Comparing Cone Beam Computed Tomography with Multi-Slice Computed Tomography in Diagnosing Osseous Defects at the Mandibular Condyle

THESIS

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

By

Elizabeth Matteson Jones

Graduate Program in Dentistry

The Ohio State University

2015

Master's Examination Committee:

Dr. Zongyang Sun, Advisor

Dr. Henry Fields

Dr. F. Michael Beck
Abstract

**Objectives:** Previous studies indicated that orthodontic-grade CBCTs are limited in displaying small defects at the mandibular condyles for diagnosis. This study investigated whether this limitation was inherent to CBCT by comparing CBCT with multislice CT, and whether it can be overcome by image segmentation and color mapping.

**Methods:** Nine fresh pig heads (18 condyles, 36 medial/lateral condylar regions) were used. Small osseous defects (diameter and depth =1.5mm) were created at condylar, medial and lateral, regions demarcated by gutta percha markers. After restoring the overlying soft tissues, the pig heads underwent orthodontic-grade CBCT scans (i-CAT, 0.4mm voxel-size) and medical-grade CT scans (GE LightSpeed, 0.625mm voxel-size) scans. Subsequently, two calibrated and blinded raters diagnosed defect number in each condylar region from CBCT and CT images using Dolphin-3D software without image segmentation, then 1-week later with proprietary image segmentation and color mapping tools of Dolphin-3D. Condylar PVS impressions were collected and evaluated by the same raters to obtain physical diagnosis. Re-diagnosis was made on randomly selected subsamples to assess reliability. Using physical diagnosis as references, the accuracy of imaging diagnosis were assessed and statistically compared among varied imaging/analysis methods.

**Results:** Image diagnoses of all imaging/analysis methods showed good or excellent intra- and inter-rater reliability except for those of segmented CBCT images, which were substantially lower. The numbers of over- and under-diagnosis per condylar region were...
not significantly differences among varied imaging/analysis methods (Wilcoxon tests, p>0.05) but those from CT tended to be lower than CBCT. Classification functions demonstrate substantial lower sensitivity and accuracy with CBCT than with CT. Logistic regression showed CT had a significantly higher probability (odds ratio 2.4) than CBCT in reaching correct diagnosis, while the use of image segmentation and color mapping did not improve the diagnostic accuracy from CBCT images.

**Conclusions:** Even at a lower voxel-size than medical CT, orthodontic-grade CBCT images of mandibular condyles are inherently more difficult than CT images for diagnosing small condylar defects, and this limitation cannot be overcome by using image segmentation and color mapping.
Dedication

Dedicated to my wonderful husband, Richard, and my parents, Nancy and Silas.

Thank you for your unwavering love and support.
Acknowledgments

I would like to thank the following:

- My thesis committee, Drs. Zongyang Sun, Henry Fields and Michael Beck for their feedback and guidance during this project.
- Dr. Melissa Papio for her diligent work on the image analysis.
- Boon Ching Tee for all of her help with specimen collection, data compilation and statistical analysis.
- The faculty, residents and staff at The Ohio State University, College of Dentistry, Section of Orthodontics for giving me an invaluable education.
- My family for their unwavering love, support and guidance.
- The Delta Dental Foundation for the financial support for this research project.
Vita

November 27, 1984..................................Born – Arlington, VA

2007..........................................................B.S. Biology, The College of William and Mary, Williamsburg, VA

2011..........................................................D.D.S., Virginia Commonwealth University, Richmond, VA

2012-2015 ..................................................Graduate Resident, Section of Orthodontics, The Ohio State University, Columbus, OH

Publications


Fields of Study

Major Field: Dentistry

Specialty: Orthodontics
# Table of Contents

Abstract ................................................................................................................................. ii

Dedication ............................................................................................................................. iv

Acknowledgments ................................................................................................................ v

Vita ........................................................................................................................................ vi

Publications ............................................................................................................................ vi

Fields of Study ....................................................................................................................... vi

Table of Contents ................................................................................................................... vii

List of Tables ......................................................................................................................... viii

List of Figures ......................................................................................................................... ix

Chapter 1: INTRODUCTION .................................................................................................. 1

Chapter 2: MATERIALS AND METHODS ............................................................................. 11

CHAPTER 3: MANUSCRIPT ................................................................................................. 20

CHAPTER 4: RESULTS ........................................................................................................... 37

CHAPTER 5: DISCUSSION AND CONCLUSIONS ................................................................. 46

Comprehensive References ................................................................................................. 53
List of Tables

Table 1. Classification functions and their derivation ........................................... 197
Table 2. Calculation of over- and under-diagnosis rate ........................................ 28
Table 3. Inter-rater Reliability .............................................................................. 38
Table 4. Intra-Rater Reliability ............................................................................ 38
Table 5. Logistic Regression analysis for correct diagnosis ................................. 54
List of Figures

Figure 1. Views of a condyle in the correct orientation for diagnosis of defects. ............ 14
Figure 2. Views of a condyle with and without global thresholding. ..................... 15
Figure 3. PVS impressions of sample 5, right and left condyles. ......................... 17
Figure 4. Between-method comparisons for over-diagnosis (OD) and under-diagnosis (UD) for non-segmented CBCT and CT images .................................................. 40
Figure 5. Between-method comparisons for over-diagnosis (OD) and under-diagnosis (UD) for non-segmented CBCT and segmented CBCT images ......................... 41
Figure 6. Sensitivity, Specificity & Accuracy for diagnosis of small osseous condylar defects using CBCT and CT images without segmentation ................................... 43
Figure 7. Sensitivity, Specificity & Accuracy for diagnosis of small osseous condylar defects using non-segmented and segmented CBCT images ............................ 44
Chapter 1: INTRODUCTION

Temporomandibular disorders (TMD) are the most common cause of non-dental pain in the orofacial region\(^1\). TMD is a collective term that describes a group of musculoskeletal conditions that occur in the temporomandibular joint (TMJ) that are characterized by pain in the muscles of mastication, the temporomandibular joint, or both\(^2,3\). The most common signs of and symptoms of TMD are joint pain, muscle pain, reduced range of motion/opening, clicking and crepitation\(^4\).

Pain in the temporomandibular region is relatively common, occurring in approximately 10% of the population over 18. It primarily affects young and middle-aged adults, rather than children or the elderly and is twice as common in women as men\(^3\). Proposed risk factors for TMJ pain include life stress, depression and presence of multiple somatic symptoms\(^3\). Low prevalence rates in children suggest etiologic factors of TMD do not exist until after puberty or that there may be a period of exposure necessary before pain develops. Decreasing rates in prevalence after middle age indicate that persons experiencing pain earlier in life cease to experience symptoms at some point possibly due to adaptive remodeling of tissues\(^3\). Higher prevalence rates for adult women indicate biologic, behavioral and psychologic/social factors associated with the female gender may increase the experience of pain in the TMJ\(^3\).
Osteoarthritis (OA) of the TMJ is an age-related disorder characterized by destruction of the articular surfaces of the mandibular condyle and the glenoid fossa through increased loading of the joint\(^5\). TMJ OA is characterized by gradual progressive destruction of articular tissues leading to flattening, erosion, osteophytes, and sclerosis of the subchondral cortical layer\(^6,7\). Several theories of OA etiology have been suggested; most are based on articular cartilage failure, while others base the cause on extra-cartilage factors as the primary cause. Regardless, there is a general agreement that the articular cartilage is involved in osteoarthritic changes\(^8\). Diagnosis of OA of the TMJ requires the presence of pain on palpation of the joint, as well as a report of ongoing pain in the TMJ or pain in response to certain range of motion tests, and coarse crepitus\(^3\).

Structural changes to the TMJ are important characteristics in establishing a diagnosis of OA\(^7\). A radiographic exam is a routine component of the clinical assessment for conditions of TMJ dysfunction, where the main goal is to verify degenerative bone changes in the joint structures\(^9\). 42.6\% of TMD patients presented with tomographic evidence of TMJ OA changes including bone erosions and osteophytes\(^7\). Campos et al, found that degenerative changes are significantly more frequent in the condyle than the articular surface\(^10\).

Subchondral bone appears to be involved in the destructive osteoarthritic process rather early. Osseous changes start focally and are not radiographically detectable, but as the degenerative process progresses, loss of the subchondral cortical layer and bone erosions
occur creating radiographic evidence of OA. The presence of erosions indicates that the TMJ is unstable and further changes in the bony surface will occur. Because of this gradual degenerative process, patients may experience symptoms for months before bony changes can be visualized on radiographs. Radiographic changes such as flattening, osteophytes, cystic formation and decreased articular space typically appear in the later stages of OA. Thus by the time TMJ OA is perceived clinically and/or radiographically, the degeneration is at an advanced stage.

Osseous changes of the TMJ have been radiographically observed in 14-44% of TMD patients using two-dimensional imaging, including panoramic radiographs, cranial projections, and tomograms. These radiographic techniques, however, are designed to avoid anatomic superimposition by using projections that are usually oblique to the condylar axis resulting in a less accurate image of the articular surface and adjacent joint space. The inherent limitation of panoramic radiographs in demonstrating the contours of the condyle and articular fossa is likely the reason for poor agreement in interpreting such images for osteophytes and erosions of the condyle. However, these two-dimensional imaging modalities continue to be used in orthodontic practices for radiographic assessment of the TMJ because of their availability, low radiation requirement and low cost. Studies have shown that a combination of radiographic modalities and views produce a more reliable diagnosis of arthritic lesions in the TMJ. Unfortunately, using multiple modalities and views only increases the radiation load applied to the patient, making the burden of treatment high. Currently, no general
consensus has been reached as to which diagnostic imaging technique should be the gold standard in detecting these osseous lesions in the TMJ\textsuperscript{19}.

The application of conventional CT in imaging the TMJ has been the most significant for the evaluation of hard tissue or bony changes on the joint\textsuperscript{19}. Superiority of CT over panoramic or magnetic resonance imaging (MRI) in displaying features of the TMJ OA has been well-documented and is widely accepted\textsuperscript{16}. Honey found CT to be superior to plain films and MRI with 87.5-96% accuracy in detecting degenerative arthritis\textsuperscript{20}. Similarly, Ahmad found that when OA was detected on CT, 26% of panoramic images and 59% of MR images displayed positive findings of OA. Therefore, approximately 75% of CT-diagnosed OA was not detected using panoramic images and about 40% were not detected using MRI. Therefore he recommends future clinical studies should use CT when possible despite the increased radiation\textsuperscript{20}. Furthermore, Hussain found that CT had higher sensitivity, higher specificity, higher positive predictive values and higher negative predictive values than MRI\textsuperscript{11}. Finally, Cevdances reported that CT has the best positive percent agreement (84.4%) for diagnosis of TMJ OA\textsuperscript{21}. Utumi et al evaluated CT images of simulated osseous lesions in the condyle to determine the validity of images acquired using MSCT. They found that sensitivity and specificity were influenced by the size of the drill hole, rather than the acquisition method of the image\textsuperscript{22}.

CT works by producing blurring of the surrounding and/or adjacent structures, thus separating a well-defined section of the TMJ from potentially obscuring adjacent
anatomy. Tomography that uses a complex motion produces views of the TMJ with improved contrast and definition. Multislice CT (MSCT) represents a potential advance in CT imaging that allows acquisition of thinner slices with high image quality in a shorter amount of time. Thus multiple overlapping slices can be reconstructed from a single exam, permitting higher quality reconstructed images without additional patient irradiation. Because multidirectional tomography provides fine detail through thin sections, it is considered by many to be the most precise method for radiographic examination of bony changes in the TMJ. Indeed, Cara et al found that several CT imaging protocols for 1mm x 0.5mm simulated osseous defects at the mandibular condyle were accurate, however the association of axial with multiplanar reconstructed images from MSCT showed the highest accuracy. Despite high accuracy and image quality, limitations exist for routine use of CT in orthodontics, including high financial and radiation dose costs, as well as access to these large, expensive systems. Thus although historically CT, as well as arthrography and MRI, has been the method of choice for evaluating the cortical contours, joint dynamics and disk position in the TMJ, they are expensive and require high dose radiation, making them unsuitable for long-term monitoring of bony changes in the TMJ condyle.

In the past decade, cone beam CT (CBCT) has been developed as an alternative to conventional CT. CBCT technology results in images of similar quality to that of CT, but are obtained through less expensive equipment, with shorter patient exposure time and significantly lower radiation dose than required for conventional CT. Namely by
providing submillimeter spatial resolution images with markedly shorter scanning times (ranging from 10-70 seconds) and lower required radiation dosages CBCT provides two major advantages over CT imaging\textsuperscript{20}.

CBCT scanners are based on volumetric tomography, a principle that uses a two-dimensional extended detector and a three-dimension x-ray beam. Combined this allows for a single rotation around the patient’s head to produce a scan of the entire region of interest; thus reducing acquisition time for volumetric data\textsuperscript{20}. CBCT uses a cone-shaped x-ray beam rather than a collimated fan-beam that is used with conventional CT. The primary images captured can be used for further secondary reconstruction in all planes and for three-dimensional reconstruction\textsuperscript{19}. Additionally, Zain-Alabdeen et al. found in their study that CBCT images were very reader friendly and easy to use. Because of the flexibility of CBCT software, the operator could more easily define osseous changes when interpreting the images\textsuperscript{19}.

Several authors have reported high dimensional accuracy of the maxillofacial structures, including the TMJ, with CBCT images\textsuperscript{19}. Indeed, Pontual reported minimal hard tissue distortion of maxillofacial bones on three-dimensional CBCT images\textsuperscript{25}. Furthermore, Honey et al. found high detectability of CBCT images for bony morphological features of the mandibular condyle when compared with conventional tomography and helical CT\textsuperscript{20}. Similarly, Tsiklakis found that CBCT was superior to CT for visualizing boney changes in the TMJ when analyzing lateral slices in isolation or through combining coronal and
lateral slices\textsuperscript{26}. Through several studies, it has been shown that the diagnostic capabilities of CBCT are equal to or greater than those of MSCT. Therefore, Honda determined that due to the decreased cost and radiation dose, CBCT is a viable diagnostic alternative for detecting erosions in the TMJ\textsuperscript{27}.

Thus the reasonable goals of TMJ imaging with CBCT are to evaluate the integrity of the bony structures when disorders are suspected, to confirm the extent and stage of progression of disorders, and to evaluate the effects of treatment\textsuperscript{1}. To assess whether CBCT can be used to fulfill this goal, our group recently examined the ability of CBCT in diagnosing small osseous defects at the mandibular condyle using fresh pig specimens\textsuperscript{28}. We found extremely small defects (1-2mm) to be difficult to detect, with 1 out of 3 defects likely missed from diagnosis if 0.4mm voxel size CBCT scans are used. While this problem was less significant (sensitivity improved to 90\%) when higher CBCT scan resolution (0.2mm-voxel size instead of 0.4mm-voxel size) was used, higher radiation associated with higher scan resolution discourages its use for clinical patients.

These data raised two questions. One is whether CBCT has the ability of detecting 1-2mm condylar defects comparable to conventional MSCT? A number of studies have attempted to address this question recently, but these studies have either omitted the soft tissue factor or focused on relatively large defects (5-10mm). More specifically, with the soft tissue removed, CBCT may be superior to helical CT in displaying hard tissue in the maxillofacial region\textsuperscript{29,30}. Honey et al demonstrated high detectability of CBCT images
for bony morphological features compared with conventional CT, but their sample included large defects of 5-10mm on dry human skulls with latex balloons filled with water to provide soft tissue attenuation\textsuperscript{20}. Hintze et al found no significant differences in diagnostic accuracy between the two modalities detecting osseous changes in the condyle, but they included defects of undefined size on dry skulls without any soft tissue simulation\textsuperscript{24}. Honda et al also found no significant difference between CBCT and helical CT in detecting erosions and osteophytes, but used embalmed cadaver specimens with undefined size of condylar defects\textsuperscript{24,27}. Controversy over the presence/absence of soft-tissue in CBCT scans exists in the literature: some groups suggest the soft-tissue reduces the precision of CBCT measurement, while others report no difference. It has been speculated that without soft-tissue, tissue contrast may be increased and scatter may be reduced, although these speculations have not been proven\textsuperscript{31}.

The current diagnosis protocol has been based on visual observation of the CT slices, a method that is subjective and insensitive to very small defects which only contain a few CT voxels. Diagnosis of density changes in all common extraoral radiographic techniques is based on darkness and brightness of images, expressed with Hounsfield Units (HU) in CT scans and with gray-scale in CBCT scans\textsuperscript{32}. HUs are standard numbers originating from CT imaging and represent the relative density of body tissues according to a calibrated gray-level scale based on values for air, water and bone density\textsuperscript{33}. Like HU in CT, gray-scale (or voxel value) represents the degree of x-ray attenuation in CBCT. Although CBCT manufacturers and software providers present gray-scales as the
HUs, they are not true HUs. Furthermore, gray-scales are not the same among devices because they are influenced by noise levels, scattered radiation, high heel effect and beam hardening artifacts\textsuperscript{32}. Whether a linear relationship exists between gray-scale values and HUs is controversial. Razi et al found a linear relationship between HUs derived from both CBCT and CT images; however Silva found that the average HU value obtained via CBCT was overestimated by 33.51\% when compared with the MSCT values\textsuperscript{32,33}. Therefore, simply comparing voxel units is not consistent or necessarily representative of the density (or lack thereof) of the condylar bone.

This inconsistency in x-ray attenuation units leads to the second question of whether CBCT and conventional CT diagnosis can benefit from global thresholding. Global thresholding where bone is segmented in both CBCT and MSCT images using a single threshold to segment the whole object everywhere in the image. The threshold value is defined for the CBCT and MSCT images individually by a histogram analysis based on the region rather than the gradient through an estimation of intensity probability density functions over the image\textsuperscript{34,35}. Global thresholding can be chosen by the operator and determined from the physical experiments\textsuperscript{34}. It may be possible that it can be used in this experiment to help objectively distinguish between the CT and CBCT images for defects among sound bone.

Specific Aims:
1) Compare the diagnostic ability of CBCT and MSCT for detecting small
(<2mm) mandibular condyle defects using visual observation.

2) Compare the diagnostic ability of CBCT in detecting small (<2mm)
mandibular condyle defects using visual observation and global thresholding.

Hypotheses:

1) Based on visual observation, the diagnostic ability of detecting condylar osseous
defects is not statistically different between CBCT images and those of MSCT.

2) The diagnostic ability of condylar osseous defects from CBCT images based on
global thresholding is better than that based on visual observation.
Chapter 2: MATERIALS AND METHODS

Study Model

Sample size calculation was based on the main hypothesis that visual diagnosis of condylar defects (presence/absence) on CBCT and CT images will be less accurate than the physical diagnosis. These diagnoses when compared with the physical truth generated Receiver Operating Characteristic (ROC) curves. Sample size calculation is based on comparison of the areas under the curve (AUC). Based on our previous study\textsuperscript{28}, we hypothesized that the visual observation will have an AUC of 0.65. By using global thresholding we hypothesized that the AUC can be improved to 0.85. With a plan to create positive:negative defect sites at a 1:1 ratio and multiple sites in each condyle, we determined that 18 condyles (9 pigs) were required to achieve a power above 80\% for detecting a 0.2 difference in AUC (MedCalc for Windows, version 12.7.0).

The target population of this research was the portion of the human population afflicted with diseases of the TMJ. Due to radiation safety and limited ability to verify osseous defects, living human subjects were not chosen for this project. Human cadaver mandibles were not selected for two reasons: the first because they have limited availability; the second because soft tissue would not be available for proper simulation. Fresh un-embalmed pig cadavers were selected because their TMJ and masticatory system have similar size, anatomy and bone metabolism as compared with those of
humans. Additionally both CBCT and CT scans could be completed with normal soft
tissue attachment, creating a more realistic image for analysis.

Experimental materials, study design and sample size determination

Nine fresh un-embalmed pig cadaver heads (4-6 months old) were used in this study.
Small holes (diameter and depth <1.5mm) were created in the condylar surfaces to
simulate osseous defects. Each condyle was divided into medial and lateral regions, with
0-4 defects created in each region for a total of 144 experimental sites.

Sample Preparation

As in our previous study, following euthanasia nine pig cadaver heads were collected
from The Ohio State University Animal Laboratory. Each condyle was dissected to
expose the condylar head, taking care to preserve the soft tissue attachment, followed by
demarcating the condyle into medial and lateral regions by small gutta percha pieces
super-glued to the medial, lateral, and posterior sections of the condyle. Overall, a total
of 36 medial or lateral condylar regions were resulted. One operator (ZS) created 0-4
small osseous defects (diameter and depth <1.5mm) in each half with a dental handpiece
(Brasseler USA®, NSK, Volvere, Vmax) using an end-cutting 1.5mm round bur
(Brasseler USA®,1.5mm). The number of defects in each region of each animal was
predetermined using a random table.
CBCT Analysis

Once the defects were created and the soft tissue was repositioned, each specimen underwent two scans: an orthodontic-grade CBCT scan (120kV, 5mA i-CAT, Imaging Science International, Hatfield, PA) at a clinically typical 0.4mm voxel-size; a medical-grade CT scan (GE LightSpeed 8-multidetector helical CT scanner, Buckinghamshire, United Kingdom) at a clinically typical 0.625mm voxel-size. Two independent, calibrated raters (EJ, MP) who were blinded of defect information and true defect number in each condylar region analyzed the CBCT and CT scans for diagnoses using Dolphin 3-D software (Patterson Supply, Inc.). Both raters received training on using the software including proper orientation of the images and advancing image slices in all three planes. Prior to the images being analyzed for diagnoses, the independent raters were calibrated. Each rater independently analyzed three randomly selected condyles. Their qualitative analyses of these condyles were compared to ensure proper defect identification and image orientation.

For each condyle, the condylar neck and posterior ramus were oriented parallel to the floor in the sagittal view with the coronal line bisecting the posterior gutta percha marker (Figure 1a). In the coronal view, the condyle was further oriented to allow the axial line to bisect the medial and lateral gutta percha markers and the sagittal line to bisect the posterior gutta percha (Figure 1b). Finally in the axial view, the coronal line bisected the medial and lateral gutta percha markers and the sagittal line bisected the posterior gutta percha marker (Figure 1c). Once the images were correctly oriented, the raters evaluated
each slice (0.5mm thickness) in all three views for the presence/absence of defects from the mandibular ramus to the superior aspect of the condyle. When defects were identified, the raters noted the region of the condyle in which the defect was present. The total number of defects in each region was then recorded.

Figure 1. A, Sagittal, B, Coronal, and C, Axial views of a condyle in the correct orientation for diagnosis of defects.
To test whether global thresholding improves diagnostic ability, one week after the CBCT and CT scans were analyzed, the same two independent, blinded and calibrated raters analyzed each scan with the help of proprietary image segmentation designed by Dolphin 3-D. The image orientation protocol was the same as that used for the non-segmentation method described above (Figure 2). The raters viewed all slices (0.5mm thickness) in all three views and detected defects in each condylar region.

Figure 2. A, Sagittal view with global thresholding, B, Sagittal view without global thresholding; C, Coronal view with global thresholding, D, Coronal view without global thresholding, and E, Axial view with global thresholding, F, Axial view without global thresholding.
To test intra-rater reliability, four weeks after the initial image evaluation by both raters, eight image sets were randomly selected for repeated evaluation following the same protocols described above.

**Physical Diagnosis**

Once the scans were obtained, the soft tissue was completely removed from the condyles and the condyles were carefully sectioned from the posterior ramus. Light-bodied polyvinyl siloxane (PVS) impressions were made of each condylar head with care taken to include the gutta percha markers for region identification (Figure 3). The halves were further delineated and labeled with an indelible marker. These PVS impressions served as the gold standard for this study. Our two blinded, independent raters evaluated each impression for the presence/absence of defects in each region and recorded their findings.
Figure 3. PVS impressions of sample 5, right and left condyles. Indelible markers were used to delineate medial and lateral regions.

Statistical Analysis

Diagnostic reliability was assessed by repeating the evaluation of 8 randomly selected image sets in both the standard view and with image segmentation. Inter-rater and intra-rater reliability were assessed by weighted Kappa tests and further qualified by 95% confidence intervals.

The diagnosis ability of CBCT and CT images were then analyzed statistically in two ways. First, CBCT and CT diagnoses were compared with physical diagnoses from the PVS impressions to derive the occurrences of over- or under-diagnosis for each condylar region (Table 1). These numbers were subsequently compared between different imaging methods using Wilcoxon-matched-pairs signed-ranks tests with each condylar region.
considered a sample. Specifically, the following 4 comparisons were made according to the hypotheses:

1. Over-diagnosis between CBCT and CT images analyzed without thresholding (for Hypothesis 1)

2. Under-diagnosis between CBCT and CT images analyzed without thresholding (for Hypothesis 1).

3. Over-diagnosis between CBCT images with and without thresholding (for Hypothesis 2)

4. Under-diagnosis between CBCT images with and without thresholding (for Hypothesis 2).

Step-down Bonferroni method of Holm was used for adjustment of the p-values for multiple comparisons.

<table>
<thead>
<tr>
<th>Number of physical defects (a)</th>
<th>Number of defects diagnosed from images (b)</th>
<th>Quantity of over-diagnosis</th>
<th>Quantity of under-diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>If b=a</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>If b&lt;a</td>
<td>0</td>
<td>a-b</td>
</tr>
<tr>
<td></td>
<td>If b&gt;a</td>
<td>b-a</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1. Calculation of over- and under-diagnosis numbers

Next, the numbers of true/false positive or true/false negative diagnoses of all condylar regions were then pooled together for calculating the overall classification functions including sensitivity, specificity, and accuracy (Table 2). More specifically, sensitivity
was tested by finding the True Positive Rate (TPR), which is defined as the number of true positives divided by the sum of the true positives and false negatives. Specificity was derived by finding the True Negative Rate (TNR), which is the number of true negatives divided by the sum of true negatives and false positives. Finally, accuracy was calculated by dividing the sum of the true positive and true negatives by the sum of true positives, true negatives, false positives and false negatives.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>True Positive Rate (TPR)</th>
<th>TP/(TP+FN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>True Negative Rate (TNR)</td>
<td>TN/(FP+TN)</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>TP+TN/(TP+TN+FP+FN)</td>
</tr>
</tbody>
</table>

Table 2. Classification functions and their derivation

A logistic regression analysis was performed to assess the relative ability in achieving accurate diagnoses among three imaging/analysis methods (non-segmented CBCT, non-segmented CT images and segmented CBCT images) as proposed in the main hypotheses.
CHAPTER 3: MANUSCRIPT

Comparing CBCT to Multislice CT in Diagnosing Small Osseous Condylar Defects

Elizabeth M. Jones\textsuperscript{a}, Melissa Papio\textsuperscript{a}, Boon Ching Tee\textsuperscript{b}, Frank M. Beck\textsuperscript{c}, Henry W. Fields\textsuperscript{d}, Zongyang Sun\textsuperscript{e}\textsuperscript{*}

\textsuperscript{a}Resident, Division of Orthodontics, College of Dentistry, Ohio State University, Columbus, OH
\textsuperscript{b}Research Associate, Division of Orthodontics, College of Dentistry, Ohio State University, Columbus, OH
\textsuperscript{c}Assistant professor, Division of Oral Biology, College of Dentistry, Ohio State University, Columbus, OH
\textsuperscript{d}Professor, Division of Orthodontics, College of Dentistry, Ohio State University, Columbus, OH
\textsuperscript{e}Associate professor, Division of Orthodontics, College of Dentistry, Ohio State University, Columbus, OH

*Correspondence to:
Zongyang Sun, Rm 4088 Postle Hall, 305 W 12\textsuperscript{th} Ave, Columbus, OH 43210. Phone: 614-247-7944, Fax: 614-688-3077. Email: sun.254@osu.edu
Acknowledgement: Financial support provided by the Dental Master’s Thesis Award Program sponsored by Delta Dental Foundation, philanthropic affiliate of Delta Dental of Michigan, Ohio and Indiana. Debra Pack, Division of Oral and Maxillofacial Pathology & Radiology provided technical support.
ABSTRACT:

Objectives: Previous studies indicated that orthodontic-grade CBCTs are limited in displaying small defects at the mandibular condyles for diagnosis. This study investigated whether this limitation was inherent to CBCT by comparing CBCT with CT, and whether it can be overcome by image segmentation and color mapping.

Methods: Nine fresh pig heads (18 condyles, 36 medial/lateral condylar regions) were used. Small osseous defects (diameter and depth <1.5mm) were created at condylar, medial and lateral, regions demarcated by gutta percha markers. After restoring the overlying soft tissues, the pig heads underwent orthodontic-grade CBCT scans (i-CAT, 0.4mm voxel-size) and medical-grade CT scans (GE LightSpeed, 0.625mm voxel-size) scans. Subsequently, two calibrated and blinded raters diagnosed defect number in each condylar region from CBCT and CT images using Dolphin-3D software without image segmentation, then 1-week later with proprietary image segmentation of Dolphin-3D. Condylar PVS impressions were collected and evaluated by the same raters to obtain physical diagnosis. Re-diagnosis was made on randomly selected subsamples to assess reliability. Using physical diagnosis as references, the accuracy of imaging diagnosis were assessed and statistically compared among varied imaging/analysis methods.

Results: Image diagnoses of all imaging/analysis methods showed good or excellent intra- and inter-rater reliability except for those of segmented CT images, which were substantially lower. The numbers of over- and under-diagnosis per region were not significantly different among varied imaging/analysis methods (Wilcoxon test, p>0.05) but those from CT tended to be lower than CBCT. Classification functions demonstrate
substantial lower sensitivity and accuracy with CBCT than with CT. Logistic regression showed CT had a significantly higher probability (odds ratio 2.4) than CBCT in reaching correct diagnosis, while the use of image segmentation and color mapping did not significantly improve the diagnostic accuracy from CBCT images.

**Conclusions:** Even at a lower voxel-size than medical CT, orthodontic-grade CBCT images of mandibular condyles are inherently more difficult than CT images for diagnosing small condylar defects, and this limitation cannot be overcome by using image segmentation.
INTRODUCTION:

Evaluation of TMJ structure and function is an important component of orthodontic diagnosis and treatment planning. For decades, mainly two-dimensional (2-D) imaging tools such as panoramic radiographs, cranial projections, and tomograms [1] have been used to assess osseous structures of the TMJ, which revealed that osseous changes can occur in 14-44% of patients with TMD. These tools, however, are inherently limited in displaying the contours of the condyle and articular fossa, thus making diagnosis of pathologies like osteophytes and erosions difficult and unreliable [2]. Compared to 2-dimensional imaging tools, multi-slice computed tomographs (CT) of the TMJ have significantly improved image quality to accurately display osseous changes of the TMJ affected by pathologies [2]. Though representing a significant advance in TMJ imaging, however, CT is not suitable for common use in orthodontics because of the high radiation dose and financial burden associated with CT [3]. As an alternative 3-D imaging tool, cone beam CT (CBCT) has become increasingly popular in orthodontics in recent years [4]. Besides displaying the structures needed for conducting routine orthodontic diagnosis and treatment planning, full field-of-view CBCT images, which often also include the TMJ, make it possible to evaluate TMJ skeletal components without additional radiographs. The accuracy and reliability of using CBCT for diagnosing osseous pathologies of the TMJ, however, greatly depends on the size of the pathology, the settings of the scan, or even the machine used[5]. More specifically, while large osseous abnormalities can be accurately displayed and diagnosed by CBCT [6] [7], small erosive defects, which may indicate early stage osseous changes[8, 9], are significantly more
difficult to diagnose [10]. This is especially the case when large field of view (FOV) or large voxel-size CBCT scans were used [11]. Patel et al further demonstrated that for CBCT images of common scan settings (0.4mm voxel size, full field-of-view), 1/3 of defects with dimensions smaller than 2mm can be overlooked.[12]

These findings raised two questions. One was whether the difficulty involved in diagnosing small condylar osseous defects was a problem inherent to relatively large voxel size or to the CBCT visual diagnostic technique. If it were the former, multi-slice CT images scanned under similar voxel sizes would show the same level of inaccuracy as orthodontic-grade CBCT images in detecting small condylar defects (<2mm). In support of this conjecture, a number of recent studies have compared CBCT with CT and found that these two imaging tools had similar ability in diagnosing condylar defects [3, 4, 6, 13-15]. It is worth noting, however, these previous studies have either omitted the soft tissue factor or focused on relatively large defects (5-10mm), thus leaving the question for small defects unanswered. The other question was whether the inaccuracy involved in diagnosing small condylar defects from CBCT images can be improved by changing the analysis methods, such as image segmentation based on grey levels and visual enhancement using color mapping. These tools are often provided by image analysis software programs and image segmentation has also been used before by others to assess condylar morphological changes and resorption[16, 17].
This study was undertaken to address these two questions. Based on finding from studies on large condylar defects, we hypothesized that the diagnostic ability of detecting condylar osseous defects would not be statistically different between orthodontic-grade CBCT images and medical-grade CT images scanned under large voxel sizes[6, 15]. We also hypothesized that the diagnostic ability of condylar osseous defects from CBCT images would be improved by using global thresholding and color mapping.

**METHODS & MATERIALS:**

A total of nine cadaver heads from 4-6 month old pigs were used in this study, which were collected from the xxx immediately following euthanasia of the animals. Each condyle was dissected to expose the condylar head with care to preserve the soft tissue attachment. Then the condyles were demarcated into medial and lateral regions by small gutta percha pieces which were super-glued to the medial, lateral, and posterior sections of the condyle. Overall, a total of 36 medial or lateral condylar regions was created. One operator (ZS) created small osseous defects (diameter and depth <1.5mm) in each medial or lateral region with a dental handpiece (Brasseler USA®, NSK, Volvere, Vmax) using an end-cutting 1.5mm round bur (Brasseler USA®,1.5mm). The number of defects in each region of each animal was predetermined using a random table.

Once the defects were created and the soft tissue was repositioned, each specimen underwent two scans: an orthodontic-grade CBCT scan (120kV, 5mA i-CAT, Imaging Science International, Hatfield, PA) at a clinically typical 0.4mm voxel-size; a medical-grade CT scan (GE LightSpeed 8-multidetector helical CT scanner, Buckinghamshire,
United Kingdom) at a clinically typical 0.625mm voxel-size. Once the scans were obtained, the soft tissue was completely removed from the condyles, which were subsequently sectioned from the posterior ramus. Light-bodied polyvinyl siloxane (PVS) impressions were made of each condylar head with care taken to include the gutta percha markers for region identification (Figure 1).

Two independent, calibrated raters (EJ, MP) who were blinded of defect location and number in each condylar region analyzed the CBCT and CT scans, then later the PVS impressions. The calibration process for CBCT and CT imaging analysis involved training using the Dolphin 3-D software (Patterson Supply, Inc.), a program most commonly used in clinical orthodontics. During analysis, CBCT or CT images were first reoriented. Specifically, the condylar neck and posterior ramus were oriented parallel to the floor in the sagittal view with the coronal line bisecting the posterior gutta percha marker (Figure 1B). In the coronal view, the condyle was further oriented to allow the axial line to bisect the medial and lateral gutta percha markers and the sagittal line to bisect the posterior gutta percha (Figure 1C). Finally in the axial view, the coronal line bisected the medial and lateral gutta percha markers and the sagittal line bisected the posterior gutta percha marker (Figure 1D). Once the images were correctly oriented, the raters evaluated each slice (0.5mm thickness) in all three views for the presence/absence of defects from the mandibular ramus to the superior aspect of the condyle. When defects were identified, the raters noted the region of the condyle in which the defect was present. The total number of defects in each region was consequently recorded.
To test whether global thresholding improves diagnostic ability, one week after the
CBCT and CT scans were analyzed, the same two independent, blinded and calibrated
raters analyzed each scan with the help of proprietary image segmentation designed by
Dolphin 3-D. The image orientation protocol was the same as that used for the non-
segmentation method described above (Figure 1 B-D). The raters viewed all slices
(0.5mm thickness) in all three views and detected defects in each condylar region.

To test intra-rater reliability, four weeks after the initial image evaluation by both raters,
eight image sets were randomly selected for repeated evaluation following the same
protocols described above.

Finally, the same two raters analyzed the PVS impressions (Fig. 1F), on which the halves
were further delineated and labeled with an indelible marker. The number of defects in
each region was counted and recorded, serving as the gold standard diagnosis.

Statistical Analysis
Inter-rater and intra-rater reliability of CBCT or CT images were assessed by Kappa tests.
The diagnostic accuracy of CBCT and CT images were analyzed statistically in two
ways. First, CBCT and CT diagnoses were compared with physical diagnoses from the
PVS impressions to derive the occurrences of over- or under-diagnosis for each condylar
region (Table 1). These values were subsequently compared between different imaging
methods (CBCT vs. CT, non-segmented CBCT vs. segmented CBCT) using Wilcoxon matched-pairs signed-ranks tests. The step-down Bonferroni method of Holm was used for adjustment of the p-values for multiple comparisons. Next, the numbers of true/false positive or true/false negative diagnoses of all condylar regions were pooled together for calculating the overall classification functions including sensitivity, specificity, and accuracy (Table 2). Subsequently, a logistic regression analysis was performed to assess the relative ability in achieving accurate diagnoses among three imaging/analysis methods (non-segmented CBCT, non-segmented CT images and segmented CBCT images).

RESULTS:

Reliability of Diagnosis:
The inter-rater reliability (kappa, $\kappa$) for all images/analysis methods were equal or above 0.75 except for that of segmented CBCT images, which was below 0.6 (Table 3). The intra-rater reliability had the same trend for both raters, all showing high $\kappa$ values except for segmented CBCT images, which was only near 0.6.

Comparison of Diagnostic Abilities
Since the overall diagnoses from the two raters demonstrated similar trends and were reliable as presented above, data from Rater 2 were used for comparison of diagnostic
abilities between non-segmented CBCT and CT images, and between segmented CBCT with non-segmented CBCT images, to test the two main research hypotheses.

The summarized measurements of over- or under-diagnosis per condylar region for each imaging tool/analytic method are shown in Fig. 2. Non-segmented CBCT images showed a tendency of more over- and under-diagnosis than non-segmented CT images, but the differences were insignificant. Segmentation of CBCT images tended to slightly reduce the amount of over-diagnosis but increased the amount of under-diagnosis; both differences were statistically insignificant.

Classification function parameters calculated based on pooled data are presented in Fig. 3A. All three (sensitivity, specificity and accuracy) demonstrated that non-segmented CT images were better than non-segmented CBCT images by 25%, 3%, and 10%, respectively. Comparison between non-segmented CBCT and segmented CBCT images is presented in Fig. 3B. Overall, all three classification function parameters were minimally changed by image segmentation. Representative images showing the quality of CT and CBCT images of the same condyle are presented in Fig. 4.

The results of logistic regression of imaging tools/analysis methods (non-segmented CBCT, non-segmented CT images and segmented CBCT images) are presented in Table 4. Relative to that of non-segmented CBCT images, the likelihood of obtaining accurate
Diagnosis from non-segmented CT images was significantly higher (odds ratio 2.41, p<0.04), while the likelihood from segmented CBCT images was only slightly and insignificantly higher (odds ratio 1.43, p>0.27).

**DISCUSSION:**

Small osseous defects in the mandibular condyle may indicate early stage osseous changes of the TMJ [8, 9]. While the advent of CBCT technology in orthodontics presents an opportunity to better screen these small osseous defects during routine orthodontic planning and treatment, a previous study by Patel et al. had indicated that orthodontic-grade CBCT imaging has a considerable level of diagnostic inaccuracy[12]. The present study was performed to further investigate whether this limitation is inherent to CBCT or to large scan voxel size by comparing CBCT with CT images, and to evaluate whether enhanced visual observation through imaging segmentation improves CBCT diagnosis.

We first assessed the diagnostic reliability. Based on common kappa value interpretations[19], our data demonstrated that overall there is good or even excellent inter-rater and intra-rater reliability for non-segmented CBCT and CT images, as well as segmented CT images, but the reliability for segmented CBCT images was only fair (Table 3). These data suggest that small osseous defects (<2mm) can be visualized and diagnosed from non-segmented CT and CBCT images in a relatively reliable, although
not necessarily accurate fashion. Our data were also consistent with previous studies, in which the effect of varied defect depth/size and CBCT scan voxel size on the diagnosis of condylar defects was investigated [12]. To our surprise, however, global segmentation and color mapping reduced rather than improved the diagnostic reliability of CBCT images. This may be caused by the relatively poor tissue contrast at the periphery of the condylar head on CBCT images, where global thresholding may in fact make the diagnosis less determinable, thus producing less reproducible diagnosis. This explanation is indirectly supported by the fact that segmentation and color mapping did not affect the reliability of CT images, which are known to better reflect tissue contrast[20]. To the author’s knowledge, to date no other studies have reported the diagnostic reliability of segmented CBCT images, thus these data suggest that precautions may be necessary when using the segmentation and color mapping tools for TMJ diagnosis.

In terms of the diagnostic accuracy, although based on average missed diagnosis (under or over diagnosis) per condylar region, there was no differences between CBCT and CT. However, the pooled data showed significantly better accuracy with CT than with CBCT (Fig. 3, Fig. 4 and Table 4). When each region was considered individually, regardless of the imaging tools, less than 1 defect per condylar region (mesial or lateral half) was over- or under-diagnosed. Two specific points concerning these data are worth noting. First, compared to that of non-segmented CBCT images, the amount of over or under-diagnosis from CT images showed a tendency of being lower. Secondly, the number of defects created in each region was randomly assigned, (ranging from 0-4), thus the difficulty in
reaching completely accurate diagnosis is different among regions. Combined, these data refute our first hypothesis that orthodontic-grade CBCT and medical-grade CT, both scanned under large voxel size settings, have similar diagnostic ability in detecting small condylar osseous defects.

Specifically for the classification function parameters, our data are in line with those reported by Zain-Alabdeen [15] but are in contrast to those reported by Honda et al. [6]. In regard to Honda et al.'s finding that CBCT produced images of better contrast and resulted in diagnosis of higher sensitivity and specificity than CT, it is beyond our ability to provide a definitive explanation. Likely, their use of CBCT images with higher scan resolutions, defects of undefined size, and cadaver specimens without soft tissues may have contributed to their findings.

Nevertheless, data from the present study and the Zain-Alabdeen et al.’s study demonstrated that diagnosis from CT images has substantially better sensitivity than that from CBCT, while the specificity is generally good for both techniques. Our CBCT sensitivity data are also consistent with our previous reports based on different specimens and raters [12]. Together, these data suggest that when orthodontic-grade CBCT images are used to screen small osseous defects in the mandibular condyles, there can be a substantial probability of false negative diagnosis while the risk of false positive diagnosis is relatively low. In addition, this limitation appears to be inherent to the CBCT technique rather than to the voxel size, because when we used medical-grade CT images
taken at an even large voxel size (0.625 mm), the diagnostic sensitivity was much better than CBCT.

Unfortunately, this limitation cannot be overcome by using image segmentation and color mapping of the CBCT images. Essentially, our data demonstrated that there was barely any improvement in diagnostic sensitivity and specificity when CBCT images were segmented and color mapped (Fig. 3B), thus refuting our second hypothesis. These data suggest that although modules for image segmentation and color mapping are often available in many software programs used for analyzing 3-D radiographic files, the benefit of these modules may be limited especially when the original images were of relatively low-quality. More specifically, with color mapping, the images become easier to read visually, but the overall quality of the images were not improved as pixel assignment is determined by the scan methods (CT or CBCT) and scan settings such as voxel size and field of view. In fact, considering that image segmentation and color mapping tend to reduce diagnostic reliability as discussed above (Table 3), one may want to avoid using these auxiliary tools for diagnosing condylar defects from orthodontic-grade CBCT images.

In summary, our present data further confirmed that there is a considerable level of inaccuracy, especially false negative diagnosis, involved in using orthodontic-grade CBCT images for screening small osseous defects at the mandibular condyles when compared to CT. It is important to clarify that this is not a limitation of using CBCT for
orthodontic purposes, but simply a limitation that clinicians need to be aware of when trying to extend the use of orthodontic-grade CBCT to the TMJ area. If it is clinically desired for such diagnosis, one may want to consider changing the scan methods, i.e., using CT instead of CBCT, or using CBCTs of small voxel-size and small field of view (FOV) containing the TMJ area as confirmed previously [11][12]. Clearly, both remedy options come with a higher level of radiation exposure, which should be carefully taken into consideration for choosing radiographic tools for orthodontic patients. The hope that a single CBCT scan is able to render all required 2-D images for orthodontics and meanwhile provide accurate diagnosis of the condyles, is unrealistic based on our data. Therefore, considering that CBCT already has a higher radiation than conventional 2-D imaging techniques, which are usually sufficient for orthodontic diagnosis and treatment planning, it is more reasonable to choose 2-D imaging over CBCT for routine orthodontic patients to reduce radiation exposure. Then, for patients with suspected TMJ issues or pathologies, an additional small voxel-size and FOV CBCT, or even CT may be prescribed.

**CONCLUSIONS:**

1. Compared to medical-grade CT images (0.625 mm voxel-size), orthodontic-grade CBCT images (0.4mm voxel-size) images have significantly inferior quality for the diagnosis of small mandibular condylar osseous defects. Using such images
for screening defects at the mandibular condyles has a relatively high probability (~30%) of making false negative diagnosis.

2. This limitation associated with the use of orthodontic-grade CBCT images for screening small condylar defects cannot be overcome by using image segmentation and color mapping of the condyles.

3. Precautions are warranted when using orthodontic-grade CBCT images for both orthodontic purposes and diagnosis of small osseous condylar defects at the TMJ.
CHAPTER 4: RESULTS

Reliability of Diagnosis:
The inter-rater reliability for non-segmented CBCT and CT images was 0.75 and 0.78 (weighted Kappa), respectively; for segmented CBCT and CT images the inter-rater reliability was 0.59 and 0.87 (Table 3). These weighted Kappa scores indicate excellent agreement for both the CBCT and CT images, moderate agreement for the image segmented CBCT image and almost perfect agreement for segmented CT images based on published criteria. The intra-rater reliability is presented in Table 4. Rater 1 and Rater 2 followed the same trends showing excellent agreement for both non-segmented CBCT and CT images as well as for segmented CT images, but only moderate agreement for the segmented CBCT images.
### Inter-rater Reliability

<table>
<thead>
<tr>
<th></th>
<th>Weighted Kappa</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CBCT-B(^1)</strong></td>
<td>0.75146</td>
<td>0.5822</td>
<td>0.9207</td>
</tr>
<tr>
<td><strong>CT-B(^1)</strong></td>
<td>0.7803</td>
<td>0.6504</td>
<td>0.9102</td>
</tr>
<tr>
<td><strong>CBCT-H</strong></td>
<td>0.58549</td>
<td>0.2994</td>
<td>0.8716</td>
</tr>
<tr>
<td><strong>CT-H</strong></td>
<td>0.86909</td>
<td>0.7856</td>
<td>0.9526</td>
</tr>
</tbody>
</table>

**Table 3. Inter-rater Reliability**

\(^1\)B: Images without segmentation; \(^2\)H: Images with segmentation

---

### Intra-rater Reliability

<table>
<thead>
<tr>
<th></th>
<th>Weighted Kappa</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rater 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCT-B(^1)**</td>
<td>0.8182</td>
<td>0.6702</td>
<td>0.9661</td>
</tr>
<tr>
<td>CT-B</td>
<td>0.7885</td>
<td>0.5468</td>
<td>1</td>
</tr>
<tr>
<td>CBCT-H(^2)**</td>
<td>0.6154</td>
<td>0.2133</td>
<td>1</td>
</tr>
<tr>
<td>CT-H</td>
<td>0.9591</td>
<td>0.8973</td>
<td>1</td>
</tr>
<tr>
<td><strong>Rater 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCT-B</td>
<td>0.8697</td>
<td>0.7779</td>
<td>0.9614</td>
</tr>
<tr>
<td>CT-B</td>
<td>0.8728</td>
<td>0.7666</td>
<td>0.979</td>
</tr>
<tr>
<td>CBCT-H</td>
<td>0.5882</td>
<td>0.2874</td>
<td>0.889</td>
</tr>
<tr>
<td>CT-H</td>
<td>0.8669</td>
<td>0.7385</td>
<td>0.9954</td>
</tr>
</tbody>
</table>

**Table 4. Intra-Rater Reliability**

\(^1\)B: Images without segmentation; \(^2\)H: Images with segmentation
Comparison of Diagnostic Abilities

Because overall the diagnoses from the two raters demonstrated similar trends and were reliable as presented above, data from Rater 2 were used for comparison of diagnostic abilities between non-segmented CBCT and CT images, and between CBCT images with and without segmentation, to test the two research hypotheses outlined in Chapter 1.

The summarized measurements of over- or under-diagnosis for each imaging tool/analytic method are shown in Fig. 4-5. Non-segmented CBCT images showed a tendency of more over- and under-diagnosis than non-segmented CT images, but the differences were insignificant. Segmentation of CBCT images tended to slightly reduce the amount of over-diagnosis but increase the amount of under-diagnosis, although again both differences were insignificant.
Figure 4. Between-method comparisons for over-diagnosis (OD) and under-diagnosis (UD) for non-segmented CBCT and CT images
Figure 5. Between-method comparisons for over-diagnosis (OD) and under-diagnosis (UD) for non-segmented CBCT and segmented CBCT images
Measurements of all regions were subsequently pooled together for calculation of classification function parameters. Comparison between non-segmented CBCT and CT images is presented in Figure 1. All three (sensitivity, specificity and accuracy) demonstrated that non-segmented CT images were better than non-segmented CBCT images by 25%, 3%, and 10%, respectively. Comparison between non-segmented CBCT and segmented CBCT images is presented in Figure 2. Overall, all three classification function parameters were minimally changed by image segmentation.
Figure 6. Sensitivity, Specificity & Accuracy of Rater 2 for diagnosis of small osseous condylar defects using CBCT and CT images without segmentation

- Non-segmented CBCT
- Non-segmented CT
The results of logistic regression of imaging tools/analysis methods (non-segmented CBCT, non-segmented CT images and segmented CBCT images) are presented in Table 7. Relative to that of non-segmented CBCT images, the likelihood of obtaining accurate diagnosis from non-segmented CT images was significantly higher (odds ratio 2.41, p<0.04), while the likelihood from segmented CBCT images was only slightly and insignificantly higher (odds ratio 1.43, p=0.27).
| Effect          | METHOD                  | Degree of freedom | t Value | Pr > |t| | Estimate | Limits |
|-----------------|-------------------------|-------------------|---------|-------|---|--------|--------|
| Intercept       |                         | 35                | -0.56   | 0.58  |   |        |        |
| METHOD D        | CBCT-H$^2$              | 35                | 1.13    | 0.27  |   | 1.431  | 0.751  | 2.725  |
| METHOD D        | CT-B                    | 35                | 2.11    | 0.04  |   | 2.413  | 1.034  | 5.634  |
| METHOD D        | CBCT-B*,$^{1,}$        |                   |         |       |   |        |        |

*CBCT-B is the referent, $^1$: Images without segmentation; $^2$: Images with segmentation

**Table 5 Logistic regression analysis for correct diagnosis**
CHAPTER 5: DISCUSSION AND CONCLUSIONS

Structural changes to the TMJ are important characteristics in establishing a diagnosis of OA. A radiographic exam is a routine component of clinical assessment for conditions of TMJ dysfunction, where the main goal is to verify degenerative bone changes in the joint structure. Osseous changes start focally and are not radiographically detectable, but as the degenerative process progresses, bone erosions occur creating radiographic evidence of OA. Although currently no general consensus as to a gold standard imaging technique for detecting osseous lesions in the TMJ has been reached, it has become increasingly popular to use CBCT for diagnosing TMJ osseous pathology. For small osseous defects (<2mm) on the condyle, however, our previous study has indicated that CBCT imaging has a considerable level of diagnostic inaccuracy. The present study was performed to further investigate whether there is an inherent limitation of CBCT imaging by comparing CBCT with CT images, and to evaluate whether enhanced visual observation through imaging segmentation improves CBCT diagnosis.

We first assessed the reliability of diagnosing small defects form mandibular condyles. Based on common kappa value interpretations, our data demonstrated that overall there is substantial inter-rater reliability for non-segmented CBCT and CT images, as well as segmented CT images (Table 3). The reliability of segmented CBCT images was lower
but still fair. The intra-rater reliability data demonstrated the same trend (Table 4). These data suggest that small osseous defects (<2mm) can be visualized and diagnosed from non-segmented CT and CBCT images in a relatively reliable fashion. These data were also consistent with findings in a previous study, in which we examined the effect of varied defect depth/size and CBCT scan voxel size on the diagnosis of condylar defects. To our surprise, however, global segmentation reduced rather than improved the diagnostic reliability of CBCT images. This may be caused by the relatively poor tissue contrast at the periphery of the condylar head, where global thresholding may in fact make the diagnosis less determinable, thus less reproducible diagnosis. To the author’s knowledge, at the time of publication, no other studies have reported the diagnostic reliability of segmented images.

In terms of the diagnostic quality, when each condylar region was treated as an individual sample and the amount of over-diagnosis or under-diagnosis from each region was compared between imaging tools/analysis methods, no statistically significant differences were found (Fig. 5). On average, regardless of the variations in imaging tools/analysis methods, less than 1 defect per region was over- or under-diagnosed. While these data indicate that the imaging tools did not make a difference to the diagnostic accuracy, and overall most regions were correctly diagnosed, a few specific points concerning these data are worth noting. First, compared to that of non-segmented CBCT images, the amount of over or under-diagnosis from CT images showed a tendency of being lower, although the difference was statistically insignificant which may be caused by
insufficient power. Next, the number of defects created in each region was randomly assigned, which were in a range of 0-4, thus the difficulty in reaching completely accurate diagnosis is different among regions. Finally, these over and under-diagnosis measurements were somewhat difficult to be interpreted in a clinically relevant way and compare to other studies.

To address these issues, measurements from all regions were subsequently pooled together to calculate the diagnostic sensitivity, specificity and accuracy, followed by conducting a logistic regression test. These results clearly showed that the diagnostic ability of non-segmented CT images were superior to that of non-segmented CBCT images (Fig. 4, Table 7), and employing image segmentation to CBCT images failed to improve the diagnostic quality (Fig. 5, Table 7). Thus, the two hypotheses that we initially designed were not supported by our data and have to be rejected.

In regard to the difference between CBCT and CT images, several studies have investigated it before. Honda et al. demonstrated a higher sensitivity with CBCT diagnosis than CT\textsuperscript{27}, which is in contrast to our finding. In their study, higher resolution CBCT images were used, the defects they were diagnosing were of undefined size, and there was no attempt to recreate soft tissue around the cadaver specimens, all of which may be potential causes for the differences of findings. In another study, Zain-Alabdeen reported lower sensitivity values for both CBCT and CT images than those of our study,
but similar to our findings, they indeed found that the CBCT diagnoses were less sensitive than those of the CT\textsuperscript{19}. They suggested that the overall low sensitivity values were due to the mild surface osseous changes, which are also reflected in the small osseous defects of our specimens.

For specificity, again we found slightly higher specificity with CT images than with CBCT (93\% versus 91\%). Honda, et al. reported perfect specificity (100\%) for both CBCT and CT images\textsuperscript{27}. Zain-Alabdeen found slightly higher specificity for both CBCT and CT (83-90\% and 85-88\% respectively) than our study and slightly better specificity with CBCT than CT\textsuperscript{19}. Nevertheless, despite the small inconsistencies among the studies, it appears that neither CT nor CBCT images are at high risks for making false positive diagnoses.

Diagnostic accuracy is a measure of the overall ability to make correct diagnosis, which is influenced by both sensitivity and specificity. Not surprisingly, our data demonstrated that the accuracy of CT was better than CBCT, while Honda et al. found the opposite\textsuperscript{27}. Again, the inconsistency may be contributed to the difference in research design including higher resolution of the CBCT images and presumed larger defect sizes, and the lack of soft tissue compensation. It has been noted that images of dry skulls without soft tissue compensation may have less noise making it easier to detect osseous changes with higher sensitivity and specificity\textsuperscript{39}. Thus, our study represents an improvement in research design by using fresh specimens with in-tact soft tissue, which is more indicative
of real-life situations and diagnosis. Our findings were also further confirmed by a logistic regression which was lacking from Honda et al.’s study, which also found a non-significant p-value (P = 0.286).

Our study, for the first time, directly compared whether segmentation of CBCT images improves the diagnosis of osseous defect at the condyles. Our data demonstrated that the improvement was rather small and insignificant. Modules for image segmentation are often available in many software programs used for analyzing 3-D radiographic files, but the benefit of these modules may be limited. With color mapping, the images become easier to read, but the overall quality of the images were not improved by the color mapping as it is determined by the scan methods (CT or CBCT) and scan settings such as voxel size, field of view. In this study we used standard clinical orthodontic voxel sizes (0.4 mm) and 12-inch FOV for CBCT scans, in order to further understand the limitations of using CBCT scans taken for orthodontic purposes for screening pathologies at the mandibular condyles, and seek potential remedies.

Once again, our present data suggested that there is a considerable amount of diagnostic inaccuracy involved in using orthodontic-grade CBCT images for screening small osseous defects at the mandibular condyles, which was previously documented by Patel et al. By comparing with CT images, which were taken at an even large voxel size (0.625 mm), we further confirmed that this is an inherent limitation of the CBCT images of this grade. Unfortunately, this limitation cannot be overcome by post-scan
manipulation of the images using digital segmentation. Thus, the remedy (if desired) for this limitation, has to be through changing the scan methods, i.e., using CT instead of CBCT, or changing the scan setting, i.e., using CBCTs of small voxel-size and small field of view (FOV), which has been confirmed by Librizzi et al.\textsuperscript{40} and Patel et al\textsuperscript{28}.

Clearly, both remedy options come with a higher level of radiation exposure, which should be carefully weighed for choosing radiographic tools for orthodontic patients. The hope that a single CBCT scan is able to render all required 2-D images for orthodontics and meanwhile provide accurate diagnosis of the condyles, is unrealistic based on our data. Therefore, considering that CBCT already has a higher radiation than conventional 2-D imaging techniques, which are often sufficient for orthodontic diagnosis and treatment planning, it is more reasonable to choose 2-D imaging over CBCT for routine orthodontic patients to reduce radiation exposure. Then, for patients with suspected TMJ issues or pathologies, an additional small voxel-size and FOV CBCT, or CT may be prescribed.

Conclusions:
1. Compared to medical-grade CT images (0.625 mm voxel-size), orthodontic-grade CBCT images (0.4mm voxel-size, 12 inch-FOV) images have significantly inferior quality to for the diagnosis of small mandibular condylar osseous defects. Using such images for screening defects at the mandibular condyles only has moderate diagnostic sensitivity 69\%, specificity 91\% and accuracy 84\%.
2. This limitation associated with orthodontic-grade CBCT images cannot be overcome by using image segmentation before performing diagnosis.
Comprehensive References


