Factors Affecting Alveolar Bone Height Measurements from CBCT Images

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Ryan LaDell Wood, DDS

Graduate Program in Dentistry

The Ohio State University

2012

Master's Examination Committee:

Dr. Zongyang Sun, Advisor

Dr. Do-Gyoon Kim

Dr. Jahanzeb Chaudhry

Dr. Binnaz Leblebicioglu

ABSTRACT

Objectives: Cone beam computed tomography (CBCT) is routinely used to measure alveolar bone dimensions. This study investigated factors likely affecting the accuracy of such measurements. **Methods:** Marker holes, apical to maxillary and mandibular molar roots, and mesio-distal molar occlusal reference grooves were created in fresh pig heads (n=26 quadrants), followed by CBCT scans at 0.4 and 0.2 mm voxel sizes before and after soft tissue removal. Subsequently, bucco-lingual sections bisecting the marker holes were cut, from which physical alveolar bone height and thickness were measured. Two blinded raters, using Dolphin 3D and OsiriX software, independently collected alveolar bone height measurements from CBCT images. Differences between CBCT and physical measurements ($D_{CBCT-Phy}$) were calculated. The mean $D_{CBCT-Phy}$ and limit of agreement (LOA, ±1.96SD) for each factor were depicted by Bland-Altman methods. The influence of each factor was examined by repeated measures ANOVAs. **Results:** The rater factor affected the mean CBCT measurements in both jaws, with mean $D_{CBCT-Phy}$ from a more-experienced rater significantly ($p<0.05$) closer to "0" than from a less-experienced rater, while their LOA ranges were similar (Maxilla, 0.17mm±2.09mm vs.-0.48mm±2.02mm; Mandible, 0.01mm±1.04mm vs. -0.70mm±1.21mm). The software factor also affected the mean CBCT measurements in both jaws, with mean $D_{CBCT-Phy}$ from Osirix

significantly ($p<0.05$) closer to "0" than from Dolphin-3D despite similar LOA ranges (Maxilla, -0.09mm±2.18mm vs. -0.22mm±2.11mm; Mandible, -0.30mm±1.23mm vs. -0.39mm±1.42mm). The soft-tissue factor only affected the mandible significantly ($p<0.05$), with mean $D_{CBCT-Phy}$ from soft-tissue-absent specimens closer to "0" than those from soft-tissue-present specimens (Maxilla, -0.13mm±2.29mm vs. -0.19mm±2.00mm; Mandible, -0.25mm±1.32mm vs. -0.44mm±1.32mm). The voxel-size factor had an insignificant effect. Generally, mandibular CBCT measurements exhibited smaller LOA ranges than maxillary, and generally the mandibular alveolar bone was thicker.

Conclusions: Individual measurements of buccal alveolar bone from CBCT images can under- or overestimate buccal bone height by a clinically significant (>1mm) amount. Factors such as rater, software, and the presence/absence of soft tissue can potentially create systematic sub-millimeter differences between the measurements of buccal alveolar bone from CBCT images and physical measurements.

Dedicated to my perfect wife, Erin, and to our wonderful children.
Thank you for your unwavering support.

ACKNOWLEDGMENTS

I would like to thank my thesis committee, Dr. Zongyang Sun, Dr. Do-Gyoon Kim, Dr. Jahanzeb Chaudhry, and Dr. Binnaz Leblebicioglu for their time and great help.

I would like to thank Greg England for his diligent work in helping with the specimen preparation and measurements.

I would like to thank Boon Ching Tee for all her help with data compilation and analysis and for tutoring me in Excel.

I would like to thank the faculty, residents and staff of The Ohio State University Section of Orthodontics.

I would like to recognize the financial support for this research provided by the Dental Master's Thesis Award Program sponsored by Delta Dental Foundation which is the philanthropic affiliate of Delta Dental of Michigan, Ohio, and Indiana.

Lastly, I would like to thank my family for their love and support over the years.

VITA

March 26, 1980………………………...Born- Gooding, Idaho

2005……………………………….....B.A. Brigham Young University, Provo, UT

2009…………………………………DDS University of North Carolina, Chapel Hill, NC

2009-present………………………...Graduate Resident in Orthodontics
The Ohio State University
Columbus, Ohio

FIELDS OF STUDY

Major Field: Dentistry

Specialty: Orthodontics

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Computed Tomography (CT) has been used in the diagnosis and treatment of medical problems since the 1970's. However, the three dimensional radiographic images that it produces require high radiation exposure to patients. In an attempt to lower radiation exposure but still maintain sufficiently high resolution, Cone Beam Computed Tomography (CBCT) was invented. By forming the x-ray beam into a cone shape instead of in a flat, fan-shaped beam, as is done in medical CT's, the same three dimensional area can be visualized with approximately one tenth the radiation dose.

This decrease in radiation dosage and the three dimensional visualization of internal structures has caused many in the dental community to consider the usefulness of CBCT in dentistry. As a result, research has been done to evaluate CBCT's ability to accurately visualize carious lesions[1, 2], periodontal structures[3], alveolar bone[4], temporomandibular joints[5], airways[6], and various forms of pathology of the head and neck[7].

Orthodontics, in particular, has seen a large increase in the use of CBCT's to evaluate impacted teeth and root resorption[8], visualize root position and angulation[9], screen for pathology[10], and assess changes in the alveolar bone[11]. The last application, however, is controversial because of the unique nature of alveolar bone. Generally, compared to the basal bone portions of the maxilla and mandible, the alveolar bone is quite thin. It also has a thin cortical plate, which makes it difficult to identify the boundaries of the bone in radiographic images. Much of the alveolar bone surface is also penetrated by gingival fibers, which may further mask the boundary between hard and soft tissues.

In addition to the properties of the alveolar bone itself, several other factors can also potentially affect the accuracy of measuring alveolar bone from CBCT images. Factors such as the rater (the person making the measurements), the software used to view and measure the CBCT images, the presence/absence of soft tissue at and around the site of interest, and voxel size of the CBCT scan may also affect the accuracy of linear measurements of alveolar bone from CBCT images.

To date, the effect of rater, software, soft tissue, or voxel size on the accuracy of linear measurements of alveolar bone from CBCT images have not been examined together in any study. Since these factors are present in the clinical use of CBCT imaging, it is important to know the effect that these factors may have on the accuracy of linear measurements of alveolar bone from CBCT images. This knowledge may help us better understand how to most effectively use CBCT imaging to evaluate alveolar bone in clinical and research applications.

REVIEW OF THE LITERATURE

To date, few studies have looked at the effect that these factors may have on the accuracy of linear measurements of alveolar bone from CBCT images.

Rater:

Any effect of the rater on the accuracy of linear measurements of alveolar bone from CBCT images has only been explored indirectly. For example, Ganguly et al, in their study assessing the accuracy of linear measurements from CBCT images, used intraclass correlation tests to assess intra-rater reliability[16]. They found no statistically significant difference between measurements performed at the same location. Timock et al assessed their intra- and inter-rater reliability by calculating concordance correlation coefficients and Pearson correlation coefficients[4]. They found high intra- and inter-rater reliability with concordance correlation coefficients of 0.98 and 0.90 for buccal bone height and thickness measurements, respectively. Sun et al assessed intra- and inter-rater agreement values with intraclass correlations, finding excellent intra-rater ($r = 0.89$-0.98) and good inter-rater ($r = 0.64$-0.90) repeatability values for bone height measurements in CBCT images[17].

Intra- and inter-rater reliability tests were found in most studies looking at the accuracy of linear measurements of alveolar bone from CBCT images but these reliability tests may not fully uncover a systematic error which can exist between raters. No studies were found that looked specifically at the rater factor by comparing the results of various raters.

Software:

There was a variety of software used in other studies such as Image J[1], Iluma[2], New Tom[15], Sirona[16], and Dolphin[4,8,17] software, with Dolphin software being the most commonly used. No studies have used OsiriX software or have assessed the effect of software on the accuracy of linear measurements of alveolar bone from CBCT images.

Soft Tissue:

Several studies have looked at the accuracy of linear measurements from CBCT images of dry human skulls[13-15] or skulls with soft tissue substitute[14]. Only two studies were found that assessed the accuracy of linear measurements of alveolar bone from CBCT images with the specimen's soft tissue intact.

In one, Ganguly et al placed fiduciary markers on the buccal and lingual alveolar plates of six embalmed cadaver heads[16]. They then compared the accuracy of measuring the alveolar bone height from the fiduciary markers, directly and in CBCT images. They found no statistically significant difference in the accuracy of imaging and direct measurements but did not assess the effect, if any, that the presence of soft tissue had on the accuracy of the measurements.

In the other, Timock et al assessed the accuracy and reliability of buccal bone height and thickness measurements from CBCT images of embalmed cadaver heads with their soft tissue intact and without fiduciary markers. They found a mean difference between direct and CBCT measurements of 0.30 mm ± 0.79 mm for height and 0.13 mm ± 0.35 mm for thickness. However, that study did not specifically evaluate the effect that the presence of soft tissue has on the accuracy of linear measurements of alveolar bone from CBCT images.

In short, no studies were found that assessed the effect of soft tissue presence/absence on the accuracy of linear measurements of alveolar bone from CBCT images.

Voxel Size:

Wenzel et al used an in vitro model to assess the effect of voxel size (0.125 mm vs. 0.25 mm) related to the diagnosis of horizontal root fractures[12]. They found that the higher resolution images resulted in an increase in sensitivity without a decrease in specificity for detection of transverse root fractures in extracted teeth.

Kamburoglu et al used an in vitro model to assess, among other things, the effect of CBCT voxel size (0.1 mm, 0.2 mm, and 0.3 mm) on the diagnosis of occlusal caries[2]. They found that voxel size made no statistically significant difference in relative treatment effect values for all caries levels.

Sun et al used maxillary segments of 6-month-old pigs to evaluate the effect of bone thickness on alveolar bone height measurements from CBCT images. Among other things, they found that measurement inaccuracies were substantially improved when they decreased the voxel size of the scan from 0.40 mm to 0.25 mm.

Statement of the Problem

There are many factors that can potentially affect the accuracy of linear measurements taken from CBCT images, such as the rater, the software used, the presence/absence of soft tissue, and the resolution setting of the CBCT machine. The purpose of this study is to assess the effect that each of these factors has on the accuracy of linear measurements of alveolar bone from CBCT images by comparing them to physical measurements. Such assessment may help us better understand CBCT imaging related to both clinical and research applications.

Specific Aims

1. Assess the effect of the rater on the accuracy of linear measurements of alveolar bone from CBCT images.

2. Assess the effect of the software used (Dolphin-3D vs. OsiriX) on the accuracy of linear measurements of alveolar bone from CBCT images.

3. Assess the effect of soft tissue attachment on the accuracy of linear measurements of alveolar bone from CBCT images.

4. Assess the effect of CBCT voxel size (0.4 mm vs. 0.2 mm voxels) on the accuracy of linear measurements of alveolar bone from CBCT images.

Null Hypothesis

There will be no effect of rater, software, soft tissue attachment, or voxel size on the accuracy of linear measurements of alveolar bone from CBCT images.

References

1. Tsuchida R, Araki K, Okano T. Evaluation of a limited cone-beam volumetric imaging system: comparison with film radiography in detecting incipient

proximal caries. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007; 104:412-6.

2. Kamburoglu K, Murat S, Yuksel SP, Cebeci AR, Paksoy CS. Occlusal caries detection by using a cone-beam CT with different voxel resolutions and a digital intraoral sensor. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2010; 109:e63-9.
3. Mohan R, Singh A, Gundappa M. Three-dimensional imaging in periodontal diagnosis - Utilization of cone beam computed tomography. J Indian Soc Periodontol 2011; 15:11-7.
4. Timock AM, Cook V, McDonald T, Leo MC, Crowe J, Benninger BL, Covell DA, Jr. Accuracy and reliability of buccal bone height and thickness measurements from cone-beam computed tomography imaging. Am J Orthod Dentofacial Orthop 2011; 140:734-44.
5. Honda K, Arai Y, Kashima M, Takano Y, Sawada K, Ejima K, Iwai K. Evaluation of the usefulness of the limited cone-beam CT (3DX) in the assessment of the thickness of the roof of the glenoid fossa of the temporomandibular joint. Dentomaxillofac Radiol 2004; 33:391-5.
6. Stratemann S, Huang JC, Maki K, Hatcher D, Miller AJ. Three-dimensional analysis of the airway with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 140:607-15.

7. Alexiou K, Stamatakis H, Tsiklakis K. Evaluation of the severity of temporomandibular joint osteoarthritic changes related to age using cone beam computed tomography. Dentomaxillofac Radiol 2009; 38:141-7.

8. Oberoi S, Knueppel S. Three-dimensional assessment of impacted canines and root resorption using cone beam computed tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;

9. Bouwens DG, Cevidanes L, Ludlow JB, Phillips C. Comparison of mesiodistal root angulation with posttreatment panoramic radiographs and cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 139:126-32.

10. Caglayan F, Tozoglu U. Incidental findings in the maxillofacial region detected by cone beam CT. Diagn Interv Radiol 2011;

11. Corbridge JK, Campbell PM, Taylor R, Ceen RF, Buschang PH. Transverse dentoalveolar changes after slow maxillary expansion. Am J Orthod Dentofacial Orthop 2011; 140:317-25.

12. Wenzel A, Haiter-Neto F, Frydenberg M, Kirkevang LL. Variable-resolution cone-beam computerized tomography with enhancement filtration compared with intraoral photostimulable phosphor radiography in detection of transverse root fractures in an in vitro model. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009; 108:939-45.

13. Veyre-Goulet S, Fortin T, Thierry A. Accuracy of linear measurement provided by cone beam computed tomography to assess bone quantity in the posterior maxilla: a human cadaver study. Clin Implant Dent Relat Res 2008; 10:226-30.

14. Suomalainen A, Vehmas T, Kortesniemi M, Robinson S, Peltola J. Accuracy of linear measurements using dental cone beam and conventional multislice computed tomography. Dentomaxillofac Radiol 2008; 37:10-7.
15. Lascala CA, Panella J, Marques MM. Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom). Dentomaxillofac Radiol 2004; 33:291-4.
16. Ganguly R, Ruprecht A, Vincent S, Hellstein J, Timmons S, Qian F. Accuracy of linear measurement in the Galileos cone beam computed tomography under simulated clinical conditions. Dentomaxillofac Radiol 2011; 40:299-305.
17. Sun Z, Smith t, Kortam S, Kim D-Y, Tee B, Fields HW. Effect of bone thickness on alveolar bone-height measurements from cone-beam computed tomography images. Amer. J. Orthod. Dentofac. Orthop. 2011; 139(2):e117-27.
18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007; 39:175-191.

CHAPTER 2

MATERIALS AND METHODS

Sample Size and Power Analysis

Seven fresh pig heads (age 3-6 months) were used in this study. Because of the lack of existing knowledge on many factors to be studied in this project, our initial power analysis was based on one of our previous studies[17] in which alveolar bone height measurement without soft tissue attached had an average overestimation of 0.64 mm at the 0.4 mm resolution level, and the standard deviation for the differences among all specimens/locations was 0.74. In this study we estimate that keeping the soft tissue attached will at least cause the same amount of overestimation with the same standard deviation. Using these parameters, a power analysis by G*Power 3.0 [18], and based on paired t-tests, we found that an 80% power requires a sample size is 13. To be conservative, a sample size of 14 was determined for this study, which has a projected power of 85%.

Post hoc power analysis was performed after the data was collected based on repeated measures (between factors) ANOVAs (detailed in the Statistical Analysis).

Preparation and Imaging

Specimen preparation is shown in Figure 1. Reference marker holes were drilled in the buccal alveolar bone at the apices of the first and second maxillary molars (four holes per maxillary quadrant) and at the apices of the first mandibular molar (three holes per mandibular quadrant) using a slow speed dental handpiece (Volvere VMax, Brasseler USA, Savannah, Georgia) with a flat-end straight fissure cross-cut carbide bur (H31-21-009, Brasseler USA, Savannah, Georgia). Marker holes were made intraorally through the buccal gingiva, leaving all soft tissue of the head intact. Reference grooves bisecting the occlusal tables of all molars in a mesio-distal direction were also created with the same handpiece and bur type.

After the reference marker holes and occlusal grooves were made, each head was scanned with an iCAT CBCT machine (120 kVp, 5mA) at 0.4 mm voxel size and 0.2 mm voxel size resolutions. Subsequently, all soft tissue, including gingiva, was removed via manual dissection, followed by re-scanning at 0.4 mm voxel size and 0.2 mm voxel size. This resulted in a total of 28 CBCT image files (104 usable quadrants because one of the heads did not receive mandibular marker holes prior to imaging).

Sectioning and Physical Measurements

Each maxillary and mandibular quadrant was then separated from the head using a circular saw. Each quadrant was sectioned through each marker hole, using an IsoMet Low Speed Saw (Beuhler, Lake Bluff, Illinois, USA), perpendicular to the molar occlusal groove. Each section was then labeled and stored separately. Alveolar bone height and thickness at the marker hole sites were measured directly from the sections by two trained raters (R.W. and Z.S.) using the following method. With digital calipers, the raters measured height and thickness on both the mesial and distal halves of each sectioned marker hole. Bone height was measured from the occlusal border of the marker hole to the crest of the alveolar bone. Bone thickness was measured at 0.5 mm and 1.0 mm from the crest of the alveolar bone. The values were recorded to the nearest 0.01 mm. The average of the values from the mesial and distal halves was used for the final height and thickness measurement for each site.

CBCT Measurements

The CBCT images were first placed in random order and relabeled to blind the raters. Two raters (R.W. and G.E.), with one (R.W., an orthodontic resident) generally more experienced in dental radiography than the other (G.E., a dental student), were trained for 2 sessions by an experienced dental radiologist on how to take linear measurements from CBCT images. Each of the raters then independently measured the height of the buccal alveolar bone in each CBCT image, using the following protocol. No calibration between the raters was done before their measurements to simulate clinical situations that clinicians of different levels of previous experience may perform such measurements independently.

The resolution of the computer screen was normalized to 1600 x 1200 pixels. The CBCT image was imported into the software (Dolphin-3D or OsiriX) as a DICOM file and the section thickness was set to 0.5 mm in all three views (coronal, sagittal, and axial). The 3D image was then oriented using the 4 equal layout view (Figure 2) such that the specimen's occlusal plane was parallel to the axial plane in the sagittal and coronal views, and the sagittal plane line bisected the occlusal groove of the molars in the axial view. The coronal view was then chosen to scroll through the sections of the image. The rater could enlarge the image with the zoom function, as needed, to better visualize each site.

The software's linear measurement tool was then used to measure the buccal alveolar bone height at each marker hole, on three consecutive sections, from the occlusal border of the marker hole to the crest of the alveolar bone (Figure 3). The value of each measurement, to the nearest 0.01mm, was recorded on an Excel sheet and the average of the three measurements was used as the final bone height measurement for that site. Orientation was performed for each quadrant prior to measuring the sites in that quadrant, to ensure that the section of the image would be made perpendicular to that quadrant's occlusal groove. This protocol was followed when using both the Dolphin-3D software and the OsiriX software.

CBCT re-measurements to assess intra-rater reliability

Eight CBCT image DICOM files were randomly chosen for re-measurement at least 6 weeks after the initial set of measurements were taken. The same protocol was followed when re-measuring.

Statistical Analysis

Intra- and inter-rater reliability in measuring alveolar bone height from CBCT images was assessed by intra-class correlation tests of data obtained during the first and second sets of measurements.

The differences between each set of CBCT measurements and physical measurements ($D_{CBCT\text{-}Phy}$) were calculated. The influence of rater, software, presence/absence of soft tissue and CBCT voxel size on the mean $D_{CBCT\text{-}Phy}$ was tested by repeated-measures analysis of variance (ANOVA). Measurements from each individual marker hole were tested first; then measurements were pooled for the maxilla (4 marker holes per quadrant) and the mandible (3 marker holes per quadrant) and tested. The mean $D_{CBCT\text{-}Phy}$ and limit of agreement (LOA, ±1.96SD) for each factor were depicted by Bland-Altman methods.

References

17. Sun Z, Smith t, Kortam S, Kim D-Y, Tee B, Fields HW. Effect of bone thickness on alveolar bone-height measurements from cone-beam computed tomography images. Amer. J. Orthod. Dentofac. Orthop. 2011; 139(2):e117-27.

18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007; 39:175-191.

CHAPTER 3

MANUSCRIPT

Factors Affecting Alveolar Bone Height Measurements from CBCT Images

Authors (Tentative):

Ryan Wood[1], Greg England[2], Boon Ching Tee[3], Jahanzeb Chaudhry[4], Binnaz Leblebicioglu[5], Do-Gyoon Kim[6], Zongyang Sun[7]

[1]Resident, Division of Orthodontics, The Ohio State University, College of Dentistry.

[2]Dental student, The Ohio State University, College of Dentistry.

[3]Research assistant, The Ohio State University, College of Dentistry.

[4]Assistant Professor, Division of Primary Care, The Ohio State University, College of Dentistry.

[5]Associate Professor, Division of Periodontology, The Ohio State University, College of Dentistry.

[6]Assistant Professor, Research Program Director, Division of Orthodontics, The Ohio State University, College of Dentistry.

[7]Assistant Professor, Division of Orthodontics, The Ohio State University, College of Dentistry.

ABSTRACT

Objectives: Cone beam computed tomography (CBCT) is routinely used to measure alveolar bone dimensions. This study investigated factors likely affecting the accuracy of such measurements. **Methods:** Marker holes, apical to maxillary and mandibular molar roots, and mesio-distal molar occlusal reference grooves were created in fresh pig heads (n=26 quadrants), followed by CBCT scans at 0.4 and 0.2 mm voxel sizes before and after soft tissue removal. Subsequently, bucco-lingual sections bisecting the marker holes were cut, from which physical alveolar bone height and thickness were measured. Two blinded raters, using Dolphin 3D and OsiriX software, independently collected alveolar bone height measurements from CBCT images. Differences between CBCT and physical measurements ($D_{CBCT\text{-}Phy}$) were calculated. The mean $D_{CBCT\text{-}Phy}$ and limit of agreement (LOA, ±1.96SD) for each factor were depicted by Bland-Altman methods. The influence of each factor was examined by repeated measures ANOVAs. **Results:** The rater factor affected the mean CBCT measurements in both jaws, with mean $D_{CBCT\text{-}Phy}$ from a more-experienced rater significantly ($p<0.05$) closer to "0" than from a less-experienced rater, while their LOA ranges were similar (Maxilla, 0.17mm±2.09mm vs.-0.48mm±2.02mm; Mandible, 0.01mm±1.04mm vs. -0.70mm±1.21mm). The software factor also affected the mean CBCT measurements in both jaws, with mean $D_{CBCT\text{-}Phy}$ from Osirix significantly ($p<0.05$) closer to "0" than from Dolphin-3D despite similar LOA ranges (Maxilla, -0.09mm±2.18mm vs. -0.22mm±2.11mm; Mandible, -0.30mm±1.23mm vs. -0.39mm±1.42mm). The soft-tissue factor only affected the mandible significantly ($p<0.05$), with mean $D_{CBCT\text{-}Phy}$ from soft-tissue-absent specimens closer to "0" than those from soft-tissue-present specimens (Maxilla, -0.13mm±2.29mm vs. -0.19mm±2.00mm; Mandible, -0.25mm±1.32mm vs. -0.44mm±1.32mm). The voxel-size factor had an insignificant effect. Generally, mandibular CBCT measurements exhibited smaller LOA ranges than maxillary, and generally the mandibular alveolar bone was thicker. **Conclusions:** Individual measurements of buccal alveolar bone from CBCT images can under- or overestimate buccal bone height by a clinically significant (>1mm) amount. Factors such as rater, software, and the presence/absence of soft tissue can potentially create systematic sub-millimeter differences between the measurements of buccal alveolar bone from CBCT images and physical measurements.

KEY WORDS: CBCT, linear measurement, alveolar bone, accuracy

INTRODUCTION

Computed Tomography (CT) has been used in the diagnosis and treatment of medical problems since the 1970's. However, the three dimensional radiographic images it produces are high in both resolution and radiation exposure. In an attempt to lower radiation exposure but still maintain sufficiently high resolution, Cone Beam Computed Tomography was invented. By forming the x-ray beam into a cone shape instead of in a flat, fan-shaped beam, as is done in medical CT's, the same three dimensional area can be visualized with approximately one tenth the radiation dose.

This decrease in radiation dosage and the three dimensional visualization of internal structures has caused many in the dental community to consider the usefulness of CBCT in dentistry. As a result, research has been done to evaluate CBCT's ability to accurately visualize carious lesions[1, 2], periodontal structures[3], alveolar bone[4], temporomandibular joints[5], airways[6], and various forms of pathology of the head and neck[7].

Orthodontics, in particular, has seen a large increase in the use of CBCT's to evaluate impacted teeth and root resorption[8], visualize root position and angulation[9], screen for pathology[10], and assess changes in the alveolar bone[11]. The last application, however, is controversial because of the unique nature of alveolar bone. Generally, compared to the basal bone portions of the maxilla and mandible, the alveolar bone is quite thin. It also has a thin cortical plate, which makes it difficult to identify the boundaries of the bone in radiographic images. Much of the alveolar bone surface is also penetrated by gingival fibers, which may further mask the boundary between hard and soft tissues.

In addition to the properties of the alveolar bone itself, several other factors can also potentially affect the accuracy of measuring alveolar bone from CBCT images. Factors such as the rater (the person making the measurements), the software used to view the CBCT images and make measurements on them, the presence/absence of soft tissue at and around the site of interest, and voxel size of the CBCT scan, may affect the accuracy of linear measurements of alveolar bone from CBCT images.

To date, few studies have looked at the effect that these factors may have on the accuracy of linear measurements of alveolar bone from CBCT images. One study assessed the effect of voxel size related to the diagnosis of horizontal root fractures[12] and another assessed its effect on the diagnosis of occlusal caries[2], but neither assessed the effect of voxel size on measurements of alveolar bone.

Similarly, several studies have looked at the accuracy of linear measurements from CBCT images of dry human skulls[13-15] or skulls with soft tissue substitute[14] or skulls with fiduciary markers in place[16]. Only one study was found that assessed the accuracy and reliability of buccal bone height and thickness measurements from CBCT images of specimens with their soft tissue intact and without fiduciary markers[4]. However, that study did not specifically evaluate the effect that the presence of soft tissue has on the accuracy of linear measurements of alveolar bone from CBCT images.

Any effect of the rater on the accuracy of linear measurements of alveolar bone from CBCT images has only been explored indirectly. Intra- and inter-rater reliability tests were found in most studies looking at the accuracy of linear measurements of alveolar bone from CBCT images but these reliability tests may not fully uncover a systematic error which can exist between raters. No studies were found that looked specifically at the rater factor by comparing the results of various raters.

In short, there are many factors that can potentially affect the accuracy of linear measurements taken from CBCT images, such as the rater, the software used, the presence/absence of soft tissue, and the voxel size setting of the CBCT machine. The purpose of this study is to assess the effect that each of these factors has on the accuracy of linear measurements of alveolar bone from CBCT images by comparing them to direct physical measurements. Such assessment may help us better understand CBCT imaging related to both clinical and research applications.

Materials & Methods

Sample Size and Power Analysis

Seven fresh pig heads (age 3-6 months) were used in this study. Our initial power analysis was based on paired t-tests using measurements obtained from a previous studies[17], and a G*Power 3.0 program[18],. With these, we found that an 80% power requires a sample size is 13. To be conservative, a sample size of 14 was determined for this study, which has a projected power of 85%.

Preparation and Imaging

Specimen preparation is shown in Figure 1. Reference marker holes were drilled in the buccal alveolar bone at the apices of the first and second maxillary molars (four holes per maxillary quadrant) and at the apices of the first mandibular molar (three holes per mandibular quadrant) using a slow speed dental handpiece (Volvere VMax, Brasseler USA, Savannah, Georgia) with a flat-end straight fissure cross-cut carbide bur (H31-21-009, Brasseler USA, Savannah, Georgia). Marker holes were made intraorally through the buccal gingiva, leaving all soft tissue of the head intact. Reference grooves bisecting the occlusal tables of all molars in a mesio-distal direction were also created with the same handpiece and bur type.

After the reference marker holes and occlusal grooves were made, each head was scanned with an iCAT CBCT machine (120 kVp, 5mA) at 0.4 mm voxel size and 0.2 mm voxel size resolutions. Subsequently, all soft tissue, including gingiva, was removed via manual dissection, followed by re-scanning at 0.4 mm voxel size and 0.2 mm voxel size. This resulted in a total of 28 CBCT image files (104 usable quadrants because one of the heads did not receive mandibular marker holes prior to imaging).

Sectioning and Physical Measurements

Each maxillary and mandibular quadrant was then separated from the head using a circular saw. Each quadrant was sectioned through each marker hole, using an IsoMet Low Speed Saw (Beuhler, Lake Bluff, Illinois, USA), perpendicular to the molar occlusal groove. Each section was then labeled and stored separately. Alveolar bone height and thickness at the marker hole sites were measured directly from the sections by two trained raters (R.W. and Z.S.) using the following method. With digital calipers, the raters measured height and thickness on both the mesial and distal halves of each sectioned marker hole. Bone height was measured from the occlusal border of the marker hole to the crest of the alveolar bone. Bone thickness was measured at 0.5 mm and 1.0 mm from the crest of the alveolar bone. The values were recorded to the nearest 0.01 mm. The average of the values from the mesial and distal halves was used for the final height and thickness measurement for each site.

CBCT Measurements

The CBCT images were first placed in random order and relabeled to blind the raters. Two raters (R.W. and G.E.) were trained by an experienced dental radiologist on how to take linear measurements from CBCT images. Each of the raters then independently measured the height of the buccal alveolar bone in each CBCT image, using the following protocol.

The resolution of the computer screen was normalized to 1600 x 1200 pixels. The CBCT image was imported into the software (Dolphin-3D or OsiriX) as a DICOM file and the section thickness was set to 0.5 mm in all three views (coronal, sagittal, and axial). The 3D image was then oriented using the 4 equal layout view (Figure 2) such that the specimen's occlusal plane was parallel to the axial plane in the sagittal and coronal views, and the sagittal plane line bisected the occlusal groove of the molars in the axial view. The coronal view was then chosen to scroll through the sections of the image. The rater could enlarge the image with the zoom function, as needed, to better visualize each site.

The software's linear measurement tool was then used to measure the buccal alveolar bone height at each marker hole, on three consecutive sections, from the occlusal border of the marker hole to the crest of the alveolar bone (Figure 3). The value of each measurement, to the nearest 0.01mm, was recorded on an Excel sheet and the average of the three measurements was used as the final bone height measurement for that site. Orientation was performed for each quadrant prior to measuring the sites in that quadrant, to ensure that the section of the image would be made perpendicular to that quadrant's occlusal groove. This protocol was followed when using both the Dolphin-3D software and the OsiriX software.

CBCT re-measurements to assess intra-rater reliability

Eight CBCT image DICOM files were randomly chosen for re-measurement at least 6 weeks after the initial set of measurements were taken. The same protocol was followed when re-measuring.

Statistical Analysis

Intra- and inter-rater reliability was assessed by intra-class correlation tests.

The differences between each set of CBCT measurements and physical measurements ($D_{CBCT\text{-}Phy}$) were calculated. The influence of rater, software, presence/absence of soft tissue and CBCT voxel size on $D_{CBCT\text{-}Phy}$ was tested by repeated-measures analysis of variance (ANOVA). Measurements from each individual marker hole were tested first; then measurements were pooled for the maxilla (4 marker holes per quadrant) and the mandible (3 marker holes per quadrant) and tested. The mean $D_{CBCT\text{-}Phy}$ and limit of agreement (LOA, ±1.96SD) for each factor were depicted by Bland-Altman methods.

Results

Both intra- and inter-rater reliability were found to be excellent. Intra-rater reliability values were all ≥ 0.957, except one that was 0.867 (Figure 4). Inter-rater reliability values were all between 0.875-0.973, except one that was 0.768 (Figure 5).

The rater factor affected the mean $D_{CBCT\text{-}Phy}$ in both jaws, with mean $D_{CBCT\text{-}Phy}$ from a more-experienced rater significantly ($p<0.05$) closer to "0" than from a less-experienced rater, while the LOA ranges from both raters were similar (Maxilla, 0.17mm±2.09mm vs.-0.48mm±2.02mm; Mandible, 0.01mm±1.04mm vs. -0.70mm±1.21mm). (Figure 6).

The software factor affected the mean $D_{CBCT\text{-}Phy}$ in both jaws, with the mean $D_{CBCT\text{-}Phy}$ from OsiriX significantly ($p<0.05$) closer to "0" than that from Dolphin 3D, while the LOA ranges of both software programs were similar (Maxilla, -0.09mm±2.18mm vs. -0.22mm±2.11mm; Mandible, -0.30mm±1.23mm vs. -0.39mm±1.42mm). (Figure 7)

The soft-tissue factor only affected the mean $D_{CBCT\text{-}Phy}$ of the mandible significantly ($p<0.05$), with mean $D_{CBCT\text{-}Phy}$ from soft-tissue-absent specimens closer to "0" than those from soft-tissue-present specimens (Maxilla, -0.13mm±2.29mm vs. -0.19mm±2.00mm; Mandible, -0.25mm±1.32mm vs. -0.44mm±1.32mm). The LOA ranges from specimens with or without soft tissue were similar (Figure 8)

The voxel-size factor had an insignificant effect on the mean $D_{CBCT\text{-}Phy}$ and the LOA ranges from both voxel levels were similar compared to other factors (Figure 9).

Summarized mean $D_{CBCT\text{-}Phy}$ and LOA of each factor for the mandible is shown in Table 1. As mentioned above, the rater, software and soft tissue factors all had significant effect on the mean $D_{CBCT\text{-}Phy}$ from the mandible. In addition, compared to the maxilla, generally, mandibular CBCT measurements exhibited smaller LOA ranges. Overall, the mandibular alveolar bone was also found to be significantly thicker than the maxillary alveolar bone (Figure 10).

Discussion

Methodology

The methods of specimen preparation, imaging, and data collection used were to ensure that the results of the study were as consistent and reliable as possible. Fresh pig heads were used to simulate clinical patients, who will also have all their soft tissue intact at the time of imaging. In a further effort to make the results of our study as clinically relevant as possible we chose not to place any radiopaque markers for measurement. Instead, marker holes were placed in the buccal alveolar bone apical to molar roots prior to imaging to provide a reference point that was clear and consistent, but as natural as possible, from which to measure alveolar bone height both directly and in CBCT images. Grooves were placed in the occlusal surfaces of the molar teeth to assist in accurate and consistent orientation both when sectioning the specimens and when measuring the height of the alveolar bone. This helped ensure consistent site selection when taking measurements, physically and in the CBCT images.

Soft tissue was removed by blunt physical dissection only. Other methods of soft tissue removal (chemical treatment, boiling, etc.) were not used because of concern that these methods might inadvertently change the mineralization, morphology, or dimensions of the bone to be measured, thereby rendering the pre- and post-soft tissue removal specimens incomparable.

Several more measures were taken to minimize common systemic biases in addition to the factors tested in this study. All CBCT images (DICOM files) were listed in random order and were given random labels to blind the raters. Computer screen resolution was set to the standard screen resolution of 1600 x 1200 pixels to ensure that the screen resolution would be consistent between raters and to make the results as generalizable as possible. Raters were trained at the same time by a dental radiologist. The bone height and thickness at each site was determined by taking the average of multiple measurements at that site.

Factors affecting CBCT measurements

The rater factor showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for both jaws. Specifically, overall the orthodontic resident had a significantly higher mean $D_{CBCT\text{-}Phy}$ than the dental student (0.17mm vs. -0.48mm for the maxilla; 0.01mm vs. -0.70mm for the mandible), suggesting that different raters may indeed produce different alveolar bone height measurements from CBCT images. First, this result contradicts the results of excellent inter-rater reliability tested by intra-class correlation (Figure 5), suggesting that intra-class correlations may not be sensitive enough to reflect the actual differences between raters. Second, it is important to understand what the differences mean. Basically, the differences of the mean $D_{CBCT\text{-}Phy}$ between the two raters suggest that the dental student tended to consistently underestimate the height of the alveolar bone, a problem not shared by the orthodontic resident. In this study, both raters were trained by the same radiologist at the same time, but did differ in experience of viewing radiographs, including CBCTs. We cannot ascertain that the difference was indeed caused by a varied level of experience between residents and dental students rather than individual variation in habits or perception of images, but reasonably this problem can be solved by doing a calibration process in addition to standard training sessions. A calibration process has not been recommended by most similar studies involving multiple raters due to generally excellent inter-rater reliability tested by intra-class correlations. Based on our findings, such a calibration seems necessary for future studies aiming at gaining mean $D_{CBCT\text{-}Phy}$ values close to "0" from CBCT images.

Third, it is also important to realize that a significant difference in gaining a mean $D_{CBCT\text{-}Phy}$ between the raters does not indicate one rater is less variable around the physical truth than the other. In fact, judging from the similar LOA ranges obtained by the two raters, one can easily say that for any individual measurement, both raters can deviate up to 2 mm from physical truth. Taking 1 mm deviation as clinically acceptable, this finding of similar LOA ranges suggests that both raters in fact can obtain similar unacceptable measurements. We think this mostly speaks to the nature of CBCT images rather than a problem associated with the raters. Specifically, for a particular CBCT image, even an experienced rater is likely to obtain a rather inaccurate measure of the alveolar bone. Does an experienced rater have a higher probability of making clinically acceptable measurements (within 1 mm of physical truth) than an inexperienced rater? This is a plausible argument and we are currently investigating our data in that aspect.

The software factor also showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for both jaws. In this case the OsiriX software had a mean $D_{CBCT\text{-}Phy}$ that was significantly closer to the physical mean than the Dolphin 3D software (-0.09 mm vs. -0.22 mm for the maxilla; -0.30 mm vs. -0.39 mm for the mandible). This might lead one to believe that OsiriX software is more accurate than Dolphin 3D software at making linear measurements of alveolar bone from CBCT images but it should be noted that the deviation of the software means from the physical mean were all <0.40 mm; an amount that most clinicians would consider clinically insignificant. So while there was a statistically significant difference between the means of the two software programs tested, it should not be forgotten that this difference was quite small and, therefore, of questionable clinical relevance. In addition, the LOA of both software programs were very similar and well above 1.0 mm for individual measurements (Table 1) suggesting that, while the mean $D_{CBCT\text{-}Phy}$ for each software was <0.40 mm from the physical mean, it is possible that any individual measurement could be inaccurate by 1.0 mm or more, regardless of which software is used. This implies that one software program is not clearly superior to the other from a clinical standpoint.

Then why was a difference noted in the accuracy of the linear measurements taken with each software program? The difference in mean $D_{CBCT\text{-}Phy}$ between the two softwares could be due to the method that each of the software programs uses to assign gray values to the voxels at the bone-soft tissue interface. If the method is different, it could result in a different apparent border of the alveolar bone and, thereby, different linear measurement values from each software. This is a factor that may affect the accuracy of measuring alveolar bone from CBCT images and we are planning to explore this factor in future studies.

The presence of soft tissue showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for the mandible only, with the soft tissue absent specimens having a significantly lower mean $D_{CBCT\text{-}Phy}$ (-0.25 mm vs. -0.44 mm). This may suggest that the presence of soft tissue on CBCT specimens makes measuring their alveolar bone less accurate. But that suggestion has been contradicted by another recent study that looked at the accuracy of linear measurements of alveolar bone in specimens that had their soft tissue intact. In that study, Timock et al[4] used embalmed cadaver heads with soft tissue intact to evaluate the accuracy and reliability of measuring the height and thickness of buccal alveolar bone from CBCT images. They measured alveolar bone height and thickness directly and from CBCT images and compared the results. They reported no significant difference between direct measurements and those taken from CBCT images and no pattern for over- or under-estimation of measurement values. These findings support the conclusion that the presence of soft tissue has no significant effect on the accuracy of linear measurements of alveolar bone from CBCT images.

So then why was there a difference found only in the mandible between the soft tissue present and absent specimens in our study and no difference found between the specimens in the Timock et al study? The answer may lie in the difference in thickness of the buccal alveolar bone being measured. In our specimens, the maxillary buccal bone was, on average, thinner than the mandibular buccal bone (Figure 10). The thin maxillary buccal bone could have acted as a confounder, with a possible effect of soft tissue presence/absence in the maxilla being masked by the difficulty of identifying the apex of the thin maxillary buccal bone. This theory is supported by the fact that the difference in means of the soft tissue present and the soft tissue absent maxillary segments is even greater than that for the mandibular segments, despite a non-significant statistical result (Table 1). Though Timock et al did not report the buccal bone thickness values of their specimens it is possible that the buccal alveolar bone of their specimens, who were all adult humans, was thicker than the buccal alveolar bone of our specimens, who were all 3-6 month old pigs. If that is true, it would help explain why they found no statistically significant difference in mean $D_{CBCT\text{-}Phy}$ for their specimens while we did find a difference.

This is further supported by the findings of Sun et al[17], who showed that alveolar bone became less visible in CBCT images as the thickness was reduced. Also, it should be noted that, although a statistically significant difference was found between the mean $D_{CBCT\text{-}Phy}$ for soft tissue present and absent specimens in the mandible, the mean $D_{CBCT\text{-}Phy}$ were all <0.45 mm while the LOA ranges were all >1.3 mm (Table 1). Such small differences in mean values were very likely not meaningful from a clinical perspective. However, when individual measurements are considered, soft tissue presence or absence can have approximately the same amount of inaccuracy as was found with other factors.

The voxel size factor had an insignificant effect on the accuracy of the linear measurements. This finding was surprising to the raters because it seemed that some of the CBCT images had noticeably greater clarity and sharpness than other images and, as a result, the crest of the alveolar bone seemed much easier to find in those images. It was thought that the improved clarity and sharpness was due to a decreased voxel size of the CBCT image but since the raters were blinded while making measurements, it was impossible to confirm at the time.

The insignificant effect of voxel size was an interesting finding because other studies have shown that voxel size may affect the accuracy of linear measurements and the diagnostic quality of CBCT images. As mentioned previously, Sun et al[17] found that measurement inaccuracies of alveolar bone height were substantially improved when they decreased the voxel size of the CBCT scan from 0.40 mm to 0.25 mm.

Likewise, a study by Wenzel et al[12] compared the abilities of photostimulable phosphor (PSP) plates with those of CBCT at 0.125mm and 0.25mm voxel size in the diagnosis of transverse root fractures. Voxel size was found to have a significant effect, with the smaller voxel size (higher resolution) improving the rater's ability to diagnose transverse root fractures.

Despite the fact that the voxel size factor had no statistically significant effect in this study, the same trends are noted in the mean $D_{CBCT\text{-}Phy}$ and LOA data as are noted with other factors (Table 1); namely, mean $D_{CBCT\text{-}Phy}$ values <0.40 mm and LOA values >1.0 mm, suggesting that the mean values are clinically insignificant while any individual measurements may be inaccurate by a clinically significant amount of 1.0 mm or more.

Though it was not a formal factor in the study it was observed that, generally, mandibular CBCT measurements exhibited smaller LOA ranges than maxillary measurements. This indicates that the raters were consistently able to measure the height of the buccal alveolar bone of the mandible more accurately than that of the maxilla. As mentioned previously, a possible explanation for this is that the buccal alveolar bone of the mandible was, on average, thicker (Figure 10). This could have made it easier to find the crest of the alveolar bone in the mandible and, thereby, to measure it with greater accuracy and consistency. The Timock et al[4] study pooled their maxilla and mandible measurement data together so it was not possible to see if the same affect was noted in their study but, as previously mentioned, they did use specimens from adult humans, which may have thicker alveolar bone in general than our specimens collected from growing pigs. Also, as previously noted, Sun et al[17] found in a separate study that alveolar bone can become less visible when the thickness was reduced, which provides indirect evidence for this speculation. To ascertain it, however, more studies involving both adult and adolescent specimens with varied bone thickness are needed.

Limitations of this Study

It should be noted, however, that these artificial occlusal grooves and marker holes are not present in clinical patients. Such ideal set up may overrate the measurement accuracy that can be expected from CBCT images clinically. The amount of factors included in this study is also large, which may involve unwanted confounding factors and complicated interactions among factors.

Future Studies & Approaches

Based on the results of this study, future studies will try to find a method to improve the accuracy of measuring alveolar bone, especially in the maxilla, with CBCT images taken at a clinically standard resolution (0.4mm). Tentatively, this new method will be using a gray level change to detect the borders of the alveolar bone. It is hoped that this method will enable clinicians and researchers to identify the borders of the alveolar bone with greater accuracy and consistency. Other studies may assess the accuracy of using the cusp tip or incisal edge as the reference point when measuring alveolar bone height. Future studies should also involve calibration of the raters, as described above, to ensure the accuracy of their measurements as much as possible.

Conclusions

1) Factors such as rater, software, and the presence of soft tissue can potentially create systematic sub-millimeter differences between the measurements of buccal alveolar bone from CBCT images and physical measurements.
2) Individual measurements of buccal alveolar bone from CBCT images can under- and overestimate buccal bone height by a clinically significant amount.

<u>References Cited</u>

1. Tsuchida R, Araki K, Okano T. Evaluation of a limited cone-beam volumetric imaging system: comparison with film radiography in detecting incipient

proximal caries. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007; 104:412-6.

2. Kamburoglu K, Murat S, Yuksel SP, Cebeci AR, Paksoy CS. Occlusal caries detection by using a cone-beam CT with different voxel resolutions and a digital intraoral sensor. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2010; 109:e63-9.

3. Mohan R, Singh A, Gundappa M. Three-dimensional imaging in periodontal diagnosis - Utilization of cone beam computed tomography. J Indian Soc Periodontol 2011; 15:11-7.

4. Timock AM, Cook V, McDonald T, Leo MC, Crowe J, Benninger BL, Covell DA, Jr. Accuracy and reliability of buccal bone height and thickness measurements from cone-beam computed tomography imaging. Am J Orthod Dentofacial Orthop 2011; 140:734-44.

5. Honda K, Arai Y, Kashima M, Takano Y, Sawada K, Ejima K, Iwai K. Evaluation of the usefulness of the limited cone-beam CT (3DX) in the assessment of the thickness of the roof of the glenoid fossa of the temporomandibular joint. Dentomaxillofac Radiol 2004; 33:391-5.

6. Stratemann S, Huang JC, Maki K, Hatcher D, Miller AJ. Three-dimensional analysis of the airway with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 140:607-15.

7. Alexiou K, Stamatakis H, Tsiklakis K. Evaluation of the severity of temporomandibular joint osteoarthritic changes related to age using cone beam computed tomography. Dentomaxillofac Radiol 2009; 38:141-7.

8. Oberoi S, Knueppel S. Three-dimensional assessment of impacted canines and root resorption using cone beam computed tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;

9. Bouwens DG, Cevidanes L, Ludlow JB, Phillips C. Comparison of mesiodistal root angulation with posttreatment panoramic radiographs and cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 139:126-32.

10. Caglayan F, Tozoglu U. Incidental findings in the maxillofacial region detected by cone beam CT. Diagn Interv Radiol 2011;

11. Corbridge JK, Campbell PM, Taylor R, Ceen RF, Buschang PH. Transverse dentoalveolar changes after slow maxillary expansion. Am J Orthod Dentofacial Orthop 2011; 140:317-25.

12. Wenzel A, Haiter-Neto F, Frydenberg M, Kirkevang LL. Variable-resolution cone-beam computerized tomography with enhancement filtration compared with intraoral photostimulable phosphor radiography in detection of transverse root fractures in an in vitro model. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009; 108:939-45.

13. Veyre-Goulet S, Fortin T, Thierry A. Accuracy of linear measurement provided by cone beam computed tomography to assess bone quantity in the posterior maxilla: a human cadaver study. Clin Implant Dent Relat Res 2008; 10:226-30.

14. Suomalainen A, Vehmas T, Kortesniemi M, Robinson S, Peltola J. Accuracy of linear measurements using dental cone beam and conventional multislice computed tomography. Dentomaxillofac Radiol 2008; 37:10-7.

15. Lascala CA, Panella J, Marques MM. Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom). Dentomaxillofac Radiol 2004; 33:291-4.

16. Ganguly R, Ruprecht A, Vincent S, Hellstein J, Timmons S, Qian F. Accuracy of linear measurement in the Galileos cone beam computed tomography under simulated clinical conditions. Dentomaxillofac Radiol 2011; 40:299-305.

17. Sun Z, Smith t, Kortam S, Kim D-Y, Tee B, Fields HW. Effect of bone thickness on alveolar bone-height measurements from cone-beam computed tomography images. Amer. J. Orthod. Dentofac. Orthop. 2011; 139(2):e117-27.

18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007; 39:175-191.

CHAPTER 4

DISCUSSION AND CONCLUSIONS

Methodology

The methods of specimen preparation, imaging, and data collection used were to ensure that the results of the study were as consistent and reliable as possible. Fresh pig heads were used to simulate clinical patients, who will also have all their soft tissue intact at the time of imaging. In a further effort to make the results of our study as clinically relevant as possible we chose not to place any radiopaque markers for measurement. Instead, marker holes were placed in the buccal alveolar bone apical to molar roots prior to imaging to provide a reference point that was clear and consistent, but as natural as possible, from which to measure alveolar bone height both directly and in CBCT images. Grooves were placed in the occlusal surfaces of the molar teeth to assist in accurate and consistent orientation both when sectioning the specimens and when measuring the height of the alveolar bone. This helped ensure consistent site selection when taking measurements, physically and in the CBCT images.

Soft tissue was removed by blunt physical dissection only. Other methods of soft tissue removal (chemical treatment, boiling, etc.) were not used because of concern that these methods might inadvertently change the mineralization, morphology, or dimensions of the bone to be measured, thereby rendering the pre- and post-soft tissue removal specimens incomparable.

Several more measures were taken to minimize common systemic biases in addition to the factors tested in this study. All CBCT images (DICOM files) were listed in random order and were given random labels to blind the raters. Computer screen resolution was set to the standard screen resolution of 1600 x 1200 pixels to ensure that the screen resolution would be consistent between raters and to make the results as generalizable as possible. Raters were trained at the same time by a dental radiologist. The bone height and thickness at each site was determined by taking the average of multiple measurements at that site.

Factors affecting CBCT measurements

The rater factor showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for both jaws. Specifically, overall the orthodontic resident had a significantly higher mean $D_{CBCT\text{-}Phy}$ than the dental student (0.17mm vs. -0.48mm for the maxilla; 0.01mm vs. -0.70mm for the mandible), suggesting that different raters may indeed produce different alveolar bone height measurements from CBCT images. First, this result contradicts the results of excellent inter-rater reliability tested by intra-class correlation (Figure 5), suggesting that intra-class correlations may not be sensitive enough to reflect the actual differences between raters. Second, it is important to understand what the differences mean. Basically, the differences of the mean $D_{CBCT\text{-}Phy}$ between the two raters suggest that the dental student tended to consistently underestimate the height of the alveolar bone, a problem not shared by the orthodontic resident. In this study, both raters were trained by the same radiologist at the same time, but did differ in experience of viewing radiographs, including CBCTs. We cannot ascertain that the difference was indeed caused by a varied level of experience between residents and dental students rather than individual variation in habits or perception of images, but reasonably this problem can be solved by doing a calibration process in addition to standard training sessions. A calibration process has not been recommended by most similar studies involving multiple raters due to generally excellent inter-rater reliability tested by intra-class correlations. Based on our findings, such a calibration seems necessary for future studies aiming at gaining mean $D_{CBCT\text{-}Phy}$ values close to "0" from CBCT images.

Third, it is also important to realize that a significant difference in gaining a mean $D_{CBCT\text{-}Phy}$ between the raters does not indicate one rater is less variable around the physical truth than the other. In fact, judging from the similar LOA ranges obtained by the two raters, one can easily say that for any individual measurement, both raters can deviate up to 2 mm from physical truth. Taking 1 mm deviation as clinically acceptable, this finding of similar LOA ranges suggests that both raters in fact can obtain similar unacceptable measurements. We think this mostly speaks to the nature of CBCT images rather than a problem associated with the raters. Specifically, for a particular CBCT image, even an experienced rater is likely to obtain a rather inaccurate measure of the alveolar bone. Does an experienced rater have a higher probability of making clinically acceptable measurements (within 1 mm of physical truth) than an inexperienced rater? This is a plausible argument and we are currently investigating our data in that aspect.

The software factor also showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for both jaws. In this case the OsiriX software had a mean $D_{CBCT\text{-}Phy}$ that was significantly closer to the physical mean than the Dolphin 3D software (-0.09 mm vs. -0.22 mm for the maxilla; -0.30 mm vs. -0.39 mm for the mandible). This might lead one to believe that OsiriX software is more accurate than Dolphin 3D software at making linear measurements of alveolar bone from CBCT images but it should be noted that the deviation of the software means from the physical mean were all <0.40 mm; an amount that most clinicians would consider clinically insignificant. So while there was a statistically significant difference between the means of the two software programs tested, it should not be forgotten that this difference was quite small and, therefore, of questionable clinical relevance. In addition, the LOA of both software programs were very similar and well above 1.0 mm for individual measurements (Table 1) suggesting that, while the mean $D_{CBCT\text{-}Phy}$ for each software was <0.40 mm from the physical mean, it is possible that any individual measurement could be inaccurate by 1.0 mm or more, regardless of which software is used. This implies that one software program is not clearly superior to the other from a clinical standpoint.

Then why was a difference noted in the accuracy of the linear measurements taken with each software program? The difference in mean $D_{CBCT\text{-}Phy}$ between the two softwares could be due to the method that each of the software programs uses to assign gray values to the voxels at the bone-soft tissue interface. If the method is different, it could result in a different apparent border of the alveolar bone and, thereby, different linear measurement values from each software. This is a factor that may affect the accuracy of measuring alveolar bone from CBCT images and we are planning to explore this factor in future studies.

The presence of soft tissue showed a significant impact on the mean $D_{CBCT\text{-}Phy}$ for the mandible only, with the soft tissue absent specimens having a significantly lower mean $D_{CBCT\text{-}Phy}$ (-0.25 mm vs. -0.44 mm). This may suggest that the presence of soft tissue on CBCT specimens makes measuring their alveolar bone less accurate. But that suggestion has been contradicted by another recent study that looked at the accuracy of linear measurements of alveolar bone in specimens that had their soft tissue intact. In that study, Timock et al[4] used embalmed cadaver heads with soft tissue intact to evaluate the accuracy and reliability of measuring the height and thickness of buccal alveolar bone from CBCT images. They measured alveolar bone height and thickness directly and from CBCT images and compared the results. They reported no significant difference between direct measurements and those taken from CBCT images and no pattern for over- or under-estimation of measurement values. These findings support the conclusion that the presence of soft tissue has no significant effect on the accuracy of linear measurements of alveolar bone from CBCT images.

So then why was there a difference found only in the mandible between the soft tissue present and absent specimens in our study and no difference found between the specimens in the Timock et al study? The answer may lie in the difference in thickness of the buccal alveolar bone being measured. In our specimens, the maxillary buccal bone was, on average, thinner than the mandibular buccal bone (Figure 10). The thin maxillary buccal bone could have acted as a confounder, with a possible effect of soft tissue presence/absence in the maxilla being masked by the difficulty of identifying the apex of the thin maxillary buccal bone. This theory is supported by the fact that the difference in means of the soft tissue present and the soft tissue absent maxillary segments is even greater than that for the mandibular segments, despite a non-significant statistical result (Table 1). Though Timock et al did not report the buccal bone thickness values of their specimens it is possible that the buccal alveolar bone of their specimens, who were all adult humans, was thicker than the buccal alveolar bone of our specimens, who were all 3-6 month old pigs. If that is true, it would help explain why they found no statistically significant difference in mean $D_{CBCT\text{-}Phy}$ for their specimens while we did find a difference.

This is further supported by the findings of Sun et al[17], who showed that alveolar bone became less visible in CBCT images as the thickness was reduced. Also, it should be noted that, although a statistically significant difference was found between the mean $D_{CBCT\text{-}Phy}$ for soft tissue present and absent specimens in the mandible, the mean $D_{CBCT\text{-}Phy}$ were all <0.45 mm while the LOA ranges were all >1.3 mm (Table 1). Such small differences in mean values were very likely not meaningful from a clinical perspective. However, when individual measurements are considered, soft tissue presence or absence can have approximately the same amount of inaccuracy as was found with other factors.

The voxel size factor had an insignificant effect on the accuracy of the linear measurements. This finding was surprising to the raters because it seemed that some of the CBCT images had noticeably greater clarity and sharpness than other images and, as a result, the crest of the alveolar bone seemed much easier to find in those images. It was thought that the improved clarity and sharpness was due to a decreased voxel size of the CBCT image but since the raters were blinded while making measurements, it was impossible to confirm at the time.

The insignificant effect of voxel size was an interesting finding because other studies have shown that voxel size may affect the accuracy of linear measurements and the diagnostic quality of CBCT images. As mentioned previously, Sun et al[17] found that measurement inaccuracies of alveolar bone height were substantially improved when they decreased the voxel size of the CBCT scan from 0.40 mm to 0.25 mm.

Likewise, a study by Wenzel et al[12] compared the abilities of photostimulable phosphor (PSP) plates with those of CBCT at 0.125mm and 0.25mm voxel size in the diagnosis of transverse root fractures. Voxel size was found to have a significant effect, with the smaller voxel size (higher resolution) improving the rater's ability to diagnose transverse root fractures.

Despite the fact that the voxel size factor had no statistically significant effect in this study, the same trends are noted in the mean $D_{CBCT\text{-}Phy}$ and LOA data as are noted with other factors (Table 1); namely, mean $D_{CBCT\text{-}Phy}$ values <0.40 mm and LOA values >1.0 mm, suggesting that the mean values are clinically insignificant while any individual measurements may be inaccurate by a clinically significant amount of 1.0 mm or more.

Though it was not a formal factor in the study it was observed that, generally, mandibular CBCT measurements exhibited smaller LOA ranges than maxillary measurements. This indicates that the raters were consistently able to measure the height of the buccal alveolar bone of the mandible more accurately than that of the maxilla. As mentioned previously, a possible explanation for this is that the buccal alveolar bone of the mandible was, on average, thicker (Figure 10). This could have made it easier to find the crest of the alveolar bone in the mandible and, thereby, to measure it with greater accuracy and consistency. The Timock et al[4] study pooled their maxilla and mandible measurement data together so it was not possible to see if the same affect was noted in their study but, as previously mentioned, they did use specimens from adult humans, which may have thicker alveolar bone in general than our specimens collected from growing pigs. Also, as previously noted, Sun et al[17] found in a separate study that alveolar bone can become less visible when the thickness was reduced, which provides indirect evidence for this speculation. To ascertain it, however, more studies involving both adult and adolescent specimens with varied bone thickness are needed.

Limitations of this Study

It should be noted, however, that these artificial occlusal grooves and marker holes are not present in clinical patients. Such ideal set up may overrate the measurement accuracy that can be expected from CBCT images clinically. The amount of factors included in this study is also large, which may involve unwanted confounding factors and complicated interactions among factors.

Future Studies & Approaches

Based on the results of this study, future studies will try to find a method to improve the accuracy of measuring alveolar bone, especially in the maxilla, with CBCT images taken at a clinically standard resolution (0.4mm). Tentatively, this new method will be using a gray level change to detect the borders of the alveolar bone. It is hoped that this method will enable clinicians and researchers to identify the borders of the alveolar bone with greater accuracy and consistency. Other studies may assess the accuracy of using the cusp tip or incisal edge as the reference point when measuring alveolar bone height. Future studies should also involve calibration of the raters, as described above, to ensure the accuracy of their measurements as much as possible.

Conclusions

1) Factors such as rater, software, and the presence of soft tissue can potentially create systematic sub-millimeter differences between the measurements of buccal alveolar bone from CBCT images and physical measurements.
2) Individual measurements of buccal alveolar bone from CBCT images can under- and overestimate buccal bone height by a clinically significant amount.

LIST OF REFERENCES

1. Tsuchida R, Araki K, Okano T. Evaluation of a limited cone-beam volumetric imaging system: comparison with film radiography in detecting incipient proximal caries. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007; 104:412-6.

2. Kamburoglu K, Murat S, Yuksel SP, Cebeci AR, Paksoy CS. Occlusal caries detection by using a cone-beam CT with different voxel resolutions and a digital intraoral sensor. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2010; 109:e63-9.

3. Mohan R, Singh A, Gundappa M. Three-dimensional imaging in periodontal diagnosis - Utilization of cone beam computed tomography. J Indian Soc Periodontol 2011; 15:11-7.

4. Timock AM, Cook V, McDonald T, Leo MC, Crowe J, Benninger BL, Covell DA, Jr. Accuracy and reliability of buccal bone height and thickness measurements from cone-beam computed tomography imaging. Am J Orthod Dentofacial Orthop 2011; 140:734-44.

5. Honda K, Arai Y, Kashima M, Takano Y, Sawada K, Ejima K, Iwai K. Evaluation of the usefulness of the limited cone-beam CT (3DX) in the

assessment of the thickness of the roof of the glenoid fossa of the temporomandibular joint. Dentomaxillofac Radiol 2004; 33:391-5.

6. Stratemann S, Huang JC, Maki K, Hatcher D, Miller AJ. Three-dimensional analysis of the airway with cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 140:607-15.

7. Alexiou K, Stamatakis H, Tsiklakis K. Evaluation of the severity of temporomandibular joint osteoarthritic changes related to age using cone beam computed tomography. Dentomaxillofac Radiol 2009; 38:141-7.

8. Oberoi S, Knueppel S. Three-dimensional assessment of impacted canines and root resorption using cone beam computed tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2011;

9. Bouwens DG, Cevidanes L, Ludlow JB, Phillips C. Comparison of mesiodistal root angulation with posttreatment panoramic radiographs and cone-beam computed tomography. Am J Orthod Dentofacial Orthop 2011; 139:126-32.

10. Caglayan F, Tozoglu U. Incidental findings in the maxillofacial region detected by cone beam CT. Diagn Interv Radiol 2011;

11. Corbridge JK, Campbell PM, Taylor R, Ceen RF, Buschang PH. Transverse dentoalveolar changes after slow maxillary expansion. Am J Orthod Dentofacial Orthop 2011; 140:317-25.

12. Wenzel A, Haiter-Neto F, Frydenberg M, Kirkevang LL. Variable-resolution cone-beam computerized tomography with enhancement filtration compared with intraoral photostimulable phosphor radiography in detection of transverse root fractures in an in vitro model. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009; 108:939-45.

13. Veyre-Goulet S, Fortin T, Thierry A. Accuracy of linear measurement provided by cone beam computed tomography to assess bone quantity in the posterior maxilla: a human cadaver study. Clin Implant Dent Relat Res 2008; 10:226-30.

14. Suomalainen A, Vehmas T, Kortesniemi M, Robinson S, Peltola J. Accuracy of linear measurements using dental cone beam and conventional multislice computed tomography. Dentomaxillofac Radiol 2008; 37:10-7.

15. Lascala CA, Panella J, Marques MM. Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom). Dentomaxillofac Radiol 2004; 33:291-4.

16. Ganguly R, Ruprecht A, Vincent S, Hellstein J, Timmons S, Qian F. Accuracy of linear measurement in the Galileos cone beam computed tomography under simulated clinical conditions. Dentomaxillofac Radiol 2011; 40:299-305.

17. Sun Z, Smith t, Kortam S, Kim D-Y, Tee B, Fields HW. Effect of bone thickness on alveolar bone-height measurements from cone-beam computed tomography images. Amer. J. Orthod. Dentofac. Orthop. 2011; 139(2):e117-27.

18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007; 39:175-191.

APPENDIX A: FIGURES & TABLES

Figure 1: CBCT image showing marker holes adjacent to molar roots (4 holes in the maxilla and 3 holes in the mandible)

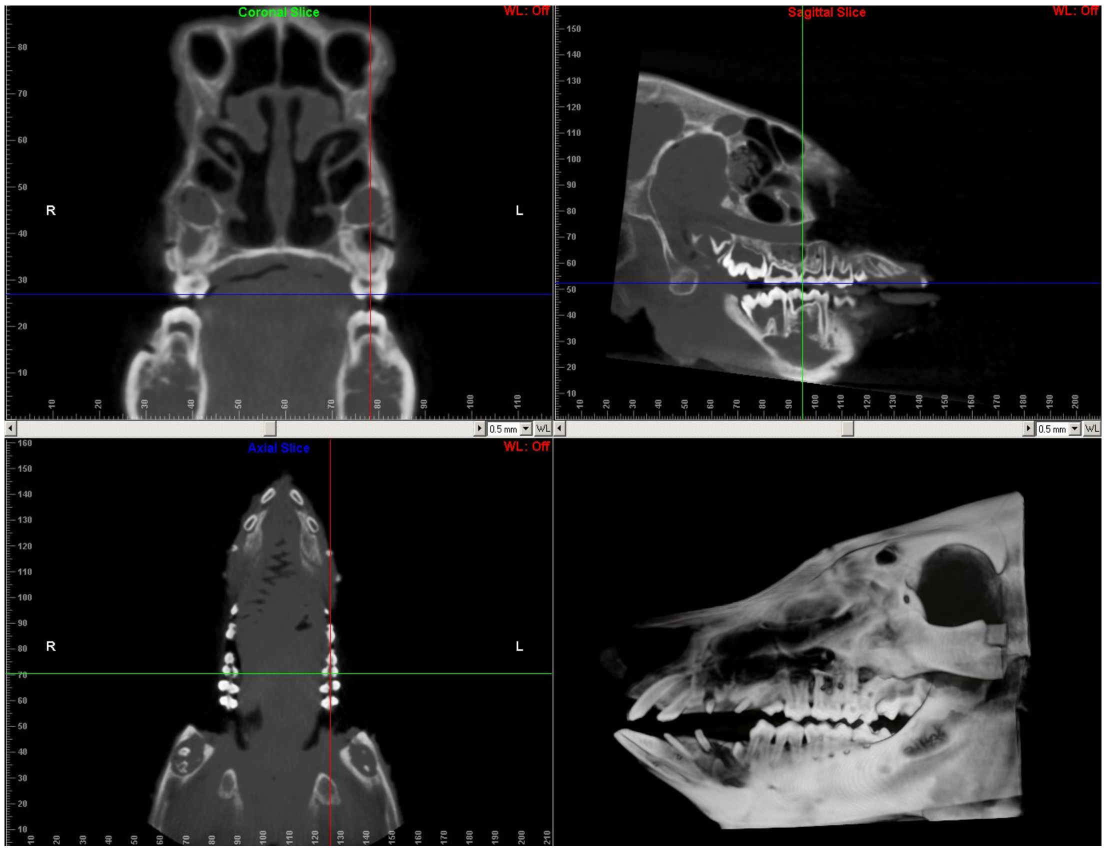

Figure 2: Four equal layout view of CBCT image, showing specimen orientation prior to measuring buccal bone height in the coronal view.

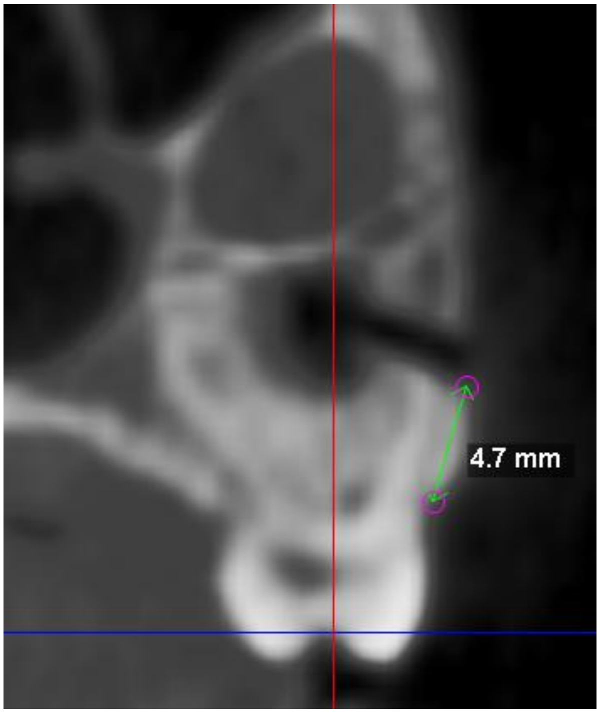

Figure 3: Linear measurement of alveolar bone height.

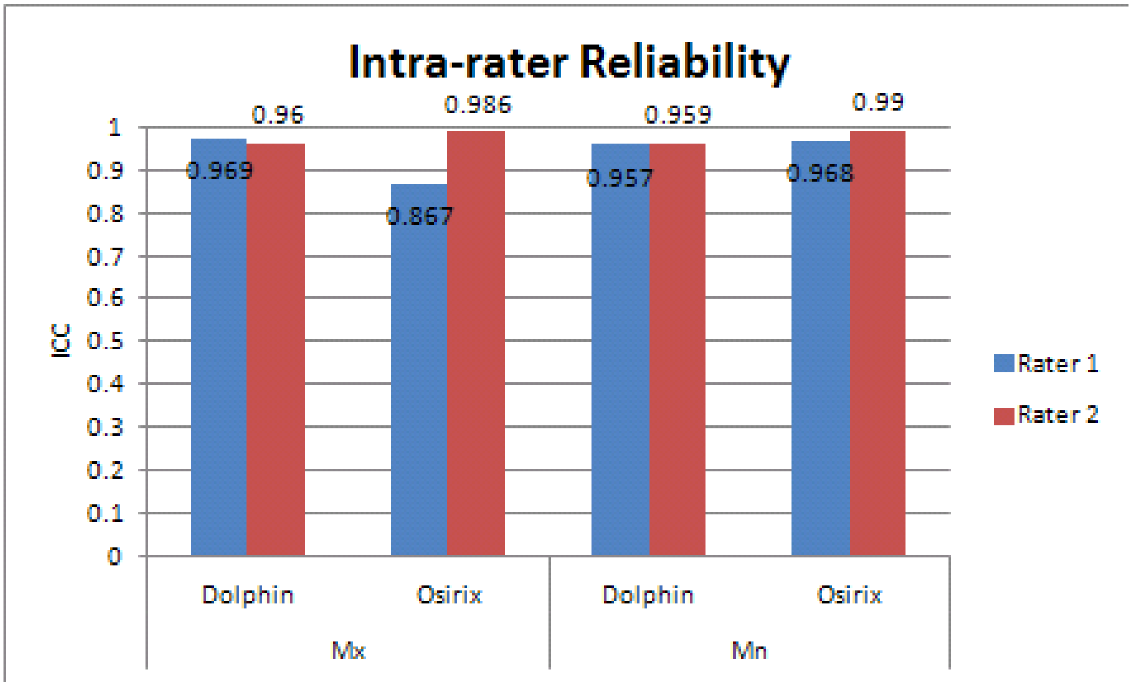

Figure 4: Intra-rater reliability was excellent for both raters.

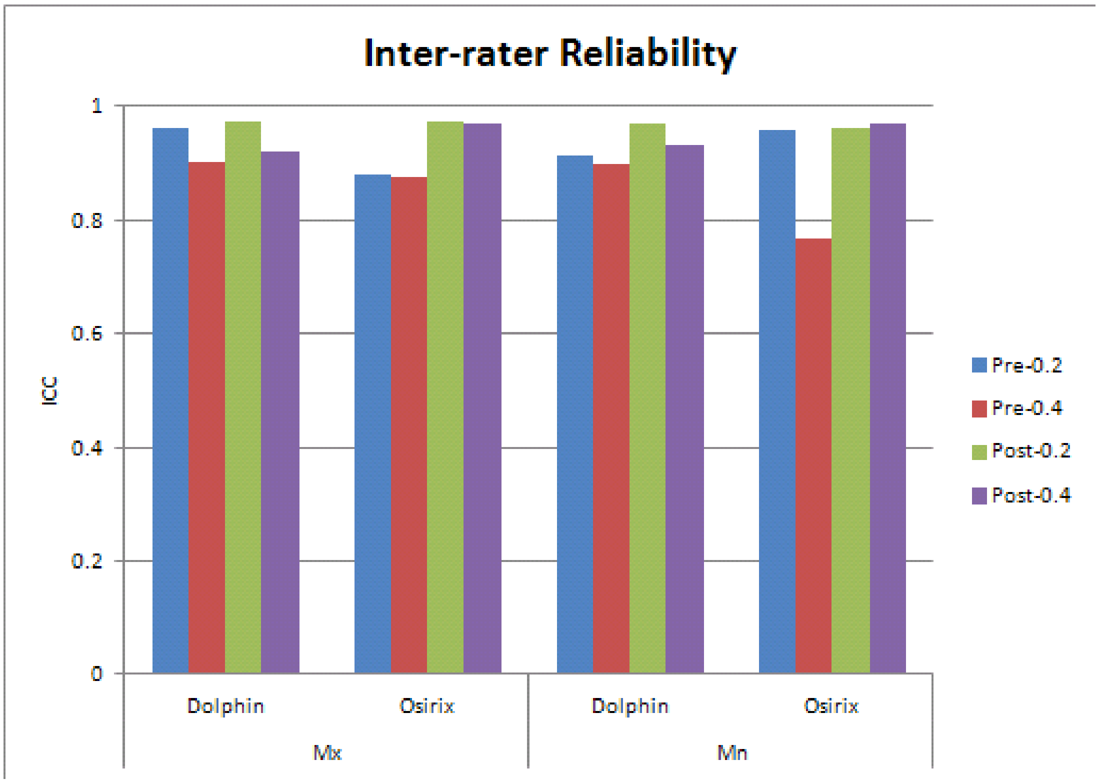

Figure 5: Inter-rater reliability values were all between 0.875-0.973, except one that was 0.768.

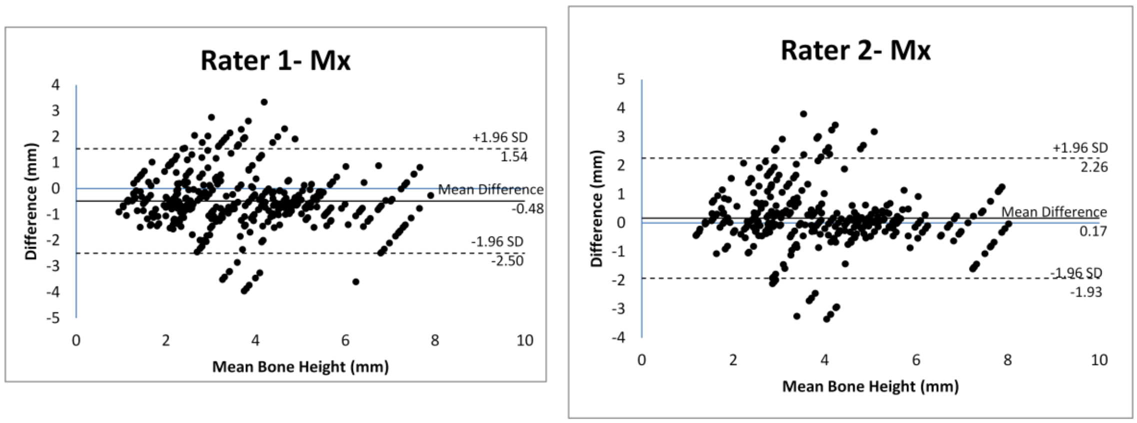

Figure 6: Bland-Altman Plots assessing the Rater factor at sites of the maxilla. Each dot represents the difference in millimeters between the rater's measurement of a site from the CBCT image and the physical measurement of that same site.

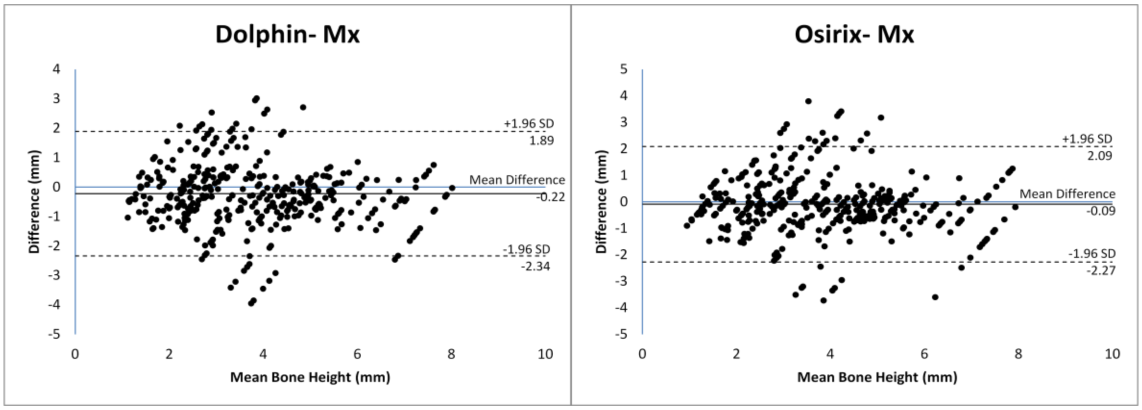

Figure 7: Bland-Altman Plots assessing the Software factor at sites of the maxilla. Each dot represents the difference in millimeters between the measurement of a site from the CBCT image with a given software and the physical measurement of that same site.

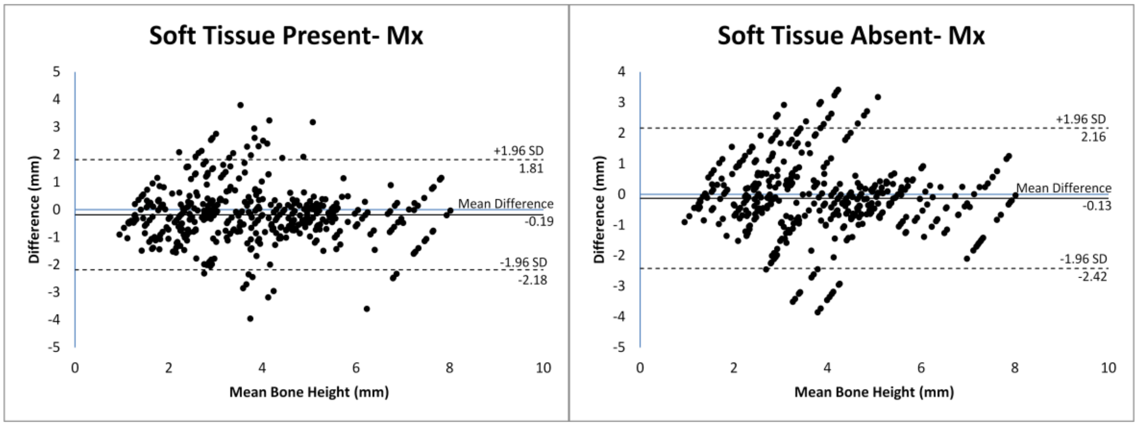

Figure 8: Bland-Altman Plots assessing the Soft Tissue factor at sites of the maxilla. Each dot represents the difference in millimeters between the measurement of a site from the CBCT image with/without soft tissue and the physical measurement of that same site.

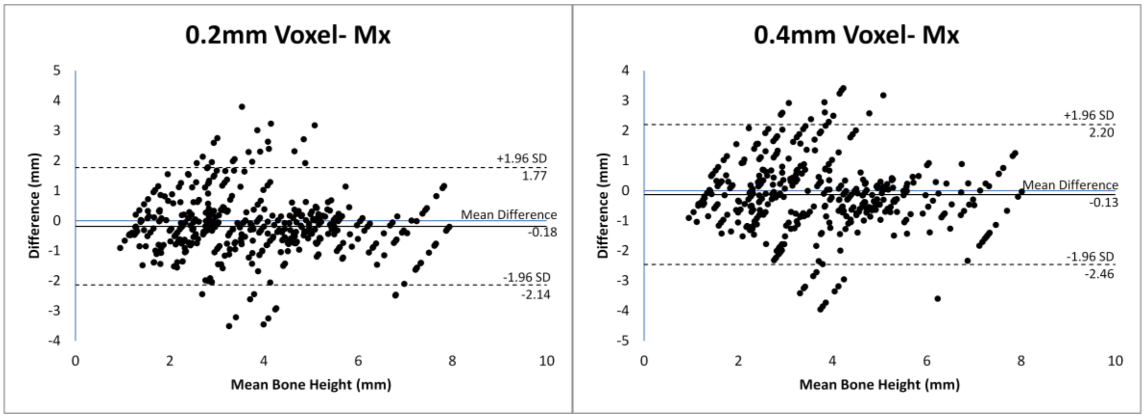

Figure 9: Bland-Altman Plots assessing the Voxel Size factor at sites of the maxilla. Each dot represents the difference in millimeters between the measurement of a site from the CBCT image with a given voxel size and the physical measurement of that same site.

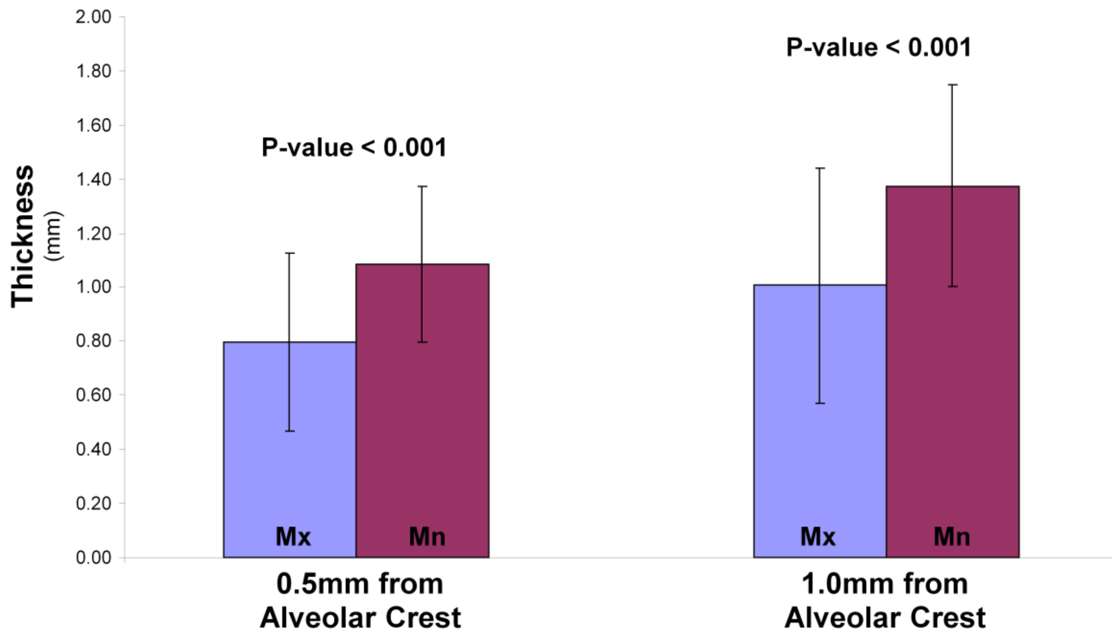

Figure 10: The buccal alveolar bone of the mandible was significantly thicker than that of the maxilla.

Possible Factor	Maxilla Mean ± LOA	Mandible Mean ± LOA	Diff. in LOA Mx - Mn
Rater			
Rater #1	-0.48mm ± 2.02mm	-0.70mm ± 1.21mm	0.81mm
Rater #2	0.17mm ± 2.09mm	0.01mm ± 1.04mm	1.05mm
Software			
Dolphin	-0.22mm ± 2.11mm	-0.39mm ± 1.42mm	0.69mm
OsiriX	-0.09mm ± 2.18mm	-0.30mm ± 1.23mm	0.95mm
Soft Tissue			
Present	-0.19mm ± 2.00mm	-0.44mm ± 1.32mm	0.68mm
Absent	-0.13mm ± 2.29mm	-0.25mm ± 1.32mm	0.97mm
Resolution			
0.20mm voxel	-0.18mm ± 1.95mm	-0.31mm ± 1.20mm	0.75mm
0.40mm voxel	-0.13mm ± 2.33mm	-0.38mm ± 1.45mm	0.88mm

Table 1: The Limits of Agreement (LOA)- the boundaries which contain 96.8% of the plotted points- were consistently smaller in the mandible, for all factors considered. This suggests that the raters were consistently able to measure the buccal alveolar bone of the mandible with greater accuracy.