarks such as *ectomolare* by integrating the need for uniqueness and seeing the excellent reliability of dental locations.[9] In fact, similar to the ICC's of traditional landmarks, those newly defined ones such as *medial* and *lateral condyles* have exhibited excellent intra- and inter-examiner

reliabilities of >0.95 in all coordinates as well.[11] Not all non-traditional

landmarks prove to have good consistency. One study found that bilateral

maxillary cant point had one of the poorest consistencies.[12] Nevertheless,

there appears to be a respectable reason for this continuing interest and

innovation in 3D cephalometric landmark choices and tactics for interpretation.

The creation of new landmarks and modification of conventional 2D

cephalometric landmarks for use in 3D volumetric imaging requires detailed

definitions. If a landmark were to be simply placed on the 3D reconstruction view,

one must also be cognizant that it also occupies a definite corresponding position

in sagittal, coronal and axial multiplanar views. When comparing the reliability of

a landmark from one study to the next, it becomes essential to note the

differences in these definitions in all x-, y- and z-coordinates to ensure that they

can truly be compared. Otherwise generalized interpretations could be made

under false pretenses.

Limitations

It is important to recognize that landmarks chosen for analysis were

predominately based on findings of a SR conducted by these authors prior to

initiating this study. The purpose of the SR was to provide an update to the

examination of all available scientific literature concerning landmarking in 3D

imaging –in hopes to discover past landmarks that have been deemed highly

reliable to locate. Given the vast interest in this area, a considerable number of

landmarks have been previously assessed and thus have met this standard set.

Due to the sheer number of possible landmarks for inclusion, the 46 landmarks

evaluated in our intra-reliability study only offers a snapshot of reliable landmarks in 3D imaging. These were not necessarily newly created landmarks *per se*, but rather a collection of those suggested before. They were new in the sense that several of these chosen landmarks are typically not used in traditional cephalometric analyses. Based on the results of this study, the decision was made to eliminate 6 landmarks with more variation relative to their counterparts for studies that followed. These skeletal landmarks (bilateral *orbitale*, bilateral *jugal point* and bilateral *antegonion*) eliminated from this assessment either had a mean error of greater than 1.0mm in at least one coordinate and/or had an alternative landmark that demonstrated superior to it in identifying a structure. This tactic was exercised in anticipation to limit potential examiner fatigue without sacrificing discovery.

Moreover, the choice in landmark selection was not entirely objective in nature. There was the inclusion of a few landmarks, *B-point and bilateral porion*, in the inter-reliability study even though they had a mean error of greater than 1.5mm in one coordinate. The reasoning for this was because alternative landmarks to represent these structural descriptions would be even more difficult and due to their importance in defining traditional 2D cephalometric planes. This would allow for future comparison of measurements made from 2D and 3D spatial relationships. Omission of other landmarks were made either because there were alternative landmarks available to represent the structure, or variation is believed to pose minimal benefit to be investigated. Some of these alternative

68

landmarks are non-traditional in a sense, and provides the opportunity to explore new ideas.

Finally, the utilization and manipulation of 3D imaging software is likely to pose a significant learning curve and thus not all observers may share the same experience with it. This could explain the minor differences in reliability seen between observers, whereas a lesser variation is seen within an observer expressed as a slightly higher ICC as opposed to their colleagues.

Conclusions

The use of a combination of both 3D-virtual and multiplanar reconstruction views concurrently in 3D landmark placement has proven to have excellent for intra-examiner reliability. Although the analysis of identification has shown that landmark placement in 3D is highly reliable within and between observers, it is pertinent to point out that small discrepancies do exist. Variations between examiners may be attributed to limited calibration training and possibly insufficient clarification of steps required to locate landmarks in the imaging software.

As the routine use of 3D landmarking in CBCT is still relatively new, reliability studies allow clinicians the ability to accurately identify craniofacial landmarks that show great consistency from one image to the next. In such a way, enhanced care can be provided by expressing craniofacial relationships in all three planes of space for an individualized patient.

Many traditional and non-traditional landmarks in 3D have proven excellent reliability in 3D. Of notability, non-traditional landmarks such as some

craniofacial foramina show particular promise. More research is needed to use

what we know about the reliability of landmarks in 3D to express important

relationships that can be used to better classify orthodontic patient populations

for diagnosis and treatment planning unlike ever before.

References

1. Gribel, B.F., et al., Accuracy and reliability of craniometric measurements on lateral cephalometry and 3D measurements on CBCT scans. (1945-7103 (Electronic)).
2. Lou, L., et al., Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2007. **104**(3): p. 402-411.
3. Kragskov, J., et al., Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans. (1055-6656 (Print)).
4. Lagravere, M.O., et al., Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images. (1528-395X (Electronic)).
5. Portney, L. and M. Watkins, Foundations of clinical research: Applications to practice. 2009, Upper Saddle River, NJ: Pearson/Prentice Hall
6. Zamora, N., et al., A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis. Medicina Oral, Patologia Oral y Cirugia Bucal, 2012. **17**(4): p. e678-e688.
7. Chien, P.C., et al., Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. Dentomaxillofacial Radiology, 2009. **38**(5): p. 262-273.
8. Strachan, A., In the brain of the beholder: the neuropsychological basis of aesthetic preferences. Harvard Brain 2000: p. 7.
9. Lagravere, M.O., et al., Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments. (0003-3219 (Print)).
10. Ludlow, J.B., et al., Precision of cephalometric landmark identification: Cone-beam computed tomography vs conventional cephalometric views. American Journal of Orthodontics and Dentofacial Orthopedics, 2009. **136**(3): p. 312.e1-312.e10.
11. Naji, P., N.A. Alsufyani, and M.O. Lagravere, Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist, 2014. **84**(5): p. 762-772.

12. Schlicher, W., et al., Consistency and precision of landmark identification in three-dimensional cone beam computed tomography scans. European journal of orthodontics, 2012. **34**(3): p. 263-275.

Appendices

Appendix 3.1: Description of skeletal and dental anatomic landmarks used in reliability

Order	Type	Landmark	Definition	3D Reconstruction	Sagittal View (YZ)	Axial View (XY)	Coronal View (XZ)
1	Skeletal	Sella	Centre-most point of pituitary fossa				
2	Skeletal	Nasion	Most anterior point of intersection of fronto-nasal suture with inter-nasal suture				
3	Skeletal	Basion	Most antero-inferior point of foramen magnum				
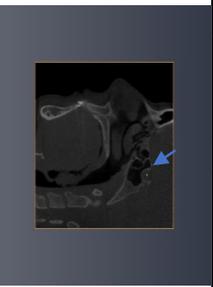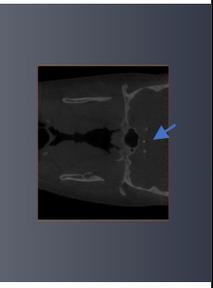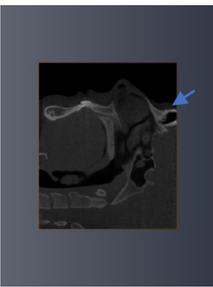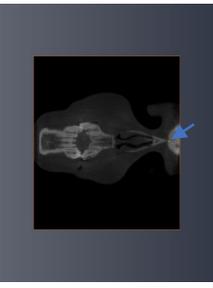
4	Skeletal	A-point	Most posterior point of maxillary concavity, located between anterior nasal spine and prosthion			
5	Skeletal	B-point	Most posterior point of mandibular concavity, located between infradentale and pogonion			
6	Skeletal	Pogonion	Most anterior point of mandibular symphysis			
7	Skeletal	Menton	Most inferior point of mandibular symphysis			
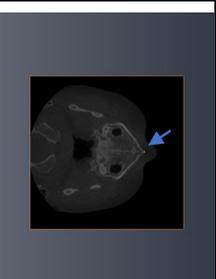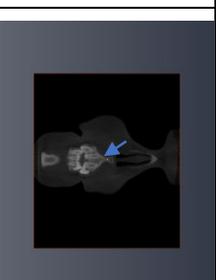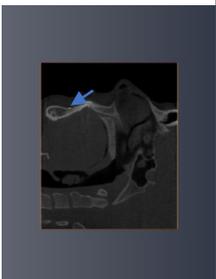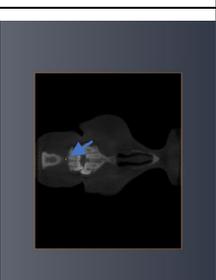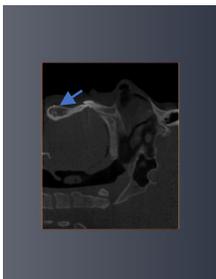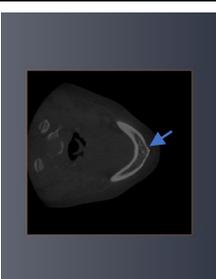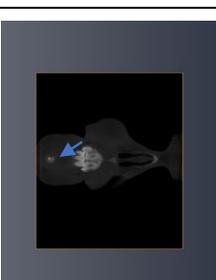
8	Skeletal	Anterior nasal spine	Tip of anterior nasal spine, located above A-point				
9	Skeletal	Posterior nasal spine	Tip of posterior nasal spine				
10	Skeletal	Orbitale (right)	Most antero-inferior point of right orbit				
11	Skeletal	Orbitale (left)	Most antero-inferior point of left orbit				
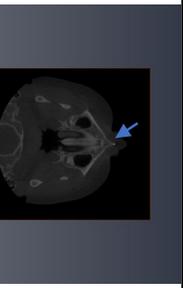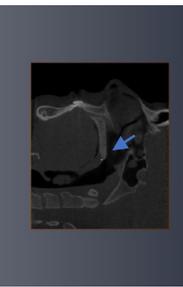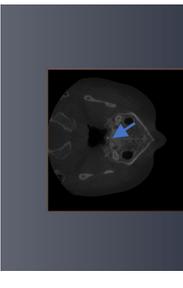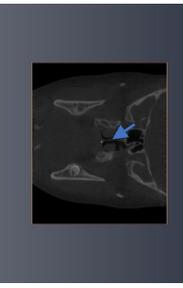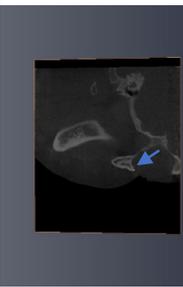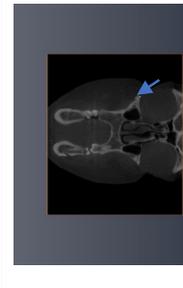
#		Landmark	Definition			
12	Skeletal	Infraorbital foramen (right)	Most antero-superior point of right infraorbital foramen opening			
13	Skeletal	Infraorbital foramen (left)	Most antero-superior point of left infraorbital foramen opening			
14	Skeletal	Incisive foramen	Most postero-superior point of incisive foramen opening			
15	Skeletal	Mental foramen (right)	Most antero-inferior point of right mental foramen opening			
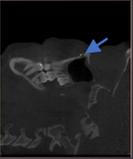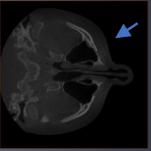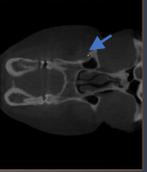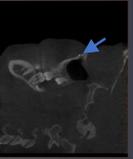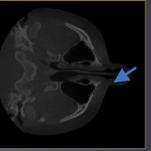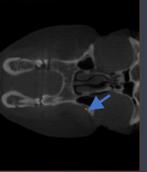
16	Skeletal	Mental foramen (left)	Most antero-inferior point of right mental foramen opening				
17	Skeletal	Superior condyle (right)	Most superior point of right mandibular condyle				
18	Skeletal	Superior condyle (left)	Most superior point of right mandibular condyle				
19	Skeletal	Posterior condyle (right)	Most posterior point of right mandibular condyle				
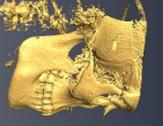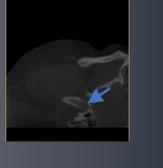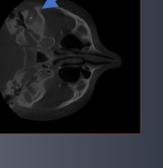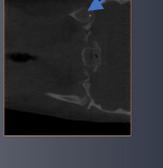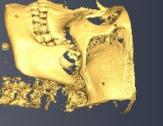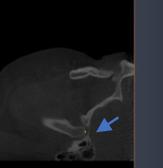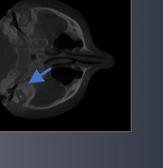
#	Type	Landmark	Description			
20	Skeletal	Posterior condyle (left)	Most posterior point of left mandibular condyle			
21	Dental	1.1 incisal edge	Most inferior point of incisal edge maxillary right central incisor			
22	Dental	1.1 root apex	Most superior point of root apex maxillary right central incisor			
23	Dental	4.1 incisal edge	Most superior point of incisal edge mandibular right central incisor			
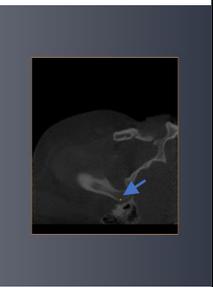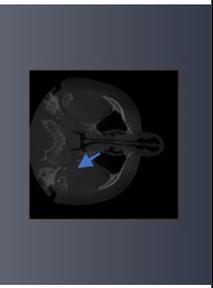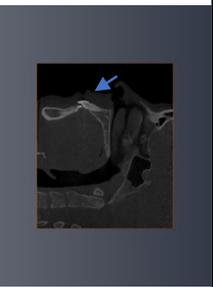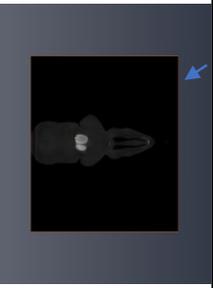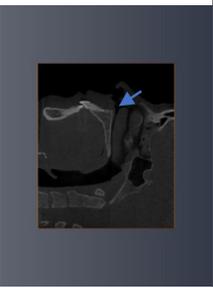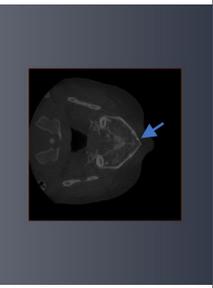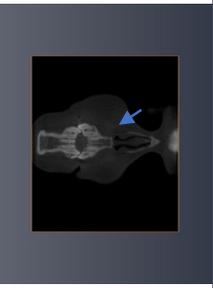
24	Dental	4.1 root apex	Most inferior point of root apex mandibular right central incisor			
25	Dental	1.6 buccal	Most mid- and anterior- point of buccal surface maxillary right first molar			
26	Dental	1.6 root apex	Most superior point of mesio- buccal root apex maxillary right first molar			
27	Dental	4.6 buccal	Most mid- and anterior- point of buccal surface mandibular right first molar			
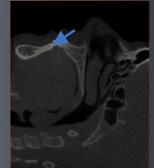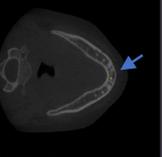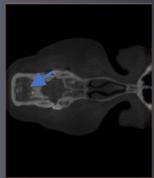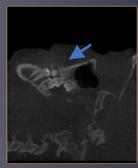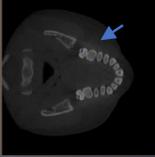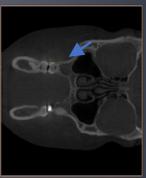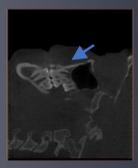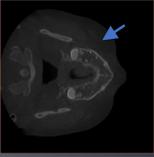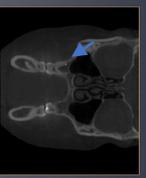
28	Dental	4.6 root apex	Most inferior point of mesio-buccal root apex mandibular right first molar			
29	Dental	2.1 incisal edge	Most inferior point of incisal edge maxillary left central incisor			
30	Dental	2.1 root apex	Most superior point of root apex maxillary left central incisor			
31	Dental	3.1 incisal edge	Most superior point of incisal edge mandibular left central incisor			
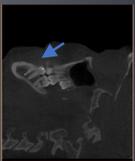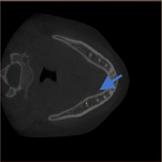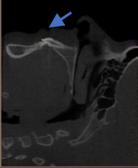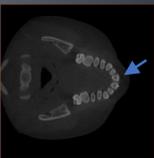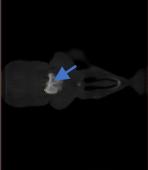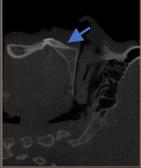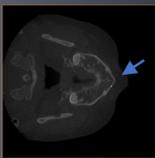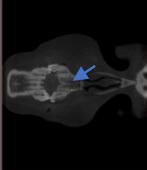
32	Dental	3.1 root apex	Most inferior point of root apex mandibular left central incisor			
33	Dental	2.6 buccal	Most mid- and anterior- point of buccal surface maxillary left first molar			
34	Dental	2.6 root apex	Most superior point of mesio- buccal root apex maxillary left first molar			
35	Dental	3.6 buccal	Most mid- and anterior- point of buccal surface mandibular left first molar			
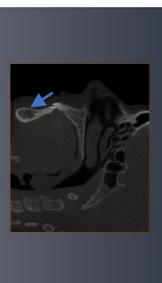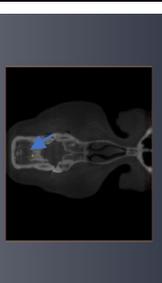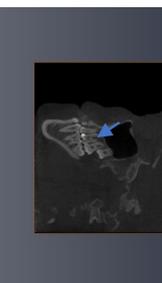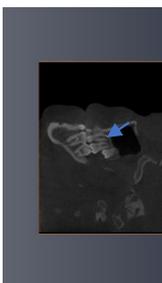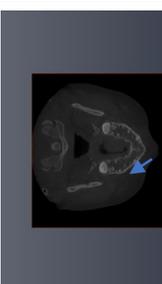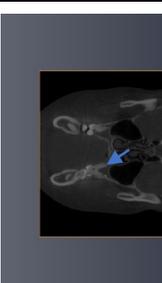
#	Type	Landmark	Description			
36	Dental	3.6 root apex	Most inferior point of mesio-buccal root apex mandibular left first molar			
37	Skeletal	Porion (right)	Most superior point of right external acoustic meatus			
38	Skeletal	Porion (left)	Most superior point of left external acoustic meatus			
39	Skeletal	Jugal point (right)	Deepest point of right jugal process of maxilla			
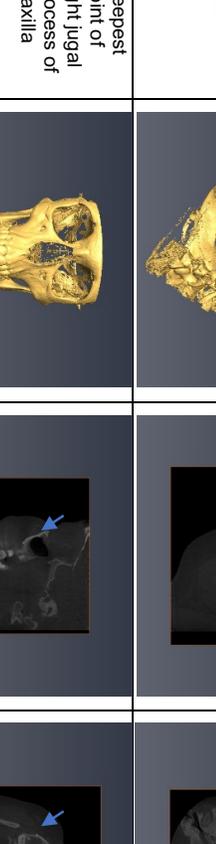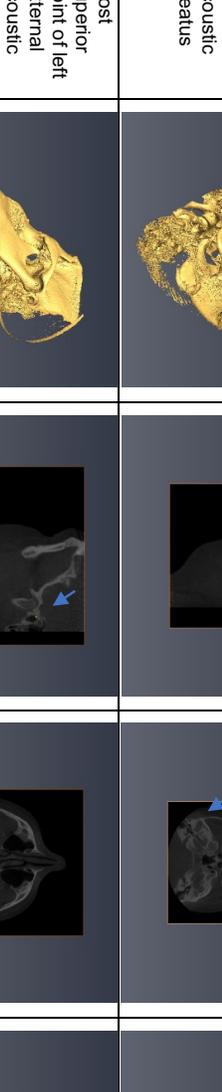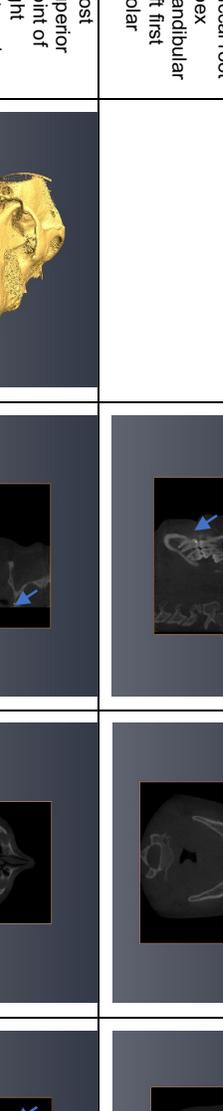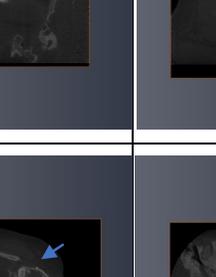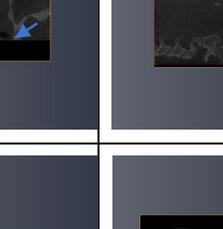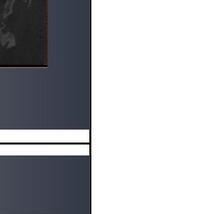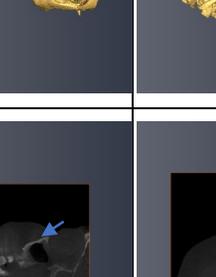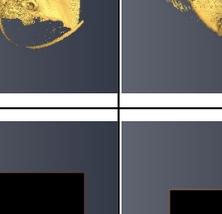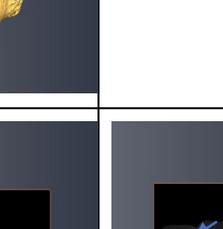
40	Skeletal	Jugal point (left)	Deepest point of left jugal process of maxilla				
41	Skeletal	Antegonion (right)	Most superior point of right antegonial notch of mandible				
42	Skeletal	Antegonion (left)	Most superior point of left antegonial notch of mandible				
43	Skeletal	Gonion (right)	Most postero-inferior point at angle of right mandible				
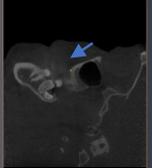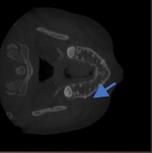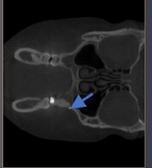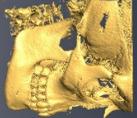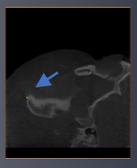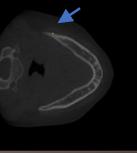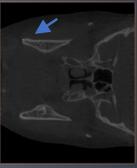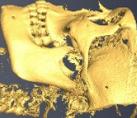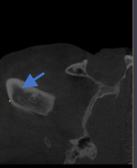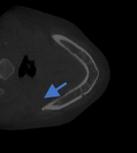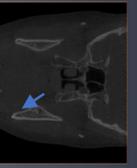
44	Skeletal	Gonion (left)	Most postero-inferior point at angle of left mandible				
45	Skeletal	Ramus point (right)	Most posterior point of ramus of right mandible				
46	Skeletal	Ramus point (left)	Most posterior point of ramus of left mandible				
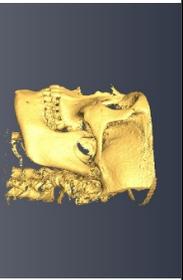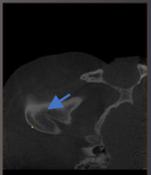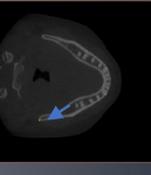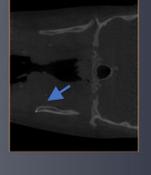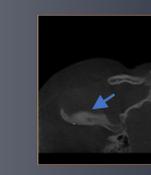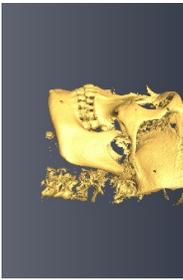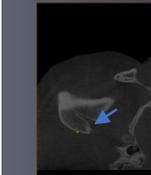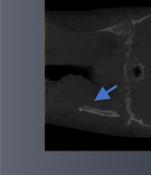
Appendix 3.2: Mean errors and standard deviations (mm) for Avizo software 3D landmark intra-examiner reliability testing.

		x-coordinate		y-coordinate		z-coordinate	
		Mean error	Standard deviation	Mean error	Standard deviation	Mean error	Standard deviation
1	Sella	0.87	0.71	0.35	0.34	0.31	0.18
2	Nasion	0.40	0.34	0.47	0.71	**1.01**	1.58
3	Basion	0.80	0.76	0.56	0.50	0.35	0.12
4	A-point	0.53	0.42	0.27	0.34	**1.01**	0.39
5	B-point	0.33	0.35	0.20	0.32	**1.53**	0.99
6	Pogonion	0.33	0.35	0.13	0.28	**1.05**	0.73
7	Menton	0.40	0.47	0.47	0.32	0.25	0.10
8	ANS	0.33	0.35	0.47	0.45	0.23	0.15
9	PNS	0.47	0.45	0.33	0.35	0.27	0.24
10	Orbitale (R)	**1.27**	0.73	0.60	0.38	0.20	0.11
11	Orbitale (L)	0.98	0.71	0.53	0.53	0.19	0.10
12	Infraorbital (R)	0.40	0.84	0.13	0.28	0.29	0.39
13	Infraorbital (L)	0.99	1.30	0.60	1.06	0.85	0.99
14	Incisive foramen	0.13	0.28	0.60	1.24	0.67	0.74
15	Mental foramen (R)	0.33	0.47	0.20	0.45	0.19	0.33
16	Mental foramen (L)	0.67	0.21	0.20	0.32	0.13	0.15
17	Superior condyle (R)	0.40	0.47	0.67	0.31	0.12	0.14
18	Superior condyle (L)	0.33	0.35	0.36	0.33	0.56	1.05
19	Posterior condyle (R)	0.53	0.28	0.09	0.22	0.36	0.14
20	Posterior condyle (L)	0.37	0.33	0.13	0.28	0.59	0.56
21	1.1 incisal edge	0.40	0.34	0.13	0.28	0.21	0.09
22	1.1 root apex	0.47	0.32	0.33	0.35	0.45	0.21
23	4.1 incisal edge	0.60	0.38	0.00	0.00	0.15	0.10
24	4.1 root apex	0.53	0.42	0.40	0.47	0.57	0.49

25	1.6 buccal	0.53	0.28	0.53	0.28	0.39	0.25
26	1.6 root apex	0.80	0.61	0.53	0.53	0.65	0.46
27	4.6 buccal	0.53	0.82	0.47	0.45	0.52	0.24
28	4.6 root apex	0.93	1.14	0.53	0.61	0.57	0.61
29	2.1 incisal edge	0.67	0.77	0.27	0.47	0.18	0.09
30	2.1 root apex	0.67	0.77	0.47	0.45	0.59	0.35
31	3.1 incisal edge	0.47	0.71	0.20	0.32	0.19	0.09
32	3.1 root apex	0.40	0.72	0.40	0.47	0.62	0.61
33	2.6 buccal	0.42	0.32	0.53	0.28	0.35	0.20
34	2.6 root apex	0.40	0.34	0.27	0.34	0.55	0.24
35	3.6 buccal	0.62	1.23	0.60	0.21	0.62	0.38
36	3.6 root apex	0.80	0.61	0.40	0.47	0.36	0.15
37	Porion (R)	**1.67**	1.23	0.50	0.53	0.38	0.33
38	Porion (L)	**1.62**	1.29	0.92	0.55	0.27	0.20
39	Jugal point (R)	0.73	0.38	**2.07**	1.39	**1.28**	0.49
40	Jugal point (L)	0.71	0.42	**2.33**	0.72	**1.41**	0.65
41	Antegonion (R)	0.47	0.45	**1.47**	0.98	0.51	0.42
42	Antegonion (L)	0.51	0.61	**1.33**	1.34	0.57	0.57
43	Gonion (R)	0.53	0.28	**1.42**	0.88	**1.22**	0.70
44	Gonion (L)	0.37	0.46	0.49	0.53	0.93	0.72
45	Ramus point (R)	0.40	0.47	0.68	0.42	**1.31**	1.21
46	Ramus point (L)	0.43	0.45	0.38	0.33	**1.11**	0.78

Appendix 3.3: ICC's for Avizo software 3D volumetric landmark intra-examiner reliability testing, coordinates, for average measures.

	x-coordinate			y-coordinate			z-coordinate		
	ICC	95% confidence interval		ICC	95% confidence interval		ICC	95% confidence interval	
		Lower bound	Upper bound		Lower bound	Upper bound		Lower bound	Upper bound
1	0.999	0.998	1.000	1.000	1.000	1.000	1.00	0.999	1.000
2	1.000	1.000	1.000	1.000	0.999	1.000	0.962	0.893	0.990
3	0.999	0.998	1.000	1.000	1.000	1.000	1.000	0.999	1.000
4	1.000	0.999	1.000	1.000	1.000	1.000	0.994	0.983	0.998
5	1.000	1.000	1.000	1.000	1.000	1.000	0.988	0.966	0.997
6	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.986	0.999
7	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.999	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
9	1.000	0.999	1.000	1.000	1.000	1.000	0.999	0.998	1.000
10	0.999	0.997	1.000	1.000	0.999	1.000	1.000	0.999	1.000
11	0.999	0.998	1.000	1.000	0.999	1.000	1.000	0.999	1.000
12	0.999	0.998	1.000	1.000	1.000	1.000	0.999	0.997	1.000
13	0.999	0.996	1.000	0.999	0.999	1.000	0.991	0.974	0.998
14	1.000	1.000	1.000	0.999	0.998	1.000	0.991	0.974	0.998
15	1.000	0.999	1.000	1.000	1.000	1.000	0.999	0.998	1.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
17	1.000	0.999	1.000	1.000	0.999	1.000	1.000	1.000	1.000
18	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.992	0.999
19	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.999	1.000
20	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	1.000
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
22	1.000	0.999	1.000	1.000	0.999	1.000	0.999	0.996	1.000
23	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
24	1.000	0.999	1.000	1.000	1.000	1.000	0.998	0.994	0.999
25	1.000	0.999	1.000	1.000	1.000	1.000	0.999	0.997	1.000
26	0.999	0.998	1.000	1.000	0.999	1.000	0.997	0.991	0.999
27	0.999	0.999	1.000	1.000	0.999	1.000	0.999	0.996	1.000
28	0.999	0.997	1.000	1.000	0.999	1.000	0.998	0.994	0.999
29	0.999	0.998	1.000	1.000	1.000	1.000	1.000	0.999	1.000
30	0.999	0.998	1.000	1.000	0.999	1.000	0.998	0.992	0.999
31	1.000	0.999	1.000	1.000	1.000	1.000	1.000	0.999	1.000
32	1.000	0.999	1.000	1.000	1.000	1.000	0.997	0.992	0.999
33	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	1.000
34	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.995	0.999
35	0.999	0.997	1.000	1.000	0.999	1.000	0.998	0.994	0.999
36	0.999	0.999	1.000	1.000	0.999	1.000	0.999	0.998	1.000

37	0.998	0.994	0.999	1.000	0.999	1.000	1.000	0.99	1.00
38	0.997	0.993	0.999	1.000	0.999	1.000	1.000	0.999	1.000
39	1.000	0.999	1.000	0.998	0.990	0.999	0.989	0.969	0.997
40	1.000	0.999	1.000	0.998	0.993	0.999	0.985	0.956	0.996
41	1.000	0.999	1.000	0.999	0.996	1.000	0.999	0.997	1.000
42	1.000	0.999	1.000	0.999	0.997	1.000	0.999	0.996	1.000
43	1.000	0.999	1.000	0.999	0.997	1.000	0.996	0.989	0.999
44	1.000	0.999	1.000	1.000	0.999	1.000	0.997	0.993	0.999
45	1.000	0.999	1.000	1.000	0.999	1.000	0.994	0.982	0.998
46	1.000	0.999	1.000	1.000	1.000	1.000	0.996	0.987	0.999

Chapter 4 : Three-dimensional skeletal and dental relationship differences in orthodontic patients with Class I and Class II Division 1 malocclusions

Alycia Sam, Manuel Lagravere-Vich, Carlos Flores-Mir, Carlos Flores-Mir, Heesoo Oh, Dan Romanyk

Introduction

Over the past decade within the realm of orthodontics, technological advances have offered practitioners the ability to produce radiographic images of increased potential diagnostic value than ever before. The dramatic evolution of radiographic imaging was first demonstrated by the transition from conventional films to digital imaging –with the added benefits of being immediately processed, having image enhancement and magnification capabilities and ability to be digitally archived.[1] More recently, this was followed by the advent of cone beam computed tomography (CBCT) and its implementation into the dental field.

One of the advantages of CBCT imaging is the ability to obtain various radiographic images from a single scan. For instance, the generation of a customary two-dimensional (2D) lateral cephalometric radiograph can now be matched by the ability to create a CBCT-generated lateral cephalometric radiograph.[2] Although one must be mindful of limitations in both scenarios as they are indeed 2D representations of a vaster three-dimensional (3D) object.

Cephalometric radiography is a standardized technique employed to provide a better understanding of an individual's craniofacial structures in three planes of space: antero-posteriorly, vertically, and transversely. The relationship between the dentition and underlying skeletal base can also be understood in

terms of these spatial relationships, in addition to the soft-tissue drape that overlies it. They are particularly useful when considering surgical-orthodontic treatment approaches, when growth predictions are needed, and for cephalometric superimpositions of pre-, during- and post-treatment images to better illustrate therapeutic changes or simply for the purposes of monitoring craniofacial growth.[3]

Each type of cephalometric film (lateral, posterior-anterior, and basilar) depicts only two of these three dimensions and consequently, the landmarks analyzed to classify "normal" relationships may not necessarily portray the entire picture at hand. 3D imaging overcomes limitations of 2D such as image distortion and magnification, while also making the identification of some landmarks around curvatures and of bilateral structures more accurate and reliable.[4, 5] This raises the question as to how traditional cephalometric landmarks and analyses can be extrapolated and applied in interpreting these 3D images, and whether specific "new" landmarks and subsequent measurements can provide enhanced diagnostic value to better categorize various malocclusions with "new" norms.

1. The first objective of this study was to investigate the difference between 2D normative values and 3D skeletal and dental measurements, for those linear and angular measurements that can be obtained with both methods of imaging.

2. The second objective of this study was to investigate if differences between measurements taken from right and left sides in the two malocclusion groups could be noted.

3. A final aim was determining the existence of 3D skeletal and dental relationships to adequately categorize orthodontic malocclusions, specifically Class I and Class II Division 1 (Class II-1).

Methods

Pre-orthodontic CBCT patient scans (iCAT, Imaging Science International, Hatfield, PA, USA) taken for routine orthodontic diagnosis and treatment planning were obtained from an existing University of Alberta graduate orthodontic program database with the same methodology as outlined in Chapter 3: *Reliability of three-dimensional skeletal and dental landmarks*.

Allocation of patients' full-FOV CBCT scans into a category

One individual, a graduate orthodontic resident (AS), evaluated each patient's scan in sequence beginning from January 2013 onward. Intra- and extra-oral photographs, in addition to study models (plaster or digital depending on availability) and CBCT- generated cephalogram were also available for the allocation process. Review of information allowed for orthodontic patients to be placed into one of various classifications: Class I malocclusion, Class II-1 malocclusion, Class II-2 malocclusion, Class III malocclusion, and fulfills exclusion criteria. The comprehensive checklist used for malocclusion classification during initial patient reviews can be found in *Appendix 4.1*. Patients that met any exclusion criteria were immediately eliminated. Exclusion criteria included craniofacial deformities (e.g. cleft lip and/or plate), facial asymmetries (e.g. chin deviations >2mm) and subdivision malocclusions.

Inclusion criteria for each Class I and Class II-1 malocclusion specified

specific skeletal, dental and soft-tissue drape characteristics that had to be withheld for its classification. Class II-2 and Class III malocclusions were systematically recorded solely for reference. It should be noted that evaluations were qualitative in nature, meaning that cephalometric measurements and values were not presented to ensure the evaluator stayed impartial when making classification decisions. In this way, group classification remained relatively independent of 2D cephalometric norm values during group allocations. Another note is that a case may not necessarily fulfill all criteria specified for a classification to be made. For example, a patient with a relatively straight soft-tissue profile may have still been allocated into the Class II-1 group due to an increased overjet.   The single examiner made decisions on cases that presented as clear definitions on extremes. For any cases that fell in the grey area, an additional two examiners (ML, CF) were sought to review it. Cephalometric measurements and values were now available for reference to assist in classification if needed now. A consensus of all evaluators for a classification choice had to be made for each scan to enable its use. The selected scans of patients for use in the remainder of the study had the following age and gender demographics: Class I subjects (ages 9.1-28.6 years, mean age = 15.5 years, 18 females and 12 males), Class II subjects (ages 8.1-45.3 years, mean age = 13.9 years, 13 females and 17 males). Sixty patient scans, 30 Class I and 30 Class II-1, were randomly selected after group classification for use in the remainder of the study. Continued analysis of these two malocclusion group samples allowed us to pursue our study objectives.

Statistical analysis

A calculation of sample size was not performed as a total of 60 subjects ($n_1 = 30$, $n_2 = 30$) were deemed an adequate sample in inferring significance from the results. Standard statistical software package (SPSS Statistics 24 for Mac, IBM) was used to analyze study data in its entirety. To test for significance, a significance level œ = 0.05 was chosen.

To address our first objective, a multivariate analysis was first used to compare mean vectors of an estimated 3D population (estimated from a sample) and standard 2D norm population for significant differences in various cephalometric values. The null hypothesis was that there is no difference among 6 mean measurements jointly among estimated 3D population mean and standard 2D norm population mean. This was conducted separately for Class I and Class II-1 malocclusion groups. These measurements consisted of 2 distances (*S-N*, *Ba-N*), 3 angles (*SNA*, *SNB*, *ANB*) and 1 ratio (*N-ANS : N-Me*) as labelled in *Table 4.1*.

Table 4.1: Six linear, angular and ratio cephalometric measurements used in sample vs. standard norm multivariate analysis

Abbreviation	Measurement	Type	2D standard norm
S-N	Sella – Nasion	Distance (mm)	68
Ba-N	Basion – Nasion	Distance (mm)	100.4
SNA	Sella – Nasion – A-point	Angle (˚)	82
SNB	Sella – Nasion – B-point	Angle (˚)	80
ANB	SNA-SNB	Angle (˚)	2
N-ANS : N-Me	Nasion – Anterior nasal spine / Nasion – Menton	Ratio (%)	0.45 or 45%

A multivariate analysis was also used to compare mean measurements taken from right and left vantage points of various measurements, for significant differences in various cephalometric measurement values between sides of Class I and Class II-1 malocclusion groups. The null hypothesis was that there is no difference among 9 mean measurement values jointly among right and left measurement mean. A total of 9, one right and one left, measurements were compared in this way. The distance and angle measurements with descriptions of these are found in *Table 4.2*.

Table 4.2: Avizo measurement descriptions for those used in right vs. left multivariate analysis

Abbreviation	Measurement	Type
U1-APog (right)	1.1 incisal edge – A-point-Pogonion plane	Distance (mm)
U1-APog (left)	2.1 incisal edge – A-point-Pogonion plane	Distance (mm)
CoSup-A (right)	Superior condyle (right) – A-point	Distance (mm)
Co-Sup-A (left)	Superior condyle (left) – A-point	Distance (mm)
CoSup-Pog (right)	Superior condyle (right) – Pogonion	Distance (mm)
CoSup-Pog (left)	Superior condyle (left) – Pogonion	Distance (mm)
U1-SN (right)	1.1 long axis – Sella-Nasion plane	Angle (˚)
U1-SN (left)	2.1 long axis – SN plane (S-N)	Angle (˚)
U1-PP (right)	1.1 long axis – Palatal plane (ANS-PNS)	Angle (˚)
U1-PP (left)	2.1 long axis – Palatal plane (ANS-PNS)	Angle (˚)
U1-L1 (right)	1.1 long axis – 4.1 long axis	Angle (˚)
U1-L1 (left)	2.1 long axis – 3.1 long axis	Angle (˚)
MP-SN (right)	Mandibular plane (Gonion-Menton) – SN plane (S-N), right	Angle (˚)
MP-SN (left)	Mandibular plane (Gonion-Mention) – SN plane (S-N), left	Angle (˚)
L1-MP (right)	4.1 long axis – Mandibular plane (Gonion-Menton), right	Angle (˚)
L1-MP (left)	3.1 long axis – Mandibular plane (Gonion-Menton), left	Angle (˚)

PP-MP (right)	Palatal plane (ANS-PNS) – Mandibular plane (Gonion-Mention), right	Angle (˚)
PP-MP (left)	Palatal plane (ANS-PNS) – Mandibular plane (Gonion-Mention), left	Angle (˚)

Sample vs. standard norm and right vs. left multivariate analyses model assumptions are found in *Appendix 4.2-4.5* for the prior and *Appendix 4.10-4.13* for the latter.

Two statistical models can determine which 3D linear and angular measurements discriminate 'best' between Class I and Class II-1 malocclusion classifications: discriminant function analysis (DA) and one-way MANOVA. However, DA is superior as it can determine how many dimensions we need to express this relationship and degree to which discriminator variables can be used distinguish between these two groups. As such, DA was chosen for this study. Follow-up analyses were conducted for the development of a computed index to use in predicting malocclusion classifications for future observations, based on these chosen continuous variables. These will be discussed further in the results section.

Model assumptions were assessed prior to the analysis and found in *Appendix 4.18-4.22*. There was violation of multivariate normality, as assessed by inspection of boxplots. However, DA is robust to this assumption if groups are of nearly equal size that holds true for this data set. The assumption of homogeneity of variance-covariance matrices was met as Box's M-test yielded a significance level œ >0.001. To assess whether there was multicollinearity among predictor variables, a pooled within-groups correlation matrices was

generated. There were two correlations that equaled or exceeded 0.7 in absolute

value, so there was also violation of multicollinearity. Nevertheless, as structures

of the craniofacial complex are undoubtedly related, the decision was made to

continue despite this. The linearity assumption was reasonably met based on a

total of 96 scatter plots for all pairs of predictor variables, separated by

classification group. Lastly, visual inspection of boxplots was suggestive of no

univariate outliers although the chi-square Q-Q plot of the Mahalanobis distance

revealed the presence of multivariate outliers, particularly in the Class I group. As

such, the test was run with and without outliers included in analysis to determine

their effects on the results.

The DA was performed to uncover the dimensions of cephalometric

measurements that can differentiate between two groups of malocclusion

classifications: Class I, Class II-1. Eight predictor variables, M1 through M8

(*Table 4.3*), were used to assess how well these two groups could be predicted

from such discriminating variables. These predictor variables were

predetermined; arbitrarily chosen by the authors based on findings of Chapter 3:

Reliability of three-dimensional skeletal and dental landmarks and measurements

postulated to be particularly useful. The null hypothesis was that in the

population, group malocclusion classification is not related to or cannot be

predicted from measurement values on the discriminant function.

Table 4.3: Avizo measurement descriptions for those used in DA

Abbreviation	Measurement	Type
M1	Infraorbital (right) – Mental foramen (right)	Distance (mm)
M2	Infraorbital (left) – Mental foramen (left)	Distance (mm)

M3	1.6 buccal – Infraorbital (right) – Infraorbital (left)	Angle (˚)
M4	Infraorbital (right) – Infraorbital (left) – 2.6 buccal	Angle (˚)
M5	1.6 root apex – Infraorbital (right) – 1.6 buccal	Angle (˚)
M6	2.6 root apex – Infraorbital (left) – 2.6 buccal	Angle (˚)
M7	1.6 buccal – A-point – 2.6 buccal	Angle (˚)
M8	3.6 buccal – B-point – 4.6 buccal	Angle (˚)

Results

Sample vs. standard norm t-tests found evidence of mean measurement value differences for two measurements (*S-N, ANB*) in the Class I malocclusion group and three measurements (*S-N*, *SNB*, *ANB*) in the Class II-1 malocclusion group between the estimated population mean (estimated from the sample) and standard 2D norm population mean, at an œ level of 0.05. Descriptive statistics and significance for these measurements are in *Table 4.4*. The 3D *S-N* distance was smaller than 2D in both Class I and Class II-1 sample populations.

Table 4.4: Descriptive statistics and significance for those measurements used in sample vs. standard norm multivariate analysis -differences

Class I			
Measurement	Mean	Std. deviation	Sig.
S-N	-3.71	3.48	.000*
Ba-N	-0.74	5.30	.452
SNA	-0.11	3.30	.858
SNB	0.93	2.92	.094
ANB	0.82	1.71	.014*
N-ANS : N-Me	0.00	0.02	.848
Class II-1			
Measurement	Mean	Std. deviation	Sig.
S-N	-2.98	3.61	.000*
Ba-N	1.10	5.67	.299
SNA	0.39	3.76	.575
SNB	-3.22	3.90	.000*
ANB	3.61	1.85	.000*
N-ANS : N-Me	0.00	0.03	.467

*Denotes statistical significance

Right vs. left t-tests found evidence of mean measurement value
differences for three measurements (*U1-APog*, *MP-SN*, *PP-MP*) in the Class I
malocclusion group and two measurements (*U1-APog*, *MP-SN*) in the Class II-1
malocclusion group, at an œ level of 0.05. Descriptive statistics and significance
for these measurements are in *Table 4.5*.

Table 4.5: Descriptive statistics and significance for those measurements used in right vs. left multivariate analysis -differences

Class I			
Measurement	Mean	Std. deviation	Sig.
U1-APog	0.86	1.32	.001*
CoSup-A	0.18	1.66	.560
CoSup-Pog	0.12	1.77	.707
U1-SN	0.54	3.69	.432
U1-PP	0.54	3.69	.431
U1-L1	0.20	4.43	.807
MP-SN	3.46	3.06	.000*
L1-MP	-0.97	3.25	.114
PP-MP	1.29	2.58	.011*
Class II-1			
Measurement	Mean	Std. deviation	Sig.
U1-APog	0.464	1.10	.028*
CoSup-A	-0.629	2.24	.134
CoSup-Pog	-0.657	2.16	.107
U1-SN	1.207	4.22	.128
U1-PP	1.206	4.21	.128
U1-L1	-1.536	4.72	.085
MP-SN	2.076	3.50	.003*
L1-MP	-0.539	3.16	.358
PP-MP	-0.767	3.16	.194

*Denotes statistical significance

Descriptive statistics, including mean and standard deviation, was computed for each of the eight predictor variables used in DA. As there were only two groups, a single canonical discriminant function was created. It had a canonical correlation of 0.474 that is suggestive of the discriminant function being moderately related to group malocclusion classification. Eigenvalue and Wilks' Lambda values presented contrasting findings. The prior demonstrated that the discriminant function was strongly related to group malocclusion classification while the latter alluded that group malocclusion classification is not related to (cannot be predicted from) measurement values on the discriminant function (Wilks' Lambda = 0.775, χ^2 (10) = 13.734, p > 0.05). Nonetheless, overall prediction of group malocclusion classification was very good. Twenty-three percent ($\eta2 = 1 - \lambda = 1 - 0.775$) of the variance in measurement scores was due to between-group difference. The differences between Class I and Class II-1 groups on the combined dependent variables was not statistically significant.

The standardized canonical discriminant function coefficients and functions at group centroids are found in *Tables 4.6 and 4.7*, respectively. Discriminant scores ranged from – 0.529 to 0.529. The optimal cut-off value to classify future observations into one of the two groups was 0.000, which is the weighted mean of the two centroids. Negative discriminant scores were more likely to be associated with Class I malocclusion classifications, whereas positive ones with Class II-1 malocclusion classifications. Misclassification of original grouped cases was more prevalent in Class I malocclusions. Although few,

misclassification of original grouped Class II-1 cases arose near the cut-off value.

Extreme cases were seen in both malocclusion classifications. Graphical

presentation of discriminant scores from the function can be visualized in *Figure*

4.1.

Table 4.6: Standardized canonical discriminant function coefficients

M1	Function 1
M2	-1.027
M3	0.446
M3	-0.557
M4	0.708
M5	0.078
M6	0.322
M7	0.163
M8	0.043

Table 4.7: Functions at group centroids

Classification	Function 1
Class I	-0.529
Class II-1	0.529
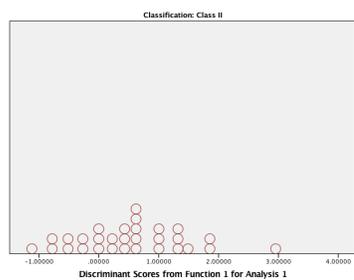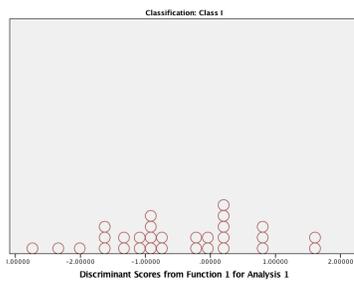
Figure 4.1: Graphical presentation of discriminant scores from the function

Since the discriminant function was found not statistically significant, further interpretation of its coefficients is not always necessary. DA with an overall non-significant result typically suggests that none of the predictor variables is related to group membership, though not always. As such, the decision was made by the authors to proceed with follow-up analyses as situations do arise where one or more predictors variable is in fact be significantly related to membership of a group when others are not. With that said, Analysis of variance (ANOVA) was used to explore the differences among group means of each predictor variable.

Five of the eight predictor variables demonstrated significance using the univariate ANOVA model, and together comprise the predictor variables with the highest structure coefficients from the set. These are *M1*, *M2*, *M6*, *M4* and *M5*. Though, it should be noted that these structure coefficients are high only relatively speaking as in actuality they demonstrate moderate to high correlations ($0.532 > d_1 > 0.737$). Positive structure coefficients relate with larger and negative structure coefficients relate with smaller measurement values. Subjects with larger Infraorbital (right) – Mental foramen (right) distance; M1, larger Infraorbital (left) – Mental foramen (left) distance; M2, smaller 2.6 root apex – Infraorbital (left) – 2.6 buccal angle; M6, smaller Infraorbital (right) – Infraorbital (left) – 2.6 buccal angle; M4 and smaller 1.6 root apex – Infraorbital (right) – 1.6 buccal angle; M5 will be classified as a Class I malocclusion patient. Subjects with smaller Infraorbital (right) – Mental foramen (right) distance; M1, smaller Infraorbital (left) – Mental foramen (left) distance; M2, larger 2.6 root apex –

Infraorbital (left) – 2.6 buccal angle; M6, larger Infraorbital (right) – Infraorbital

(left) – 2.6 buccal angle; M4 and larger 1.6 root apex – Infraorbital (right) – 1.6

buccal angle; M5 will be classified as a Class II-1 malocclusion patient. A

summary of structure coefficients and tests of equality of group means is found in

Table 4.8. Mean measurements for the five predictor variables, for each Class I

and Class II-1 malocclusion groups, demonstrating significance using the

univariate ANOVA model is found in *Table 4.9*.

Table 4.8: Structure coefficients and tests of equality of means

Variable	d_1	Wilks' Lambda (p-value)
M1	-0.737	0.864 (0.004)
M2	-0.723	0.869 (0.004)
M6	0.583	0.910 (0.020)
M4	0.555	0.918 (0.027)
M5	0.532	0.924 (0.033)
M7	0.259	0.981 (0.293)
M3	0.185	0.990 (0.451)
M8	0.102	0.997 (0.667)

Table 4.9: Mean measurements for Class I and Class II-1, for those demonstrating significance using the univariate ANOVA model

Measurement	Class I (mean)	Class II (mean)
M1	63.41 mm	59.63 mm
M2	63.08 mm	59.14 mm
M6	11.08˚	13.73˚
M4	93.89˚	96.33˚
M5	11.59˚	13.40˚

Construction of formulas using Fisher's linear discriminant functions may

be of benefit for classifying future cases *(Table 4.10)*. When evaluating

classification functions, cross-validation showed that 66.7% of Class I and

73.3.7% of Class II-1 grouped cases were correctly classified; thus, the

discriminant function did a better job of classifying Class II-1 group subjects.

Cross-validated classification results are summarized in *Table 4.11*.

Table 4.10: Fisher's linear discriminant functions

Classification	Class I	Class II-1
M1	8.113	7.889
M2	-1.020	-0.928
M3	5.201	5.054
M4	3.348	3.529
M5	2.334	2.353
M6	1.325	1.404
M7	3.210	3.251
M8	2.945	2.955
(Constant)	-917.227	-917.99

Table 4.11: Classification results

	Count	Classification	Predicted Group Membership		Total
			Class I	Class II-1	
Cross-validated	Count	Class I	20.0	10	30
		Class II-1	8	22	30
	%	Class I	66.7	33.3	100.0
		Class II-1	26.7	73.3	100.0

Discussion

This research project began with the intention of developing a novel 3D cephalometric analysis for classifying orthodontic patients into Class I and Class II-1 malocclusions with more accuracy and precision than possible before. The main limitation identified is the initial categorization of patient CBCTs from the database into classification groups, as the analyses placed a huge emphasis on this data being correct (gold standard). As hundreds of these scans were considered, only a single examiner reviewed the bulk of them to make classification choices. There may exist some individual bias in this process, although an attempt was made to mitigate the magnitude of this by using

checklists with clear definitions (*Appendix 4.1*) and a second review by two experienced orthodontists for those cases that fell somewhere in between. Also, there are likely many more discriminator measurements in 3D that are important in classifying these patients into appropriate groups since only 8 measurements were evaluated in this study. However, we believe that we met our objective of creating a novel analysis of this sort and look forward to further research in this area to build upon these findings.

There are associated challenges with the classification of Class II malocclusions. Simply grouped, they can be divided into Division 1 characterized by increased overjet and Division 2 characterized by retroclined upper incisors masking the severity or lack thereof of the antero-posterior jaw discrepancy. Further subdivision of Class II Division 2 (Class II-2) by Van der Linden into Types A/B/C is based the severity of the incisor relationship.[6]

One should be reminded that though this simplicity makes classification of this malocclusion group relatively easy, most cases do not fit these clear-cut definitions. For instance, this is complicated by Lavergne and Petrovic's finding that when children are grouped based on morphogenetic and morphophysiologic features, similarities in tissue level growth potential and response to functional appliances are shared between members.[7] Thus, if using a servosystem theory to describe craniofacial growth, a multitude of malocclusion possibilities emerge. McNamara's study of Class II malocclusions also depicted many combinations, based on skeletal and dental measurements.[8] Even yet, Moyer has found that

Class II malocclusions can be described as a range of horizontal and vertical

characteristics, with subtypes exhibiting distinguishing features.[9]

<u>Investigation of the difference between 2D normative values and 3D skeletal and</u>

<u>dental measurements</u>

It could be very interesting to use the data from this research in extension

studies, to investigate the difference between 2D normative values and 3D

skeletal and dental measurements obtained with both methods of imaging for all

the measurements included in this study in its entirety. This was in fact objective

#1 of this study and achieved to some degree, though it could only be carried out

directly for 6 of 24 measurements. Each of these 2 distances (*S-N*, *Ba-N*), 3

angles (*SNA, SNB, ANB*) or ratio (*N-ANS : N-Me*) measurements required only

two or three coordinates, or landmarks, in construction of a linear line or angle.

The simplicity in their measurements allowed for the direct comparison with 2D

normative values.

The reason this assessment was not completed for the remaining 18

measurements are due to the way in which values were obtained from the 3D

information –it did not allow for the direct comparison with 2D normative values.

For these measurements, such as *SN-MP (right)*, 3D coordinates of specific

landmarks were used to obtain the angle taken from the point of intersection

between constructed *S-N* line and *Go-Me* right lines. These same landmark

coordinates could then be projected onto a constructed midsagittal plane, then

used to create two new planes to determine the same type of angle seen on a

traditional lateral cephalogram. In this way, a direct comparison between "new" 3D and standard 2D norms could be compared during cephalometric analyses.

Another consideration is that there is a continuous increase in length of the anterior cranial base (*S-N*) until adulthood, due to the change of position of both *Sella* and *nasion* that seems to accompany size of changes of the frontal sinus. As the sample of patients in this study were primarily adolescents (Class I average age: 15.5 years, Class II-1 average age: 13.9), there was an attempt to account for the *S-N* 2D standard norm used for comparison by authors; 68mm was used instead of an average of 75mm for adults based on the University of Alberta Analysis. Thus, variables such as age and gender of subjects likely have cofounding effects on the results. A worthy note is that *S-N* measurement in 3D was found to be smaller in both Class I and Class II-1 malocclusion groups than expected in 2D. It would be interesting to see if this same finding could be substantiated in the future.

The other 2D standard norms used for comparison were obtained from the same source, except for cranial base length (*Ba-N)* that is from Scott's research of the cranial base. *SNA*, *SNB* and *ANB* values are based on Steiner's analysis. Upper face height to total face height, *N-ANS : N-Me*, ratio or percentage is based on Wylie and Johnson's analysis. Besides the source of standard norm value, it is important to point out that values are very much age and gender specific.

Investigation of the difference between measurements taken from right and left sides in the two malocclusion groups

Instead, a right vs. left multivariate analysis was carried out for the remaining 18 measurements, allowing for 9 direct comparisons. Authors of this study trust that this was an insightful addition to this body of research as some have suggested a perceptual and cognitive preference for aesthetics in a left to right direction, which is complicated by the fact that some degree of facial asymmetry exists within every individual. Thus, it is by no surprise that differences in craniofacial measurements made from right and left sides have also been affected in such a way. These measurement differences were seen for, but not limited to *U1-APog* and *MP-SN* for both malocclusion groups. These directionality differences were interestingly also seen for *PP-SN* only in the Class I malocclusion group.

Variations of measurements taken from right and left sides can be attributed to different reasons. Although the *A-Pog* plane runs mid-sagittal for instance, the *U1* position on right and left sides is dependent on not only antero-posterior and vertical (as in 2D cephalometric tracings) but transverse relationships also. Moreover, *U1* position is often taken from the most proclined and protruded upper incisor in 2D to overcome their superimposition in those images.

Efforts for the elimination of gross facial asymmetry in the sample population were made during the initial exclusion criteria (e.g. chin deviations >2mm). Even so, it is interesting that this study confirms that some facial asymmetry still exists within individuals. Since the *M-P plane* is constructed from landmarks situated in lower third of the face where asymmetry is most marked, it

is not surprisingly to find that two of three measurements exhibiting this right vs.

left difference included this craniofacial plane.

<u>Investigation of the existence of 3D skeletal and dental relationships to</u>

<u>adequately categorize orthodontic malocclusions</u>

When making the decision to choose using DA or one-way MANOVA to

analyze the relationship between the eight measurements (continuous variables)

in *Table 4.3* and malocclusion classification (categorical variable) into Class I or

Class II-1 (two levels), the benefits of each analysis was compared and weighed.

Both could determine which of these measurements could discriminate 'best'

between Class I and Class II-1 malocclusion classifications, although DA had two

distinct advantages. Firstly, it could determine how many dimensions we need to

express this relationship and degree to which these predetermined

measurements can be used to discriminate between these groups. Secondly, it

would allow the development of a computed index to use in prediction of a

malocclusion classification based on predetermined measurements, for future

observations.

The discriminate function in this study was found not statistically

significant, however, it is possible that one or more of these eight measurements

to be still significantly connected to group membership. An objective of this study

was to evaluate "new" 3D measurements that have rarely, if at all, been tested

before to create a novel 3D analysis for classification of patient malocclusions.

Since this entailed the assessment of "new" 3D measurements that were in one

way or another subjectively chosen, it is likely that not all the 8 predictor

measurements are critical variables for this task. The addition of any trivial measurement choices in the model analysis may have led to an error degrees of freedom reduction, and corresponding reduction in statistical power.

When looking at individual mean values for these 3D measurements, linear distance *infraorbital – menton* was found to be greater in the Class I than Class II-1 malocclusion group bilaterally. This was a particularly interesting finding as it could highlight the importance of vertical consideration in Class I malocclusion patients.

Conclusions

Findings of this research is allusive to the existence of measurement value differences between an estimated 3D population of Class I malocclusion patients and standard 2D norms generally used in orthodontic cephalometric interpretation for this group. They were found to include linear and angular measurements: *S-N* and *ANB*. Interestingly, 3D *S-N* distances were smaller than 2D. Although measurements differences do exist, they may be within the confines of acceptability from the clinical standpoint. Some variation is typically expected and established standard deviations remain. In the 3D context, there also appears to be differences in measurement values taken between right and left sides of an individual's craniofacial skeleton. This does not necessarily hold true for all measurements and classification groups, but has been found to be influential on *U1-APog*, *MP-SN and PP-MP.*

Discriminant analysis produced a non-significant overall result, although the authors chose to continue investigate the relationship between predictor

variables for group membership into Class I and Class II-1 in more detail. Even

though the discriminant function was found not statistically significant, it still has

the ability like one-way MANOVA to determine which measurements discriminate

'best' between Class I and Class II-1 malocclusion classifications. It is likely that

if different or different combination of predictor variables were used, different

results will be seen.

The follow-up analysis DA using the univariate ANOVA model was

suggestive of just that –some linear measurements can in fact discriminate 'best'

between Class I and Class II-1 malocclusion classification groups. Demonstrating

moderate to high correlations, ranging from 0.532 to 0.737, *Infraorbital (right) –*

Mental foramen (right) distance, *Infraorbital (left) – Mental foramen (left)*

distance, *2.6 root apex – Infraorbital (left) – 2.6 bucca*l angle, *Infraorbital (right) –*

Infraorbital (left) – 2.6 buccal angle, and *1.6 root apex – Infraorbital (right) – 1.6*

buccal angle are stronger discriminator variables relative to others tested. The

relationship of the 8 distance and angle measurements can be expressed in a

single dimension and exhibited a correct classification of 70.0% of cross-

validated group cases overall, for both malocclusion groups together.

Further research into 3D measurement values, traditional and those newly

defined, may enhance the development of this novel 3D cephalometric analysis

to classify future orthodontic patients using the appropriate, and sometimes even

"new", normal measurement values. Moreover, the reveal of such discriminator

3D measurements and classification functions for classifying future cases may

enable clinicians to adequately classify patients in a non-traditional way. Specific

skeletal and dental patterns only depicted in 3D volumetric CBCT scans may one

day hold the key to adequately classify Class I, Class II-1, Class II-2, Class III

and even subdivision cases like never.

References

1. Forsyth, D.B., S. Shaw Wc Fau - Richmond, and S. Richmond, *Digital imaging of cephalometric radiography, Part 1: Advantages and limitations of digital imaging.* (0003-3219 (Print)).
2. Silva, M.A., et al., *Cone-beam computed tomography for routine orthodontic treatment planning: a radiation dose evaluation.* (1097-6752 (Electronic)).
3. Proffit, W.R., H.W. Fields, and D.M. Sarver, *Contemporary Orthodontics-E-Book.* 2014: Elsevier Health Sciences.
4. Lagravere, M.O., et al., *Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments.* (0003-3219 (Print)).
5. Lagravere, M.O., et al., *Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images.* (1097-6752 (Electronic)).
6. Van der Linden, F.P.G.M., *Development of the dentition.* Vol. 1. 1983: Quintessence Pub Co.
7. Lavergne, J. and A. Petrovic, *Discontinuities in occlusal relationship and the regulation of facial growth. A cybernetic view.* The European Journal of Orthodontics, 1983. **5**(4): p. 269-278.
8. McNamara Jr, J.A., *Components of Class II malocclusion in children 8–10 years of age.* The Angle Orthodontist, 1981. **51**(3): p. 177-202.
9. Moyers, R., *Auxologic categorization and chronobiologic specification for the choice of appropriate orthodontic treatment.* American Journal of Orthodontics and Dentofacial Orthopedics, 1994. **105**(2): p. 192-205.

Appendices

Appendix 4.1: Checklist for initial patient reviews

Checklist for initial patient reviews

Patient code:

Year: _____ Appointment date:
_____

Exclusion criteria (check all that apply):

☐ Craniofacial deformities (e.g. cleft lip and/or palate)
☐ Asymmetries (e.g. chin deviations >2mm)
☐ One or more permanent first molar(s) absent

SKELETAL

Maxilla	Orthognathic	Prognathic	Retrognathic
Mandible	Orthognathic	Prognathic	Retrognathic
Growth pattern	Normal	Vertical	Horizontal
Interincisal angle	Normal	Decreased	Increased

Skeletal interpretation:

☐ Class I   ☐ Class II-1   ☐ Class II-2   ☐ Class III

DENTAL

Molar relationship (R)	Class I	Class II	Class III
Molar relationship (L)	Class I	Class II	Class III
Overjet	Within acceptable range (1-3mm)	Decreased (<1mm)	Increased (>3mm)
Overbite	Within acceptable range (20-40%)	Decreased (<20%)	Increased (>40%)

Dental impression:

☐ Class I   ☐ Class II-1   ☐ Class II-2   ☐ Class III

Subdivision malocclusion?
☐ No   ☐ Yes

Page 1 of 2

Type of Class II-2 malocclusion, if applicable:

☐ Type A *(All four Mx incisors tipped palatally, Mx canines well-aligned)*

☐ Type B *(Mx central incisors tipped palatally, Mx lateral incisors tipped labially)*

☐ Type C *(All four Mx incisors tipped palatally, Mx canines buccally positioned outside arch)*

SOFT TISSUE DRAPE

Growth pattern	Mesocephalic	Dolichocephalic	Brachycephalic
Facial type	Orthognathic	Prognathic	Retrognathic
Profile type	Straight to Mild convex	Convex	Concave

Soft tissue impression:

☐ Class I   ☐ Class II   ☐ Class III

_____

OVERALL CLASSIFICATION (check all that apply):

☐   **Class I**

☐   **Class II-1**

☐   **Class II-2**   Type A   Type B   Type C

☐   **Class III**

☐   **Subdivision malocclusion**

☐   **Fulfills exclusion criteria**

Page 2 of 2

Appendix 4.2: Sample vs. standard norm multivariate analysis verification: boxplots
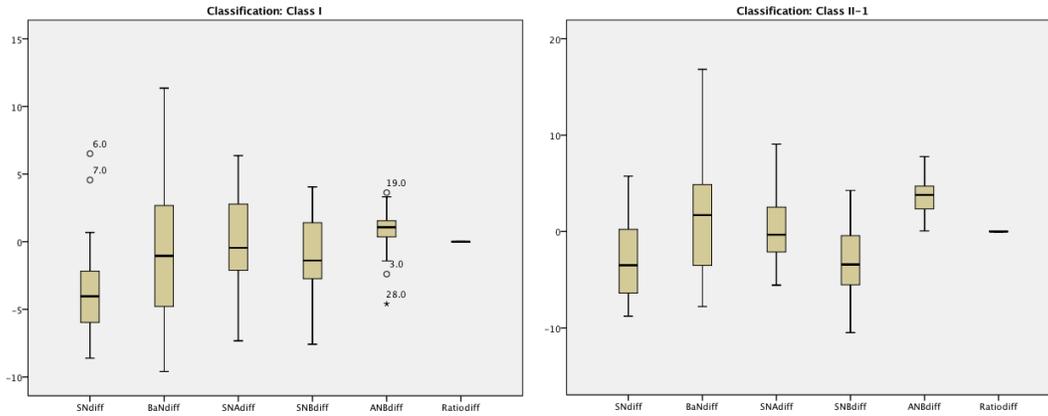
Appendix 4.3: Sample vs. standard norm multivariate analysis verification: Mahalanobis D^2 vs. X2^2 Q-Q plot
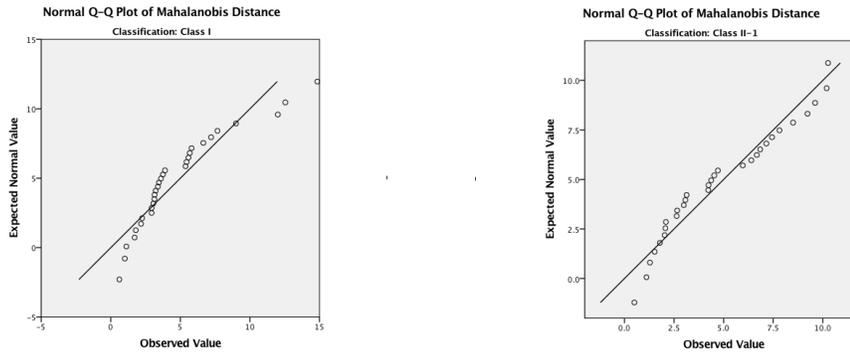
Appendix 4.4: Sample vs. standard norm multivariate analysis verification: correlations

Correlations[a]

		SNdiff	BaNdiff	SNAdiff	SNBdiff	ANBdiff	Ratiodiff
SNdiff	Pearson Correlation	1	.837[**]	−.292	−.239	−.155	.129
	Sig. (2−tailed)		.000	.118	.203	.415	.496
	N	30	30	30	30	30	30
BaNdiff	Pearson Correlation	.837[**]	1	−.193	−.110	−.185	.260
	Sig. (2−tailed)	.000		.306	.563	.327	.165
	N	30	30	30	30	30	30
SNAdiff	Pearson Correlation	−.292	−.193	1	.856[**]	.467[**]	−.131
	Sig. (2−tailed)	.118	.306		.000	.009	.490
	N	30	30	30	30	30	30
SNBdiff	Pearson Correlation	−.239	−.110	.856[**]	1	−.057	−.055
	Sig. (2−tailed)	.203	.563	.000		.766	.773
	N	30	30	30	30	30	30
ANBdiff	Pearson Correlation	−.155	−.185	.467[**]	−.057	1	−.159
	Sig. (2−tailed)	.415	.327	.009	.766		.400
	N	30	30	30	30	30	30
Ratiodiff	Pearson Correlation	.129	.260	−.131	−.055	−.159	1
	Sig. (2−tailed)	.496	.165	.490	.773	.400	
	N	30	30	30	30	30	30

[**]. Correlation is significant at the 0.01 level (2−tailed).

a. Classification = Class I

Correlations[a]

		SNdiff	BaNdiff	SNAdiff	SNBdiff	ANBdiff	Ratiodiff
SNdiff	Pearson Correlation	1	.851**	-.414*	-.519**	.254	.154
	Sig. (2-tailed)		.000	.023	.003	.176	.417
	N	30	30	30	30	30	30
BaNdiff	Pearson Correlation	.851**	1	-.294	-.335	.109	.372*
	Sig. (2-tailed)	.000		.115	.071	.566	.043
	N	30	30	30	30	30	30
SNAdiff	Pearson Correlation	-.414*	-.294	1	.884**	.169	.008
	Sig. (2-tailed)	.023	.115		.000	.373	.966
	N	30	30	30	30	30	30
SNBdiff	Pearson Correlation	-.519**	-.335	.884**	1	-.311	.106
	Sig. (2-tailed)	.003	.071	.000		.094	.577
	N	30	30	30	30	30	30
ANBdiff	Pearson Correlation	.254	.109	.169	-.311	1	-.208
	Sig. (2-tailed)	.176	.566	.373	.094		.271
	N	30	30	30	30	30	30
Ratiodiff	Pearson Correlation	.154	.372*	.008	.106	-.208	1
	Sig. (2-tailed)	.417	.043	.966	.577	.271	
	N	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Classification = Class II-1

Appendix 4.5: Sample vs. standard norm multivariate analysis verification: matrix
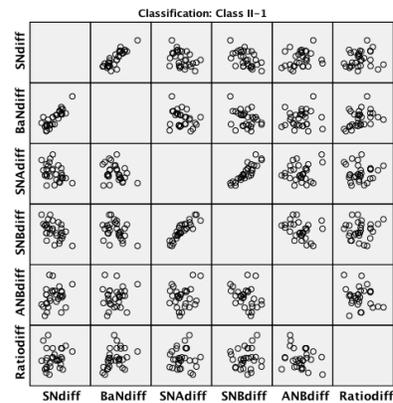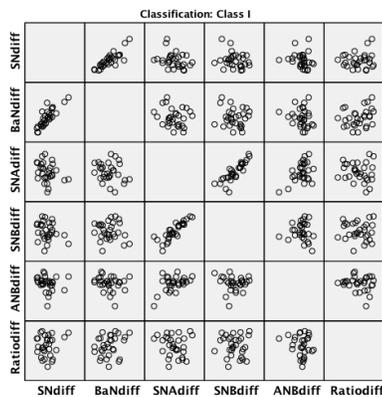
Appendix 4.6: Descriptive statistics including mean, standard deviation and number of cases for each measurement in sample vs. standard norm multivariate analysis

Descriptive Statistics[a]

	Mean	Std. Deviation	N
SNdiff	-3.7084	3.47983	30
BaNdiff	-.7366	5.29537	30
SNAdiff	-.1087	3.30171	30
SNBdiff	-.9246	2.92428	30
ANBdiff	.8159	1.70784	30
Ratiodiff	-.0007	.01923	30

a. Classification = Class I

Descriptive Statistics[a]

	Mean	Std. Deviation	N
SNdiff	-2.9819	3.61415	30
BaNdiff	1.0951	5.66838	30
SNAdiff	.3891	3.75558	30
SNBdiff	-3.2240	3.89543	30
ANBdiff	3.6131	1.84558	30
Ratiodiff	.0038	.02805	30

a. Classification = Class II-1

Appendix 4.7: Multivariate tests for sample vs. standard norm multivariate analysis

Multivariate Tests[a,b]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.805	20.604[c]	5.000	25.000	.000	.805
	Wilks' Lambda	.195	20.604[c]	5.000	25.000	.000	.805
	Hotelling's Trace	4.121	20.604[c]	5.000	25.000	.000	.805
	Roy's Largest Root	4.121	20.604[c]	5.000	25.000	.000	.805

a. Classification = Class I

b. Design: Intercept

c. Exact statistic

Multivariate Tests[a,b]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.930	66.725[c]	5.000	25.000	.000	.930
	Wilks' Lambda	.070	66.725[c]	5.000	25.000	.000	.930
	Hotelling's Trace	13.345	66.725[c]	5.000	25.000	.000	.930
	Roy's Largest Root	13.345	66.725[c]	5.000	25.000	.000	.930

a. Classification = Class II-1

b. Design: Intercept

c. Exact statistic

Appendix 4.8: Univariate tests for sample vs. standard norm multivariate analysis

Tests of Between-Subjects Effects[a]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	SNdiff	.000[b]	0	.	.	.	.000
	BaNdiff	.000[b]	0	.	.	.	.000
	SNAdiff	.000[b]	0	.	.	.	.000
	SNBdiff	.000[b]	0	.	.	.	.000
	ANBdiff	.000[c]	0	.	.	.	.000
	Ratiodiff	.000[b]	0	.	.	.	.000
Intercept	SNdiff	412.568	1	412.568	34.071	.000	.540
	BaNdiff	16.276	1	16.276	.580	.452	.020
	SNAdiff	.355	1	.355	.033	.858	.001
	SNBdiff	25.647	1	25.647	2.999	.094	.094
	ANBdiff	19.970	1	19.970	6.847	.014	.191
	Ratiodiff	1.390E-5	1	1.390E-5	.038	.848	.001
Error	SNdiff	351.167	29	12.109			
	BaNdiff	813.187	29	28.041			
	SNAdiff	316.137	29	10.901			
	SNBdiff	247.990	29	8.551			
	ANBdiff	84.585	29	2.917			
	Ratiodiff	.011	29	.000			
Total	SNdiff	763.735	30				
	BaNdiff	829.463	30				
	SNAdiff	316.492	30				
	SNBdiff	273.638	30				
	ANBdiff	104.555	30				
	Ratiodiff	.011	30				
Corrected Total	SNdiff	351.167	29				
	BaNdiff	813.187	29				
	SNAdiff	316.137	29				
	SNBdiff	247.990	29				
	ANBdiff	84.585	29				
	Ratiodiff	.011	29				

a. Classification = Class I

b. R Squared = .000 (Adjusted R Squared = .000)

c. R Squared = .000 (Adjusted R Squared = .000)

Tests of Between-Subjects Effects[a]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	SNdiff	.000[b]	0	.	.	.	.000
	BaNdiff	.000[c]	0	.	.	.	.000
	SNAdiff	.000[b]	0	.	.	.	.000
	SNBdiff	.000[d]	0	.	.	.	.000
	ANBdiff	.000[d]	0	.	.	.	.000
	Ratiodiff	.000[e]	0	.	.	.	.000
Intercept	SNdiff	266.753	1	266.753	20.422	.000	.413
	BaNdiff	35.975	1	35.975	1.120	.299	.037
	SNAdiff	4.542	1	4.542	.322	.575	.011
	SNBdiff	311.819	1	311.819	20.549	.000	.415
	ANBdiff	391.624	1	391.624	114.975	.000	.799
	Ratiodiff	.000	1	.000	.543	.467	.018
Error	SNdiff	378.801	29	13.062			
	BaNdiff	931.784	29	32.130			
	SNAdiff	409.027	29	14.104			
	SNBdiff	440.057	29	15.174			
	ANBdiff	98.779	29	3.406			
	Ratiodiff	.023	29	.001			
Total	SNdiff	645.554	30				
	BaNdiff	967.759	30				
	SNAdiff	413.568	30				
	SNBdiff	751.875	30				
	ANBdiff	490.403	30				
	Ratiodiff	.023	30				
Corrected Total	SNdiff	378.801	29				
	BaNdiff	931.784	29				
	SNAdiff	409.027	29				
	SNBdiff	440.057	29				
	ANBdiff	98.779	29				
	Ratiodiff	.023	29				

a. Classification = Class II–1
b. R Squared = .000 (Adjusted R Squared = .000)
c. R Squared = .000 (Adjusted R Squared = .000)
d. R Squared = .000 (Adjusted R Squared = .000)
e. R Squared = .000 (Adjusted R Squared = .000)

Appendix 4.9: Parameter estimates for sample vs. standard norm multivariate

Parameter Estimates[a]

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
SNdiff	Intercept	−3.708	.635	−5.837	.000	−5.008	−2.409	.540
BaNdiff	Intercept	−.737	.967	−.762	.452	−2.714	1.241	.020
SNAdiff	Intercept	−.109	.603	−.180	.858	−1.342	1.124	.001
SNBdiff	Intercept	−.925	.534	−1.732	.094	−2.017	.167	.094
ANBdiff	Intercept	.816	.312	2.617	.014	.178	1.454	.191
Ratiodiff	Intercept	−.001	.004	−.194	.848	−.008	.007	.001

a. Classification = Class I

Parameter Estimates[a]

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
SNdiff	Intercept	−2.982	.660	−4.519	.000	−4.331	−1.632	.413
BaNdiff	Intercept	1.095	1.035	1.058	.299	−1.022	3.212	.037
SNAdiff	Intercept	.389	.686	.567	.575	−1.013	1.791	.011
SNBdiff	Intercept	−3.224	.711	−4.533	.000	−4.679	−1.769	.415
ANBdiff	Intercept	3.613	.337	10.723	.000	2.924	4.302	.799
Ratiodiff	Intercept	.004	.005	.737	.467	−.007	.014	.018

a. Classification = Class II–1

Appendix 4.10: Right vs. left multivariate analysis assumptions verification: boxplots
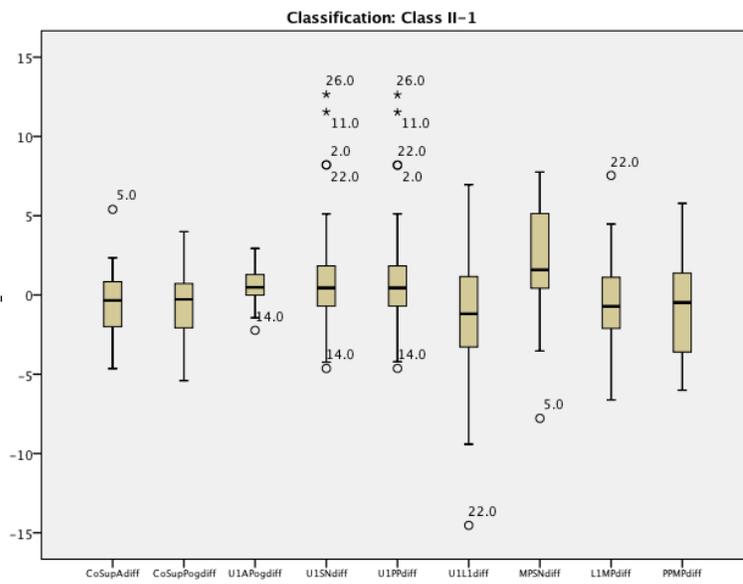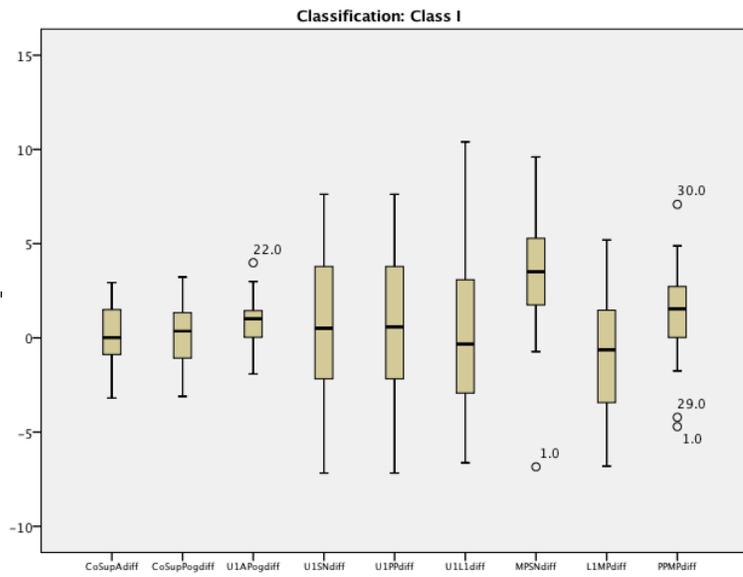
Appendix 4.11: Right vs. left multivariate analysis assumptions verification: Mahalanobis D^2 vs. X2^2 Q-Q plot

Normal Q–Q Plot of Mahalanobis Distance
Classification: Class I

Normal Q–Q Plot of Mahalanobis Distance
Classification: Class II–1
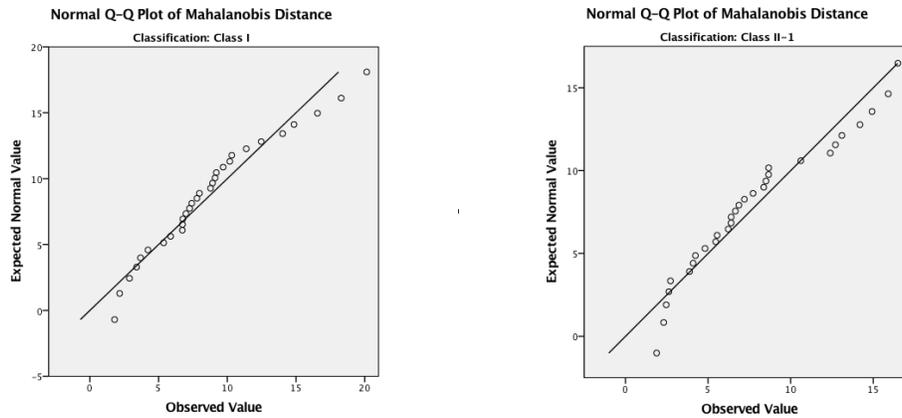
Appendix 4.12: Right vs. left multivariate analysis assumptions verification: correlations

Correlations[a]

		CoSupAdiff	CoSupPogdiff	U1Apogright diff	U1SNdiff	U1PPdiff	U1L1diff	MPSNdiff	L1MPdiff	PPMPdiff
CoSupAdiff	Pearson Correlation	1	.639[**]	.097	.009	.009	.112	-.208	-.027	-.068
	Sig. (2-tailed)		.000	.612	.961	.961	.557	.270	.886	.723
	N	30	30	30	30	30	30	30	30	30
CoSupPogdiff	Pearson Correlation	.639[**]	1	.406[*]	.059	.057	.036	.220	-.239	.235
	Sig. (2-tailed)	.000		.026	.758	.763	.849	.242	.203	.212
	N	30	30	30	30	30	30	30	30	30
U1Apogrightdiff	Pearson Correlation	.097	.406[*]	1	.519[**]	.518[**]	-.461[*]	.392[*]	.027	.295
	Sig. (2-tailed)	.612	.026		.003	.003	.010	.032	.888	.114
	N	30	30	30	30	30	30	30	30	30
U1SNdiff	Pearson Correlation	.009	.059	.519[**]	1	1.000[**]	-.656[**]	.115	-.183	.005
	Sig. (2-tailed)	.961	.758	.003		.000	.000	.546	.332	.979
	N	30	30	30	30	30	30	30	30	30
U1PPdiff	Pearson Correlation	.009	.057	.518[**]	1.000[**]	1	-.657[**]	.115	-.183	.005
	Sig. (2-tailed)	.961	.763	.003	.000		.000	.546	.333	.979
	N	30	30	30	30	30	30	30	30	30
U1L1diff	Pearson Correlation	.112	.036	-.461[*]	-.656[**]	-.657[**]	1	-.163	-.473[**]	-.091
	Sig. (2-tailed)	.557	.849	.010	.000	.000		.389	.008	.632
	N	30	30	30	30	30	30	30	30	30
MPSNdiff	Pearson Correlation	-.208	.220	.392[*]	.115	.115	-.163	1	-.327	.695[**]
	Sig. (2-tailed)	.270	.242	.032	.546	.546	.389		.078	.000
	N	30	30	30	30	30	30	30	30	30
L1MPdiff	Pearson Correlation	-.027	-.239	.027	-.183	-.183	-.473[**]	-.327	1	-.241
	Sig. (2-tailed)	.886	.203	.888	.332	.333	.008	.078		.200
	N	30	30	30	30	30	30	30	30	30
PPMPdiff	Pearson Correlation	-.068	.235	.295	.005	.005	-.091	.695[**]	-.241	1
	Sig. (2-tailed)	.723	.212	.114	.979	.979	.632	.000	.200	
	N	30	30	30	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Classification = Class I

		CoSupAdiff	CoSupPogdiff	U1Apogright diff	U1SNdiff	U1PPdiff	U1L1diff	MPSNdiff	L1MPdiff	PPMPdiff
CoSupAdiff	Pearson Correlation	1	.749**	.015	.055	.055	-.201	-.234	.320	.072
	Sig. (2-tailed)		.000	.938	.774	.772	.287	.214	.085	.705
	N	30	30	30	30	30	30	30	30	30
CoSupPogdiff	Pearson Correlation	.749**	1	.056	-.028	-.028	-.182	.135	.102	.320
	Sig. (2-tailed)	.000		.768	.881	.882	.336	.478	.592	.084
	N	30	30	30	30	30	30	30	30	30
U1Apogrightdiff	Pearson Correlation	.015	.056	1	.699**	.699**	-.440*	.347	-.371*	.193
	Sig. (2-tailed)	.938	.768		.000	.000	.015	.060	.043	.307
	N	30	30	30	30	30	30	30	30	30
U1SNdiff	Pearson Correlation	.055	-.028	.699**	1	1.000**	-.813**	-.029	.038	.005
	Sig. (2-tailed)	.774	.881	.000		.000	.000	.879	.841	.978
	N	30	30	30	30	30	30	30	30	30
U1PPdiff	Pearson Correlation	.055	-.028	.699**	1.000**	1	-.813**	-.029	.038	.005
	Sig. (2-tailed)	.772	.882	.000	.000		.000	.879	.842	.978
	N	30	30	30	30	30	30	30	30	30
U1L1diff	Pearson Correlation	-.201	-.182	-.440*	-.813**	-.813**	1	.031	-.410*	-.157
	Sig. (2-tailed)	.287	.336	.015	.000	.000		.870	.024	.406
	N	30	30	30	30	30	30	30	30	30
MPSNdiff	Pearson Correlation	-.234	.135	.347	-.029	-.029	.031	1	-.707**	.472**
	Sig. (2-tailed)	.214	.478	.060	.879	.879	.870		.000	.008
	N	30	30	30	30	30	30	30	30	30
L1MPdiff	Pearson Correlation	.320	.102	-.371*	.038	.038	-.410*	-.707**	1	-.094
	Sig. (2-tailed)	.085	.592	.043	.841	.842	.024	.000		.622
	N	30	30	30	30	30	30	30	30	30
PPMPdiff	Pearson Correlation	.072	.320	.193	.005	.005	-.157	.472**	-.094	1
	Sig. (2-tailed)	.705	.084	.307	.978	.978	.406	.008	.622	
	N	30	30	30	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. Classification = Class II-1

Appendix 4.13: Right vs. left multivariate assumptions verification: matrix scatterplot

Classification: Class I

Classification: Class II–1
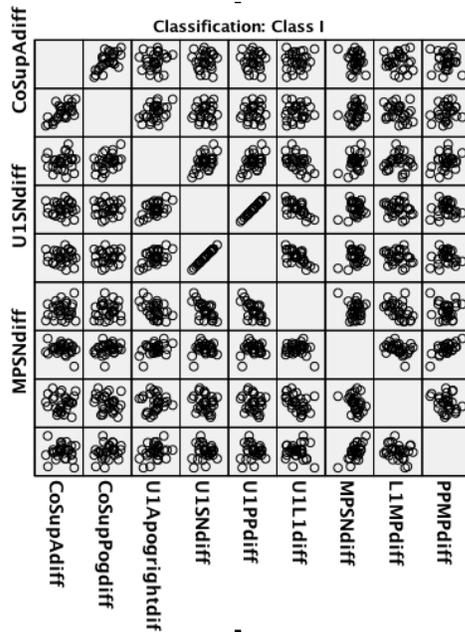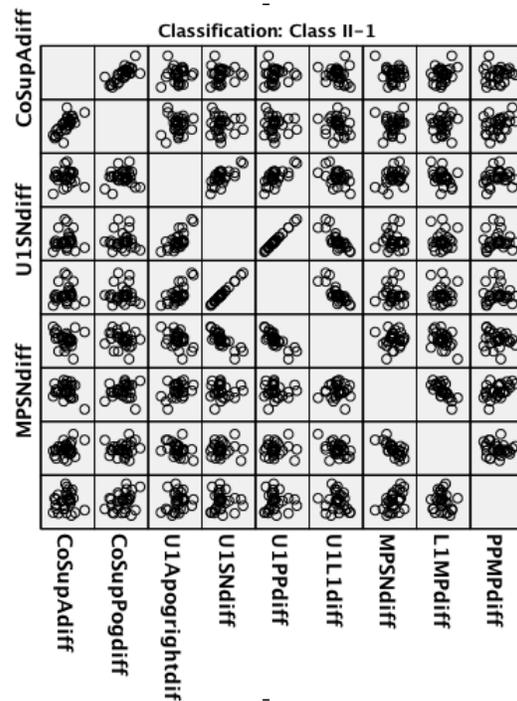
Appendix 4.14: Descriptive statistics including mean, standard deviation and number of cases for each measurement in right vs. left multivariate analysis

Descriptive Statistics[a]

	Mean	Std. Deviation	N
CoSupAdiff	.1793	1.66489	30
CoSupPogdiff	.1231	1.77442	30
U1Apogrightdiff	.8627	1.32254	30
U1SNdiff	.5377	3.69391	30
U1PPdiff	.5372	3.68506	30
U1L1diff	.1989	4.42510	30
MPSNdiff	3.4615	3.06400	30
L1MPdiff	-.9662	3.25075	30
PPMPdiff	1.2868	2.57962	30

a. Classification = Class I

Descriptive Statistics[a]

	Mean	Std. Deviation	N
CoSupAdiff	-.6288	2.23559	30
CoSupPogdiff	-.6568	2.16044	30
U1Apogrightdiff	.4635	1.09639	30
U1SNdiff	1.2074	4.22155	30
U1PPdiff	1.2064	4.21420	30
U1L1diff	-1.5360	4.71648	30
MPSNdiff	2.0763	3.49775	30
L1MPdiff	-.5385	3.15930	30
PPMPdiff	-.7669	3.16165	30

a. Classification = Class II–1

Appendix 4.15: Multivariate tests for right vs. left multivariate analysis

Multivariate Tests[a,b]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.716	5.890[c]	9.000	21.000	.000	.716
	Wilks' Lambda	.284	5.890[c]	9.000	21.000	.000	.716
	Hotelling's Trace	2.524	5.890[c]	9.000	21.000	.000	.716
	Roy's Largest Root	2.524	5.890[c]	9.000	21.000	.000	.716

a. Classification = Class I

b. Design: Intercept

c. Exact statistic

Multivariate Tests[a,b]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.620	3.800[c]	9.000	21.000	.006	.620
	Wilks' Lambda	.380	3.800[c]	9.000	21.000	.006	.620
	Hotelling's Trace	1.629	3.800[c]	9.000	21.000	.006	.620
	Roy's Largest Root	1.629	3.800[c]	9.000	21.000	.006	.620

a. Classification = Class II–1

b. Design: Intercept

c. Exact statistic

Appendix 4.16: Univariate tests for right vs. left multivariate analysis

Tests of Between-Subjects Effects[a]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	CoSupAdiff	.000[b]	0	.	.	.	.000
	CoSupPogdiff	.000[c]	0	.	.	.	.000
	U1Apogrightdiff	.000[b]	0	.	.	.	.000
	U1SNdiff	.000[b]	0	.	.	.	.000
	U1PPdiff	.000[b]	0	.	.	.	.000
	U1L1diff	.000[b]	0	.	.	.	.000
	MPSNdiff	.000[d]	0	.	.	.	.000
	L1MPdiff	.000[e]	0	.	.	.	.000
	PPMPdiff	.000[b]	0	.	.	.	.000
Intercept	CoSupAdiff	.964	1	.964	.348	.560	.012
	CoSupPogdiff	.454	1	.454	.144	.707	.005
	U1Apogrightdiff	22.329	1	22.329	12.766	.001	.306
	U1SNdiff	8.675	1	8.675	.636	.432	.021
	U1PPdiff	8.659	1	8.659	.638	.431	.022
	U1L1diff	1.187	1	1.187	.061	.807	.002
	MPSNdiff	359.450	1	359.450	38.288	.000	.569
	L1MPdiff	28.008	1	28.008	2.650	.114	.084
	PPMPdiff	49.675	1	49.675	7.465	.011	.205
Error	CoSupAdiff	80.383	29	2.772			
	CoSupPogdiff	91.308	29	3.149			
	U1Apogrightdiff	50.724	29	1.749			
	U1SNdiff	395.705	29	13.645			
	U1PPdiff	393.810	29	13.580			
	U1L1diff	567.863	29	19.581			
	MPSNdiff	272.254	29	9.388			
	L1MPdiff	306.455	29	10.567			
	PPMPdiff	192.979	29	6.654			
Total	CoSupAdiff	81.347	30				
	CoSupPogdiff	91.763	30				
	U1Apogrightdiff	73.053	30				
	U1SNdiff	404.379	30				
	U1PPdiff	402.469	30				
	U1L1diff	569.049	30				
	MPSNdiff	631.704	30				
	L1MPdiff	334.463	30				
	PPMPdiff	242.654	30				

Corrected Total	CoSupAdiff	80.383	29			
	CoSupPogdiff	91.308	29			
	U1Apogrightdiff	50.724	29			
	U1SNdiff	395.705	29			
	U1PPdiff	393.810	29			
	U1L1diff	567.863	29			
	MPSNdiff	272.254	29			
	L1MPdiff	306.455	29			
	PPMPdiff	192.979	29			

a. Classification = Class I

b. R Squared = .000 (Adjusted R Squared = .000)

c. R Squared = .000 (Adjusted R Squared = .000)

d. R Squared = .000 (Adjusted R Squared = .000)

e. R Squared = .000 (Adjusted R Squared = .000)

Tests of Between-Subjects Effects[a]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	CoSupAdiff	.000[b]	0	.	.	.	.000
	CoSupPogdiff	.000[c]	0	.	.	.	.000
	U1Apogrightdiff	.000[d]	0	.	.	.	.000
	U1SNdiff	.000[c]	0	.	.	.	.000
	U1PPdiff	.000[c]	0	.	.	.	.000
	U1L1diff	.000[c]	0	.	.	.	.000
	MPSNdiff	.000[c]	0	.	.	.	.000
	L1MPdiff	.000[c]	0	.	.	.	.000
	PPMPdiff	.000[c]	0	.	.	.	.000
Intercept	CoSupAdiff	11.861	1	11.861	2.373	.134	.076
	CoSupPogdiff	12.942	1	12.942	2.773	.107	.087
	U1Apogrightdiff	6.444	1	6.444	5.361	.028	.156
	U1SNdiff	43.737	1	43.737	2.454	.128	.078
	U1PPdiff	43.663	1	43.663	2.459	.128	.078
	U1L1diff	70.777	1	70.777	3.182	.085	.099
	MPSNdiff	129.329	1	129.329	10.571	.003	.267
	L1MPdiff	8.698	1	8.698	.871	.358	.029
	PPMPdiff	17.643	1	17.643	1.765	.194	.057
Error	CoSupAdiff	144.937	29	4.998			
	CoSupPogdiff	135.357	29	4.667			
	U1Apogrightdiff	34.860	29	1.202			
	U1SNdiff	516.823	29	17.821			
	U1PPdiff	515.024	29	17.759			
	U1L1diff	645.111	29	22.245			
	MPSNdiff	354.794	29	12.234			
	L1MPdiff	289.454	29	9.981			
	PPMPdiff	289.885	29	9.996			
Total	CoSupAdiff	156.799	30				
	CoSupPogdiff	148.299	30				
	U1Apogrightdiff	41.304	30				
	U1SNdiff	560.560	30				
	U1PPdiff	558.687	30				
	U1L1diff	715.888	30				
	MPSNdiff	484.122	30				
	L1MPdiff	298.152	30				
	PPMPdiff	307.528	30				

Corrected Total	CoSupAdiff	144.937	29				
	CoSupPogdiff	135.357	29				
	U1Apogrightdiff	34.860	29				
	U1SNdiff	516.823	29				
	U1PPdiff	515.024	29				
	U1L1diff	645.111	29				
	MPSNdiff	354.794	29				
	L1MPdiff	289.454	29				
	PPMPdiff	289.885	29				

a. Classification = Class II–1

b. R Squared = .000 (Adjusted R Squared = .000)

c. R Squared = .000 (Adjusted R Squared = .000)

d. R Squared = .000 (Adjusted R Squared = .000)

*Appendix 4.17: Parameter estimates for right vs. left multivariate analysis*

Parameter Estimates[a]

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
CoSupAdiff	Intercept	.179	.304	.590	.560	-.442	.801	.012
CoSupPogdiff	Intercept	.123	.324	.380	.707	-.539	.786	.005
U1Apogrightdiff	Intercept	.863	.241	3.573	.001	.369	1.357	.306
U1SNdiff	Intercept	.538	.674	.797	.432	-.842	1.917	.021
U1PPdiff	Intercept	.537	.673	.799	.431	-.839	1.913	.022
U1L1diff	Intercept	.199	.808	.246	.807	-1.453	1.851	.002
MPSNdiff	Intercept	3.461	.559	6.188	.000	2.317	4.606	.569
L1MPdiff	Intercept	-.966	.594	-1.628	.114	-2.180	.248	.084
PPMPdiff	Intercept	1.287	.471	2.732	.011	.324	2.250	.205

a. Classification = Class I

Parameter Estimates[a]

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
CoSupAdiff	Intercept	-.629	.408	-1.541	.134	-1.464	.206	.076
CoSupPogdiff	Intercept	-.657	.394	-1.665	.107	-1.464	.150	.087
U1Apogrightdiff	Intercept	.463	.200	2.315	.028	.054	.873	.156
U1SNdiff	Intercept	1.207	.771	1.567	.128	-.369	2.784	.078
U1PPdiff	Intercept	1.206	.769	1.568	.128	-.367	2.780	.078
U1L1diff	Intercept	-1.536	.861	-1.784	.085	-3.297	.225	.099
MPSNdiff	Intercept	2.076	.639	3.251	.003	.770	3.382	.267
L1MPdiff	Intercept	-.538	.577	-.934	.358	-1.718	.641	.029
PPMPdiff	Intercept	-.767	.577	-1.329	.194	-1.947	.414	.057

a. Classification = Class II–1

Appendix 4.18: Discriminative analysis assumptions verification: boxplots

Appendix 4.19: Discriminative analysis assumptions verification: Mahalanobis D2 vs. D22 Q-Q plot
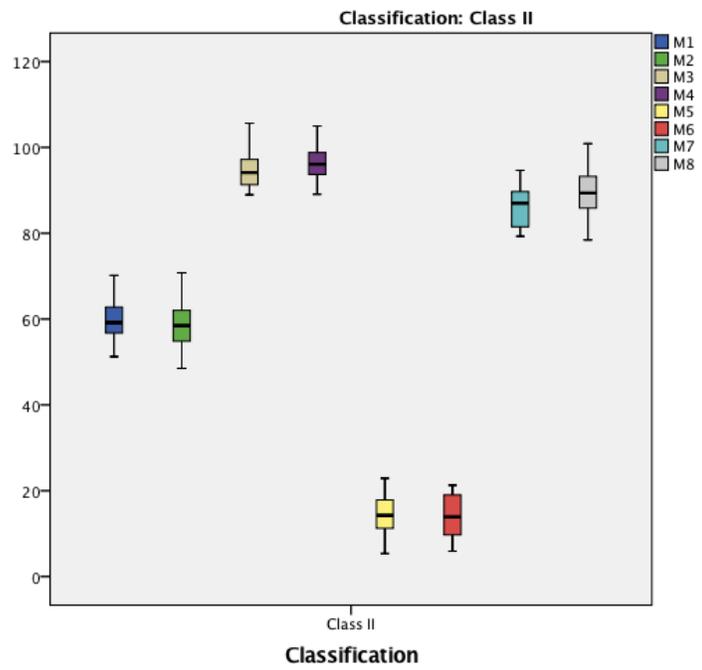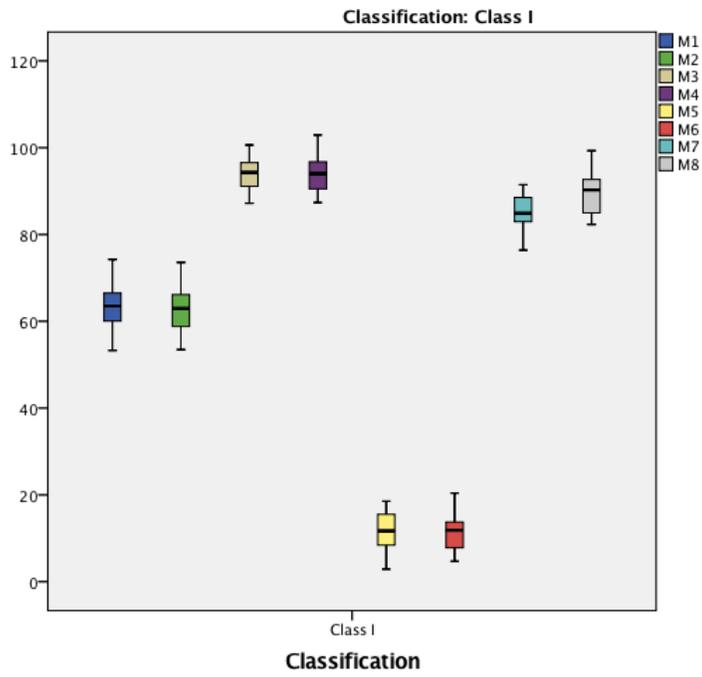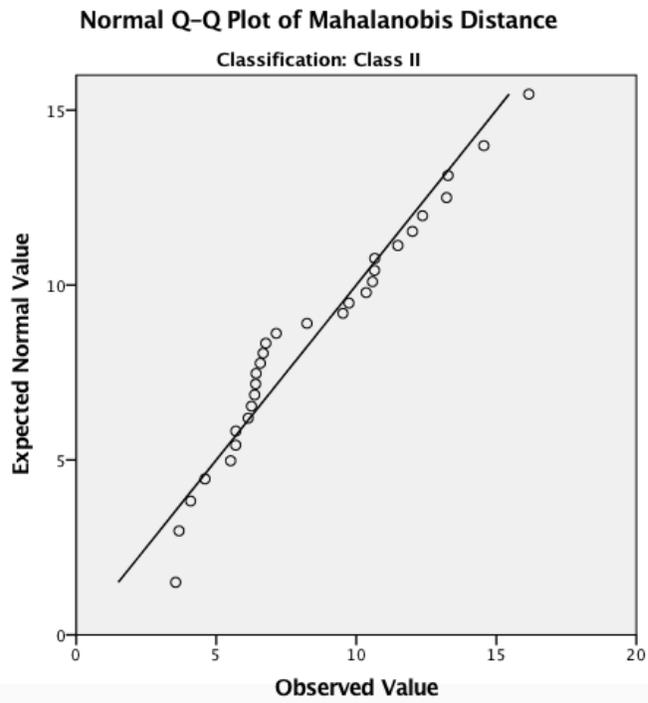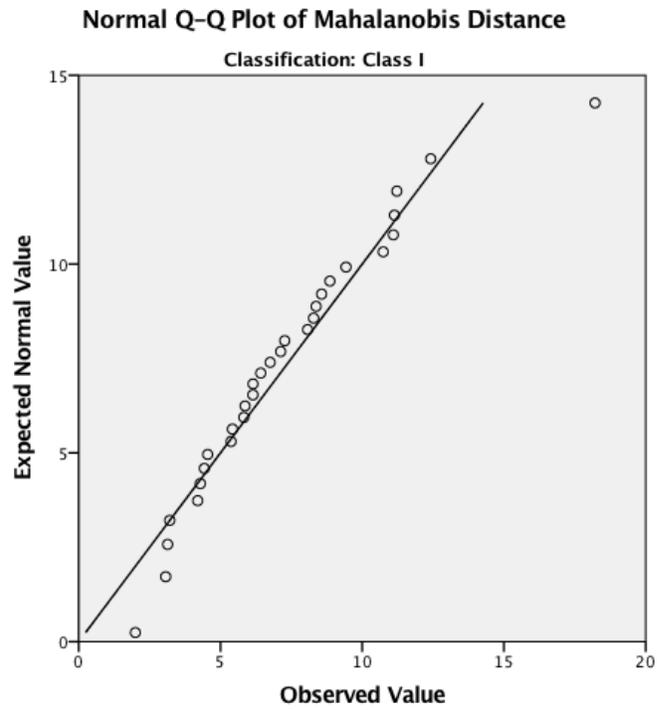
Appendix 4.20: Discriminative analysis assumptions verification: correlations

		M1	M2	M3	M4	M5	M6	M7	M8
M1	Pearson Correlation	1	.949**	-.372**	-.273*	-.652**	-.432**	-.243	-.019
	Sig. (2-tailed)		.000	.003	.035	.000	.001	.062	.886
	N	60	60	60	60	60	60	60	60
M2	Pearson Correlation	.949**	1	-.296*	-.292*	-.642**	-.468**	-.236	.006
	Sig. (2-tailed)	.000		.022	.024	.000	.000	.069	.964
	N	60	60	60	60	60	60	60	60
M3	Pearson Correlation	-.372**	-.296*	1	.571**	.309*	.082	.327*	-.022
	Sig. (2-tailed)	.003	.022		.000	.016	.533	.011	.870
	N	60	60	60	60	60	60	60	60
M4	Pearson Correlation	-.273*	-.292*	.571**	1	.101	.164	.261*	-.044
	Sig. (2-tailed)	.035	.024	.000		.441	.212	.044	.740
	N	60	60	60	60	60	60	60	60
M5	Pearson Correlation	-.652**	-.642**	.309*	.101	1	.700**	.123	.159
	Sig. (2-tailed)	.000	.000	.016	.441		.000	.349	.226
	N	60	60	60	60	60	60	60	60
M6	Pearson Correlation	-.432**	-.468**	.082	.164	.700**	1	-.055	.104
	Sig. (2-tailed)	.001	.000	.533	.212	.000		.676	.429
	N	60	60	60	60	60	60	60	60
M7	Pearson Correlation	-.243	-.236	.327*	.261*	.123	-.055	1	.223
	Sig. (2-tailed)	.062	.069	.011	.044	.349	.676		.086
	N	60	60	60	60	60	60	60	60
M8	Pearson Correlation	-.019	.006	-.022	-.044	.159	.104	.223	1
	Sig. (2-tailed)	.886	.964	.870	.740	.226	.429	.086	
	N	60	60	60	60	60	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 4.21: Discriminative analysis assumptions verification: matrix scatterplot
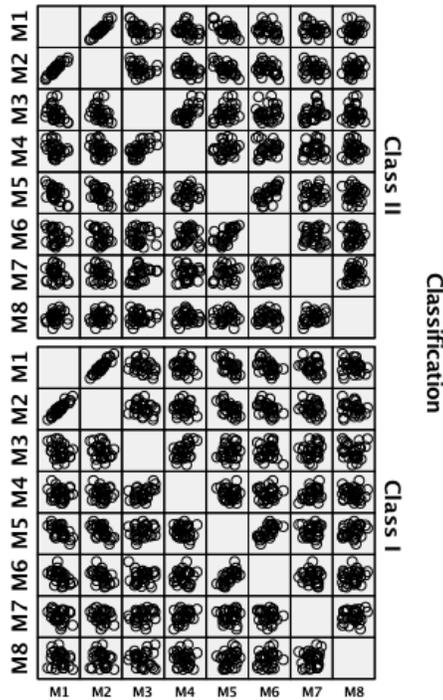
Appendix 4.22: Discriminative analysis assumptions verification: Box's Test of Equality of Covariance Matrices

Box's M		39.733
F	Approx.	.941
	df1	36
	df2	11319.368
	Sig.	.571

Tests null hypothesis of equal
population covariance
matrices.

Appendix 4.23: Descriptive statistics including mean, standard deviation and number of cases for each measurement in discriminative analysis

	N	Minimum	Maximum	Mean	Std. Deviation
M1	60	51.2542681	74.2348301	61.5198949	5.16915882
M2	60	48.5065975	73.5679958	61.1071857	5.47939624
M3	60	87.1785437	105.613273	94.5388795	4.02049441
M4	60	87.3572115	104.971017	95.1130486	4.30038969
M5	60	2.87540966	22.8905067	12.7932622	4.41983756
M6	60	4.69630575	21.2864280	12.4037330	4.46030840
M7	60	76.3970659	94.6644495	85.8672697	4.27139405
M8	60	78.4367795	100.851839	89.7284393	4.91726605
Valid N (listwise)	60				

Chapter 5 : General discussion & major conclusions

Alycia Sam, Manuel Lagravere-Vich, Carlos Flores-Mir, and Heesoo Oh

General discussion

This thesis was designed to encompass studies that delved into the scope of 3D skeletal and dental radiographic landmarking, using such information to gain a better understanding of and possibly even uncover distinct relationships shared within different malocclusion groups. The ultimate end goal of this research was to develop a novel 3D craniofacial analysis to adequately categorize orthodontic malocclusions, specifically Class I and Class II-1, using skeletal and dental relationships collectively from all three planes of space. The first question that came to mind was: how are we going to go about doing this?

To begin, an initial patient chart review of the entire CBCT database from the University of Alberta Graduate Orthodontic clinic was done. This comprised of over 750 CBCTs taken between January 2013 and December 2016 as a routine part of orthodontic diagnosis and treatment planning. They were all reviewed to isolate pretreatment full field of view (F-FOV) CBCTs from the database. Nearly 300 F-FOV CBCTs were acquired from this process. Next, each of these CBCT scans were reviewed using the checklist for initial patient reviews previously described in *Appendix 4.1* by year beginning with 2013. The purpose of this was to verify that choice for of Class I and Class II-1 malocclusion groups followed our justification that these classifications were the most prevalent in the general population. This trend was in fact consistent from 2013

through 2015 with the following results: 88 Class I, 91 Class II-1, 20 Class II-2, 34 Class III, and 16 Fulfills exclusion criteria. Note, 2016 was not reviewed in this second step as this pattern was already demonstrated in the previous three years. These Class I and Class II CBCT scans served as the subject pools for the research studies in this manuscript.

The main challenge in this undertaking was the mere fact that this was a novel 3D analysis we wanted to create –to our knowledge, a detailed protocol as to how to go about such a task does not exist. Forerunners like *Lee et. al*[1] have been successful in introducing newly defined 3D landmarks and using them to statistically differentiate between patients with Class I and Class II skeletal problems, although there were distinct limitations to their proposal. Indeed, it was a systematic method for using 3D landmarks as the basis for a novel skeletal analysis, but landmark location was based on an average of only two of three planes (sagittal and axial) and measurements taken from them were compared individually between the two classification groups. These were a couple of things kept in mind when shaping the research objectives for this thesis.

Three research objectives for this thesis, with major conclusions of each:

Objective #1: To evaluate the reliability, intra-examiner, of landmark identification in 3D cephalometric imaging.

In 2014, Lisboa *et al.* published a systematic review (SR) appraising the scientific literature available concerning the reliability and reproducibility of 3D cephalometric landmarks using CBCT.[2] Before beginning the greater part of this thesis, an updated search was conducted in October 2017. An additional 3

articles for that met inclusion criteria, while also deciding to eliminate any articles that used human dry skulls as a sample population. A total of 13 articles related to the topic of interested were generated.

Included studies underwent an assessment using a newly developed methodological scoring for reliability articles, adapted from Lagravere *et al*.[3] with modifications based on Bialocerkowski et al.[4] The majority had methodological limitations and were of moderate quality. Due to their heterogeneity, key data of each could not be combined and was reported qualitatively. Midsagittal plane followed by bilateral landmarks demonstrated the highest reliability in 3D. Those with the lowest reliability were located on anatomic structures with prominent curvatures without definitive borders, such as those marked on the condyle.

Reliability of landmark identification is often reported with ICC's or millimeters of variation, either demonstrated within (intra-) or between (inter-) observers. Majority of 3D landmarks demonstrated excellent intra- and inter-examiner reliabilities (ICC>0.9) albeit they tended to be slightly better for the prior. Although a minimum number of dental landmarks were reported on, most are recommended for use. Both traditional and non-traditional cephalometric landmarks were equally reliable. No consensus for thresholds used to determine landmarking variations with clinical significance was found, ranging anywhere from 0.5-1.5mm. For the most part, landmarks with a variation <0.5mm were deemed not clinically significant.

Objective #2:

a) To examine the potential differences between 2D normative values and 3D novel skeletal and dental measurements, for those linear and angular measurements that can be obtained with both methods of imaging.

A 3D sample vs. 2D standard norm multivariate analysis was performed to compare measurements taken from these two different imaging modalities. In the Class I malocclusion group, *S-N* and *ANB* had mean measurement value differences between the 3D estimate population mean and 2D standard norm population. The same was seen in the Class II-1 malocclusion group in addition to *SNB*. All 6 measurements tested were constructed along the midsagittal plane, and it would be beneficial for future studies to focus on making these comparisons using para-medial components as well. This type of follow-up will require the development and use of mathematical coding using numerical analysis software to allow for direct comparison in 2D and 3D for a complete analysis of craniofacial measurements in its totality.

b) To examine the potential differences between 3D measurements taken from right and left sides in the two malocclusion groups.

A right vs. left multivariate analysis multivariate analysis was performed to compare measurements taken from two different sides of an individual. Notable differences were not observed for all measurements tested, but were apparent for 3 comparisons in the Class I malocclusion group (*U1-APog, MP-SN, PP-SN*) and 2 comparisons in Class II malocclusion group (*U1-APog, MP-SN*). It is evident that both the specific measurement in question and type of malocclusion

130

group may have substantial influences on whether right and left differences be existent.

Objective #3: To investigate the existence of alternative 3D skeletal and dental relationships to adequately categorize orthodontic malocclusions, specifically Class I and Class II Division 1 (Class II-1).

A discriminant analysis (DA) was done to evaluate how well 8 predictor measurements could classify patients into malocclusion groups, Class I and Class II-1. This statistical model was originally chosen over one-way MANOVA because it can determine which of these predictor measurements could discriminate best between these two classifications. However, there was evidence that group malocclusion could not be predicted from these measurement values on the discriminant functions and so follow-up analyses were not technically required. We decided to perform them anyways, and it gave some indication that 5 of these predictor measurements (*Infraorbital (right) – Mental foramen (right)* distance, *Infraorbital (left) – Mental foramen (left)* distance, *2.6 root apex – Infraorbital (left) – 2.6 buccal* angle, *Infraorbital (right) – Infraorbital (left) – 2.6 buccal angle*, and *1.6 root apex – Infraorbital (right) – 1.6 buccal angle)* could be stronger discriminators than the others. Without a significant overall result for DA, these findings could not be substantiated. For this study and associated outcome, the follow-up analyses to DA were conducted for more of interest than a means of evidence *per se*.

Limitations

3D landmark selections based on SR

The articles obtained from the most recent SR obtained an extensive compilation of traditional and non-traditional 3D cephalometric landmarks that have undergone reliability testing. Firstly, these in no way reflect all those that are possible since one could create a completely new landmark or change the definition for a previously named one also. Secondly, it was difficult to use a rigid cut-off point for identification variations with clinical significance and typically ranged anywhere from 0.5-1.5mm. We were also purposeful in incorporating a relatively equal number of skeletal and dental landmarks that also may be useful in constructing important craniofacial measurements. With that said, ultimately final choices in landmark selection were somewhat subjective even though we tried to avoid it as much as possible.

Intra-reliability testing

Intra-reliability testing was carried out by the principal investigator of this study, evaluating the reliability of 46 landmarks. These landmarks were chosen based on both findings of the SR and authors' experience in their potential importance in analysis of craniofacial form in 3D. The quantity was kept to a limited number to prevent examiner fatigue even for an experienced examiner, especially since landmarking of patient CBCT scans were repeated three times and thus affecting the results. This sheer number of landmarks was balanced by the lower number of subjects tested, 10.

The principal investigator had significant experience with Avizo Version 8.1 software (Visualization Sciences Group, Burlington, MA, USA) and familiarly

with the precise landmark definitions. This may have given them an advantage in use and manipulation of CBCT in multiplanar reconstruction views (MPRV) and 3D virtual reconstruction views (3D-VRV) during landmark placement.

Inter-examiner reliability testing was originally planned, however authors decided to omit this during the execution stage. It was structured to compare the reliability of landmark placement in 3D between three separate examiners. Two examiners in addition to the principal investigator had previous experience with the software, although no additional hands-on training or calibration was provided. The intra-examiner component was deemed more relevant to the conduction of subsequent studies and for this reason, its omission was not deemed significant in meeting our objectives.

Forty of the 46 landmarks demonstrated excellent intra-examiner reliability and were then selected to be used in further studies.

Examining potential differences in skeletal and dental measurements

As many of the 40 landmarks set in 3D would have been ideally used to construct measurements for a direct comparison with 2D normative values. Unfortunately, this was not a straight-forward task for all measurements, especially when an angle was to be taken between two distinct planes. Each plane would be constructed from 2 landmarks (for a total of 4) that complicates matters. It would require landmarks be individually projected onto a pre-defined sagittal plane, construction of the planes of interest, where only then appropriate measurements could be taken. Such a feat would require the development and use of a new mathematical algorithm, and so was not included in this manuscript.

However, these authors are currently working on its development for a side-project.

Direct comparisons of 2D normative values and 3D skeletal and dental measurements was possible for 6 measurements of interest. For the remainder of the 3D measurements we were interested in (and where direct comparison of 2D and 3D not possible), comparisons between right and left sides was performed. This bilateral comparison was completed for the 9 remaining measurements.

3D measurement selections & use in novel cephalometric analysis

Without solid evidence that a given 3D measurement can distinguish between different malocclusion groups, it will and has been challenging to combine it with others to carry-out a discriminant function analysis with predictor variables entered in one step. It was used to try and assess how good malocclusion classifications could be predicted by 8 discriminant variables, in our case 8 predictor measurements. The issue with DA is that is a combination effect; one or more of these predictor measurements may be significantly related to membership of a group where others are not.

Based on the findings of such research, it could be prudent to rethink the classification of various orthodontic malocclusions. Divergence from the traditional clear-cut definitions for Class I, Class II-1, Class II-2, and Class III malocclusion groups may be necessary for a more progressive approach to classifying a patient's craniofacial complex and understanding their treatment

needs. Perhaps consideration of a more unique approach to addressing

conditions and needs is necessitated.

Future recommendations

- To compare 2D normative values and 3D skeletal and dental

  measurements that incorporate four or more landmarks in its construct, as

  to finding the angle between two distinct planes, 3D landmarks need to be

  projected on a sagittal plane and a mathematical algorithm developed to

  allow for direct comparison with 2D. The authors of this chapter are in the

  process of working on this as a side-project.

- Create and evaluate the reliability of newly defined 3D landmarks. With

  many limitations of 2D imaging removed, the possibilities are numerous.

  Once reliability is proven, seek to use them in the construct of

  measurements that describe the craniofacial complex. However, not all

  measurements taken in 3D will be important skeletal and dental predictors

  in orthodontic malocclusion classifications. It will be essential for an

  investigator to critically evaluate the potential usefulness of a

  measurement in question and its ability to differentiate between different

  malocclusion groups.

- After the identification of reliable 3D landmarks and use in taking

  measurements, a comprehensive library of 3D normative values for

  different malocclusion groups could be helpful in explaining craniofacial

  relationships in three planes of space. For example, we would expect that

  some linear measurements have greater values when taken at a diagonal

spatially in comparison to the same measurement taken from imaging from a single perspective dimensionally.

- Although plausible, it would be quite complex to apply an algorithm to existing 2D cephalometric longitudinal growth studies to convert the data into 3D, as to avoid needing to take a CBCT scan altogether. The craniofacial complex is not a one-size fits all unit.

- Further research is needed for continued development and testing of normative 3D data, to allow a novel 3D cephalometric analysis to be used as a diagnostic aid at diagnosis and treatment planning stages. This could enable clinicians to replace the use of 2D imaging to confidently identify patient malocclusion classifications using an alternate modality. It would permit the use of information from a CBCT scan in its totality, rather than needing to generate CBCT-generated cephalograms and then basing analyses on measurements taken from two planes of space.

References

1. Lee, M., G. Kanavakis, and R.M. Miner, *Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review.* Angle Orthodontist, 2015. **85**(1): p. 3-10.
2. Lisboa, C.d.O., et al., *Reliability and reproducibility of three-dimensional cephalometric landmarks using CBCT: a systematic review.* Journal of Applied Oral Science, 2015. **23**(2): p. 112-119.
3. Lagravère, M.O. and P.W. Major, *Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography.* American Journal of Orthodontics and Dentofacial Orthopedics, 2005. **128**(5): p. 657-660.
4. Bialocerkowski, A., N. Klupp, and P. Bragge, *How to read and critically appraise a reliability article.* International journal of therapy and rehabilitation 2010. **17**(3): p. 112-114.

Bibliography

1. Ruellas, A.C., et al., *Common 3-dimensional coordinate system for assessment of directional changes.* (1097-6752 (Electronic)).
2. Rischen, R.J., et al., *Records needed for orthodontic diagnosis and treatment planning: a systematic review.* (1932-6203 (Electronic)).
3. Durao, A.R., et al., *Validity of 2D lateral cephalometry in orthodontics: a systematic review.* (2196-1042 (Electronic)).
4. Karatas, O.H. and E. Toy, *Three-dimensional imaging techniques: A literature review.* (1305-7456 (Print)).
5. Machado, G.L., *CBCT imaging - A boon to orthodontics.* (1013-9052 (Print)).
6. Nalcaci, R., O. Ozturk F Fau - Sokucu, and O. Sokucu, *A comparison of two-dimensional radiography and three-dimensional computed tomography in angular cephalometric measurements.* (0250-832X (Print)).
7. Devereux, L., et al., How important are lateral cephalometric radiographs in orthodontic treatment planning? (1097-6752 (Electronic)).
8. Ludlow, J.B., et al., Precision of cephalometric landmark identification: Cone-beam computed tomography vs conventional cephalometric views. American Journal of Orthodontics and Dentofacial Orthopedics, 2009. **136**(3): p. 312.e1-312.e10.
9. Lagravere, M.O., et al., Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments. (0003-3219 (Print)).
10. Lagravère, M.O., et al., Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. American Journal of Orthodontics and Dentofacial Orthopedics, 2010. **137**(5): p. 598-604.
11. van Vlijmen, O.J.C., et al., A comparison between 2D and 3D cephalometry on CBCT scans of human skulls. International journal of oral and maxillofacial surgery, 2010. **39**(2): p. 156-160.
12. Naji, P., N.A. Alsufyani, and M.O. Lagravere, Reliability of anatomic structures as landmarks in three-dimensional cephalometric analysis using CBCT. Angle Orthodontist, 2014. **84**(5): p. 762-772.
13. Lagravere M.O., G.J.M., Flores-Mir C., Carey J., Heo G., Major P.W., Cranial base foramen location accuracy and reliability in cone-beam computerized tomography. Am J Orthod Dentofacial Orthop, 2011. **139**(3): p. 203-10.
14. Lisboa, C.d.O., et al., Reliability and reproducibility of three-dimensional cephalometric landmarks using CBCT: a systematic review. Journal of Applied Oral Science, 2015. **23**(2): p. 112-119.
15. Knobloch, K., P.M. Yoon U Fau - Vogt, and P.M. Vogt, Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. (1878-4119 (Electronic)).
16. van Vlijmen, O.J., et al., Evidence supporting the use of cone-beam computed tomography in orthodontics. (1943-4723 (Electronic)).

17. Bialocerkowski, A., N. Klupp, and P. Bragge, How to read and critically appraise a reliability article. International journal of therapy and rehabilitation 2010. **17**(3): p. 112-114.

18. Da Neiva, M.B., et al., Evaluation of cephalometric landmark identification on CBCT multiplanar and 3D reconstructions. Angle Orthodontist, 2015. **85**(1): p. 11-17.

19. Lee, M., G. Kanavakis, and R.M. Miner, Newly defined landmarks for a three-dimensionally based cephalometric analysis: a retrospective cone-beam computed tomography scan review. Angle Orthodontist, 2015. **85**(1): p. 3-10.

20. Ghoneima, A., et al., Accuracy and reliability of landmark-based, surface-based and voxel-based 3D cone-beam computed tompography superimposition methods Orthod Craniofac Res., 2017. **20**(4): p. 227-236.

21. Zamora, N., et al., A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis. Medicina Oral, Patologia Oral y Cirugia Bucal, 2012. **17**(4): p. e678-e688.

22. Frongia, G., et al., Assessment of the reliability and repeatability of landmarks using 3-D cephalometric software. Cranio, 2012. **30**(4): p. 255-263.

23. Schlicher, W., et al., Consistency and precision of landmark identification in three-dimensional cone beam computed tomography scans. European journal of orthodontics, 2012. **34**(3): p. 263-275.

24. Hassan, B., et al., Precision of identifying cephalometric landmarks with cone beam computed tomography in vivo. European journal of orthodontics, 2013. **35**(1): p. 38-44.

25. Chien, P.C., et al., Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. Dentomaxillofacial Radiology, 2009. **38**(5): p. 262-273.

26. de Oliveira, A.E.F., et al., Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2009. **107**(2): p. 256-265.

27. Strachan, A., In the brain of the beholder: the neuropsychological basis of aesthetic preferences. Harvard Brain 2000: p. 7.

28. Mah, J.K., J.C. Huang, and H. Choo, Practical applications of cone-beam computed tomography in orthodontics. The Journal of the American Dental Association, 2010. **141**: p. 7S-13S.

29. Imaging, V., Amira 5: Amira User's Guide. 1999-2009. p. 40-42.

30. Adams, G.L., et al., *Comparison between traditional 2-dimensional cephalometry and a 3-dimensional approach on human dry skulls.* American Journal of Orthodontics and Dentofacial Orthopedics, 2004. **126**(4): p. 397-409.

31. Lagravère, M.O. and P.W. Major, *Proposed reference point for 3-dimensional cephalometric analysis with cone-beam computerized tomography.* American Journal of Orthodontics and Dentofacial Orthopedics, 2005. **128**(5): p. 657-660.

32. Olszewski, R., et al., *3D CT-based cephalometric analysis: 3D cephalometric theoretical concept and software.* Neuroradiology, 2006. **48**(11): p. 853-862.

33. Subramanyan, K., M. Palomo, and M. Hans. *Creation of three dimensional craniofacial standards from CBCT images*.

34. Lou, L., et al., *Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2007. **104**(3): p. 402-411.

35. Periago, D.R., et al., *Linear accuracy and reliability of cone beam CT derived 3-dimensional images constructed using an orthodontic volumetric rendering program.* Angle Orthodontist, 2008. **78**(3): p. 387-395.

36. Schlueter, B., et al., *Cone beam computed tomography 3D reconstruction of the mandibular condyle.* Angle Orthodontist, 2008. **78**(5): p. 880-888.

37. Brown, A.A., et al., *Linear accuracy of cone beam CT derived 3D images.* Angle Orthodontist, 2009. **79**(1): p. 150-157.

38. Hassan, B., P. Van Der Stelt, and G. Sanderink, *Accuracy of three-dimensional measurements obtained from cone beam computed tomography surface-rendered images for cephalometric analysis: Influence of patient scanning position.* European journal of orthodontics, 2009. **31**(2): p. 129-134.

39. Moreira, C.R., et al., *Assessment of linear and angular measurements on three-dimensional cone-beam computed tomographic images.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2009. **108**(3): p. 430-436.

40. Couceiro, C.P. and O. de Vasconcellos Vilella, *2D / 3D Cone-Beam CT images or conventional radiography: Which is more reliable?* Dental Press Journal of Orthodontics, 2010. **15**(5): p. 40-41.

41. Kim, J., D. Lee, and S. Lee, *Comparison of the observer reliability of cranial anatomic landmarks based on cephalometric radiograph and three-dimensional computed tomography scans.* Journal of the Korean Association of Oral and Maxillofacial Surgeons, 2010. **36**(4): p. 262-269.

42. Kim, T.S., et al., *A comparison of cone-beam computed tomography and direct measurement in the examination of the mandibular canal and adjacent structures.* Journal of endodontics, 2010. **36**(7): p. 1191-1194.

43. Ogawa, N., et al., *Application of cone beam CT 3D images to cephalometric analysis.* Orthodontic Waves, 2010. **69**(4): p. 138-150.

44. Chen, J., et al. *A new annotation method for 3D cephalometric landmark in CBCT*.

45. Cheung, L.K., et al., *Three-dimensional cephalometric norms of Chinese adults in Hong Kong with balanced facial profile.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2011. **112**(2): p. e56-e73.

46. Damstra, J., et al., *Reliability and the smallest detectable difference of measurements on 3-dimensional cone-beam computed tomography images.* American Journal of Orthodontics and Dentofacial Orthopedics, 2011. **140**(3): p. e107-e114.

47. Gribel, B.F., et al., *From 2D to 3D: An algorithm to derive normal values for 3-dimensional computerized assessment.* Angle Orthodontist, 2011. **81**(1): p. 5-12.

48. Lagravre, M.O., et al., *Optimization analysis for plane orientation in 3-dimensional cephalometric analysis of serial cone-beam computerized tomography images.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 2011. **111**(6): p. 771-777.

49. Medelnik, J., et al., *Accuracy of anatomical landmark identification using different CBCT- and MSCT-based 3D images : An in vitro study.* Journal of Orofacial Orthopedics, 2011. **72**(4): p. 261-278.

50. Tomasi, C., et al., *Reliability and reproducibility of linear mandible measurements with the use of a cone-beam computed tomography and two object inclinations.* Dentomaxillofacial Radiology, 2011. **40**(4): p. 244-250.

51. van Vlijmen, O.J.C., et al., *Measurements on 3D models of human skulls derived from two different cone beam CT scanners.* Clinical oral investigations, 2011. **15**(5): p. 721-727.

52. Wong, R.W.K., A.C.M. Chau, and U. Hägg, *3D CBCT McNamara's cephalometric analysis in an adult southern Chinese population.* International journal of oral and maxillofacial surgery, 2011. **40**(9): p. 920-925.

53. Yildirim, E., et al., *Evaluation of differences between two and three dimensional cephalometric measurements.* Gulhane Medical Journal, 2011. **53**(1): p. 43-49.

54. Reychler, H.P., *Reproducibility of 3D cephalometric landmarks on cone-beam and low-dose computed tomography.* 20th International Conference on Oral and Maxillofacial Surgery, ICOMS 2011 Santiago Chile, 2011. **40**(10): p. 1160.

55. Frongia, G., M.G. Piancino, and P. Bracco, *Cone-beam computed tomography: Accuracy of three-dimensional cephalometry analysis and influence of patient scanning position.* Journal of Craniofacial Surgery, 2012. **23**(4): p. 1038-1043.

56. Kim, M., et al., *Evaluation of accuracy of 3D reconstruction images using multi-detector CT and cone-beam CT.* Imaging Science in Dentistry, 2012. **42**(1): p. 25-33.

57. Kim, S.G., et al., *Development of 3D statistical mandible models for cephalometric measurements.* Imaging Science in Dentistry, 2012. **42**(3): p. 175-182.

58. Patcas, R., et al., *Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: A tale of two CTs.* Dentomaxillofacial Radiology, 2012. **41**(8): p. 637-644.

59. Santos, T.d.S., et al., *Evaluation of reliability and reproducibility of linear measurements of cone-beam-computed tomography.* Indian Journal of Dental Research, 2012. **23**(4): p. 473-478.

60. Shibata, M., et al., *Reproducibility of three-dimensional coordinate systems based on craniofacial landmarks: a tentative evaluation of four systems created on images obtained by cone-beam computed tomography with a large field of view.* Angle Orthodontist, 2012. **82**(5): p. 776-784.

61. Titiz, I., et al., *Repeatability and reproducibility of landmarks--a three-dimensional computed tomography study.* European journal of orthodontics, 2012. **34**(3): p. 276-286.
62. Frongia, G., P. Bracco, and M.G. Piancino, *Three-dimensional cephalometry: a method for the identification and for the orientation of the skull after cone-bean computed tomographic scan.* Journal of Craniofacial Surgery, 2013. **24**(3): p. e308-11.
63. Gaia, B.F., et al., *Comparison of precision and accuracy of linear measurements performed by two different imaging software programs and obtained from 3D-CBCT images for le Fort i osteotomy.* Dentomaxillofacial Radiology, 2013. **42**(5).
64. Katkar, R.A., et al., *Comparison of observer reliability of three-dimensional cephalometric landmark identification on subject images from Galileos and i-CAT cone beam CT.* Dento-Maxillo-Facial Radiology, 2013. **42**(9): p. 20130059.
65. Mendonca, D.A., et al., *Comparative study of cranial anthropometric measurement by traditional calipers to computed tomography and three-dimensional photogrammetry.* Journal of Craniofacial Surgery, 2013. **24**(4): p. 1106-1110.
66. Olszewski, R., et al., *Reproducibility of three-dimensional cephalometric landmarks in cone-beam and low-dose computed tomography.* Clinical oral investigations, 2013. **17**(1): p. 285-292.
67. Farronato, G., et al., *Assessment of inter- and intra-operator cephalometric tracings on cone beam CT radiographs: comparison of the precision of the cone beam CT versus the latero-lateral radiograph tracing.* Progress in Orthodontics, 2014. **15**: p. 1.
68. Fuyamada, M., et al., *Reproducibility of maxillofacial landmark identification on three-dimensional cone-beam computed tomography images of patients with mandibular prognathism: Comparative study of a tentative method and traditional cephalometric analysis.* Angle Orthodontist, 2014. **84**(6): p. 966-973.
69. Gaia, B.F., et al., *Validity of three-dimensional computed tomography measurements for le Fort i osteotomy.* International journal of oral and maxillofacial surgery, 2014. **43**(2): p. 197-203.
70. Gaia, B.F., et al., *Accuracy and reliability of linear measurements using 3-dimensional computed tomographic imaging software for Le Fort I Osteotomy.* British Journal of Oral & Maxillofacial Surgery, 2014. **52**(3): p. 258-263.
71. Hwang, J.J., et al., *Factors influencing superimposition error of 3D cephalometric landmarks by plane orientation method using 4 reference points: 4 point superimposition error regression model.* PLoS ONE [Electronic Resource], 2014. **9**(11): p. e110665.
72. Jacquet, W., et al., *The influence of slice orientation in 3D CBCT images on measurements of anatomical structures.* IFMBE Proceedings, 2014. **41**: p. 229-232.

73. Jung, S.Y., et al., *Analysis of mandibular structure using 3D facial computed tomography.* Otolaryngology - Head and Neck Surgery (United States), 2014. **151**(5): p. 760-764.

74. Kim, H.J., et al., *Construction and validation of the midsagittal reference plane based on the skull base symmetry for three-dimensional cephalometric craniofacial analysis.* Journal of Craniofacial Surgery, 2014. **25**(2): p. 338-342.

75. Lee, K.M., H.S. Hwang, and J.H. Cho, *Comparison of transverse analysis between posteroanterior cephalogram and cone-beam computed tomography.* Angle Orthodontist, 2014. **84**(4): p. 715-719.

76. Lee, S.H., et al., *Three-dimensional architectural and structural analysis - A transition in concept and design from Delaire's cephalometric analysis.* International journal of oral and maxillofacial surgery, 2014. **43**(9): p. 1154-1160.

77. Liang, C., et al., *Norms of mcnamara's cephalometric analysis on lateral view of 3D CT imaging in adults from Northeast China.* Journal of Hard Tissue Biology, 2014. **23**(2): p. 249-254.

78. Fernandes, T.M.F., et al., *Comparison between 3D volumetric rendering and multiplanar slices on the reliability of linear measurements on CBCT images: an in vitro study.* Journal of Applied Oral Science, 2015. **23**(1): p. 56-63.

79. Metzler, P., et al., *Validity of the 3D VECTRA photogrammetric surface imaging system for cranio-maxillofacial anthropometric measurements.* (1865-1569 (Electronic)).

80. Jung, P.K., G.C. Lee, and C.H. Moon, *Comparison of cone-beam computed tomography cephalometric measurements using a midsagittal projection and conventional two-dimensional cephalometric measurements.* Korean Journal of Orthodontics, 2015. **45**(6): p. 282-288.

81. Pittayapat, P., et al., *Reproducibility of the sella turcica landmark in three dimensions using a sella turcicaspecific reference system.* Imaging Science in Dentistry, 2015. **45**(1): p. 15-22.

82. Gupta, A., et al., *A knowledge-based algorithm for automatic detection of cephalometric landmarks on CBCT images.* (1861-6429 (Electronic)).

83. Dias P Fau - Neves, L., et al., *CraMs: Craniometric Analysis Application Using 3D Skull Models.* (1558-1756 (Electronic)).

84. Nam, K.U. and J. Hong, *Is Three-Dimensional Soft Tissue Prediction by Software Accurate?* (1536-3732 (Electronic)).

85. Lemieux, G., et al., *Precision and accuracy of suggested maxillary and mandibular landmarks with cone-beam computed tomography for regional superimpositions: An in vitro study.* American Journal of Orthodontics & Dentofacial Orthopedics, 2016. **149**(1): p. 67-75.

86. Wikner, J., et al., *Linear accuracy and reliability of volume data sets acquired by two CBCT-devices and an MSCT using virtual models: a comparative in-vitro study.* (1502-3850 (Electronic)).

87. Almuzian, M., et al., *Effects of Le Fort I Osteotomy on the Nasopharyngeal Airway-6-Month Follow-Up.* (1531-5053 (Electronic)).

88. Gupta, A., et al., *Accuracy of 3D cephalometric measurements based on an automatic knowledge-based landmark detection algorithm.* (1861-6429 (Electronic)).

89. Zhang J Fau - Gao, Y., et al., *Automatic Craniomaxillofacial Landmark Digitization via Segmentation-Guided Partially-Joint Regression Forest Model and Multiscale Statistical Features.* (1558-2531 (Electronic)).

90. Gribel, B.F., et al., Accuracy and reliability of craniometric measurements on lateral cephalometry and 3D measurements on CBCT scans. (1945-7103 (Electronic)).

91. Kragskov, J., et al., Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans. (1055-6656 (Print)).

92. Portney, L. and M. Watkins, Foundations of clinical research: Applications to practice. 2009, Upper Saddle River, NJ: Pearson/Prentice Hall

93. Forsyth, D.B., S. Shaw Wc Fau - Richmond, and S. Richmond, *Digital imaging of cephalometric radiography, Part 1: Advantages and limitations of digital imaging.* (0003-3219 (Print)).

94. Silva, M.A., et al., *Cone-beam computed tomography for routine orthodontic treatment planning: a radiation dose evaluation.* (1097-6752 (Electronic)).

95. Proffit, W.R., H.W. Fields, and D.M. Sarver, *Contemporary Orthodontics-E-Book.* 2014: Elsevier Health Sciences.

96. Van der Linden, F.P.G.M., *Development of the dentition.* Vol. 1. 1983: Quintessence Pub Co.

97. Lavergne, J. and A. Petrovic, *Discontinuities in occlusal relationship and the regulation of facial growth. A cybernetic view.* The European Journal of Orthodontics, 1983. **5**(4): p. 269-278.

98. McNamara Jr, J.A., *Components of Class II malocclusion in children 8–10 years of age.* The Angle Orthodontist, 1981. **51**(3): p. 177-202.

99. Moyers, R., *Auxologic categorization and chronobiologic specification for the choice of appropriate orthodontic treatment.* American Journal of Orthodontics and Dentofacial Orthopedics, 1994. **105**(2): p. 192-205.