# EX VIVO EVALUATION OF NEW 2D AND 3D DENTAL IMAGING TECHNOLOGY FOR DETECTING CARIES

Laurence R Gaalaas

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements for the degree of Master of Science in
Oral and Maxillofacial Radiology in the School of Dentistry.

Chapel Hill
2015

Approved by:

Donald Tyndall

André Mol

Eric Everett

**ABSTRACT**

Laurence R Gaalaas: EX VIVO EVALUATION OF NEW 2D AND 3D DENTAL IMAGING TECHNOLOGY FOR DETECTING CARIES
(Under the direction of Donald Tyndall)

Proximal dental caries remains a prevalent disease with only modest detection rates by current diagnostic systems. X-ray radiography represents the most common and successful means of diagnosing early dental caries lesions, however; many new systems are available without controlled validation of diagnostic efficacy. This study evaluated the caries detection of three new dental radiographic imaging technologies: an intraoral digital detector employing an advanced sharpening filter, an extraoral "panoramic bitewing" imaging unit, and a cone beam-CT system with advanced artifact reduction. An ex vivo study design using extracted human teeth, expert observer ratings, and micro-CT ground truth analysis was employed. All modalities performed similarly in overall diagnostic accuracy yet differences were noted in selected system sensitivities and specificities. The CBCT system demonstrated the best assessment of lesion depth and lesion cavitation. Incorporating hydroxyapatite calibration phantoms allowed assessment of imaging consistency, linearity, and contrast resolution.

**ACKNOWLEDGEMENTS**

Special thanks are due to Thomas J Schaaff, DDS, who generously offered many Friday afternoons and access to his practice's Planmeca ProMax® panoramic unit. Thanks are also due to John Preisser, Jr, PhD, Research Professor in the Department of Biostatistics, Gillings School of Global Public Health and Deputy Director of the Biostatistical Core with the North Carolina Translational and Clinical Sciences (TraCS); Adane Wogu, PhD Candidate and Research Assistant, Department of Biostatistics, Gillings School of Public Health; and Ananta Bangdiwala, MS Candidate and Research Assistant, Gillings School of Public Health. Their support, advice, and timely statistical analyses proved invaluable to the project. We finally thank the Small Animal Imaging Core facility at the UNC Biomedical Imaging Research Center for providing the micro-CT imaging service. The imaging core is supported in part by an NCI cancer center Grant #P30-CA016086-35-37.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ANOVA | analysis of variance |
| $A_Z$ | area under the ROC curve |
| CBCT | cone beam computed tomography |
| CNR | contrast to noise ratio |
| DT | thesis mentor (Donald Tyndall) |
| FOV | field of view |
| HAP | hydroxyapatite |
| HD | high definition |
| IRB | institutional review board |
| MARS | metal artifact reduction software |
| µSv | micro-Sieverts |
| micro-CT | micro-computed tomography or microtomography |
| PanBW | Planmeca panoramic bitewings |
| PI | principal investigator (Laurence R Gaalaas) |
| PSP | photostimulable phosphor (bitewings) |
| ROC | receiver operating characteristic (curve) |
| ROI | region of interest |
| Schick33 | Schick 33 direct digital bitewings |
| SdNR | signal difference to noise ratio |
| UNC | University of North Carolina at Chapel Hill |
| XG3D | Sirona Orthophos XG3D cone beam CT |

## OVERALL INTRODUCTION

This thesis represents two separate but complimentary approaches to evaluating three relatively new dental radiographic imaging systems for proximal caries detection. These new technologies include the Schick 33 intraoral direct digital sensor (Sirona Dental, Salzburg, Germany) employing a new dynamic image sharpening adjustment, a Planmeca ProMax® panoramic unit (Planmeca Inc., Helsinki, Finland) operating in extraoral "panoramic bitewing" mode, and a Sirona Orthophos XG3D CBCT (Sirona Dental, Salzburg, Germany) employing Metal Artifact Reduction Software (MARS), an advanced approach to reducing artifacts in CBCT. Each of these systems is available in the dental market but with little or no controlled independent validation of caries detection efficacy.

The first approach to evaluating caries detection is clinically focused. We employed an ex vivo study design with extracted human teeth placed into a phantom human mandible. Micro-computed tomography was used as a ground truth analysis to establish the proximal caries status of the extracted teeth. Ten expert dentist observers who were blinded to the caries status of the teeth rated confidence in lesion presence, caries lesion depth, and confidence in lesion cavitation with images from each new modality as well as images from a photostimulable phosphor (PSP) "control modality" system. As a "paired-case, paired-reader" study design, all teeth were imaged on all clinical modalities, and all observers viewed all teeth. A combination of lesion presence, depth, and cavitation status assessment along with subsequent receiver operating characteristic, sensitivity/specificity, and correlation analysis methods provided a thorough and holistic view of the systems' proximal caries detection capabilities.

The second approach to evaluating system performance is fundamentally technical. We incorporated chemically defined hydroxyapatite (HAP) phantom rods in a wide density range (12.5 to 1500 mg/cc) into each image from each clinical modality. We also included these phantom rods into each micro-CT scan of the extracted teeth. Comparing the reported grayscale value of each rod in each image to its HAP density, we were able to evaluate each modality's exam-to-exam consistency, linearity of

signal response, and contrast resolution. Understanding that radiographic caries diagnosis relies heavily on displayed contrast of dental hard tissues, these evaluations allowed us to compare technical performance with caries detection observations from the first, clinical part of the thesis. This method of incorporating known-density phantoms into micro-CT scans for the purpose of hard tissue calibration is well documented in the literature, however; incorporation of such calibration phantoms into actual modern imaging modalities is relatively novel. Accordingly the authors feel that this second aspect of the thesis can stand alone as a significant contribution to the technical dental imaging literature.

## MANUSCRIPT ONE: CARIES DETECTION BY OBSERVER ANALYSIS

### Introduction

Dental caries remains the most prevalent dental disease, affecting millions of individuals world-wide. Nearly 80 percent of patients have at least one cavity by 17 years of age, and in general 20% of the population experiences at least 60% of the caries burden, all despite the caries reduction effects resulting from the widespread use of fluoride.[1] Overall caries remains a major problem for large segments of the population.[1] Radiology plays a major role in caries diagnosis, particularly for the interproximal contact regions of posterior teeth where clinical examination is much more challenging.[2, 3] Among many diagnostic modalities the bitewing radiograph stands as the most accepted and best overall approach to detection of dental caries.[2-6] The detection rate of proximal caries is far from ideal, however; as standard bitewing radiographs detect only about 60% of proximal lesions.[7-9]

The relatively recent introduction of digital dental radiographic imaging has so far failed to demonstrate any increase in caries detection rates. Numerous studies have demonstrated that photostimulable-phosphor (PSP) and current direct-digital bitewing images provide detection rates equivalent to traditional film images.[5, 6, 10-15] Digital imaging does allow for more powerful imaging processing, enhancement, and manipulation compared to traditional film.[16] It is also well understood that that caries diagnosis remains an image contrast-limited task rather than an image resolution limited task.[17, 18] Limited studies suggest that advanced contrast adjustment improves proximal caries diagnosis[19] and observer agreement.[20] Image sharpening, a variation of contrast enhancement has been explored as a means to improve caries detection, but most studies have demonstrated no increase in caries detection.[21, 22] Furthermore such image enhancements have been implicated with increased false positive diagnoses.[23] Overall there does not yet appear to be an image enhancement algorithm that delivers clear gains in caries diagnosis.[16]

Compared to intraoral imaging, extraoral imaging promises improved patient comfort and efficiency. These modalities have been evaluated for their caries detection potential with overall mixed and modest results. Standard digital panoramic imaging demonstrates inferior detection rates compared with intraoral imaging due to the superimposition of additional structures, increased image blurriness, and inconsistent opening of posterior proximal contacts.[24-26] Conversely, digitally enhanced tomography images (Cranex® TOME scanograms, Soredex, Tuusula, Finland) demonstrated similar caries detection accuracy to standard film bitewing.[27] At this time we are aware of only one published study that has evaluated panoramic bitewing images (Planmeca Promax®, Planmeca Inc., Helsinki, Finland) for proximal caries detection.[28] This study demonstrated inferior detection rates of the extraoral systems compared to standard intraoral bitewings, but use of root lesions and extensively involved proximal lesions in their sample may have influenced final results.

Cone beam CT (CBCT) imaging in 3D overcomes the fundamental limitation of superimposition of structures inherent to conventional 2D imaging. Many studies have demonstrated CBCT caries detection rates approximately equivalent to intraoral modalities if the teeth are non-restored.[15, 29-35] Other studies have demonstrated increased detection of caries lesions involving the dentin,[36, 37] more accurate lesion depth determination,[34] and improved occlusal caries detection[35] using CBCT versus conventional bitewing techniques. A recent review of the literature summarizes that currently, CBCT is equivalent to intraoral techniques at detecting clinically relevant caries lesions in minimally restored teeth but increased dose, cost, time, and artifact concerns dictate that bitewing images are still the preferred modality for proximal caries detection.[2]

It is understood that beam hardening and streak artifacts from both metal objects and dense tooth structure (enamel) are a major limitation of CBCT imaging.[38, 39] In reality beam hardening and extinction artifacts are significant contributors to streak artifacts.[38, 40] Scatter, noise, exponential edge-gradient, and aliasing effects are additionally implicated in artifact formation but to a lesser extent.[40] Various metal artifact reduction techniques are available and their strategies vary from post-processing of the artifacts within the reconstructed images, to preprocessing the raw projection data prior to reconstruction, to iterative reconstruction techniques employing statistical modeling of regions affected

by artifact.[41-44] In general, pre-processing of the raw projection data and iterative techniques demonstrate more promise in artifact reduction than post-processing approaches.[41-43] In both in vitro and in vivo imaging environments, at least one specific algorithm has demonstrated the ability to increase contrast to noise ratio.[45, 46] Despite multiple manufacturer's efforts, the combination of beam hardening and streak artifacts in thick proximal enamel and around existing restorations or metallic appliances continues to be a source of caries false positives and a major obstacle for taking advantage of 3D imaging for diagnosing caries.[30, 47]

Our best understanding of the caries disease process has identified lesion depth, activity, and cavitation status as significant indicators for the likelihood of lesion progression.[48, 49] There has been a corresponding shift in caries treatment philosophy from an aggressive approach emphasizing surgical intervention and tooth restoration, to a preventative approach emphasizing early detection and non-surgical treatment of lesions.[1, 4, 50] This non-restorative approach has been shown to reduce overall dental treatment costs as well as the need for additional dental restorative therapy.[48, 50] The new approach of identifying and treating incipient and early caries lesions places a maximum demand on radiographic imaging systems. With regard to lesion detection, this point represents our rationale for desiring increased sensitivity from new diagnostic systems with no corresponding compromise in the already high specificity rates. This point also raises the importance of a modality's ability to accurately identify lesion depth and lesion cavitation.[51]

Radiographic examination plays a role in the assessment of lesion depth, which influences treatment decisions. As a caries lesion progresses, the likelihood of successfully treating the lesion non-restoratively diminishes quickly. The likelihood of more severe dental disease such as infection of the pulp and tooth fracture correspondingly increases. The assessment of lesion depth by radiographic means is important for interproximal surface lesions where clinical assessment of lesion status is not possible. Existing research on the ability of conventional 2D radiographic modalities to accurately identify lesion depth shows mixed results. Limited studies report that 2D bitewing radiographs provide accurate depth assessment,[52] whereas most studies conclude that lesion depth assessment from bitewing radiographs is unreliable,[53, 54] or consistently underestimated.[13] Despite the mixed results, there is general consensus

that bitewing images most likely underestimate lesion depth approximately 30%.[16] Research on the ability of CBCT to assess lesion depth is also somewhat mixed with one study indicating no difference in depth accuracy between CBCT and conventional 2D modalities when assessing occlusal lesions,[55] and other studies indicating truly improved depth accuracy of proximal lesions.[34, 56]

While a single radiograph cannot provide any information about lesion activity, i.e. the current status of tooth surface demineralization versus remineralization, accurate relative depth measurements can provide an indication of lesion activity. Radiographs at different points in time can indicate whether an existing lesion has grown in size (indicating an active lesion) or remained the same size (indicating an inactive lesion). Despite concerns about reproducible image geometry, this radiographic approach to identifying lesion activity has been validated for conventional 2D bitewing imaging.[5] In summary, accurate depth assessment can provide clinically relevant information for caries assessment.

In addition to lesion depth, the presence of lesion cavitation represents a significant component in the decision process of whether to treat a caries lesion by non-surgical or surgical means.[57] A cavitated lesion demonstrates a much higher likelihood of progression because the plaque biofilm becomes physically trapped in the compromised tooth structure, preventing hygienic removal by the patient and blocking the therapeutic effects of remineralizing agents.[49] Surface cavitation status has been shown to be a significantly better predictor of caries lesion progression than lesion depth alone.[58, 59]

Unfortunately, conventional 2D imaging modalities are largely unable to differentiate the cavitation status of caries lesions.[2] Historically, studies have identified a link between lesion depth as assessed by radiographs and the likelihood of surface cavitation. A treatment-decision threshold was established based on these relationships by which only lesions assessed as extending into the dentin or deeper were selected for surgical intervention.[2] Recent work, including a comprehensive literature review, has called into question this relationship between lesion depth, cavitation, and its implications for treatment selection. More recent studies claim that according to the best available evidence, cavitation status of a lesion claims an ever-increasing role in the likelihood of a lesion's progression, whereas lesion depth may be less relevant to lesion progression than once believed.[2, 60]

The growing clinical significance of lesion cavitation status is particularly interesting because CBCT provides significantly improved detection of cavitated caries lesions with no compromise in specificity, compared to other common 2D modalities.[32, 61] There is general consensus that at an approximately 30% increase in sensitivity, this boost over conventional modalities is clinically relevant.[2] Coupled with the increased emphasis of cavitation status and its implications for treatment decisions, it appears that CBCT could play an important role in caries diagnosis.

A variety of new dental imaging technologies have recently been introduced to the dental market, each having developments that may increase caries detection. These new technologies include the Schick 33 intraoral direct digital sensor (Sirona Dental, Salzburg, Germany), a Planmeca ProMax® panoramic unit in bitewing mode (Planmeca Inc., Helsinki, Finland), and a Sirona Orthophos XG3D CBCT in High Definition (HD) with Metal Artifact Reduction Software (MARS) (Sirona Dental, Salzburg, Germany).

The Schick 33 intraoral sensor has a new dynamic image sharpening filter which alters contrast and may aid in caries visualization. This image sharpening filter directly addresses the contrast-limited task of caries identification.[17, 18] The Planmeca ProMax® panoramic bitewing images are acquired with a new image projection geometry designed to open posterior dental contacts and consequently increase caries visibility. The Sirona Orthophos XG3D CBCT in HD mode with MARS employs artifact reduction algorithms which may overcome the beam hardening and streak artifact limitation in CBCT caries imaging for both restored and non-restored teeth. Conceptually, the Sirona iteration of MARS employs a pre-processing approach that removes dense objects from the raw projection data before reconstructing the scanned volume. It then replaces the dense objects in the volume with appropriate grayscale values. Approaches similar to this strategy have shown the greatest potential for meaningful artifact reduction.[42]

Very little research has been done on these new imaging systems with respect to their caries detection ability. The Schick 33 sensor with its dynamic image processing is very new to the market and no independent scientific evaluation has been performed on the system. Furthermore, non-peer-reviewed publications promoting the proven superior caries detection abilities of Planmeca panoramic bitewing images have been circulated.[62] Sirona Orthophos XG3D CBCT images in HD mode employing

MARS demonstrate reduced beam hardening and streak artifacts, yet the impact of this technique on caries detection with and without dental restorations or appliances has yet to be evaluated.

The National Institutes of Health Consensus Development Conference Statement on Diagnosis and Management of Dental Caries Throughout Life calls for continued research on diagnostic methods, including new devices and techniques.[1] Especially because all three of these imaging systems are available in the dental market, evaluation of their caries diagnostic capability is critical. In vitro/ex vivo studies have been validated as an appropriate means of assessing radiographic caries diagnosis.[63] Accordingly, the aim of this project is to establish the diagnostic efficacy of these three new dental radiographic imaging technologies for diagnosing the presence, depth, and cavitation status of proximal caries in non-restored teeth using an ex vivo study design.

**Materials and Methods**

Institutional review board (IRB) approval was obtained to collect de-identified extracted human teeth from existing tooth repositories and to perform observer sessions at the University of North Carolina at Chapel Hill (UNC) School of Dentistry (Study #13-2843). Tooth selection criteria included human premolar or molar teeth with an unrestored status or minimally restored status with cervical (non-coronal, non-proximal) involvement only. Selection criteria also included proximal caries status of sound/no-lesion or small to moderate sized lesion as estimated by visual, tactile, and bitewing radiographic techniques. Teeth with large, cavitated coronal lesions were excluded. A total of 29 extracted teeth were selected for the sample. Three of the 29 teeth were selected to be used twice in the sample, giving a total tooth sample of 32 teeth/64 proximal surfaces. The three teeth used twice in the sample were selected because they demonstrated simultaneously classic proximal lesion morphology and challenging lesion visibility when screened with bitewing radiography.

A dry human mandible with edentulous posterior segments and residual extraction sockets was used as the ex vivo phantom for this study. Wax was used to hold extracted human teeth in the residual extraction sockets. Pairs of one premolar tooth and one molar tooth were randomly selected from the sample to establish an anatomically appropriate pair of teeth for placement in the phantom mandible.

Premolar/molar pairs were assigned either teeth "19" and "20" positions on the left side of the mandible or teeth "29" and "30" positions on the right side of the mandible. Slight adjustments were made to the randomized order to ensure that the three teeth used twice in the sample were assigned to opposite sides of the mandible. This adjustment step ensured that the duplicated teeth would be imaged in separate right and left anatomic orientations, in effect mirroring their radiographic image and minimizing the possibility for biased observer recall. The individual teeth were numbered and the order of premolar/molar pairs was recorded so that the exact tooth sequence could be replicated.

Each pair of teeth was placed in its assigned mandible site prior to imaging with each modality. Non-restored and non-carious premolars and molars were placed anteriorly and posteriorly, respectively, to the paired sample teeth. All teeth were arranged as anatomically correct as possible and all contacts were closed to best simulate real patient anatomy. Efforts were made to make contact orientations as parallel as possible in order to minimize the confounding effect of "closed contacts" on the task of interpretation.

A 1 cm layer of wax was placed around the mandible to simulate the effect of soft tissue attenuation of the x-ray beam. For the extraoral imaging modalities (the Sirona XG3D CBCT and the Planmeca panoramic bitewings) an additional water balloon and 3 mm thick by 15.5 cm diameter aluminum cylinder were placed inside and around the mandible/wax phantom, respectively, to simulate the attenuation properties of the human head relevant to extraoral imaging.

The control modality consisted of photostimulable phosphor (PSP) bitewing images were taken used as the standard UNC School of Dentistry radiology clinic system: Gendex size 2 PSP plates (Gendex, Hatfield, PA) were exposed with a Focus intraoral source (Instrumentarium Dental, Tuusula, Finland) at 70kVp, 7mA, 0.2s, and 40cm SID with standard 30cm rectangular collimation. Exposure time was selected to achieve appropriate receptor exposure. Image geometry was controlled by hand for every image resulting in open proximal contacts for each set of extracted teeth and a clinically-relevant simulation of image acquisition. The exposed plates were processed with a ScanX IO ILE scanner (Air Techniques, Melville, NY) through MiPACS Dental Enterprise Viewer 3.1.1401 operating ScanX Plugin Version 1.2.8 (Medicore Imaging, Charlotte, NC). Plates were scanned with the following settings:

Intraoral High (#2) setting, 16-bit, invert images, and Image Enhancement: Enable histogram stretch, upper histogram cut 0.3, Lower histogram cut 3.4, Gamma correction 0.7. Images were saved with no additional adjustment of window and level and no additional applied filters.

The Schick 33 (Schick33) direct digital bitewing images were taken with the same Focus intraoral source (Instrumentarium Dental, Tuusula, Finland) at 70kVp, 0.05s, and 40cm SID standard 30cm rectangular collimation. The software interface for the Schick 33 sensor was run CDR DICOM for Windows Version 5.4.1658.5883 (Sirona Dental, Salzburg, Germany). Exposure time was selected to achieve appropriate receptor exposure according to the CDR exposure indicator tool. Images were acquired with the following settings: Schick 33 High Resolution Acquisition, subtract dark image, acquire 12 bit image. Like the PSP modality, image geometry was controlled by hand for every image resulting in open proximal contacts for each set of extracted teeth and a clinically relevant simulation of image acquisition.

The panoramic bitewing (PanBW) images were taken with a Promax® Dimax 4 unit (Planmeca Inc., Helsinki, Finland), in panoramic bitewing mode operating at 72kVp and 11mA with a square average jaw shape. This particular unit was in active use in a private dental practice in the region. The mandible was placed in the machine according to standard anatomic positioning. Small adjustments were made to the phantom position to ensure open posterior contacts on the images. Dexis Version 9.0.5 imaging software (Dexis LLC, Hatfield, PA) was used to acquire and store the images. Dexis imaging software was chosen instead of Planmeca Romexis native imaging software because Dexis was used exclusively by the private office operating the unit.

The CBCT (XG3D) scans were taken with a Sirona Orthophos XG3D (Sirona Dental, Salzburg, Germany) unit operating at 85kV and 6mA with a 5x5cm FOV, and 0.1mm voxel size. Both high definition (HD) mode and metal artifact reduction software (MARS) were used. Scans were reconstructed and stored using Sidexis XG Version 2.56 software (Sirona Dental, Salzburg, Germany).

Caries ground truth status was established by micro computed tomography (micro-CT). The micro-CT scans were acquired with a SCANCO Medical µCT 40 scanner (Scanco Medical AG, Bruttisellen, Switzerland) operating at 70kVp, 0.115 mA, 200s scan time, with 0.5mm Al filtration, 5µm focal spot size, 24µm CMOS camera pitch, 1 sample per pixel, and convolution kernel 3. Extracted teeth were placed in a

2cm diameter poly-ether-imid sample tube with 0.7mm wall thickness. To confirm system accuracy, the unit underwent routine weekly bone density calibration using manufacturer supplied hydroxyapatite (HAP) phantoms ranging in density from 99-800 mg/cc. For the year prior to imaging, average variance for the 800 mg/cc calibration phantom was 12 mg/cc which was well within manufacturer tolerances. No beam hardening correction was applied as bone density calibrations were considered accurate and selected system parameters were deemed acceptable for accurate dental tissue imaging. Scans were reconstructed with a 20μm voxel size using Scanco Version 1.2a software (Scanco Medical AG, Bruttisellen, Switzerland).

A total of 10 observers were recruited from the UNC School of Dentistry Division of Oral and Maxillofacial Radiology. All observers had specialty-level training in oral and maxillofacial radiology and familiarity with diagnosing caries on intraoral, extraoral, and 3D imaging modalities. Therefore, all observers were considered to have "expert" skill and knowledge of the radiographic interpretation of caries above and beyond the general dentist level of training. Of the 10 observers, 4 were board-certified faculty members in the division of oral and maxillofacial radiology. The remaining 6 observers were graduate students in the UNC School of Dentistry's oral and maxillofacial radiology graduate training program.

All observers participated in an orientation session prior to the interpretation sessions. The orientation reviewed study objectives, details of the four different imaging modalities, and the nature of the tooth sample. Observers were instructed on how to properly interpret and respond to three interpretation tasks regarding caries presence, lesion depth, and lesion cavitation. Proper use of a 5-point scale for receiver operating characteristic (ROC) analysis was explained. Additional emphasis was placed on the nature of beam hardening artifact in CBCT images and the suspected imaging differences from real caries lesions. Example images of both beam hardening artifacts and caries lesions were shown. The software used for each experimental modality was demonstrated. Informed consent to participate in the study was obtained from each observer per IRB protocol.

The specific tasks asked of the observers for each proximal surface of each sample tooth were as follows. Task 1: rate the likelihood of caries presence on a 5-point scale where 1 = caries definitely not

present, 2 = caries probably not present, 3 = unsure, 4 = caries probably present, and 5 = caries definitely present. Task 2: rate the caries lesion depth on a 5-point scale where 1 = caries not present, 2 = caries involving the outer half of enamel, 3 = caries involving the inner half of enamel, 4 = caries involving the outer half of dentin, and 5 = caries involving the inner half of dentin. Task 3: rate the likelihood of lesion cavitation on a 5-point scale where 1 = cavitation definitely not present, 2 = cavitation probably not present, 3 = unsure, 4 = cavitation probably present, and 5 = cavitation definitely present.

Observer sessions were held in the UNC School of Dentistry's dental student radiology interpretation room. Six dual-monitor workstations with Lenovo LT2252p monitors (Lenovo, Beijing, China) as the primary diagnostic display were available. The diagnostic displays underwent TCG-18 test pattern quality control checks prior to the sessions. Adjustments to monitor brightness and contrast settings were made to confirm visualization of both the 5% and 95% contrast levels in the test pattern. Ambient lighting in the room was subdued to appropriate interpretation levels. The principle investigator (PI) was present during all observer session to troubleshoot and clarify questions.

MiPACS Dental Enterprise Viewer 3.1.1401 (Medicore Imaging, Charlotte, NC) was used for the PSP bitewing interpretation software. Observers were allowed to use brightness and contrast settings but not allowed to use any available secondary imaging filters. CDR DICOM for Windows Version 5.4.1658.5883 (Sirona Dental, Salzburg, Germany) was used for the direct digital bitewing interpretation software. The general dentistry task with default 35% sharpening was selected for initial image display. Observers were instructed to use the sharpening adjustment and allowed to use the brightness/contrast adjustments. Observers were not allowed to change the image display task. DEXview Version 10.0.2 (Dexis LLC, Hatfield, PA) was used to view native panoramic bitewing .DEX image files. Observers were allowed to use brightness/contrast adjustments and the proprietary ClearVu filter. DEXview software was chosen instead of native Planmeca Romexis software because the private office operating the Planmeca unit employed Dexis software exclusively for their interpretation needs. Galaxis Version 1.9.4368.23294 (Sirona Dental, Salzburg, Germany) was used for the CBCT interpretation software. Observers were allowed to use any slice formatting and brightness/contrast adjustment tools for their interpretation.

The first observer session began with interpretation of images of four calibration teeth. These extracted teeth were independent of the study sample. The caries status of the calibration teeth confirmed by micro-CT was known by the PI prior to the session. Observers were asked to interpret images of the calibration teeth for all four modalities according to the three specific observer tasks. After interpretation, the PI discussed the known caries status relative to the observer responses and provided guidance to correct any interpretation errors.

Following interpretation and discussion of the calibration set of images, the observers completed interpretation of images of all 32 subject teeth imaged by all four modalities. The sequence of image/modality interpretation was prescribed so that no modality was biased towards the beginning or end of the session and no set of the same extracted teeth were interpreted in succession. Observers recorded their responses on provided forms. Breaks were allowed during the session to minimize observer fatigue. After a wash-out period of 2-3 weeks, observers returned to complete a second session. The observers interpreted images of exactly half (16 of 32) of the subject teeth imaged by all four modalities.

To complete the caries ground truth analysis, each micro-CT scan was reviewed by the PI and thesis mentor (DT) to establish caries status, lesion depth, and cavitation status of every proximal surface in the sample. Decision discrepancies were resolved by consensus following a discussion of the image findings.

Caries observation scores for each observer-modality combination were compiled with the ground truth caries presence status for each tooth using Excel Version 14.0.7143.5000 (Microsoft Office, Microsoft, Redmond, WA). Receiver operating characteristic (ROC) curves were constructed using a web-based ROC Analysis tool available from Johns Hopkins University School of Medicine (www.jrocfit.org), and fitted area under the curve ($A_z$) scores were recorded. Raw sensitivity and specificity scores were calculated for each observer-modality combination where caries presence ratings 4 and 5 were considered a positive diagnosis and caries ratings 1,2, and 3 a negative diagnosis.

As an exploratory measure to evaluate whether the tooth sample was too easy to differentiate subtle diagnostic differences between modalities, 4 surfaces with a rated caries presence of 5 for 95% or

more of all observations were excluded from the initial data set. For two of these surfaces, caries presence was 5 for 100% of observations. ROC curves were constructed from this truncated data set similarly to those created from the full data set.

Observer-modality combination $A_z$ scores for both full and truncated data sets were compared with a fixed-effects main effects two-way analysis of variance (ANOVA). A p-value < 0.05 was considered statistically significant for overall test of effects. Sensitivity and specificity scores were compared with a fixed-effects Freidman's two-way nonparametric analysis of variance (ANOVA) using observer and modality as main factors. A non-parametric approach was chosen based on the non-normal behavior of sensitivity and specificity scores. A p-value < 0.05 was considered statistically significant for overall test of effects. Appropriate t-test or Wilcoxon rank sum pairwise comparisons were made between modalities when statistical significance was observed between modalities. The Bonferroni adjusted significance criterion was used for post-hoc tests. Because there were six pairwise comparisons among the four modalities, $p < 0.05/6 = 0.0083$ was considered statistically significant.

Only true positive observations were considered for depth accuracy analysis. For each observer's depth observation on each modality, the level of agreement between the depth observation and depth assessment by micro-CT ground truth were calculated. The level of agreement with the ground truth measures the accuracy of the depth scores from the modality (how close the modality score is to the ground truth score). It was assumed that images were independent and that ratings were independent. Weighted kappa statistics were calculated and depth measures that were closer to agreement were weighted higher than those that were farther. Weights were as follows for analysis: 1 for perfect agreement, 0.66 for one off, 0.33 for two off, and 0 for 3 off. Weighted kappa coefficients for each observer were compared with one-way analysis of variance (ANOVA). A p-value < 0.05 was considered statistically significant for overall test of effects. Post-hoc t-test pairwise comparisons were made between modalities when statistical significance was observed between modalities. The Bonferroni adjusted significance criterion was used for post-hoc tests. Because there were six pairwise comparisons among the four modalities, $p < 0.05/6 = 0.0083$ was considered statistically significant.

Cavitation observation scores for each observer-modality combination were compiled with the ground truth caries presence status for each tooth using Excel Version 14.0.7143.5000 (Microsoft Office, Microsoft, Redmond, WA). ROC analysis of the cavitation data was discarded because of the limited number of known cavitated lesions in the sample and the corresponding inability to construct meaningful ROC curves. Raw sensitivity and specificity scores were calculated for each observer-modality combination where a caries presence ratings of 4 and 5 was considered a positive diagnosis and a caries ratings of 1, 2, and 3 a negative diagnosis. Sensitivity and specificity scores were compared with a fixed-effects Freidman's two-way nonparametric analysis of variance (ANOVA) using observer and modality as main factors. A non-parametric approach was chosen based on the non-normal behavior of sensitivity and specificity scores. A p-value $< 0.05$ was considered statistically significant for overall test of effects. Wilcoxon rank sum pairwise comparisons were made between modalities when statistical significance was observed between modalities. Bonferroni adjusted significance criterion was used for post-hoc tests where because there were six pairwise comparisons among the four modalities, a $p < 0.05/6 = 0.0083$ was considered statistically significant.

The observation scores from the second session were used to determine the intraobserver agreement. For each of every observer's tooth surface assessment (presence, depth, and cavitation) the level of agreement between the first and second session observations were calculated. It was assumed that images were independent and that ratings were independent. Weights for analysis were as follows: 1 for perfect agreement, 0.75 for one off, 0.5 for 2 off, 0.25 for 3 off, 0 for 4 off. Weighted kappa statistics were calculated and observations that were closer to agreement were weighted higher than those that were further. Kappa values greater than or equal to 0.4 were considered moderate agreement and scores greater than or equal to 0.8 were considered strong agreement. Confidence intervals for each kappa (95%) were produced and any interval that contains 0 suggested that there is no relevant agreement at the 0.05 significance level. The median weighted kappa for each modality was determined. Weighted kappa coefficients for each observer were compared with one-way analysis of variance (ANOVA). A p-value $< 0.05$ was considered statistically significant for overall test of effects. Post-hoc t-test pairwise comparisons were made between modalities when statistical significance was observed

between modalities. Bonferroni adjusted significance criterion was used for post-hoc tests. Because there were six pairwise comparisons among the four modalities, $p < 0.05/6 = 0.0083$ was considered statistically significant. Analyses were done using SAS software (version 9.3, SAS Institute Inc., Cary, NC).

**Results**

A photograph and volumetric rendering of the phantom mandible imaging setup can been seen in Figure 1. Example images of extracted teeth and caries can be seen in Figure 2. Micro-CT analysis indicated that of 64 total proximal surfaces, 34 surfaces had caries lesions and 30 surfaces were sound. Of the 34 surfaces with lesions, 9 extended into the outer half of enamel, 9 into the inner half of enamel, 14 into the outer half of dentin, and 2 into the inner half of dentin. 12 of the 34 surfaces with lesions demonstrated breakdown of the outer layer of enamel and were classified as cavitated. These findings are summarized in Table 1.

$A_Z$ scores for both full and truncated sets are provided in Table 2 and Table 3 with accompanying compiled ROC curves per modality for full and truncated sets in Figure 3 and Figure 4, respectively. Caries detection sensitivity and specificity scores are provided in Table 4. For the full data set, fixed-effects main effects two-way ANOVA of $A_z$ scores indicated no significant difference between observers ($p=0.15$) and modalities ($p=0.10$). Subsequent analysis of the truncated data set also indicated no significant differences between observers ($p=0.22$) and modalities ($p=0.12$). Fixed-effects Friedman's two-way non-parametric ANOVA of caries presence sensitivity scores for the full data set indicated significant differences between observers ($p<0.0001$) and modalities ($p=0.0002$). Similar analysis of specificity scores indicated significant differences between observers ($p=0.0005$) and modalities ($p=0.0016$). Subsequent Wilcoxon rank sum pairwise comparisons between modalities indicated that PSP (0.57) demonstrated a higher sensitivity than Schick33 (0.48) ($p=0.0018$), XG3D demonstrated a higher sensitivity than Schick33 (0.48) ($p<0.0001$), and XG3D (0.62) demonstrated a higher sensitivity than PanBW (0.53) ($p=0.0071$). Schick33 (0.96) demonstrated higher specificity than PanBW (0.86)

(p=0.0005) and XG3D (0.97) demonstrated a higher specificity than PanBW (0.86) (p=0.0013). A summary of statistical findings for caries detection is provided in Table 5.

Of the 64 surfaces, the range of true positive depth scores was 14 to 25 for PSP, 9 to 23 for the Schick33, 8 to 25 for PanBW, and 17 to 26 for the XG3D. Weighted kappa values for each observer's depth score, based on images from each modality, are displayed in Table 6, along with median weighted kappa values by modality. Scores above 0.4 are considered moderately accurate while scores above 0.8 are highly accurate. Observers whose 95% confidence intervals for the weighted kappa includes 0 do not have significant accuracy of depth measurement for the modality. A visual comparison of the weighted kappa values among modalities is displayed in Figure 5. The median accuracy of the observer-assigned depth scores is moderately high for PSP (K = 0.40), Schick 33 (K = 0.43), and XG3D, with XG3D having the best accuracy (K = 0.63). The PanBW modality is considered fairly inaccurate (K = 0.27). One-way ANOVA of kappa coefficients indicated significant differences between modalities (p=0.003). Subsequent t-test pairwise comparisons indicated that XG3D had significantly better depth agreement than PanBW (p=0.002). A summary of statistical findings for the depth analysis is provided in Table 7.

Cavitation detection sensitivity and specificity scores are provided in Table 8. For cavitation sensitivity scores, fixed-effects Freidman's two-way nonparametric ANOVA indicated significant differences between observers (p=0.0063) and modalities (p<0.0001). In cavitation specificity scores, there were similar significant differences between modalities (p=0.0025) but no significant differences between observers (p=0.1829). Subsequent Wilcoxon rank sum pairwise comparisons indicated that XG3D (0.83) had higher sensitivity than PSP (0.54), Schick33 (0.47), and PanBW (0.38) (p=0.0011, <0.0001, <0.0001). XG3D (0.96) had significantly lower specificity than PSP (0.99) and Schick33 (0.99). Cavitation detection statistical findings are summarized in Table 9.

Similarly, the weighted kappa values for each observer's caries presence, depth, and cavitation scores, based on images from each modality, are displayed in Tables 10, 11, and 12, respectively, along with median weighted kappas by modality. Scores above 0.4 are considered moderately accurate while scores above 0.8 are highly accurate. All but one observer-modality combinations had at least moderate intraobserver agreement (K>0.40) for caries presence and depth questions. All but three observer-

modality combination had at least moderate intraobserver agreement (K>0.40) for the cavitation question. Overall, based on median weighted kappa, intraobserver agreement of caries presence is moderately high for each modality (PSP: K = 0.62; Pan BW: K = 0.62; XG3D: K = 0.65). Intraobserver agreement for caries depth is moderately high for each modality (PSP: K = 0.76; Schick 33: K = 0.74; XG3D: K = 0.73). Intraobserver agreement of caries cavitation is moderately high for each modality (PSP: K = 0.69; Schick 33: K = 0.70; XG3D: K = 0.72). Box plots of intraobserver agreement by modality for caries presence, depth, and cavitation questions are provided in Figures 6, 7, and 8. One-way ANOVA of kappa coefficients indicated no significant differences between modalities for caries presence, depth, and cavitation ratings (p=0.411, 0.376, 0.197, respectively) (Table 13).

**Discussion**

ROC analysis was chosen as the primary tool for discriminating caries detection abilities between the different modalities because this analysis removes the effect of differences in observers' decision thresholds and provides the best overall indication of diagnostic accuracy.[64-66] ROC analysis of caries diagnosis indicated no significant differences between modalities in terms of their $A_Z$ scores. This suggests that all four modalities performed equally with respect to identifying lesions. One key nuance of ROC analysis, however, is that two separate ROC curves may have very different curve shapes but the exact same $A_Z$. The difference between curve shapes occurs as a result of a tradeoff between sensitivity and specificity. Irrelevant to overall diagnostic accuracy, this tradeoff between sensitivity and specificity can be clinically relevant. For caries detection, the tradeoff between sensitivity and specificity is indeed important. It is much more beneficial to the patient and provider to maximize specificity with minimal penalty in corresponding sensitivity because a false positive lesion can result in unnecessary surgical treatment of the tooth at additional cost to the patient and significant detriment to the tooth's future restorative prognosis.[7] Furthermore because caries is increasingly a slowly-progressing disease, the small penalty in decreased sensitivity can be mitigated by detection of the lesion at a later patient recall date when the lesion is slightly larger.[7] To reconcile these points with our previously stated desires for

increased sensitivity, the authors reiterate that we desire improved sensitivity from new diagnostic systems with no corresponding penalty in the already high specificities.

According to this rationale, subsequent analysis of modality sensitivity and specificity scores elucidated potentially relevant clinical differences in performance. PSP (0.57) demonstrated a higher sensitivity than Schick33 (0.48), XG3D demonstrated a higher sensitivity than Schick33 (0.48), and XG3D (0.62) demonstrated a higher sensitivity than PanBW (0.53). Schick33 (0.96) demonstrated higher specificity than PanBW (0.86) and XG3D (0.97) demonstrated a higher specificity than PanBW (0.86). Per the discussion of ROC curves above, these differences in sensitivity and specificity between modalities can be appreciated in Figure 3 and Figure 4. Finding significant differences between observers in the caries and cavitation sensitivity/specificity analyses was not surprising because these measures of diagnostic accuracy are dependent on the individual observer's decision threshold. This observation reinforces the importance of ROC analysis as a means to assessing overall diagnostic efficacy. The specific ANOVA analyses for sensitivity/specificity results were chosen because they address observer variation before making a determination of significant modality effects.

The tendency of Schick 33 to have lower sensitivity compared to other modalities may be attributable to the dynamic image sharpening slider. Based on observer comments, it was found that different observers had different preferences for the level of sharpness. The immediate adjustment of image sharpness during interpretation may have distracted or otherwise influenced the observers into missing very subtle, difficult to detect lesions resulting in somewhat decreased sensitivity. Alternatively the filter may have truly obscured lesions, making it impossible more difficult to see. However, this study was not designed to test the specific influence of the image sharpening in the Schick 33 system, therefore no definitive conclusions about the sharpening system itself can be made at this time. Overall the Schick 33 system proved equivalent to the other modalities in the ROC analysis, and demonstrated no significant decrease in specificity. Thus, the Schick 33 system demonstrates equivalent overall diagnostic accuracy and specificity, but lower sensitivity.

The tendency of panoramic bitewing images to have lower specificity compared to other modalities may be attributable to image "ghost" artifacts from the contralateral aspect of the mandible

formed as a result of image acquisition geometry, similar to that in traditional panoramic imaging.[24-26] Observers commented that horizontal streak artifact from the contralateral body of the phantom mandible were superimposed over the teeth of interest, resulting in the perception that image interpretation was more challenging. It is possible that these artifacts may have mimicked lesions in a number of cases, resulting in an elevated false positive rate. The fact that the panoramic bitewing modality proved equivalent to other modalities in ROC analysis suggests promise as an effective and patient-comfortable imaging modality, the authors are concerned that a potentially elevated false positive rate could have detrimental effects associated with treatment decisions.

The XG3D cone beam CT system tended to have higher sensitivity than other modalities while incurring no penalty in specificity. This finding is different than previous studies which cited beam hardening artifacts from nearby enamel as a source of decreased specificity (increased false positive diagnoses).[30, 47] One possible explanation of reduced false positive lesions is that the XG3D was operating in HD mode with the MARS reconstruction algorithm. While the MARS algorithm is designed to reduce artifacts from metal restorations in a scan, it is possible that the algorithm may also reduce artifacts from dense tooth structure. An alternative explanation for the reduction in false positive lesions for the XG3D is the investigators' careful demonstration of beam hardening artifacts versus real caries lesions to the observers. This demonstration may have educated the observers to more consistently avoid false positive diagnoses of artifacts as lesions. Unfortunately, this study was designed to test the overall efficacy of the XG3D system compared to other modalities, not the specific influence of MARS with and without metal objects in the scan or the specific influence of educating observers on artifact recognition. This study also did not examine any differences between far-field and near-field artifact reduction by MARS, when it is known that far-field artifacts are more controllable yet near-field artifacts are possibly more relevant to tooth pathology diagnosis including dental caries.[67]

Regarding true positive depth analysis, the median accuracy of the observer-assigned depth scores as compared to the ground truth of the micro-CT was classified as fairly inaccurate for PanBW (0.27), and moderate at best for PSP (0.40) and Schick33 (0.43). The XG3D modality, although still only moderately high (0.63), had significantly more accurate depth scores than the PanBW modality. These

findings agree with previous studies showing that radiographic assessment of lesion depth with bitewing radiographs is only fairly accurate,[52, 53] and that CBCT systems allow better depth assessment.[34, 56]

Another analysis approach that was considered to measure the accuracy of true positive depth scores was a correlation coefficient. The weighted kappa strategy was chosen over correlations because it better captures how well the modality depth scores actually match the micro-CT scores, and not just a general relationship trend. When assessing accuracy of one categorical variable by another, when each has as few as 5 levels, scores that are off by more than 1 are not considered matches, meaning that such a score from the modality is inaccurate. For example, the ratings from observer 8 on the PSP modality gives a Spearman correlation of 0.71, which supports a linear relationship, and a weighted kappa of 0.338, which suggests that there is little agreement. In fact, a scatter plot of the PSP depth scores by the micro-CT score for this observer shows an increasing trend, thus reasonable correlation, but the markers are quite spread out, meaning that the scores do not often actually match.

The XG3D CBCT system demonstrated markedly better detection of cavitated lesions compared to the other modalities. The XG3D cavitation sensitivity of 0.83, compared to PSP (0.54), Schick 33 (0.47), and Pan BW (0.38) represents an approximately 30% increase in sensitivity over conventional 2D modalities. This finding is highly consistent with other recent studies, some of which include images influenced by metal artifact.[2, 32, 61] Our statistical analysis indicated that XG3D (0.96) had significantly lower specificity than PSP (0.99) and Schick (0.99), however; these reported specificities are all extremely high. The authors conclude that the statistically significant differences found regarding cavitation specificity are clinically irrelevant because (1) the already very high specificity of the "statistically different" XG3D system, (2) the very poor sensitivities achieved by the other modalities, and (3) the relatively limited number of cavitated lesions in the sample (12 of 64 surfaces), which biased results towards high specificities. Based on the findings in this study and other recent studies,[2, 32, 58] it seems apparent that CBCT affords better detection of cavitated proximal lesions. Evidence suggests that this relationship holds true even in the presence of some metal artifacts.[2, 58] Coupled with the increased clinical importance of lesion cavitation status as a marker for potential progression,[57-59] CBCT may represent an important tool for caries assessment.

Almost all observer-modality combination weighted-kappa correlation scores were at least moderately agreeing (K>0.40) with most scores demonstrating moderate to high agreement in Tables 10, 11, and 12 and Figures 6, 7, and 8, respectively. A review of similar multireader ROC studies cites a range of intraobserver agreement from 0.35 to 0.59,[68] confirming typical results in this study. No statistically significant differences in intraobserver correlations scores were observed between modalities for the caries presence, caries depth, and caries cavitation questions. This finding indicates that our observers performed uniformly modality-to-modality throughout the study.

CBCT imaging incurs much greater radiation dose, financial cost, and acquisition/interpretation time compared to other imaging modalities. The typical radiation dose for a CBCT scan varies widely from several micro Sieverts (μSv) up to more than 1000 μSv depending primarily on selected scan resolution and FOV.[69-71] The Sirona Orthophos XG3D protocol used in this study had an effective dose of 53 μSv per mandibular posterior quadrant scan.[71] Conversely, the dose from a standard set of four PSP or digital bitewing images is about 5 μSv, and dose from a standard panoramic image (approximately equal to a panoramic bitewing dose) is about 15 μSv.[69] The financial cost of CBCT imaging is frequently in the hundreds of dollars, compared to tens of dollars for bitewing or panoramic/panoramic bitewing exams. The total acquisition and interpretation time for a CBCT may be 10 minutes or more, compared to total acquisition and interpretation time of bitewing radiographs and panoramic/panoramic bitewings lasting 1-2 minutes or much less. These tradeoffs (or lack thereof) of diagnostic efficacy versus various costs must be carefully considered before selecting the imaging modality to address the relevant diagnostic question. With regard to caries detection, it is the authors' opinion that the marginal increase in lesion sensitivity, marginal increase in lesion depth assessment, and even the drastic improvement in lesion cavitation detection may not be worth the significant increases in dose, cost, and time. Of course, if a CBCT is taken for other clinical indications it appears prudent to assess the visible regions of teeth for caries.

Lack of statistical power is an unlikely explanation for why this study found no statistically significant modality or observer effects on $A_Z$ scores from ROC analysis. An estimated sample of 51 surfaces is recommended for a "paired-case, paired-reader" study design[72] with the following power factors: $A_Z$ analysis, moderate observer accuracy, a moderate difference in system accuracy, 1:1 case

frequency, a large level of observer variability, and ten observers.[73] Slightly elevated intraobserver variation in our study compared to this cited power calculation was likely balanced by the modest increase in number of cases to 64. Furthermore, the authors' choice of power factors in this study was to deliberately detect clinically relevant differences in detection.[73, 74] With believe elevated interobserver variation in this study was typical and expected because a review of similar multireader ROC studies found that reader variability is almost always larger than modality variation.[68]

A second explanation for why this study found no statistical significant modality or observer effects on $A_Z$ scores is that the sample of lesions was "too easy" to diagnose, resulting in all observations reaching some non-differentiating maximum level. The authors consider mean $A_Z$ scores per modality ranging from 0.780 to 0.828 as somewhat high relative to other similar proximal caries studies. One might argue that the level of lesion difficultly was not challenging enough to reveal subtle differences in diagnostic capability between modalities. The authors suggest in response that the relatively well-balanced spread of lesion size used in this study represents a clinically relevant sample of lesion sizes and accordingly our assessment of overall diagnostic capability by ROC applicable to clinically relevant lesions. In other words, intentionally increasing the sample difficulty to reveal subtle modality differences could result in a misrepresentation of clinically relevant modality differences. As an exploratory measure to investigate the influence of "easy" lesions on our sample, four readily diagnosed lesions were excluded from the sample and the ROC analysis was performed on the truncated sample. Subsequent statistical results remained unchanged, providing additional evidence that our lesion sample difficulty was not likely a significant confounding factor in the study results.

A limitation of this study was the ideal conditions established for proximal caries diagnosis. These ideal conditions were chosen to best control confounding variables and isolate the actual differences in detection efficacy between modalities, but they may not represent clinically relevant interpretation scenarios. Interpretation took place under ideal viewing conditions with subdued ambient lighting, quiet surroundings, and contrast calibrated monitors. A more relevant interpretation environment may be in a well-lit operatory in a busy clinic. Additionally, the observers were qualified experts in dental radiologic diagnosis, each having specialty training in oral and maxillofacial radiology and a detailed knowledge of

radiographic caries diagnosis. A typical caries diagnostician may be a general dentist with only general-dental level training in radiographic diagnosis. Also, all images were acquired with ideal image geometry resulting in open proximal contacts, whereas clinical imaging of real patients dictates that it is not always possible to open all posterior proximal contacts. Non-physiologic point contacts established by the random selection of extracted teeth may not mimic true clinical imaging either, because in real patients, proximal tooth contacts demonstrate an element of broadness not perfectly simulated in our experimental system. Of note, the investigators found that acquiring panoramic bitewing images with open contacts using the phantom mandible was significantly more difficult than the other modalities. Once the correct focal layer "jaw shape" was established for the phantom mandible to reliably open posterior contacts, it was observed that there were only several millimeters of tolerance in mandible position that maintained open proximal contacts. While the panoramic bitewing images indicated overall diagnostic efficacy equivalent to the other modalities, it is unclear how frequently closed contacts occur during clinical imaging and how much those closed contacts would influence caries diagnosis. Given the difficulty experienced with opening contacts with a phantom mandible under controlled conditions, the investigators suspect that closed contacts may be a significant diagnostic obstacle specific to extraoral imaging for caries detection. Additional studies are needed to further clarify how readily panoramic bitewing systems open contacts and the resulting effect on caries detection. Not surprisingly the CBCT scans were very insensitive to phantom mandible position with respect to opening contacts because isotropic 3D imaging allows investigation of any contact area regardless of tooth position.

The traditional gold-standard method to validate dental caries status is visual examination of thin section histology, however, in recent years, micro-CT has proven to be an effective alternative. Numerous studies have employed and found success with micro-CT as a ground truth caries validation tool with purposes including simple caries detection,[75, 76] evaluating lesion depth,[55] evaluating progression of lesion morphology,[77] calculating lesion mineral concentration,[78, 79] and validating caries removal techniques.[80-82] Micro-CT has also been shown to agree with transverse microradiography, another accepted gold-standard technique for the evaluation of caries.[83] One study reported that micro-CT does not agree with histological validation at different disease severities and cites small enamel lesions as

having the greatest discrepancy between histology and micro-CT results. They explain that two scenarios associated with histology (not micro-CT) may confound accurate caries identification. First, it is very difficult to create a thin section exactly at the location of a small caries lesion confined to the enamel, and second, that the procedure of sectioning the tooth may cause loss of tooth and/or lesion material. [84] Micro-CT on the other hand allowed unimpeded evaluation of the entire tooth anatomy, including very small lesions. Another study reports that micro-CT generally underestimates caries lesion depth compared to histological assessment.[85] However, tooth discoloration/stain due to infiltrative bacterial activity may confound the assessment of lesion depth with histology and may not represent the true extent of decalcification. Conversely, micro-CT evaluates mineral decalcification directly. The selection of 70 kVp tube potential in this study's micro-CT analysis is supported by the literature[86, 87] and system calibration efforts outlined in the materials and methods section support well-calibrated imaging. It is the authors' opinion that this study's micro-CT analysis functioned as an appropriate and valid ground-truth modality, based on the body of research in support of micro-CT for a caries validation technique, the problems with missed lesions or lost tooth structure associated with histology slice preparation, the possible confounding factor of tooth stain associated with histology slice analysis, and our selected system technique factors.

In conclusion, this study found equivalent overall diagnostic efficacy for posterior proximal caries detection between three new and one control dental radiographic systems: the Schick 33 intraoral direct digital sensor (Sirona Dental, Salzburg, Germany) employing a new dynamic image sharpening adjustment, a Planmeca ProMax® panoramic unit (Planmeca Inc., Helsinki, Finland) operating in extraoral "panoramic bitewing" mode, a Sirona Orthophos XG3D CBCT (Sirona Dental, Salzburg, Germany) employing Metal Artifact Reduction Software (MARS), and the PSP bitewing system utilized in the UNC School of Dentistry radiology clinics. Additional findings from this study clarify potentially important differences between these modalities with respect to their overall caries detection:

1. The Schick 33 system with its dynamic image sharpening adjustment has lower lesion detection sensitivity with no significant decrease in lesion specificity. It is unknown at this time how the extent of imaging sharpening may influence caries detection.

2. The Planmeca panoramic bitewing images show decreased specificity compared to other modalities, possibly due to confounding streak or "ghost" artifacts from the contralateral body of the mandible. The resulting increase in false positive diagnoses may be a significant concern regarding unnecessary surgical tooth procedures. Panoramic bitewings appear to be less accurate than CBCT for assessing lesion depth. Difficulty in acquiring Planmeca panoramic bitewing images with open proximal contacts during this study may be an indicator of frequently closed contacts during clinical imaging. More studies are needed on panoramic bitewing systems to clarify how often proximal contacts are closed and the influence on caries detection.
3. The Sirona Orthophos XG3D CBCT shows increased sensitivity compared to other modalities with no significant penalty in specificity. It is unclear whether this system's artifact reduction algorithms or increased observer training results in a reduced number of false positive diagnoses. The XG3D also demonstrates more accurate lesion depth assessment than the Planmeca panoramic bitewing images and substantially better detection of cavitated lesions compared to all other modalities. These diagnostic differences may be clinically relevant, but their value in overall treatment decisions and patient care must be weighed with the substantially increased radiation dose, financial cost, and acquisition/interpretation time associated with CBCT. Given these study findings, it appears prudent that if a CBCT scan is taken for other clinical reasons, proximal regions of teeth visible in the scan that are not corrupted by significant metal artifact should be evaluated for clinically significant caries lesions and cavitation. Additional studies are needed to evaluate the specific influence of MARS on image artifacts with and without the presence of metal objects in the scan (including restorations) and any differences in management of near field versus far-field artifacts.
4. This study was performed with expert observers under ideal viewing conditions with ideal image geometry. These features of study design were chosen to best evaluate the diagnostic capabilities of the experimental modalities. Accordingly, the results may not apply to all clinical scenarios and environments.

**Tables and Figures**

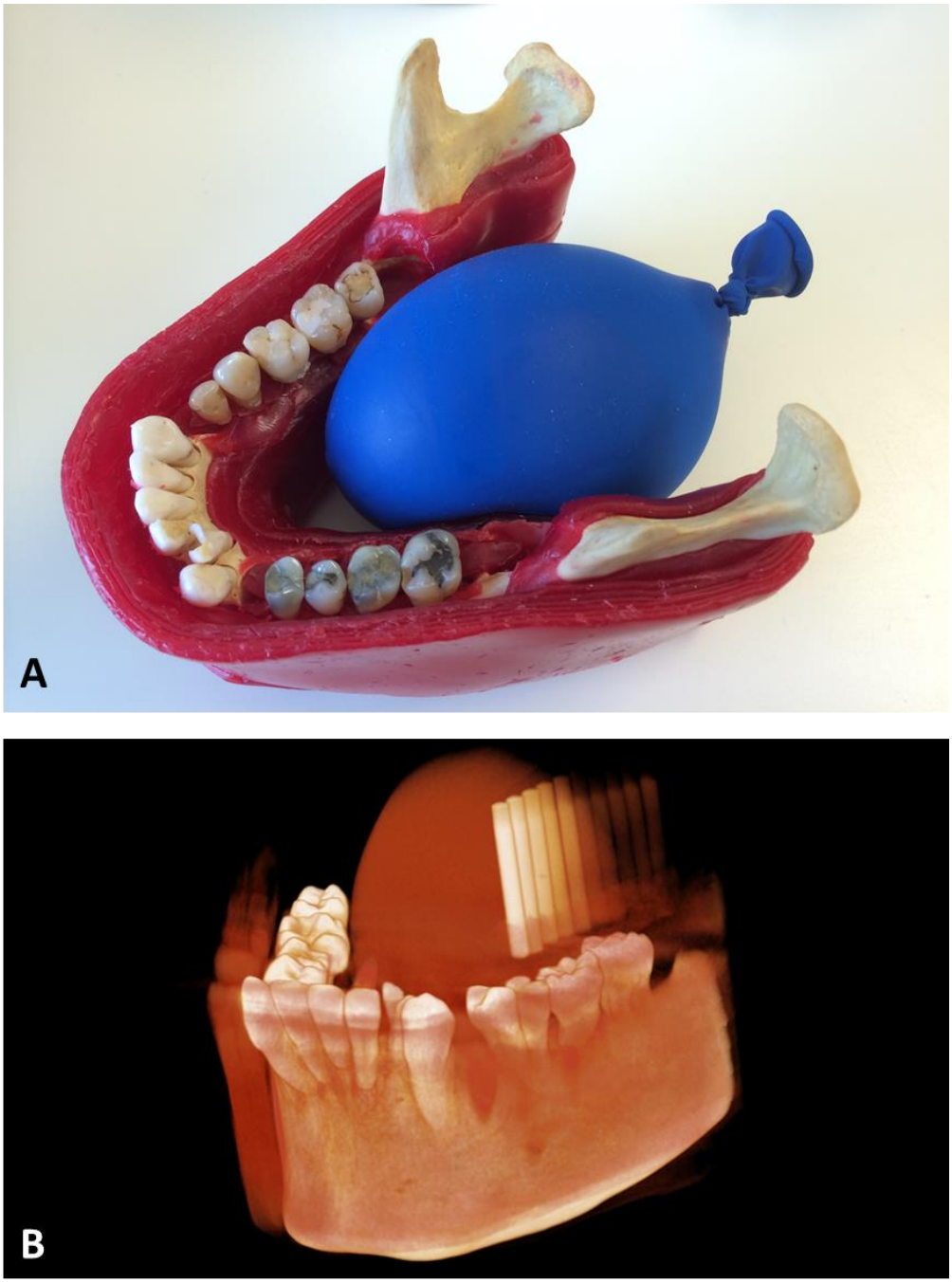


Fig. 1. Photograph of phantom mandible with wax and latex water-filled balloon (A), Volumetric rendering of large FOV CBCT scan of phantom setup (B), note there were no metal or extensive composite restorations present in the tooth sample. Rods present in the volume rendering are relevant to manuscript 2.

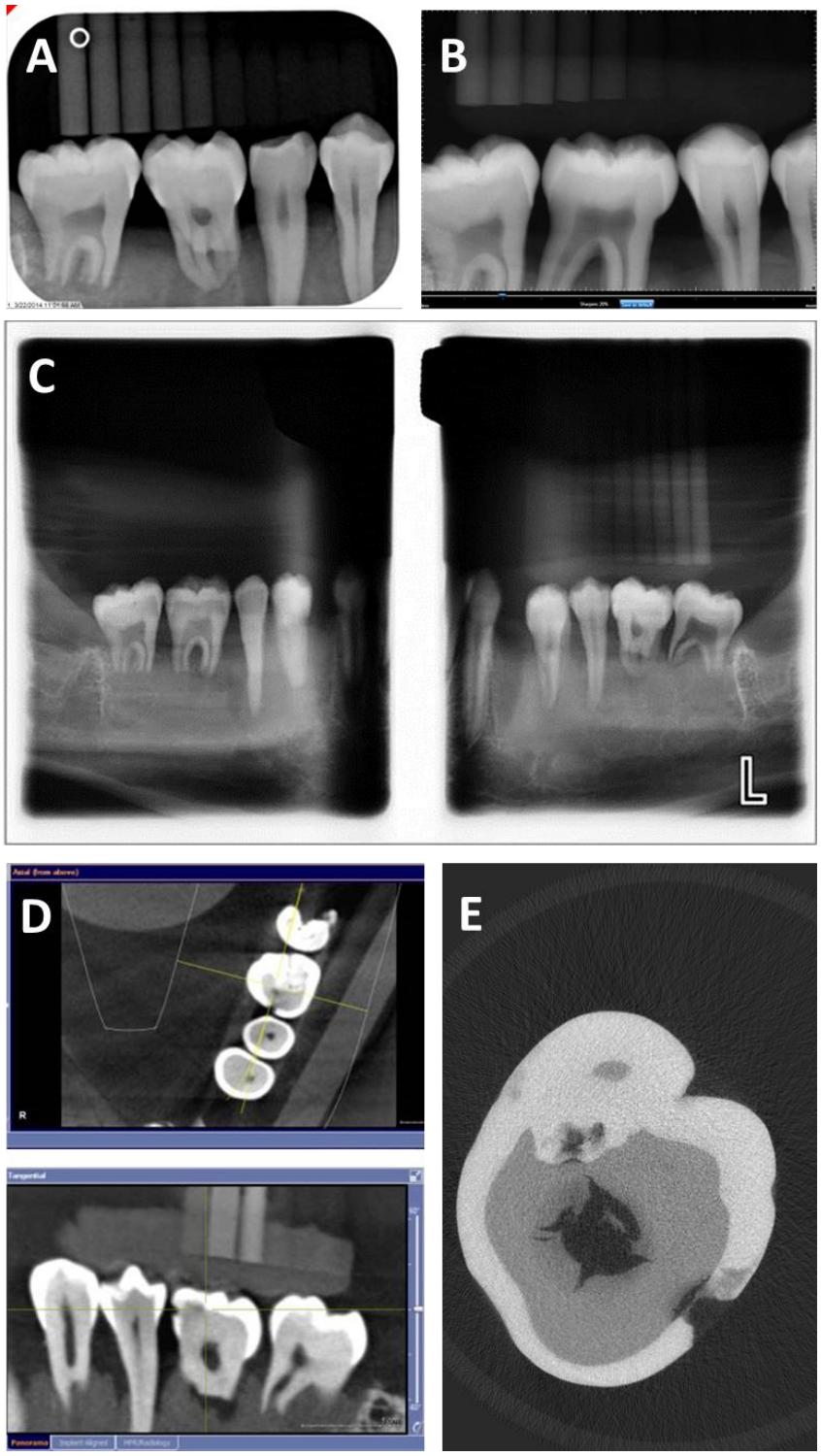


Fig. 2. Representative images of extracted teeth and caries for PSP (A), Schick 33 (B), Planmeca panoramic bitewing (C), Sirona Orthophos XG3D CBCT (D), and micro-CT ground truth (D). Calibrations rods visualized in images are relevant to manuscript 2.

**Table 1.** Proximal surface ground truth status as assessed by micro-CT

| *Caries* | *Surfaces* | *Depth** | *Surfaces* | *Cavitation* | *Surfaces* |
|---|---|---|---|---|---|
| Sound | 30 | 1 | 30 | Absent | 52 |
| Lesion | 34 | 2 | 9 | Present | 12 |
| | | 3 | 9 | | |
| | | 4 | 14 | | |
| | | 5 | 2 | | |
| **Total** | 64 | | 64 | | 64 |

*Lesion depth scale:

1 = caries not present

2 = caries involving the outer half of enamel

3 = caries involving the inner half of enamel

4 = caries involving the outer half of dentin

5 = caries involving the inner half of dentin

**Table 2.** Full set $A_Z$ scores derived from ROC curves for each modality-observer combination, mean $A_Z$ scores with standard deviations, and compiled* $A_Z$ for each modality and each observer

| | *Modality* | | | | | |
|---|---|---|---|---|---|---|
| *Observer* | *PSP* | *Schick33* | *PanBW* | *XG3D* | *Mean* | *SD* |
| 1 | 0.800 | 0.725 | 0.757 | 0.829 | 0.778 | 0.046 |
| 2 | 0.680 | 0.690 | 0.818 | 0.781 | 0.742 | 0.068 |
| 3 | 0.821 | 0.856 | 0.746 | 0.858 | 0.820 | 0.052 |
| 4 | 0.868 | 0.803 | 0.781 | 0.802 | 0.814 | 0.038 |
| 5 | 0.846 | 0.729 | 0.879 | 0.832 | 0.822 | 0.065 |
| 6 | 0.798 | 0.762 | 0.695 | 0.852 | 0.777 | 0.066 |
| 7 | 0.810 | 0.816 | 0.755 | 0.872 | 0.813 | 0.048 |
| 8 | 0.899 | 0.790 | 0.870 | 0.842 | 0.850 | 0.046 |
| 9 | 0.812 | 0.814 | 0.780 | 0.842 | 0.812 | 0.025 |
| 10 | 0.865 | 0.816 | 0.828 | 0.773 | 0.821 | 0.038 |
| *Mean* | 0.820 | 0.780 | 0.791 | 0.828 | 0.805 | 0.023 |
| *SD* | 0.060 | 0.052 | 0.058 | 0.033 | 0.031 | |
| *Compiled* | 0.795 | 0.758 | 0.763 | 0.811 | 0.782 | |

Abbreviations: ROC, receiver operating characteristic curve, $A_Z$ area under the ROC curve
*Compiled $A_Z$ scores for each modality including all observers and for each observer for all modalities

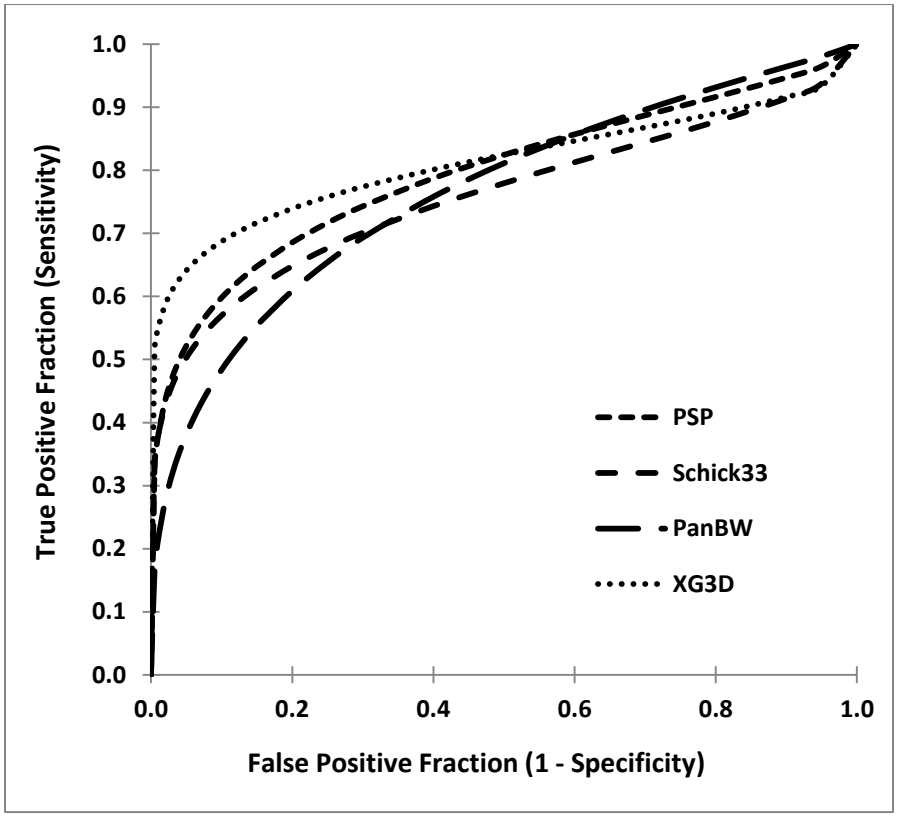


Fig. 3. Full set ROC curves for each modality's compiled observations.

**Table 3.** Truncated set $A_Z$ scores derived from ROC curves for each modality-observer combination, mean $A_Z$ scores with standard deviations, and compiled* $A_Z$ for each modality and each observer

| | *Modality* | | | | | |
|---|---|---|---|---|---|---|
| *Observer* | *PSP* | *Schick33* | *PanBW* | *XG3D* | *Mean* | *SD* |
| 1 | 0.796 | 0.688 | 0.766 | 0.817 | 0.767 | 0.057 |
| 2 | 0.640 | 0.660 | 0.793 | 0.752 | 0.711 | 0.073 |
| 3 | 0.796 | 0.837 | 0.712 | 0.839 | 0.796 | 0.059 |
| 4 | 0.855 | 0.784 | 0.756 | 0.776 | 0.793 | 0.043 |
| 5 | 0.826 | 0.693 | 0.864 | 0.814 | 0.799 | 0.074 |
| 6 | 0.771 | 0.733 | 0.662 | 0.833 | 0.750 | 0.072 |
| 7 | 0.785 | 0.794 | 0.724 | 0.855 | 0.790 | 0.054 |
| 8 | 0.886 | 0.762 | 0.853 | 0.821 | 0.831 | 0.053 |
| 9 | 0.787 | 0.789 | 0.749 | 0.821 | 0.787 | 0.029 |
| 10 | 0.848 | 0.796 | 0.805 | 0.742 | 0.798 | 0.044 |
| *Mean* | 0.799 | 0.754 | 0.768 | 0.807 | 0.782 | 0.025 |
| *SD* | 0.067 | 0.058 | 0.063 | 0.038 | 0.033 | |
| *Compiled* | 0.769 | 0.730 | 0.735 | 0.788 | | |

Abbreviations: ROC, receiver operating characteristic curve, $A_Z$ area under the ROC curve

*Compiled $A_Z$ scores for each modality including all observers and for each observer for all modalities

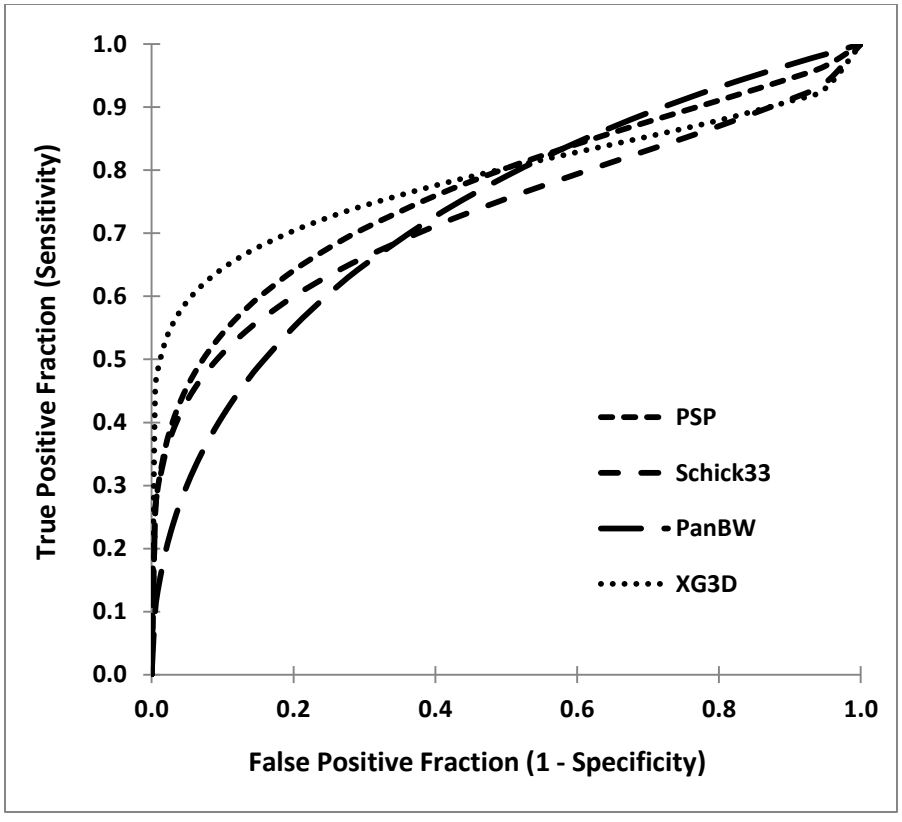


Fig. 4. Truncated set ROC curves for each modality's compiled observations.

**Table 4.** Caries detection sensitivity and specificity scores

| | *Modality* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *PSP* | | *Schick33* | | *PanBW* | | *XG3D* | | *Mean* | | *SD* | |
| *Observer* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* |
| 1 | 0.41 | 0.97 | 0.26 | 1.00 | 0.24 | 1.00 | 0.53 | 0.97 | 0.36 | 0.98 | 0.14 | 0.02 |
| 2 | 0.41 | 1.00 | 0.32 | 1.00 | 0.44 | 0.90 | 0.50 | 1.00 | 0.42 | 0.98 | 0.07 | 0.05 |
| 3 | 0.71 | 0.70 | 0.68 | 0.90 | 0.62 | 0.87 | 0.74 | 0.97 | 0.68 | 0.86 | 0.05 | 0.11 |
| 4 | 0.56 | 0.93 | 0.50 | 1.00 | 0.53 | 0.90 | 0.50 | 1.00 | 0.52 | 0.96 | 0.03 | 0.05 |
| 5 | 0.56 | 1.00 | 0.35 | 1.00 | 0.56 | 0.93 | 0.53 | 0.93 | 0.50 | 0.97 | 0.10 | 0.04 |
| 6 | 0.74 | 0.60 | 0.68 | 0.80 | 0.74 | 0.50 | 0.76 | 0.97 | 0.73 | 0.72 | 0.04 | 0.21 |
| 7 | 0.44 | 1.00 | 0.50 | 1.00 | 0.47 | 0.93 | 0.59 | 1.00 | 0.50 | 0.98 | 0.06 | 0.03 |
| 8 | 0.65 | 0.93 | 0.50 | 0.97 | 0.59 | 0.90 | 0.71 | 0.93 | 0.61 | 0.93 | 0.09 | 0.03 |
| 9 | 0.65 | 0.90 | 0.59 | 1.00 | 0.65 | 0.73 | 0.65 | 0.97 | 0.63 | 0.90 | 0.03 | 0.12 |
| 10 | 0.59 | 0.93 | 0.41 | 0.97 | 0.50 | 0.93 | 0.74 | 1.00 | 0.56 | 0.96 | 0.14 | 0.03 |
| *Mean* | 0.57 | 0.90 | 0.48 | 0.96 | 0.53 | 0.86 | 0.62 | 0.97 | | | | |
| *SD* | 0.12 | 0.14 | 0.14 | 0.07 | 0.14 | 0.14 | 0.11 | 0.03 | | | | |
| *Compiled** | 0.57 | 0.90 | 0.48 | 0.96 | 0.53 | 0.86 | 0.62 | 0.97 | | | | |

Abbreviations: SD standard deviation, TPR true positive rate (sensitivity), TNR true negative rate (specificity)

*Compiled sensitivity/specificity scores for each modality including all observers and for each observer for all modalities

**Table 5**. Summary of statistical findings comparing caries presence $A_Z$ scores for complete, truncated sets and caries sensitivity/specificity scores

| *Data set* | *Effect* | *p-value* | *Pairwise comparison* | *p-value* |
|---|---|---|---|---|
| $A_Z$ scores complete set[1] | Observer | 0.15 | NA | |
| | Modality | 0.10 | | |
| $A_Z$ scores truncated set[1] | Observer | 0.22 | NA | |
| | Modality | 0.12 | | |
| Caries sensitivity[2] | Observer | <0.0001* | PSP vs PanBW | 0.23 |
| | Modality | 0.0002* | PSP vs Schick33 | 0.0018* |
| | | | PSP vs XG3D | 0.10 |
| | | | PanBW vs Schick33 | 0.033 |
| | | | PanBW vs XG3D | 0.0071* |
| | | | Schick33 vs XG3D | <0.0001* |
| Caries specificity[2] | Observer | 0.0005* | PSP vs PanBW | 0.10 |
| | Modality | 0.0016* | PSP vs Shick33 | 0.034 |
| | | | PSP vs XG3D | 0.07 |
| | | | PanBW vs Schick33 | 0.0005* |
| | | | PanBW vs XG3D | 0.0013* |
| | | | Schick33 vs XG3D | 0.73 |

*Denotes statistically significant difference

[1]Fixed-effects main effects two-way ANOVA

[2]Fixed-effects Friedman's two-way nonparametric ANOVA and Wilcoxon rank sum pairwise comparison, Bonferroni adjustment for statistical significance

**Table 6.** Depth accuracy correlation with micro-CT assessed depth for true positive diagnoses only

| *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* | *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PSP | | | | | 0.4 | PanBW | | | | | 0.27 |
| | 1 | 0.30 | 0.09 | 0.51 | | | 1 | 0.13 | -0.20 | 0.45 | |
| | 2 | 0.36 | 0.10 | 0.63 | | | 2 | 0.38 | 0.09 | 0.66 | |
| | 3 | 0.28 | 0.09 | 0.48 | | | 3 | 0.33 | 0.11 | 0.55 | |
| | 4 | 0.53 | 0.31 | 0.75 | | | 4 | 0.23 | -0.02 | 0.47 | |
| | 5 | 0.17 | 0.00 | 0.35 | | | 5 | 0.23 | -0.03 | 0.49 | |
| | 6 | 0.43 | 0.21 | 0.64 | | | 6 | 0.27 | -0.02 | 0.56 | |
| | 7 | 0.61 | 0.23 | 0.98 | | | 7 | 0.23 | -0.06 | 0.52 | |
| | 8 | 0.34 | 0.11 | 0.56 | | | 8 | 0.26 | -0.04 | 0.57 | |
| | 9 | 0.51 | 0.25 | 0.76 | | | 9 | 0.41 | 0.17 | 0.66 | |
| | 10 | 0.57 | 0.25 | 0.88 | | | 10 | 0.43 | 0.16 | 0.70 | |
| Schick33 | | | | | 0.43 | XG3D | | | | | 0.63 |
| | 1 | 0.40 | 0.01 | 0.79 | | | 1 | 0.60 | 0.35 | 0.84 | |
| | 2 | 0.42 | 0.11 | 0.73 | | | 2 | 0.44 | 0.01 | 0.86 | |
| | 3 | 0.12 | -0.07 | 0.31 | | | 3 | 0.60 | 0.36 | 0.83 | |
| | 4 | 0.47 | 0.28 | 0.67 | | | 4 | 0.33 | 0.03 | 0.64 | |
| | 5 | 0.13 | -0.05 | 0.32 | | | 5 | 0.29 | 0.03 | 0.55 | |
| | 6 | 0.57 | 0.34 | 0.80 | | | 6 | 0.72 | 0.50 | 0.93 | |
| | 7 | 0.41 | 0.07 | 0.75 | | | 7 | 0.66 | 0.37 | 0.94 | |
| | 8 | 0.50 | 0.22 | 0.78 | | | 8 | 0.66 | 0.41 | 0.90 | |
| | 9 | 0.44 | 0.16 | 0.73 | | | 9 | 0.72 | 0.50 | 0.94 | |
| | 10 | 0.87 | 0.62 | 1.00 | | | 10 | 0.77 | 0.59 | 0.95 | |

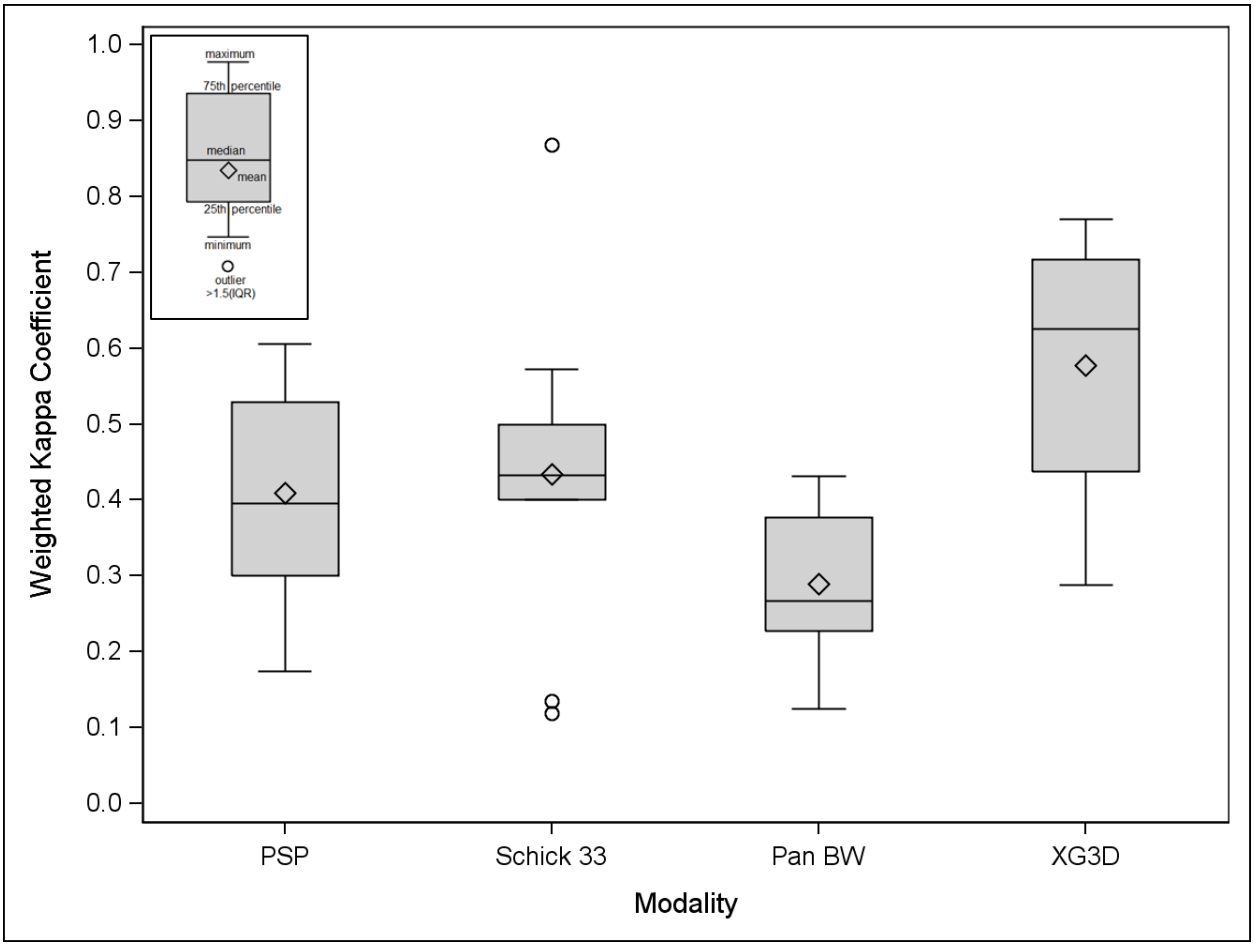


Fig. 5. Observed depth correlation with micro-CT depth boxplots by modality.

**Table 7**. Summary of statistical findings comparing depth observation weighted kappa correlation coefficients (correlation with micro-CT depth assessment)

| *Data set* | *Effect* | *p-value* | *Pairwise comparison* | *p-value* |
|---|---|---|---|---|
| Lesion depth correlation with micro-CT (kappa)[1] | | | | |
| | Modality | 0.003* | | |
| | | | PSP vs PanBW | 0.62 |
| | | | PSP vs Shick33 | 1.00 |
| | | | PSP vs XG3D | 0.15 |
| | | | PanBW vs Schick33 | 0.32 |
| | | | PanBW vs XG3D | 0.002* |
| | | | Schick33 vs XG3D | 0.30 |

*Denotes statistically significant difference

[1]One-way ANOVA with t-test pairwise comparison, Bonferroni adjustment for statistical significance

**Table 8.** Cavitation detection sensitivity and specificity scores

| | *Modality* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *PSP* | | *Schick33* | | *PanBW* | | *XG3D* | | *Mean* | | *SD* | |
| *Observer* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* | *TPR* | *TNR* |
| 1 | 0.58 | 1.00 | 0.50 | 1.00 | 0.42 | 0.94 | 0.83 | 0.94 | 0.58 | 0.97 | 0.18 | 0.03 |
| 2 | 0.58 | 1.00 | 0.42 | 1.00 | 0.33 | 0.98 | 0.92 | 0.96 | 0.56 | 0.99 | 0.26 | 0.02 |
| 3 | 0.42 | 1.00 | 0.33 | 1.00 | 0.08 | 1.00 | 0.58 | 0.96 | 0.35 | 0.99 | 0.21 | 0.02 |
| 4 | 0.83 | 0.98 | 0.83 | 1.00 | 0.58 | 0.98 | 0.75 | 0.92 | 0.75 | 0.97 | 0.12 | 0.03 |
| 5 | 0.92 | 1.00 | 0.50 | 0.98 | 0.58 | 0.94 | 0.92 | 0.88 | 0.73 | 0.95 | 0.22 | 0.05 |
| 6 | 0.25 | 1.00 | 0.42 | 1.00 | 0.58 | 0.96 | 0.92 | 0.96 | 0.54 | 0.98 | 0.28 | 0.02 |
| 7 | 0.50 | 0.98 | 0.25 | 1.00 | 0.42 | 1.00 | 0.83 | 0.98 | 0.50 | 0.99 | 0.25 | 0.01 |
| 8 | 0.33 | 1.00 | 0.50 | 1.00 | 0.17 | 1.00 | 0.92 | 0.96 | 0.48 | 0.99 | 0.32 | 0.02 |
| 9 | 0.67 | 0.96 | 0.58 | 0.96 | 0.25 | 1.00 | 0.83 | 1.00 | 0.58 | 0.98 | 0.25 | 0.02 |
| 10 | 0.33 | 1.00 | 0.33 | 1.00 | 0.42 | 1.00 | 0.83 | 1.00 | 0.48 | 1.00 | 0.24 | 0.00 |
| *Mean* | 0.54 | 0.99 | 0.47 | 0.99 | 0.38 | 0.98 | 0.83 | 0.96 | | | | |
| *SD* | 0.22 | 0.01 | 0.16 | 0.01 | 0.18 | 0.02 | 0.10 | 0.03 | | | | |
| *Compiled** | 0.54 | 0.99 | 0.47 | 0.99 | 0.38 | 0.98 | 0.83 | 0.96 | | | | |

Abbreviations: SD standard deviation, TPR true positive rate (sensitivity), TNR true negative rate (specificity)

*Compiled sensitivity/specificity scores for each modality including all observers and for each observer for all modalities

**Table 9**. Summary of statistical findings comparing cavitation sensitivity and specificity scores

| *Data set* | *Effect* | *p-value* | *Pairwise comparison* | *p-value* |
|---|---|---|---|---|
| Cavitation sensitivity[1] | | | | |
| | Observer | 0.0063* | | |
| | Modality | <0.0001* | | |
| | | | PSP vs PanBW | 0.043 |
| | | | PSP vs Schick33 | 0.16 |
| | | | PSP vs XG3D | 0.0011* |
| | | | PanBW vs Schick33 | 0.50 |
| | | | PanBW vs XG3D | <0.0001* |
| | | | Schick33 vs XG3D | <0.0001* |
| Cavitation specificity[1] | | | | |
| | Observer | 0.18 | | |
| | Modality | 0.0025* | | |
| | | | PSP vs PanBW | 0.14 |
| | | | PSP vs Schick33 | 0.78 |
| | | | PSP vs XG3D | 0.0015* |
| | | | PanBW vs Schick33 | 0.077 |
| | | | PanBW vs XG3D | 0.056 |
| | | | Schick33 vs XG3D | 0.0007* |

*Denotes statistically significant difference
[1]Fixed-effects Friedman's two-way nonparametric ANOVA and Wilcoxon rank sum pairwise comparison, Bonferroni adjustment for statistical significance

**Table 10.** Intraobserver agreement for caries presence observations

| *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* | *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PSP | | | | | 0.62 | Pan BW | | | | | 0.62 |
| | 1 | 0.57 | 0.27 | 0.86 | | | 1 | 0.45 | 0.07 | 0.83 | |
| | 2 | 0.57 | 0.34 | 0.80 | | | 2 | 0.65 | 0.48 | 0.81 | |
| | 3 | 0.46 | 0.26 | 0.66 | | | 3 | 0.66 | 0.53 | 0.79 | |
| | 4 | 0.86 | 0.75 | 0.96 | | | 4 | 0.67 | 0.41 | 0.92 | |
| | 5 | 0.83 | 0.68 | 0.98 | | | 5 | 0.88 | 0.78 | 0.98 | |
| | 6 | 0.36 | 0.13 | 0.59 | | | 6 | 0.63 | 0.40 | 0.85 | |
| | 7 | 0.74 | 0.58 | 0.89 | | | 7 | 0.52 | 0.34 | 0.71 | |
| | 8 | 0.69 | 0.52 | 0.85 | | | 8 | 0.51 | 0.35 | 0.66 | |
| | 9 | 0.60 | 0.44 | 0.77 | | | 9 | 0.59 | 0.42 | 0.76 | |
| | 10 | 0.64 | 0.46 | 0.82 | | | 10 | 0.60 | 0.43 | 0.78 | |
| Schick 33 | | | | | 0.73 | XG3D | | | | | 0.65 |
| | 1 | 0.82 | 0.64 | 0.99 | | | 1 | 0.80 | 0.61 | 0.99 | |
| | 2 | 0.74 | 0.54 | 0.93 | | | 2 | 0.63 | 0.43 | 0.83 | |
| | 3 | 0.74 | 0.56 | 0.91 | | | 3 | 0.46 | 0.24 | 0.68 | |
| | 4 | 0.82 | 0.63 | 1.00 | | | 4 | 0.71 | 0.51 | 0.91 | |
| | 5 | 0.81 | 0.65 | 0.97 | | | 5 | 0.58 | 0.35 | 0.80 | |
| | 6 | 0.55 | 0.31 | 0.79 | | | 6 | 0.57 | 0.35 | 0.79 | |
| | 7 | 0.68 | 0.51 | 0.85 | | | 7 | 0.74 | 0.59 | 0.89 | |
| | 8 | 0.54 | 0.33 | 0.75 | | | 8 | 0.62 | 0.43 | 0.82 | |
| | 9 | 0.62 | 0.44 | 0.80 | | | 9 | 0.67 | 0.53 | 0.81 | |
| | 10 | 0.72 | 0.56 | 0.89 | | | 10 | 0.75 | 0.60 | 0.91 | |

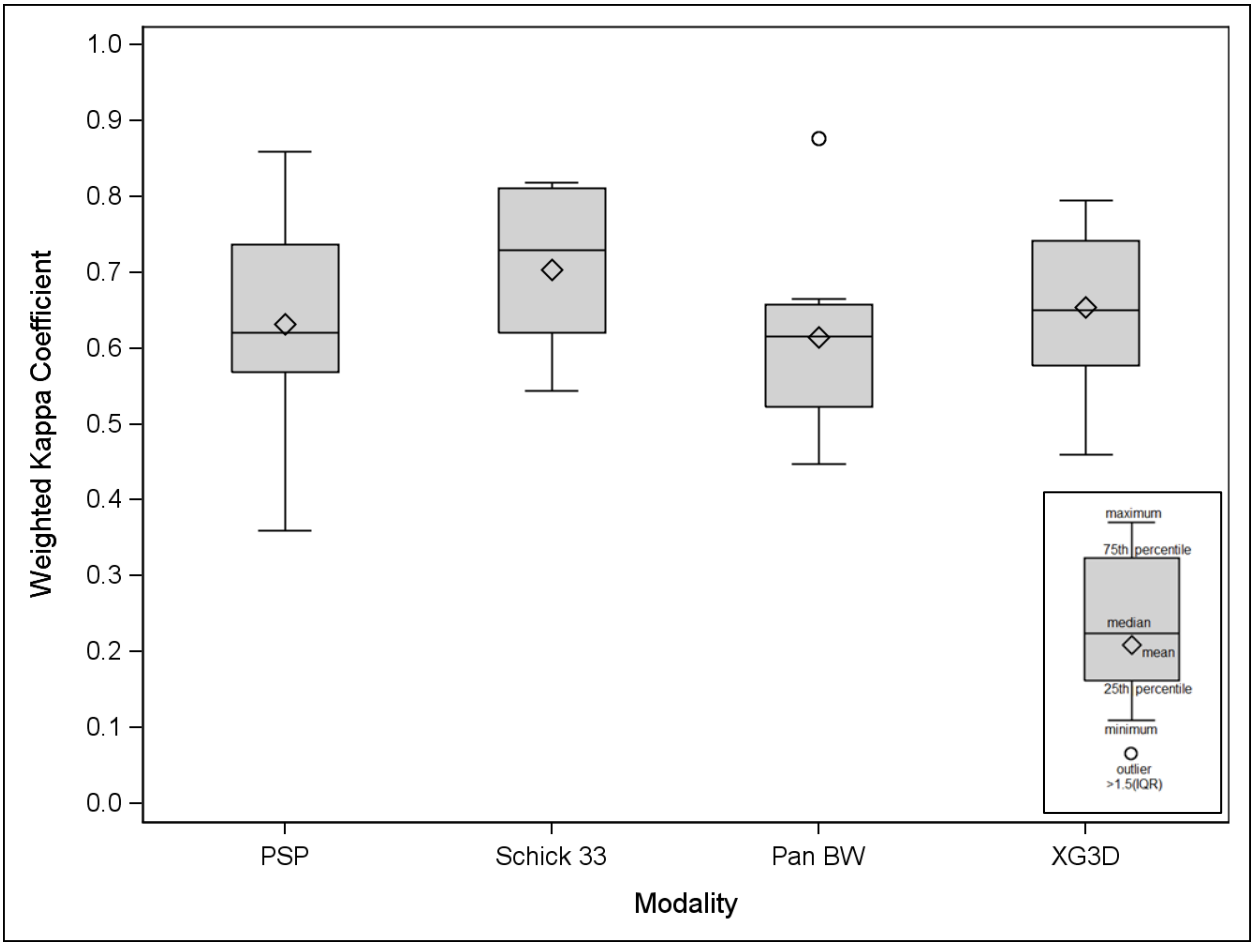


Fig. 6. Intraobserver agreement boxplots for caries presence observations by modality.

**Table 11.** Intraobserver agreement for caries depth observations

| *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* | *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PSP | | | | | 0.76 | Pan BW | | | | | 0.66 |
| | 1 | 0.80 | 0.60 | 1.00 | | | 1 | 0.47 | 0.03 | 0.91 | |
| | 2 | 0.75 | 0.48 | 1.00 | | | 2 | 0.82 | 0.66 | 0.98 | |
| | 3 | 0.45 | 0.21 | 0.68 | | | 3 | 0.73 | 0.55 | 0.90 | |
| | 4 | 0.81 | 0.63 | 0.99 | | | 4 | 0.74 | 0.47 | 1.00 | |
| | 5 | 0.77 | 0.54 | 1.00 | | | 5 | 0.92 | 0.83 | 1.00 | |
| | 6 | 0.58 | 0.32 | 0.85 | | | 6 | 0.53 | 0.29 | 0.77 | |
| | 7 | 0.83 | 0.66 | 0.99 | | | 7 | 0.65 | 0.42 | 0.88 | |
| | 8 | 0.81 | 0.66 | 0.96 | | | 8 | 0.68 | 0.51 | 0.85 | |
| | 9 | 0.59 | 0.40 | 0.78 | | | 9 | 0.47 | 0.26 | 0.68 | |
| | 10 | 0.70 | 0.49 | 0.90 | | | 10 | 0.61 | 0.41 | 0.80 | |
| Schick 33 | | | | | 0.74 | XG3D | | | | | 0.73 |
| | 1 | 0.80 | 0.60 | 1.00 | | | 1 | 0.75 | 0.55 | 0.96 | |
| | 2 | 0.89 | 0.67 | 1.00 | | | 2 | 0.70 | 0.46 | 0.94 | |
| | 3 | 0.62 | 0.40 | 0.85 | | | 3 | 0.68 | 0.46 | 0.90 | |
| | 4 | 0.79 | 0.62 | 0.96 | | | 4 | 0.68 | 0.47 | 0.90 | |
| | 5 | 0.73 | 0.45 | 1.00 | | | 5 | 0.75 | 0.52 | 0.98 | |
| | 6 | 0.73 | 0.53 | 0.94 | | | 6 | 0.73 | 0.54 | 0.91 | |
| | 7 | 0.74 | 0.51 | 0.97 | | | 7 | 0.88 | 0.72 | 1.00 | |
| | 8 | 0.66 | 0.43 | 0.89 | | | 8 | 0.73 | 0.57 | 0.88 | |
| | 9 | 0.60 | 0.39 | 0.81 | | | 9 | 0.70 | 0.52 | 0.89 | |
| | 10 | 0.81 | 0.62 | 1.00 | | | 10 | 0.82 | 0.67 | 0.97 | |

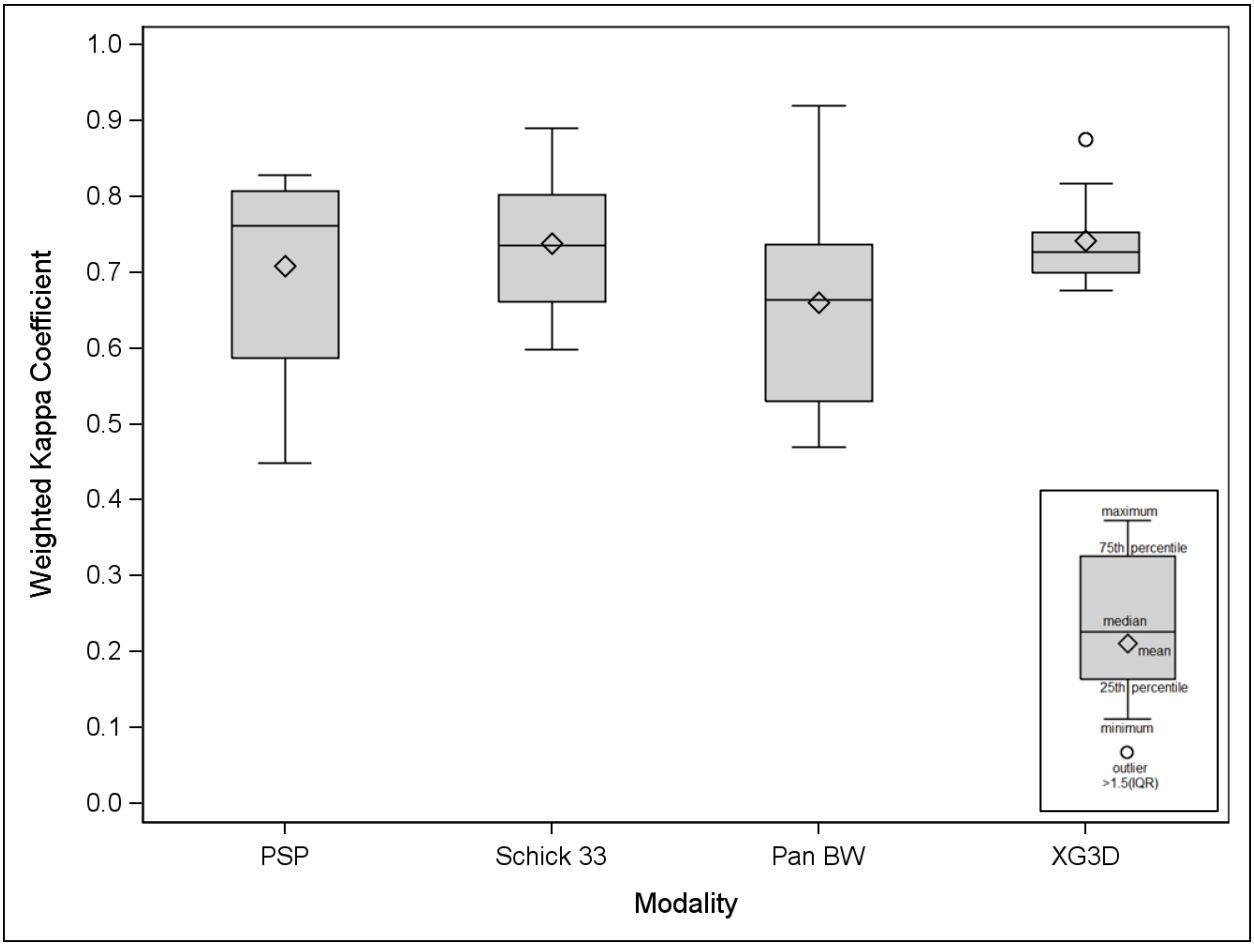


Fig. 7. Intraobserver agreement boxplots for caries depth observations by modality.

**Table 12.** Intraobserver agreement for caries cavitation observations

| *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* | *Modality* | *Observer* | *Weighted Kappa Coefficient* | *95% CI* | | *Median Kappa* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PSP | | | | | 0.69 | Pan BW | | | | | 0.60 |
| | 1 | 0.75 | 0.43 | 1.00 | | | 1 | 0.69 | 0.27 | 1.00 | |
| | 2 | 0.59 | 0.26 | 0.93 | | | 2 | 0.62 | 0.34 | 0.90 | |
| | 3 | 0.70 | 0.38 | 1.00 | | | 3 | 0.41 | 0.12 | 0.71 | |
| | 4 | 0.60 | 0.31 | 0.89 | | | 4 | 0.78 | 0.48 | 1.00 | |
| | 5 | 0.63 | 0.33 | 0.94 | | | 5 | 0.69 | 0.49 | 0.89 | |
| | 6 | 0.18 | 0.02 | 0.33 | | | 6 | 0.40 | 0.00 | 0.79 | |
| | 7 | 0.81 | 0.62 | 1.00 | | | 7 | 0.72 | 0.43 | 1.00 | |
| | 8 | 0.67 | 0.42 | 0.93 | | | 8 | 0.23 | -0.03 | 0.50 | |
| | 9 | 0.86 | 0.75 | 0.97 | | | 9 | 0.58 | 0.28 | 0.89 | |
| | 10 | 0.70 | 0.55 | 0.84 | | | 10 | 0.51 | 0.23 | 0.79 | |
| Schick 33 | | | | | 0.70 | XG3D | | | | | 0.72 |
| | 1 | 0.64 | 0.32 | 0.95 | | | 1 | 0.59 | 0.34 | 0.84 | |
| | 2 | 0.72 | 0.50 | 0.94 | | | 2 | 0.62 | 0.37 | 0.88 | |
| | 3 | 0.48 | 0.17 | 0.80 | | | 3 | 0.55 | 0.26 | 0.84 | |
| | 4 | 0.67 | 0.33 | 1.00 | | | 4 | 0.65 | 0.41 | 0.88 | |
| | 5 | 0.86 | 0.66 | 1.00 | | | 5 | 0.84 | 0.70 | 0.97 | |
| | 6 | 0.34 | 0.03 | 0.65 | | | 6 | 0.72 | 0.47 | 0.97 | |
| | 7 | 0.79 | 0.52 | 1.00 | | | 7 | 0.97 | 0.91 | 1.00 | |
| | 8 | 0.81 | 0.74 | 0.88 | | | 8 | 0.73 | 0.53 | 0.92 | |
| | 9 | 0.63 | 0.41 | 0.84 | | | 9 | 0.74 | 0.53 | 0.96 | |
| | 10 | 0.84 | 0.69 | 1.00 | | | 10 | 0.80 | 0.62 | 0.98 | |

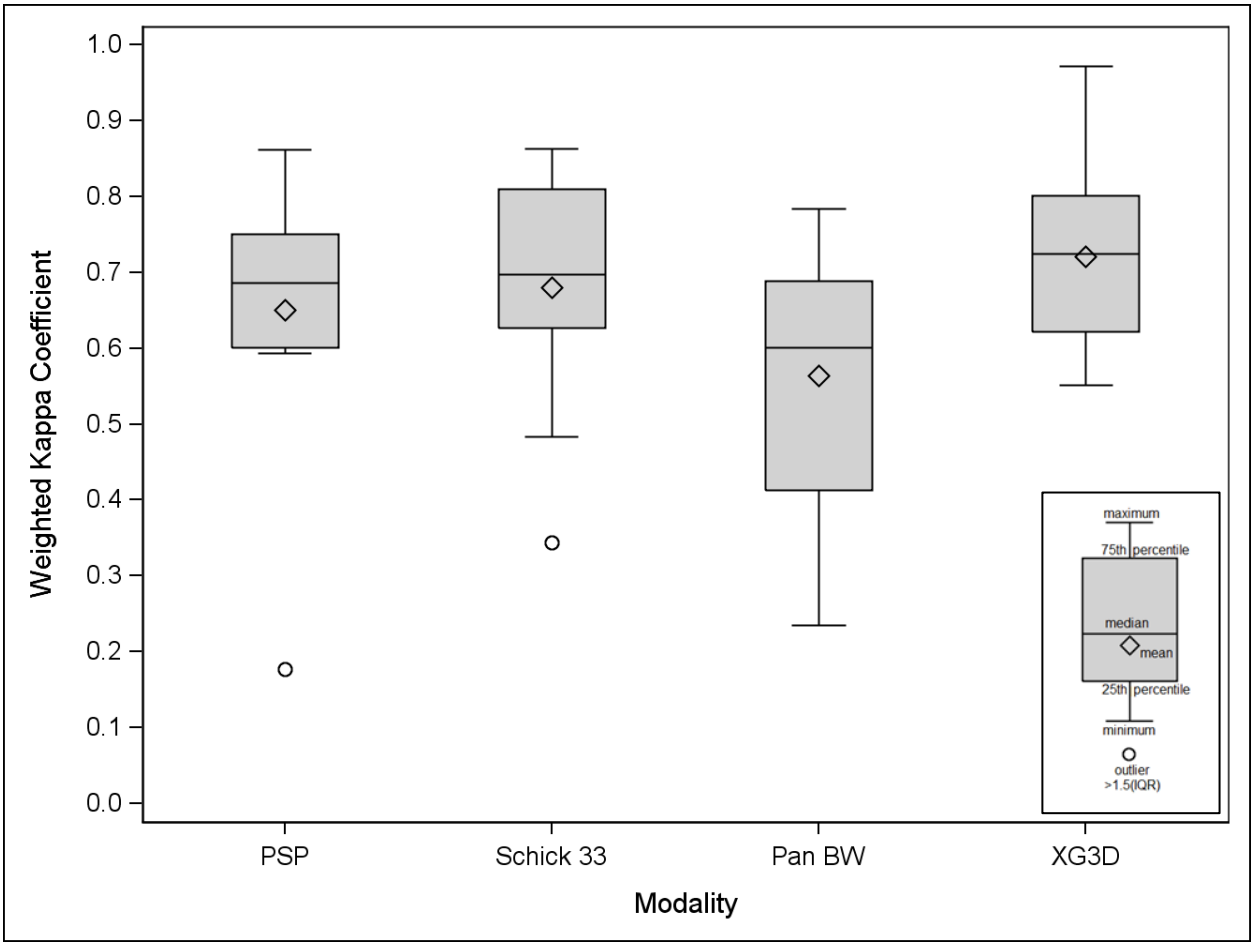


Fig. 8. Intraobserver agreement boxplots for caries cavitation observations by modality.

**Table 13**. Summary of statistical findings comparing intraobserver agreement kappa correlation coefficients

| *Data set* | *Effect* | *p-value* | *Pairwise comparison* | *p-value* |
|---|---|---|---|---|
| Intraobserver agreement caries presence[1] | Modality | 0.411 | NA | |
| Intraobserver agreement caries depth[1] | Modality | 0.376 | NA | |
| Intraobserver agreement caries cavitation[1] | Modality | 0.197 | NA | |

[1]one-way ANOVA

## MANUSCRIPT TWO: EVALUATION OF SYSTEM GRAYSCALE PERFORMANCE

### Introduction

Micro-CT, also known as micro-computed tomography or microtomography is an important research tool in dentistry for the assessment and characterization ex vivo obtained teeth and other hard tissues.[88-91] The modality provides very high resolution 3D information of specimen morphology and attenuation properties in a non-destructive manner with a relatively short amount of time and at a low cost. There has been a significant amount of research validating the capability of micro-CT to accurately assess mineral density.[88, 90, 92-98] A number of important conclusions can be made from this body of work.

First, unit calibration to known mineral densities is vital for accurate density measurements.[94, 99-101] By including phantoms of known densities researchers can construct calibration curves and accurately predict the observed mineral densities observed in the specimens. Ideally, calibration phantoms have the same or very similar chemical composition to the specimen of interest, exhibit a range of densities that is similar or greater than the range present in the specimen, and demonstrate spatial uniformity at the resolution of the scan to prevent biased measurement errors.[86, 94, 101] The most commonly used material for micro-CT calibration in hard tissue and dental imaging is hydroxyapatite (HAP), a chemical composition that mimics closely the chemical makeup and radiographic attenuation properties of mineralized tissue such as teeth and bone.[94]

Second, beam hardening artifacts must be addressed in order to achieve accurate density measurements.[86, 101] Beam hardening artifacts occur when the polychromatic x-ray sources used in most radiographic imaging systems undergo non-uniform attenuation by highly mineralized tissues such as teeth and bone. The lower energy photons in the x-ray beam are attenuated to a greater extent than higher energy photons, resulting in an increase or "hardening" of the mean beam energy. This increase in beam energy results in altered attenuation properties and a corresponding misinterpretation by the

system detector. In effect, subsequent tissues imaged with a hardened beam are computed as less attenuating than they really are.

The initial approach to minimizing the effects of beam hardening artifacts is to limit the bandwidth of the original x-ray beam while maintaining sufficiently high photon energy to significantly penetrate the specimen of interest. This goal can be achieved by maintaining a relatively high tube potential and moderate to high filtration. Previous studies have utilized 70kVp as an acceptable tube potential to balance penetrating power of tooth structure with minimal beam hardening artifacts in micro-CT dental imaging.[86, 87] Beam hardening artifacts in micro-CT are also reduced by imaging the specimen surrounded by air instead of surrounded by higher attenuation media like water or soft tissue.[102] Beam hardening correction algorithms are available for many micro-CT systems but require very specific calibration to the specimen material of interest and relative size matching between the calibration phantoms and specimen.[99, 102] A final factor related to beam hardening artifacts is the presence of truncation artifacts due to specimen attenuation from outside of the image field of view. These artifacts are readily avoided by selecting specimens that are entirely imaged by the system field of view (FOV).[102]

Third, both spatial resolution and noise levels must be adequate for the specific research question.[99] It is well accepted that spatial resolution and noise are inversely related at a fixed dose, however, long scan times and high doses for inanimate specimens such as teeth are relatively minor concerns. Furthermore, available micro-CT systems are capable of very high spatial resolution imaging – less than 10 microns – which in many clinical dental research questions is well above the threshold of resolution to identify clinically relevant tooth features. As discussed above, spatial uniformity of the calibration phantoms at the resolution of the scan is important for accurate measurement, however; phantom uniformity is more significantly dictated by manufacturing method.[101]

The traditional gold standard method to validate dental caries status is visual examination of thin section histology, however, in recent years micro-CT has proven to be an effective alternative. Numerous studies have employed and found success with micro-CT as a ground truth caries validation tool for purposes including caries detection,[75, 76] evaluating lesion depth,[55] evaluating progression of lesion morphology,[77] evaluating lesion mineral concentration,[78, 79] and validating caries removal techniques.[80-82]

Micro-CT has also been shown to agree with transverse microradiography, another accepted gold-standard technique for the evaluation of caries.[83] One study reported that micro-CT does not agree with histological validation at different disease severities and cites small enamel lesions as having the greatest discrepancy between histology and micro-CT results. They explain that two scenarios associated with histology may confound accurate caries identification. First, it is very difficult to create a thin section exactly at the exact location of a small caries lesion confined to the enamel, and second, that the procedure of sectioning the tooth may cause loss of tooth and/or lesion material.[84] These two scenarios associated with histology are actually confounding for ground truth analysis and highlight the advantages of non-destructive 3D micro-CT analysis. Another study reports that micro-CT generally underestimates caries lesion depth compared to histological assessment.[85] However, tooth discoloration/stain due to infiltrative bacterial activity may confound the assessment of lesion depth with histology and may not represent the true extent of decalcification. Conversely, micro-CT evaluates mineral decalcification directly. It is the authors' conclusion that micro-CT is an appropriate and valid ground truth modality based on the body of research in support of micro-CT for a caries validation technique, the problems associated with missed lesions or lost tooth structure associated with histology slice preparation, and the possible confounding factor of tooth stain associated with histology slice analysis.

Despite the growing acceptance of micro-CT for the assessment of dental hard tissues and the growing acceptance for use as a ground truth evaluation of dental caries, very little research has been done to assess the mineral density accuracy and reliability of clinical dental imaging systems used to diagnose dental conditions. The prevailing modality for the radiographic diagnosis of dental caries is intraoral bitewing radiography.[2-6] Other modalities such as panoramic imaging, panoramic bitewing imaging, and cone beam computer tomography (CBCT) imaging have also been evaluated for their caries diagnostic accuracy.[2]

In a related study, three new technologies in dental caries imaging were evaluated for their carries diagnostic accuracy compared to a control modality. These new technologies include the Schick 33 intraoral direct digital sensor (Sirona Dental, Salzburg, Germany), a Planmeca ProMax panoramic unit in

bitewing mode (Planmeca Inc., Helsinki, Finland), and a Sirona Orthophos XG3D CBCT in High Definition (HD) mode with Metal Artifact Reduction Software (MARS) (Sirona Dental, Salzburg, Germany). Each modality offered different technological advancements designed to aid in the clinical diagnosis of caries. A control modality was Gendex size 2 photostimulable phosphor (PSP) plates (Gendex, Hatfield, PA) exposed by a Focus intraoral source (Instrumentarium Dental, Tuusula, Finland), the standard system used in the dental school clinics at the University of North Carolina at Chapel Hill (UNC). This caries detection study used micro-CT as a ground truth modality for the assessment of dental caries.

The authors viewed this evaluation of caries diagnoses on four clinical modalities with a micro-CT ground truth as an opportunity to examine clinical imaging modalities with micro-CT calibration techniques. Because multiple image sets were acquired on each modality, the project allowed repeated measurements of the same phantoms on each modality. A variety of valuable research questions can be addressed using this imaging arrangement. First, by leveraging the repeated images taken on each modality for the caries study, a measure of exam-to-exam consistency can be evaluated. Second, with proper phantom design employing different HAP mineral densities, grayscale linearity of each modality can be reported. Third, by varying the interval between different HAP densities, a threshold or limit of contrast resolution may be observed statistically. Fourth, any differences in observed caries diagnostic accuracy between the different modalities as reported by the other study can be directly compared to the calibration phantom results reported in this study. Accordingly, the aim of this project is to assess exam-to-exam consistency, grayscale linearity, threshold of rod discrimination, and comparisons with caries diagnostic accuracy of the four new dental radiographic imaging technologies and micro-CT gold standard system.

**Materials and Methods**

Hydroxyapatite (HAP) rods (CIRS, Inc., Norfolk, VA) were obtained with known densities in the range of human tooth tissue for the purpose of being imaged in every exam of the related and previously described caries detection study. This caries study received Institutional Review Board (IRB) approval for tooth collection and observer analysis (Study #13-2843). The rod dimensions were 2.95x30-50mm. The

HAP densities selected were 0 (resin only), 12.5, 25, 50, 150, 500, 750, 1000, 1250, and 1500 mg/cc. Rods were trimmed to a uniform length of 30mm. All 10 rods were assembled into a linear phantom. Masking tape was used to arrange the linear phantom with rods in order of increasing density.

The caries study utilized a dry human mandible with edentulous posterior segments. Wax was used to hold extracted human teeth in the posterior segments for the caries study. A 1cm layer of wax surrounded the mandible to simulate the effect of soft tissue attenuation of the x-ray beam. For the extraoral imaging modalities which included the Sirona Orthophos XG3D CBCT and the Planmeca panoramic bitewings, the mandible was surrounded by a 3mm thick by 15.5cm diameter aluminum cylinder and a latex balloon filled with water placed within the mandible/wax phantom, respectively, to simulate attenuation properties of the human head relevant during extraoral imaging.

For each PSP and direct digital bitewing image, the rod phantom was stabilized upright on top of the extracted teeth and positioned so that an image of all ten rods was visible on the top portion of each bitewing. The rod phantom was oriented perpendicularly to the x-ray beam to avoid image overlap between the rods. The rod phantom was positioned upright on top of the extracted teeth similarly to the bitewing images for every CBCT exam.

Initial images of the rod phantom taken with the Planmeca panoramic bitewing modality with the mandible/teeth/wax/cylinder complex proved unusable because of horizontal tomographic "ghost" artifact from the contralateral mandible. Accordingly, the rod phantom was reimaged in the panoramic bitewing unit surrounded only by the aluminum cylinder. A non-attenuating foam prop was used to stabilize the rod phantom upright in the location of the posterior image focal trough. Similar unit parameters were used.

The rods were reorganized into a circular phantom in order to fit into a Scanco Medical (Scanco Medical AG, Bruttisellen, Switzerland) 2cm micro-CT sample vial. A segment of thin plastic tubing was used to stabilize the resin only (0 mg/cc) rod in the center and remaining rods peripherally in order of increasing concentration. The circular rod phantom was placed on top of each tooth imaged in the micro-CT unit. The micro-CT scan FOV was adjusted so that at least several millimeters of the rods were imaged in each scan.

The PSP control modality images were taken using the standard UNC School of Dentistry radiology clinic system: Gendex size 2 PSP sensors (Gendex, Hatfield, PA) were exposed with a Focus intraoral source (Instrumentarium Dental, Tuusula, Finland) at 70kVp, 7mA, 0.2s, and 40cm SID with standard 30cm rectangular collimation. The exposed sensors were processed with a ScanX IO ILE scanner (Air Techniques, Melville, NY) through MiPACS Dental Enterprise Viewer 3.1.1401 operating ScanX Plugin Version 1.2.8 (Medicore Imaging, Charlotte, NC). Sensors were scanned with the following settings: Intraoral High (#2) setting, 16-bit, invert images, and Image Enhancement: Enable histogram stretch, upper histogram cut 0.3, Lower histogram cut 3.4, Gamma correction 0.7. Images were saved with no additional adjustment of window/level and no additional applied filters.

The Schick 33 (Schick33) direct digital bitewing images were taken with the same Focus intraoral source (Instrumentarium Dental, Tuusula, Finland) at 70kVp, 0.05s, and 40cm SID standard 30cm rectangular collimation. A Schick 33 size 2 direct digital sensor was run by CDR DICOM for Windows Version 5.4.1658.5883 (Sirona Dental, Salzburg, Germany). Images were acquired with the following settings: Schick 33 High Resolution Acquisition, subtract dark image, acquire 12 bit image.

The panoramic bitewing (PanBW) images were taken with a Promax® Dimax 4 unit (Planmeca Inc., Helsinki, Finland), in panoramic bitewing mode operating at 72kVp and 11mA with a square average jaw shape. This particular unit was in active use in a private dental practice in Raleigh, NC. Dexis Version 9.0.5 imaging software (Dexis LLC, Hatfield, PA) was used to acquire and store the images. Dexis imaging software was chosen instead of Planmeca Romexis native imaging software because Dexis was used exclusively by the private office operating the unit.

The XG3D CBCT (XG3D) scans were taken with a Sirona Orthophos XG3D (Sirona Dental, Salzburg, Germany) unit operating at 85kV and 6mA with a 5x5cm FOV, and 0.1mm voxel size. Both high definition (HD) mode and metal artifact reduction software (MARS) were used. Scans were reconstructed and stored using Sidexis XG Version 2.56 software (Sirona Dental, Salzburg, Germany).

The micro-CT scans were acquired with a SCANCO Medical µCT 40 scanner (Scanco Medical AG, Bruttisellen, Switzerland) operating at 70kVp, 0.115 mA, 200s scan time, with 0.5mm Al filtration, 5µm focal spot size, 24µm CMOS camera pitch, 1 sample per pixel, and convolution kernel 3. Extracted teeth

were placed in a 2cm diameter poly-ether-imid sample tube with 0.7mm wall thickness. To confirm system accuracy, the unit underwent routine weekly bone density calibration using manufacturer supplied hydroxyapatite phantoms ranging in density from 99-800 mg/cc. For the year prior to imaging, average variance for the 800 mg/cc calibration phantom was 12 mg/cc, which was well within manufacturer tolerances. No beam hardening correction was applied as bone density calibrations were considered accurate and selected system parameters were deemed acceptable for accurate dental tissue imaging. Scans were reconstructed with a 20μm voxel size using Scanco Version 1.2a software (Scanco Medical AG, Bruttisellen, Switzerland).

Acquired images of the rod phantom totaled as follows: 16 phantom images each from the PSP bitewing, direct digital bitewing, and CBCT modalities; 8 images from the panoramic bitewing modality, and 13 scans from the micro-CT system. All images were exported and saved in DICOM format with default/native resolution, window/level, and processing settings.

Saved DICOM images were imported into ImageJ 1.48v (National Institutes of Health, USA) for grayscale value analysis. PSP bitewing, direct digital bitewing, and panoramic bitewing images were imported as single DICOM images. Because the Schick33 direct digital images were stored with an inverted grayscale, imported direct digital DICOM images were re-inverted to maintain uniform grayscale orientation with the other modalities. The ImageJ "measure" tool was used to record the mean grayscale values of three circular $1mm^2$ regions of interest (ROI's) centered on the longitudinal axis of each rod. Additional mean grayscale values were recorded from three identical $1mm^2$ ROI's from non-attenuated background in the image.

CBCT and micro CT DICOM scans were imported into ImageJ as virtual stacks. The "Z-project" tool was used to create a single 1mm thick cross-section of the phantom rods for each exam. The ImageJ "measure" tool was used to record the mean grayscale values of three $0.5mm^2$ ROI's for each rod. Similarly, additional mean grayscale values were recorded from three identical $0.5mm^2$ ROI's from non-attenuated background in the image.

For each modality, fixed-effects main effects two-way analysis of variance (ANOVA) was performed with grayscale value (averaged over the 3 ROI values for each rod/image combination) as the

response and with rod density and image as main effects. Plots of rod density by grayscale value were used to determine if interaction terms were necessary to investigate. For each test (one for each modality), a p-value less than 0.05 was considered statistically significant.

A visual examination of the plots of rod density by grayscale value (again averaged over the 3 ROI values for each rod/image combination) determined if a linear relationship was reasonable. Results from the above described 2-way ANOVA dictated that control for image when computing R-square values was appropriate. R-square values were calculated and residual plots were evaluated for even distribution. A high R-square value and even distribution of the residual plot supported the determination of a linear relationship. Regression formulas were derived for the purpose of calculating predicted grayscale values from reported mineral densities of healthy and diseased enamel and dentin, as a comparison to observed tissue grayscale values in the clinical caries detection study.

For each modality, simultaneous t-test pairwise comparisons of the mean grayscale value averaged over the three ROI's for each rod density were conducted using Tukey's method to adjust for multiple comparisons, since data are balanced within each modality and all pair-wise comparisons were of interest. For each pairwise comparison an adjusted p-value less than 0.05 suggested that the two rod densities being compared had different mean grayscale values. Background density was excluded from this analysis because of a presumed non-linear relationship with HAP density. Recognizing that the resin only rod and background may also have a non-linear relationship with the other rods, this analysis was carried out both including and excluding the resin only rod observations. The results from these pairwise analyses were used to estimate the threshold of rod discrimination or contrast resolution between rods (the minimum rod density for which the mean grayscale values are statistically significantly different). Analyses were done using SAS software (version 9.3, SAS Institute Inc., Cary, NC).

**Results**

A volumetric rendering of a micro-CT scan of an extracted tooth with caries and calibration rods is provided in Figure 9. Representative ROIs for each rod and background for all modalities are provided in Figure 10. Grayscale value means and standard deviations for each rod density across images by

modality are given in Table 14. Figures 11-15 display rod density by mean grayscale value for each modality. Concerning the 2-way ANOVA with mean grayscale value as the response and with rod density and image as main effects, statistically significant results ($p = 0.0003$ Micro-CT, $p < 0.0001$ for all others) of the test for the effect of image on mean grayscale value for each modality suggest that each modality suffers some inconsistency across exams. Of note, much of this error was attributable to variable vertical position of individual image plots. In other words, each image plot maintained remarkably similar shape but is positioned at an overall variable mean grayscale value (vertical position). Based on visual interpretation of Figures 11-15, micro-CT demonstrated the least vertical variation exam to exam, followed by XG3D, PSP, PanBW, and Schick 33, respectively. Correcting for this vertical distribution, the assessments of linear relationships were supported by the high R-square values from the 2-way ANOVA, displayed in Table 15, as well as a fairly even spread of residuals. Visually, micro-CT as well as modalities PSP, Schick 33, and XG3D can be determined to have a fairly linear relationship between rod density and grayscale value (Figures 11-15). The "check mark" shape of the PanBW modality, as well as its residual plot suggest non-linear trend, at least for densities below approximately 750mg/cc (Figure 14). Table 16 contains the minimum rod density that proved always distinguishable from all others, both ignoring the resin only rod, as well as considering it. Table 17 indicates the difference in rod densities for which the mean grayscale values are not statistically significantly different. As expected from the visual interpretation of Figures 11-15, most of the indistinguishable rods had low densities. Micro-CT appeared to have the best rod differentiation (12.5, 50 mg/cc threshold), followed by the XG3D, PSP, and Schick33 modalities (150, 150 mg/cc), and lastly the PanBW (750, 750 mg/cc) which performed the poorest with respect to threshold of rod discrimination. Summaries of reported mineral densities of healthy and diseased dental tissue plus comparisons of our regression predicted and mean observer grayscale values are outlined in Table 18. All tissue types demonstrated remarkably accurate similarity with the exception of carious dentin.

**Discussion**

The authors found that each modality demonstrated some statistically significant variation scan to scan. Surprisingly, this observation included the micro-CT modality despite the unit's extensive calibration and accepted high level of precision. We statistically inferred that much of this error was due to variable vertical distribution of each exam with an overall maintenance of curve shape. This vertical distribution is almost certainly due to overall image brightness adjustments taking place during image pre-processing protocols. It is possible that the changing human teeth present in each micro-CT, PSP, Schick33, and XG3D exam may have influenced overall image brightness/grayscale value. Interestingly, this phenomenon was also observed in the Planmeca panoramic bitewing images which included phantom rods only and no teeth or phantom mandible. We will revisit this observation below. Visual inspection of the plots for each modality revealed a trend related to imaging technology. The 3D modalities (micro-CT and XG3D) appeared to have less variation exam to exam whereas the 2D modalities (PSP, Schick33, and PanBW) had more (Figures 11-15). It is possible that by calculating 3D image information, the level of error introduced by superimposed attenuating objects is reduced. Altogether, the conclusion that all modalities including micro-CT failed a strict test of exam-to-exam consistency reinforces the conclusion that calibration phantoms are necessary in each exam if accurate grayscale/mineral density measurements are to be obtained.[86, 94, 99-101]

Once the vertical distribution of curves in each modality was controlled, the authors found remarkably high linearity for all modalities (Table 15). The PanBW demonstrated the most non-linear trend at least at low densities (Figure 14). We ascribe this deviation to either variable low signal/background pre-processing on this particular unit, or the tomographic method of image acquisition which inherently introduces a significant amount of image data from superimposed objects onto the final image. Overall we suggest that observed exam-to-exam inconsistency may not be clinically relevant because such consistent linearity is maintained. This linearity dictates that within each exam, relative (not absolute) signal differences are reasonably reliable.

The minimum distinguishable rod density observations appeared to align with modality technology (Table 16 and Table 17). As a research imaging modality, micro-CT was able to distinguish

the differences between all rod densities (12.5 and 50 mg/cc, respectively, ignoring and considering resin rod). The clinical modalities PSP, Shick33, and XG3D performed not as well, each identifying the 150 mg/cc rod (both ignoring and considering the resin rod). The PanBW modality performed the poorest with discrimination only to the 750 mg/cc rod, both ignoring and considering the resin rod. We suggest that the PanBW tomographic image acquisition geometry imparts in a significant amount of image information from superimposed tissues resulting in somewhat less reliable grayscale values.

It is notable that the most variation and non-linear error in grayscale value for all modalities occurred at the background, resin only, and low density rods. This variation is particularly notable on the PanBW modality, where plot lines assume a "check mark" pattern in the region of rod densities 12.5 to 250mg/cc (Figure 14). We suspect that image pre-processing algorithms correct and balance these low-signal regions prior to image output, with the manner of correction variable between modalities and even between exams. This observation applies to a major limitation of this study related to the rod density choice for threshold of detection determination. Our selected rod densities of 0 (resin only), 12.5, 25, 50, 150, 500, 750, 1000, 1250, and 1500 mg/cc varied non-monotonically and biased narrow intervals of density difference towards the low end of density. In other words, because the interval between rod densities is both somewhat inconsistent and tended to be larger at higher densities, the opportunity to observe the similar thresholds of differentiation at these high densities may have been missed. At this time it is unclear whether observered thresholds of detection were due to true limits of contrast resolution or variations in image processing at low signals/low grayscale values.

Our study employed an unconventional method of establishing image contrast resolution. More commonly, researchers employ a visual test phantom typically composed of round holes of variable depth and size placed into a uniform phantom.[103-107] After imaging, observers rate which holes they can visualize resulting in an observer-dictated measure of contrast or contrast-resolution. Other approaches utilize subjective observer ratings of contrast of actual anatomy.[108] Many studies compute contrast to noise ratio (CNR) or signal difference to noise ratio (SdNR) directly, using observer grayscale values, region of interest standard deviation, and a variety of related formulas.[104-107] Finally, certain approaches measure x-ray photon scatter as a surrogate measure of CNR.[109] Additional studies and analyses utilizing

these more conventional methods of calculating contrast resolution are needed to better elucidate the contrast resolution performance of the modalities in question.

During Planmeca panoramic bitewing image acquisition, it was quickly noted that signals of the phantom rods were heavily corrupted with horizontal streak or "ghost" artifact from the contralateral portion of the phantom mandible bone and teeth. We accordingly decided to reimage the phantom rods alone in the Planmeca unit without the presence of teeth, bone, or other tissue mimicking objects. Only the circular aluminum cylinder surrounded the phantom rods for imaging. As discussed above, the PanBW images still demonstrated poor exam-to-exam consistency, the worst signal linearity of all modalities, and the worst threshold of rod discrimination of all modalities. At this time it is still unclear whether variable image preprocessing steps or the tomographic image acquisition geometry are responsible for these observations.

This study found that the 3D modalities tended to perform better than 2D modalities with respect to exam consistency, linearity, and rod discrimination. The related caries detection study concluded that the XG3D system demonstrated improved lesion depth determination and lesion cavitation detection compared to the other clinical modalities. For the caries detection study, micro-CT was chosen as the method of establishing the ground truth status of tooth caries presence, depth, and cavitation status. This choice implied some prima facie confidence in the modality to accurately assess tooth structures and densities. This study's findings that the micro-CT modality demonstrates the greatest exam-to-exam consistency, the greatest linearity, and the best threshold of rod discrimination further support the choice of micro-CT as a ground truth modality.

Another expected limitation of this study was the limited range of densities selected for the HAP phantom rods. As discussed in the introduction, it is recommended that for appropriate calculation of mineral density from grayscale values, the density calibration rod density range must span the range of densities observed in the tissues of interest.[99, 101] In this study, calibration rod densities extended only as high as 1500 mg/cc, a level approximately equivalent to the mineral density of dentin and only half the mineral density of enamel.[77-79, 81, 93, 95, 96, 110, 111] Recognizing that our micro-CT system demonstrated very high linearity, we extrapolated predicted grayscale values from these reported densities using the derived

regression formula. We also collected and calculated mean grayscale values for sound and diseased enamel and dentin from micro-CT scans of the teeth relevant to the related study. Summaries of reported mineral densities of healthy and demineralized dental tissue plus comparisons of our regression predicted and mean observer grayscale values are outlined in Table 18. Predicted and observed grayscale values for sound enamel, caries enamel, and sound dentin are remarkably similar. This observation suggests that our limited range of calibration rod densities and micro-CT protocol was appropriate for extended prediction into the range of high density healthy enamel and that beam hardening artifact was likely not a significant contributor to calibration error. Only carious dentin predicted and observed grayscale values deviated substantially. We suspect that this discrepancy was not related to rod calibration error because (1) the observed densities were well within the range of our calibration rod densities, (2) observed and reported dentin demineralization varies widely, and (3) the presence of additional attenuation from residual protein in the demineralized dentin may have biased observed grayscale values upward.[81, 94]

Visual inspection of Figs 14 and 15 reveals a plot line kink at 750 mg/cc for the PanBW and XG3D modalities. This kink is rather pronounced for the XG3D. Interestingly such a pronounced deviation in linearity is not noted on the plot for the micro-CT. We suspect that there may be some preprocessing change histogram analysis and grayscale output for these two clinical modalities, possibly to present a final image with higher or more attractive overall contrast specific to dental imaging. Alternatively the kink could be an artifact imparted by the aluminum cylinder unique to these two modality imaging protocols. We do not suspect that there are deviations in the reported HAP rod densities because we would expect to see similar deviations in the micro-CT data.

As mentioned in the introduction, error from beam hardening is a significant consideration for accurate imaging of dental hard tissues. The very high linearity observed from the micro-CT data confirms that our unit calibration steps were appropriate and beam hardening was not an issue for the range of mineral densities covered by the phantom rods. Some literature supports that extensive beam hardening correction for dental imaging with micro-CT may not be as significant as previously thought. One study found at most a 6% error in grayscale value due to beam hardening artifacts.[112] Other studies suggest that because of variable size and anatomy of teeth, modeling the actual amount of beam

hardening artifact is difficult[97] and correction is not significantly beneficial.[102] A slight fanning of mean grayscale values at densities above 1000 mg/cc for the XG3D may be due to slight beam hardening artifact (Figure 15). The rods were aligned linearly for acquisition in the XG3D modality. At some point during the 360 degree acquisition of basis projections, photons traversed the length of the phantom resulting in an opportunity for optimal hardening of the x-ray beam. The variability or "fanning" out of plots for different XG3D exams may be explained by slightly different orientations of the phantom rod relative to the angles of basis images acquired by the machine. Regardless, the very high level of linearity observed for the unit suggests that beam hardening was not a major source of grayscale error for the range of mineral densities covered by the phantoms.

The corresponding caries detection study found that all clinical modalities performed with similar overall diagnostic accuracy for caries detection. There were differences in observed sensitivities and specificities between modalities with respect to caries detection. The Schick 33 demonstrated relatively decreased sensitivity, the XG3D relatively increased sensitivity, and the Planmeca panoramic bitewing unit relatively decreased specificity. With respect to depth accuracy, the Planmeca unit demonstrated decreased accuracy and the XG3D CBCT demonstrated increased accuracy. With respect to cavitated lesion detection, the XG3D CBCT demonstrated significantly increased sensitivity with no relevant compromise in specificity compared to all other modalities.

For caries overall detection, it appears that the 3D modality (XG3D CBCT) offers a slight advantage in caries sensitivity, better assessment of lesion depth, and markedly better detection of lesion cavitation compared to the other 2D modalities. Notably, the fact that micro-CT was chosen to establish ground truth lesion status in the caries study implies confidence in 3D imaging over 2D. In this study, the 3D modalities (XG3D CBCT and micro-CT) demonstrated the best exam-to-exam consistency and linearity. The tomography based modality (Planmeca panoramic bitewing images) demonstrated the worst linearity and worst threshold of rod discrimination, especially considering the images were acquired without a phantom mandible or teeth present. 3D imaging provides obvious advantages over 2D imaging when assessing complex anatomy. Considering all clinical and technical findings, however, the authors infer that the 3D imaging technologies used in these two studies also provide tangible benefits for both

clinical caries detection and technical contrast performance. We believe this concept is worth considering when evaluating future modalities for clinical and technical performance.

Conclusions:

1. Each modality demonstrated some exam-to-exam inconsistency likely due to overall image brightness processing. The micro-CT system demonstrated the best overall consistency, followed by the XG3D CBCT, PSP bitewings, Planmeca panoramic bitewings, and Schick 33 bitewings. With respect to micro-CT analysis, this finding reinforces the observation that calibration phantoms are required in each exam if accurate grayscale measurements to assess mineral density are desired.
2. Within each exam, each modality demonstrated reasonable linearity. The Planmeca panoramic bitewing images demonstrated acceptable linearity, however; visually worse than other clinical modalities, especially considering that the phantom rod images were taken with no phantom mandible or teeth in the image. Horizontal "ghost" artifact resulting from the tomographic image acquisition geometry are likely significant confounding factors preventing accurate grayscale response for this modality.
3. Additional studies are needed to clarify the true contrast resolution of these systems and the effect of low signal/background processing on threshold of grayscale discrimination.
4. Despite utilizing HAP calibration phantoms with densities in the lower portion of dental tissue densities, our micro-CT analysis appears to correctly classify healthy and demineralized enamel and dentin densities with the exception of carious dentin. Residual proteins in demineralized dentin may be a confounding factor. These findings suggest that thorough beam hardening correction and the use of very high density phantoms for enamel calibration may not be necessary for accurate dental tissue imaging.
5. Overall, the 3D imaging modalities used in this study demonstrate better contrast performance than the 2D modalities. The panoramic bitewing modality with tomographic image acquisition geometry demonstrates the worst contrast performance.

**Tables and Figures**

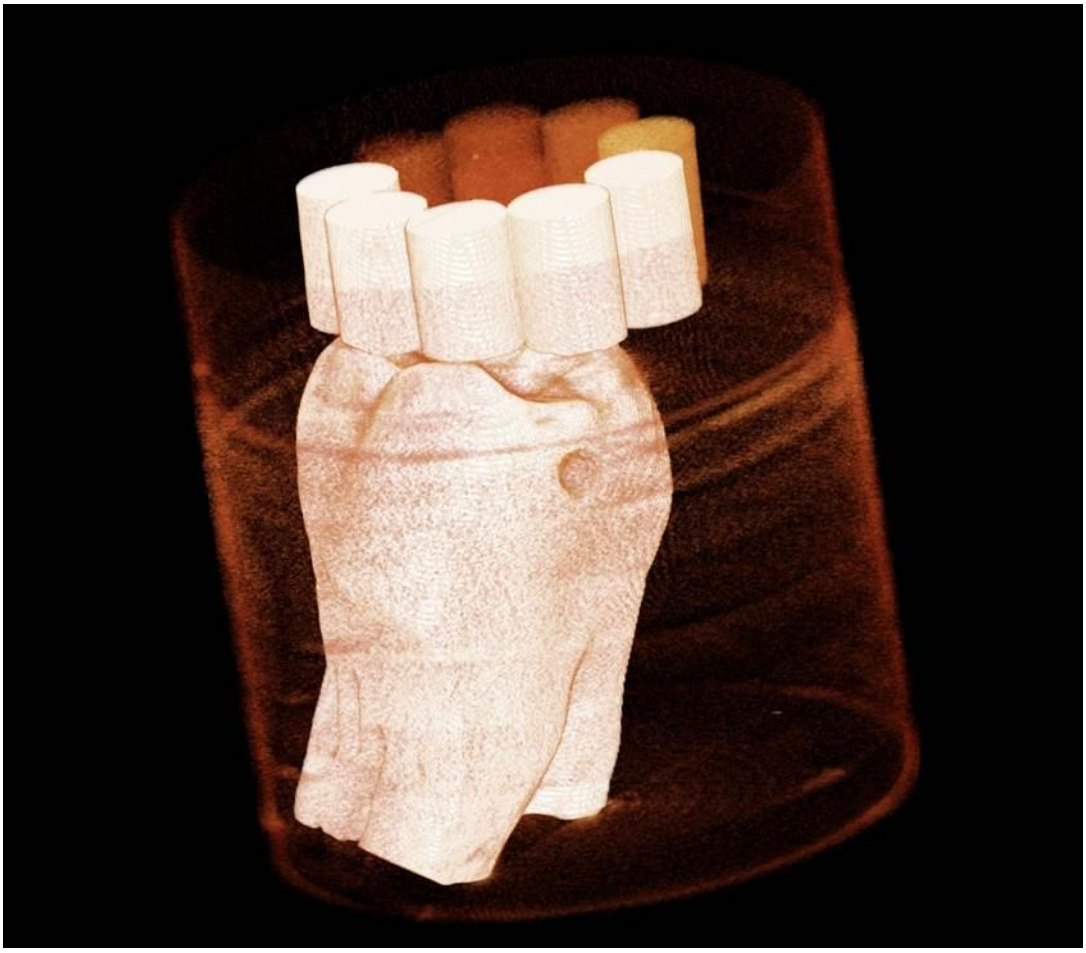

Fig. 9. Volumetric rendering of micro-CT scan of extracted tooth with caries and calibration rods.

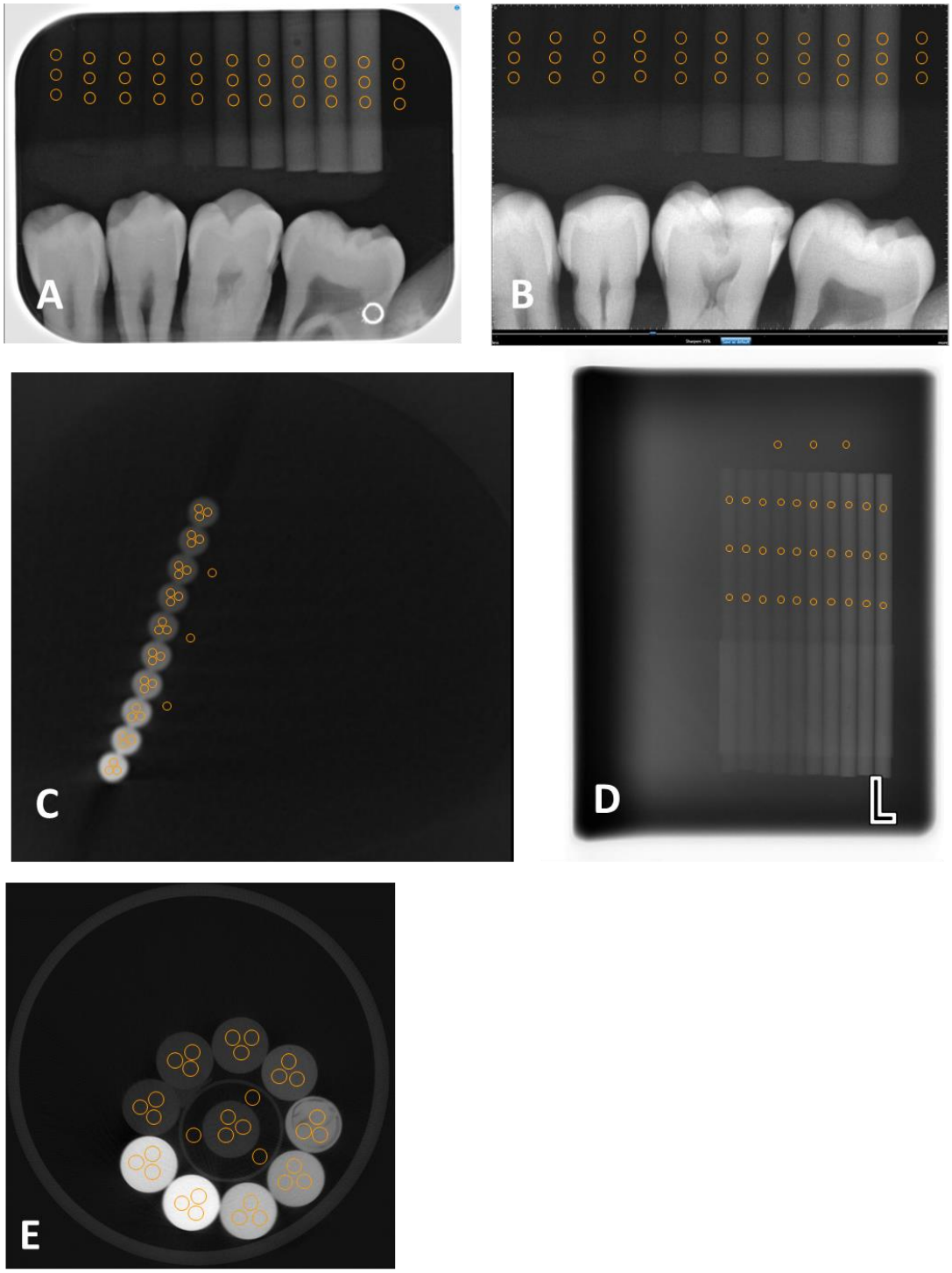


Fig. 10. Representative region of interest selections for each rod and background for PSP (A), Schick 33 (B) Sirona Orthophos XG3D CBCT (C), Planmeca panoramic bitewing images (D), and micro-CT (E).

**Table 14.** Mean and standard deviation (SD) grayscale value by rod density by modality

| *Rod Density (mg/cc)* | *Modality* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Micro-CT* | | *PSP* | | *Schick33* | | *PanBW* | | *XG3D* | |
| | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| Background | 129 | 58 | 1.2 | 0.6 | 845 | 171 | 82 | 1.7 | -368 | 4 |
| Resin | 1883 | 42 | 5.3 | 1.5 | 1043 | 153 | 94 | 1.6 | 43 | 11 |
| 12.5 | 1728 | 60 | 4.7 | 1.2 | 1040 | 156 | 92 | 1.5 | 47 | 11 |
| 25 | 1840 | 45 | 4.4 | 1.0 | 1041 | 160 | 89 | 1.5 | 52 | 11 |
| 50 | 2098 | 27 | 4.6 | 1.0 | 1048 | 163 | 86 | 1.3 | 63 | 12 |
| 150 | 3130 | 51 | 7.5 | 1.3 | 1116 | 158 | 86 | 1.7 | 131 | 13 |
| 500 | 6368 | 128 | 20.9 | 1.4 | 1333 | 153 | 92 | 1.6 | 341 | 13 |
| 750 | 8218 | 53 | 29.2 | 1.3 | 1446 | 151 | 95 | 1.5 | 473 | 20 |
| 1000 | 10666 | 55 | 40.7 | 1.8 | 1596 | 146 | 102 | 1.4 | 720 | 21 |
| 1250 | 12690 | 82 | 50.0 | 1.9 | 1707 | 142 | 106 | 1.4 | 923 | 24 |
| 1500 | 14584 | 73 | 60.2 | 2.0 | 1810 | 139 | 110 | 1.6 | 1156 | 41 |

**Table 15.** Regression and R-square by modality

| *Modality* | *Formula* | *R-square* |
|---|---|---|
| Micro-CT | y = 8.7243x + 1736.5 | 0.9999 |
| PSP | y = 0.0377x + 2.7438 | 0.9981 |
| Schick33 | y = 0.5346x + 1037.4 | 0.9972 |
| PanBW | y = 0.0143x + 86.967 | 0.9983 |
| XG3D | y = 0.7242x + 12.45 | 0.9993 |

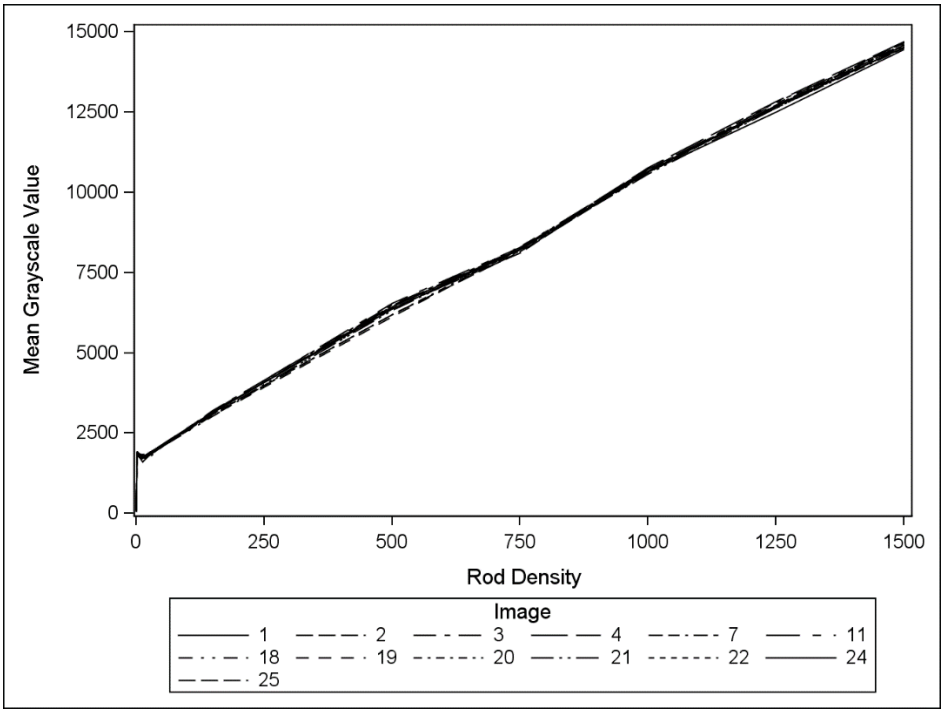


Fig. 11. Mean grayscale value by rod density (mg/cc) for micro-CT.

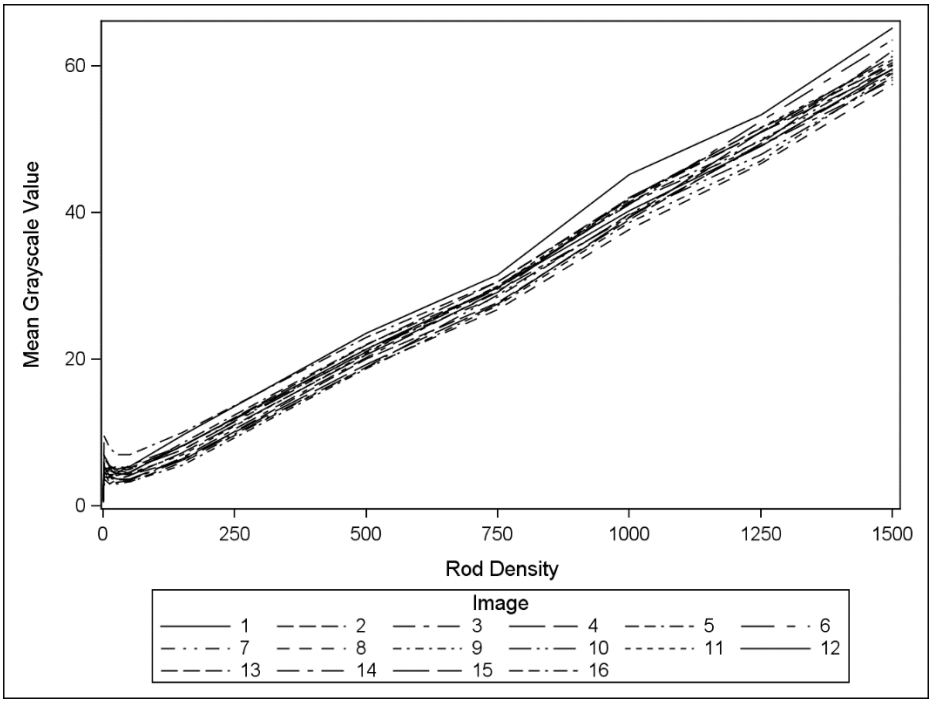


Fig. 12. Mean grayscale value by rod density (mg/cc) for PSP.

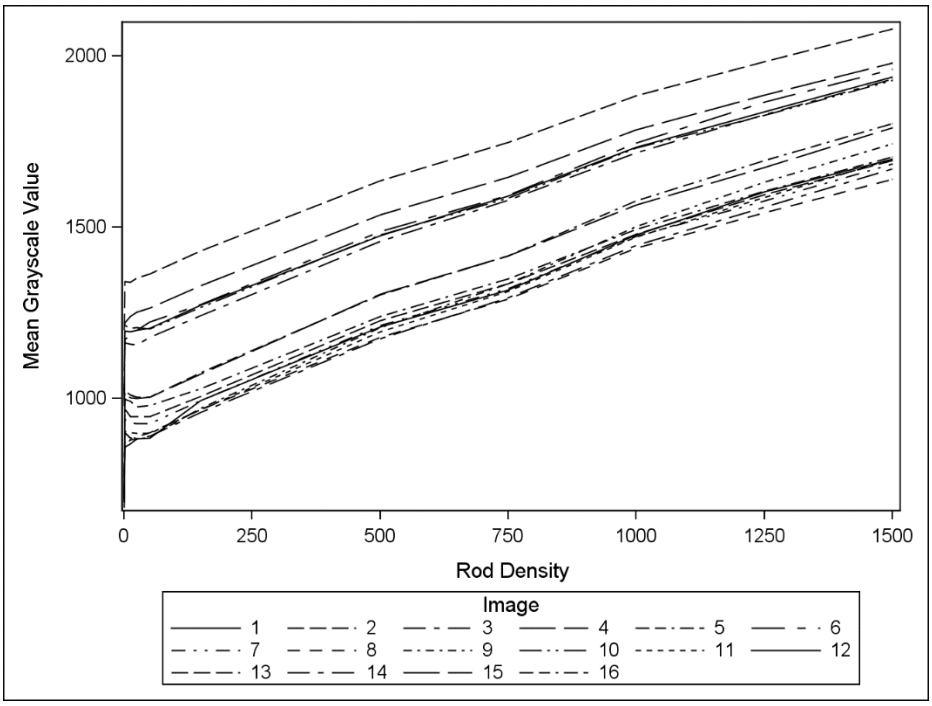


Fig. 13. Mean grayscale value by rod density (mg/cc) for Schick33.

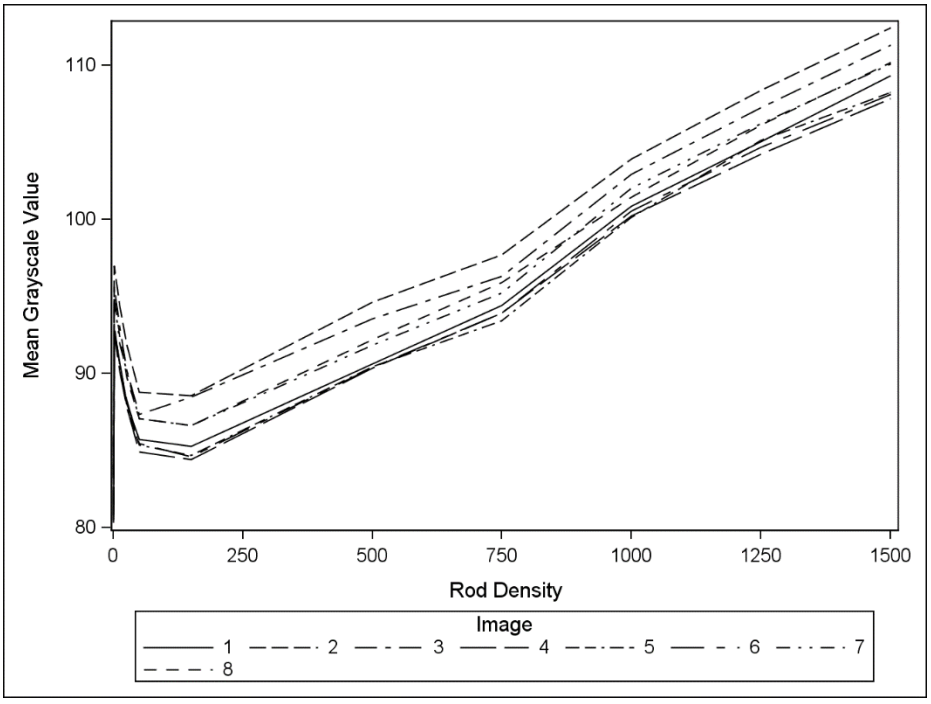


Fig. 14. Mean grayscale value by rod density (mg/cc) for PanBW.

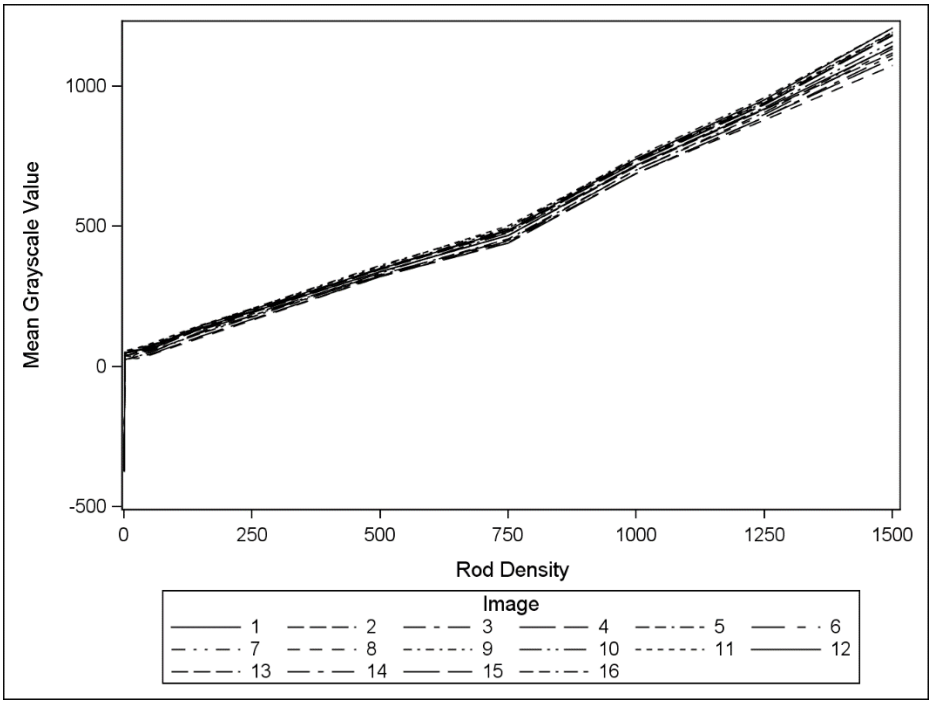


Fig. 15. Mean grayscale value by rod density (mg/cc) for XG3D.

**Table 16.** Summary of "failed" (non-significant) pairwise comparisons of average grayscale value (considering resin rod)

| *Modality* | *Rod Density Comparison* | *Difference Between Rod Densities* | *Difference between grayscale means* | *95% Confidence Limits* | |
|---|---|---|---|---|---|
| *Micro-CT* | *1 - 25* | 24.0 | 42.7 | -34.7 | 120.1 |
| *PSP* | *1 - 12.5* | 11.5 | 0.6 | -0.5 | 1.7 |
| | *1 - 25* | 24.0 | 0.8 | -0.3 | 1.9 |
| | *1 - 50* | 49.0 | 0.7 | -0.4 | 1.8 |
| | *12.5 - 25* | 12.5 | 0.2 | -0.9 | 1.3 |
| | *12.5 - 50* | 37.5 | 0.1 | -1.0 | 1.2 |
| | *25 - 50* | 25.0 | -0.2 | -1.3 | 0.9 |
| *Schick33* | *1 - 12.5* | 11.5 | 3.0 | -19.5 | 25.5 |
| | *1 - 25* | 24.0 | 1.2 | -21.3 | 23.7 |
| | *1 - 50* | 49.0 | -5.7 | -28.2 | 16.8 |
| | *12.5 - 25* | 12.5 | -1.7 | -24.2 | 20.8 |
| | *12.5 - 50* | 37.5 | -8.7 | -31.2 | 13.8 |
| | *25 - 50* | 25.0 | -7.0 | -29.5 | 15.5 |
| *PanBW* | *12.5-500* | 487.5 | -0.1 | -0.7 | 0.6 |
| | *50-150* | 100.0 | 0.3 | -0.3 | 1.0 |
| *XG3D* | *1 - 12.5* | 11.5 | -3.9 | -18.4 | 10.6 |
| | *1 - 25* | 24.0 | -9.0 | -23.5 | 5.5 |
| | *1 - 50* | 12.5 | -5.1 | -19.6 | 9.4 |
| | *25 - 50* | 25.0 | -11.2 | -25.7 | 3.3 |

**Table 17.** Minimum distinguishable rod density by modality

| *Modality* | *Ignoring resin rod* | *Considering resin rod* |
|---|---|---|
| | *Rod density (mg/cc)* | |
| *Micro-CT* | *12.5* | *50* |
| *PSP* | *150* | *150* |
| *Schick33* | *150* | *150* |
| *PanBW* | *750* | *750* |
| *XG3D* | *150* | *150* |

**Table 18.** Summary of reported mineral density of enamel and dentin density in healthy and caries states, regression predicted micro-CT grayscale values, and observed micro-CT grayscale values

| *Study* | *Mineral Density (mg/cc)* | | | |
|---|---|---|---|---|
| | *Sound Enamel* | *Caries Enamel* | *Sound Dentin* | *Caries Dentin* |
| He et al.[96] | 2228 | | | |
| He et al.[110] | 2006-2423 | | | |
| Wong et al.[95] | 2690-2920 | | | |
| Dowker et al.[77] | 2750-2780 | 900-1800 | | |
| Dowker et al.[79] | 2800 | 1800-2500 | | |
| Huang et al.[78] | 2650-2890 | 1480-2580 | | |
| Neves et al.[82] | 2890 | 270 | 1740 | |
| Clementino-Leudemann and Kunzelmann[93] | 2570-2760 | | 1450-1530 | |
| Kinney et al.[111] | | | 1290 | 550 |
| Ahmed et al. [81] | | | 1440 | 300-430 |
| Approx. mean reported density (standard deviation) | 2504 (304) | 1714 (684) | 1490 (140) | 430 (102) |
| Regression predicted micro-CT grayscale value | 23584 | 16690 | 14735 | 5459 |
| Mean observed micro-CT grayscale value (standard deviation) | 24105 (397) | 17455 (2501) | 14310 (405) | 10596 (1750) |

## REFERENCES


[1] Diagnosis and management of dental caries throughout life. National Institutes of Health Consensus Development Conference statement, March 26-28, 2001. Journal of dental education 2001 Oct;65(10):1162-8.

[2] Wenzel A. Radiographic display of carious lesions and cavitation in approximal surfaces: Advantages and drawbacks of conventional and advanced modalities. Acta odontologica Scandinavica 2014 May;72(4):251-64.

[3] Mann J, Pettigrew JC, Revach A, Arwas JR, Kochavi D. Assessment of the DMF-S index with the use of bitewing radiographs. Oral surgery, oral medicine, and oral pathology 1989 Nov;68(5):661-5.

[4] Braga MM, Mendes FM, Ekstrand KR. Detection activity assessment and diagnosis of dental caries lesions. Dental clinics of North America 2010 Jul;54(3):479-93.

[5] Wenzel A. Bitewing and Digital Bitewing Radiography for Detection of Caries Lesions. Journal of Dental Research 2004;83(suppl 1):C72-C5.

[6] White SC, Yoon DC. Comparative performance of digital and conventional images for detecting proximal surface caries. Dento maxillo facial radiology 1997 Jan;26(1):32-8.

[7] Dove SB. Radiographic diagnosis of dental caries. Journal of dental education 2001 Oct;65(10):985-90.

[8] Douglass CW, Valachovic RW, Wijesinha A, Chauncey HH, Kapur KK, McNeil BJ. Clinical efficacy of dental radiography in the detection of dental caries and periodontal diseases. Oral surgery, oral medicine, and oral pathology 1986 Sep;62(3):330-9.

[9] Bader JD, Shugars DA, Bonito AJ. Systematic reviews of selected dental caries diagnostic and management methods. Journal of dental education 2001 Oct;65(10):960-8.

[10] Khan EA, Tyndall DA, Ludlow JB, Caplan D. Proximal caries detection: Sirona Sidexis versus Kodak Ektaspeed Plus. General dentistry 2005 Jan-Feb;53(1):43-8.

[11] Wenzel A. Digital radiography and caries diagnosis. Dento maxillo facial radiology 1998 Jan;27(1):3-11.

[12] Wenzel A. Digital imaging for dental caries. Dental clinics of North America 2000 Apr;44(2):319-38, vi.

[13] Syriopoulos K, Sanderink GC, Velders XL, van der Stelt PF. Radiographic detection of approximal caries: a comparison of dental films and digital imaging systems. Dento maxillo facial radiology 2000 Sep;29(5):312-8.

[14] Dove SB, McDavid WD. A comparison of conventional intra-oral radiography and computer imaging techniques for the detection of proximal surface dental caries. Dento maxillo facial radiology 1992 Aug;21(3):127-34.

[15] Senel B, Kamburoglu K, Ucok O, Yuksel SP, Ozen T, Avsever H. Diagnostic accuracy of different imaging modalities in detection of proximal caries. Dento maxillo facial radiology 2010 Dec;39(8):501-11.

[16] Wenzel A. A review of dentists' use of digital radiography and caries diagnosis with digital systems. Dento maxillo facial radiology 2006 Sep;35(5):307-14.

[17] Kashima I. Computed radiography with photostimulable phosphor in oral and maxillofacial radiology. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 1995 Nov;80(5):577-98.

[18] Huda W, Rill LN, Benn DK, Pettigrew JC. Comparison of a photostimulable phosphor system with film for dental radiology. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 1997 Jun;83(6):725-31.

[19] Li G, Yoshiura K, Welander U, Shi XQ, McDavid WD. Detection of approximal caries in digital radiographs before and after correction for attenuation and visual response. An in vitro study. Dento maxillo facial radiology 2002 Mar;31(2):113-6.

[20] Moystad A, Svanaes DB, van der Stelt PF, Grondahl HG, Wenzel A, van Ginkel FC, et al. Comparison of standard and task-specific enhancement of Digora storage phosphor images for approximal caries diagnosis. Dento maxillo facial radiology 2003 Nov;32(6):390-6.

[21] Talaeipour A, Hafezi L, Niktash A, Mirarjomandi S. Proximal dental enamel caries diagnosis in digital radiography with and without sharpening enhancement filter (In vitro). Journal of Research in Dental Sciences 2015 January 11(Number 4 (42)):214 - 9.

[22] Koob A, Sanden E, Hassfeld S, Staehle HJ, Eickholz P. Effect of digital filtering on the measurement of the depth of proximal caries under different exposure conditions. American journal of dentistry 2004 Dec;17(6):388-93.

[23] Schweitzer DM, Berg RW. A digital radiographic artifact: A clinical report. The Journal of prosthetic dentistry 2010 Jun;103(6):326-9.

[24] Akkaya N, Kansu O, Kansu H, Cagirankaya LB, Arslan U. Comparing the accuracy of panoramic and intraoral radiography in the diagnosis of proximal caries. Dento maxillo facial radiology 2006 May;35(3):170-4.

[25] Flint DJ, Paunovich E, Moore WS, Wofford DT, Hermesch CB. A diagnostic comparison of panoramic and intraoral radiographs. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology 1998 6//;85(6):731-5.

[26] Akarslan ZZ, Akdevelioglu M, Gungor K, Erten H. A comparison of the diagnostic accuracy of bitewing, periapical, unfiltered and filtered digital panoramic images for approximal caries detection in posterior teeth. Dento maxillo facial radiology 2008 Dec;37(8):458-63.

[27] Khan EA, Tyndall DA, Caplan D. Extraoral imaging for proximal caries detection: Bitewings vs scanogram. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology 2004;98(6):730-7.

[28] Kamburoglu K, Kolsuz E, Murat S, Yuksel S, Ozen T. Proximal caries detection accuracy using intraoral bitewing radiography, extraoral bitewing radiography and panoramic radiography. Dento maxillo facial radiology 2012 Sep;41(6):450-9.

[29] Park YS, Ahn JS, Kwon HB, Lee SP. Current status of dental caries diagnosis using cone beam computed tomography. Imaging science in dentistry 2011 Jun;41(2):43-51.

[30] Rathore S, Tyndall D, Wright J, Everett E. Ex vivo comparison of Galileos cone beam CT and intraoral radiographs in detecting occlusal caries. Dento maxillo facial radiology 2012 Sep;41(6):489-93.

[31] Valizadeh S, Tavakkoli MA, Karimi Vasigh H, Azizi Z, Zarrabian T. Evaluation of Cone Beam Computed Tomography (CBCT) System: Comparison with Intraoral Periapical Radiography in Proximal Caries Detection. Journal of dental research, dental clinics, dental prospects 2012 Winter;6(1):1-5.

[32] Wenzel A, Hirsch E, Christensen J, Matzen LH, Scaf G, Frydenberg M. Detection of cavitated approximal surfaces using cone beam CT and intraoral receptors. Dento maxillo facial radiology 2013;42(1):39458105.

[33] 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 surgery, oral medicine, oral pathology, oral radiology, and endodontics 2007 Sep;104(3):412-6.

[34] Kalathingal SM, Mol A, Tyndall DA, Caplan DJ. In vitro assessment of cone beam local computed tomography for proximal caries detection. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 2007 Nov;104(5):699-704.

[35] Kayipmaz S, Sezgin OS, Saricaoglu ST, Can G. An in vitro comparison of diagnostic abilities of conventional radiography, storage phosphor, and cone beam computed tomography to determine occlusal and approximal caries. European journal of radiology 2011 Nov;80(2):478-82.

[36] Young SM, Lee JT, Hodges RJ, Chang TL, Elashoff DA, White SC. A comparative study of high-resolution cone beam computed tomography and charge-coupled device sensors for detecting caries. Dento maxillo facial radiology 2009 Oct;38(7):445-51.

[37] Haiter-Neto F, Wenzel A, Gotfredsen E. Diagnostic accuracy of cone beam computed tomography scans compared with intraoral image modalities for detection of caries lesions. Dento maxillo facial radiology 2008 Jan;37(1):18-22.

[38] Schulze R, Heil U, Gross D, Bruellmann DD, Dranischnikow E, Schwanecke U, et al. Artefacts in CBCT: a review. Dento maxillo facial radiology 2011 Jul;40(5):265-73.

[39] Draenert FG, Coppenrath E, Herzog P, Muller S, Mueller-Lisse UG. Beam hardening artefacts occur in dental implant scans with the NewTom cone beam CT but not with the dental 4-row multidetector CT. Dento maxillo facial radiology 2007 May;36(4):198-203.

[40] De Man B, Nuyts J, Dupont P, Marchal G, Suetens P. Metal streak artifacts in X-ray computed tomography: a simulation study. Nuclear Science Symposium, 1998 Conference Record 1998 IEEE; 1998 1998. p. 1860-5 vol.3.

[41] Zhang Y, Zhang L, Zhu XR, Lee AK, Chambers M, Dong L. Reducing metal artifacts in cone-beam CT images by preprocessing projection data. International journal of radiation oncology, biology, physics 2007 Mar 1;67(3):924-32.

[42] Meilinger M, Schmidgunst C, Schutz O, Lang EW. Metal artifact reduction in cone beam computed tomography using forward projected reconstruction information. Zeitschrift fur medizinische Physik 2011 Sep;21(3):174-82.

[43] Schulze RK, Berndt D, d'Hoedt B. On cone-beam computed tomography artifacts induced by titanium implants. Clinical oral implants research 2010 Jan;21(1):100-7.

[44] Dong J, Kondo A, Abe K, Hayakawa Y. Successive iterative restoration applied to streak artifact reduction in X-ray CT image of dento-alveolar region. International journal of computer assisted radiology and surgery 2011 Sep;6(5):635-40.

[45] Bechara BB, Moore WS, McMahan CA, Noujeim M. Metal artefact reduction with cone beam CT: an in vitro study. Dento maxillo facial radiology 2012 Mar;41(3):248-53.

[46] Bechara B, McMahan CA, Geha H, Noujeim M. Evaluation of a cone beam CT artefact reduction algorithm. Dento maxillo facial radiology 2012 Jul;41(5):422-8.

[47] Tyndall DA, Rathore S. Cone-beam CT diagnostic applications: caries, periodontal bone assessment, and endodontic applications. Dental clinics of North America 2008 Oct;52(4):825-41, vii.

[48] Baelum V. What is an appropriate caries diagnosis? Acta odontologica Scandinavica 2010 Mar;68(2):65-79.

[49] Kidd EA, Fejerskov O. What constitutes dental caries? Histopathology of carious enamel and dentin related to the action of cariogenic biofilms. J Dent Res 2004;83 Spec No C:C35-8.

[50] Selwitz RH, Ismail AI, Pitts NB. Dental caries. The Lancet 2007;369(9555):51-9.

[51] Wenzel A. Current trends in radiographic caries imaging. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 1995 Nov;80(5):527-39.

[52] Young DA, Featherstone JDB. Digital Imaging Fiber-Optic Trans-Illumination, F-speed radiographic film and depth of approximal lesions. The Journal of the American Dental Association 2005;136(12):1682-7.

[53] Jacobsen JH, Hansen B, Wenzel A, Hintze H. Relationship between histological and radiographic caries lesion depth measured in images from four digital radiography systems. Caries research 2004 Jan-Feb;38(1):34-8.

[54] Ekstrand KR, Luna LE, Promisiero L, Cortes A, Cuevas S, Reyes JF, et al. The reliability and accuracy of two methods for proximal caries detection and depth on directly visible proximal surfaces: an in vitro study. Caries research 2011;45(2):93-9.

[55] Kamburoglu K, Kurt H, Kolsuz E, Oztas B, Tatar I, Celik HH. Occlusal caries depth measurements obtained by five different imaging modalities. Journal of digital imaging 2011 Oct;24(5):804-13.

[56] Akdeniz BG, Grondahl HG, Magnusson B. Accuracy of proximal caries depth measurements: comparison between limited cone beam computed tomography, storage phosphor and film radiography. Caries research 2006;40(3):202-7.

[57] Bille J, Thylstrup A. Radiographic diagnosis and clinical tissue changes in relation to treatment of approximal carious lesions. Caries research 1982;16(1):1-6.

[58] Sansare K, Raghav M, Sontakke S, Karjodkar F, Wenzel A. Clinical cavitation and radiographic lesion depth in proximal surfaces in an Indian population. Acta odontologica Scandinavica 2014 Nov;72(8):1084-8.

[59] Ferreira Zandona A, Santiago E, Eckert GJ, Katz BP, Pereira de Oliveira S, Capin OR, et al. The natural history of dental caries lesions: a 4-year observational study. J Dent Res 2012 Sep;91(9):841-6.

[60] Thylstrup A, Bille J, Qvist V. Radiographic and observed tissue changes in approximal carious lesions at the time of operative treatment. Caries research 1986;20(1):75-84.

[61] Sansare K, Singh D, Sontakke S, Karjodkar F, Saxena V, Frydenberg M, et al. Should cavitation in proximal surfaces be reported in cone beam computed tomography examination? Caries research 2014;48(3):208-13.

[62] Langlais RP. 6 reasons why you should use a panoramic X-ray for bitewings. 2013 [cited 2013 May 16, 2013];

[63] Hintze H, Wenze A. Clinical and laboratory radiographic caries diagnosis. A study of the same teeth. Dento maxillo facial radiology 1996 Jun;25(3):115-8.

[64] Metz CE. Basic principles of ROC analysis. Seminars in nuclear medicine 1978 Oct;8(4):283-98.

[65] Metz CE. ROC Methodology in Radiologic Imaging. Investigative Radiology 1986;21(9):720-33.

[66] ten Bosch JJ, Angmar-Mansson B. Characterization and validation of diagnostic methods. Monographs in oral science 2000;17:174-89.

[67] Pauwels R, Stamatakis H, Bosmans H, Bogaerts R, Jacobs R, Horner K, et al. Quantification of metal artifacts on cone beam computed tomography images. Clinical oral implants research 2013 Aug;24 Suppl A100:94-9.

[68] Rockette HE, Campbell WL, Britton CA, Holbert JM, King JL, Gur D. Empiric assessment of parameters that affect the design of multireader receiver operating characteristic studies. Academic radiology 1999 Dec;6(12):723-9.

[69] Ludlow JB, Davies-Ludlow LE, White SC. Patient risk related to common dental radiographic examinations: the impact of 2007 International Commission on Radiological Protection recommendations regarding dose calculation. Journal of the American Dental Association (1939) 2008 Sep;139(9):1237-43.

[70] Ludlow JB, Ivanovic M. Comparative dosimetry of dental CBCT devices and 64-slice CT for oral and maxillofacial radiology. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 2008 Jul;106(1):106-14.

[71] Ludlow JB, Timothy R, Walker C, Hunter R, Benavides E, Samuelson DB, et al. Effective dose of dental CBCT-a meta analysis of published data and additional data for nine CBCT units. Dento maxillo facial radiology 2015;44(1):20140197.

[72] Obuchowski NA. Multireader receiver operating characteristic studies: a comparison of study designs. Academic radiology 1995 Aug;2(8):709-16.

[73] Obuchowski NA. How many observers are needed in clinical studies of medical imaging? AJR American journal of roentgenology 2004 Apr;182(4):867-9.

[74] Obuchowski NA. Sample size tables for receiver operating characteristic studies. AJR American journal of roentgenology 2000 Sep;175(3):603-8.

[75] Soviero VM, Leal SC, Silva RC, Azevedo RB. Validity of MicroCT for in vitro detection of proximal carious lesions in primary molars. Journal of dentistry 2012 Jan;40(1):35-40.

[76] Park Y-S, Bae K-H, Chang J, Shon W-J. Theory of X-ray microcomputed tomography in dental research: application for the caries research. Journal of Korean Academy of Conservative Dentistry 2011;36(2):98.

[77] Dowker SEP, Dowker SEP, Elliott JC, Elliott JC, Davis GR, Davis GR, et al. Longitudinal Study of the Three-Dimensional Development of Subsurface Enamel Lesions during in vitro Demineralisation. Caries research 2003;37(4):237-45.

[78] Huang TT, Jones AS, He LH, Darendeliler MA, Swain MV. Characterisation of enamel white spot lesions using X-ray micro-tomography. Journal of dentistry 2007 Sep;35(9):737-43.

[79] Dowker SE, Elliott JC, Davis GR, Wilson RM, Cloetens P. Synchrotron x-ray microtomographic investigation of mineral concentrations at micrometre scale in sound and carious enamel. Caries research 2004 Nov-Dec;38(6):514-22.

[80] Lo EC, Zhi QH, Itthagarun A. Comparing two quantitative methods for studying remineralization of artificial caries. Journal of dentistry 2010 Apr;38(4):352-9.

[81] Ahmed M, Davis GR, Wong FS. X-ray microtomography study to validate the efficacies of caries removal in primary molars by hand excavation and chemo-mechanical technique. Caries research 2012;46(6):561-7.

[82] Neves Ade A, Coutinho E, Vivan Cardoso M, Jaecques SV, Van Meerbeek B. Micro-CT based quantitative evaluation of caries excavation. Dental materials : official publication of the Academy of Dental Materials 2010 Jun;26(6):579-88.

[83] Hamba H, Nikaido T, Sadr A, Nakashima S, Tagami J. Enamel lesion parameter correlations between polychromatic micro-CT and TMR. J Dent Res 2012 Jun;91(6):586-91.

[84] Mitropoulos P, Rahiotis C, Stamatakis H, Kakaboura A. Diagnostic performance of the visual caries classification system ICDAS II versus radiography and micro-computed tomography for proximal caries detection: an in vitro study. Journal of dentistry 2010 Nov;38(11):859-67.

[85] Taylor AM, Satterthwaite JD, Ellwood RP, Pretty IA. An automated assessment algorithm for micro-CT images of occlusal caries. The surgeon : journal of the Royal Colleges of Surgeons of Edinburgh and Ireland 2010 Dec;8(6):334-40.

[86] Zou W, Hunter N, Swain MV. Application of polychromatic microCT for mineral density determination. J Dent Res 2011 Jan;90(1):18-30.

[87] Luedemann TNRC, Kunzelmann KH, Matsura M, Hickel R. Attenuation coefficient in X-ray Microtomography-XMT for enamel, dentin and caries. 2005.

[88] Schmitz JE, Teepe JD, Hu Y, Smith CE, Fajardo RJ, Chun YH. Estimating mineral changes in enamel formation by ashing/BSE and microCT. J Dent Res 2014 Mar;93(3):256-62.

[89] Plotino G, Grande NM, Pecci R, Bedini R, Pameijer CH, Somma F. Three-dimensional imaging using microcomputed tomography for studying tooth macromorphology. Journal of the American Dental Association (1939) 2006 Nov;137(11):1555-61.

[90] Hahn SK, Kim JW, Lee SH, Kim CC, Hahn SH, Jang KT. Microcomputed tomographic assessment of chemomechanical caries removal. Caries research 2004 Jan-Feb;38(1):75-8.

[91] Davis GR, Wong FS. X-ray microtomography of bones and teeth. Physiological measurement 1996 Aug;17(3):121-46.

[92] Elliott JC, Wong FS, Anderson P, Davis GR, Dowker SE. Determination of mineral concentration in dental enamel from X-ray attenuation measurements. Connective tissue research 1998;38(1-4):61-72; discussion 3-9.

[93] Clementino-Luedemann TN, Kunzelmann KH. Mineral concentration of natural human teeth by a commercial micro-CT. Dental materials journal 2006 Mar;25(1):113-9.

[94] Schwass DR, Swain MV, Purton DG, Leichter JW. A system of calibrating microtomography for use in caries research. Caries research 2009;43(4):314-21.

[95] Wong FS, Anderson P, Fan H, Davis GR. X-ray microtomographic study of mineral concentration distribution in deciduous enamel. Archives of oral biology 2004 Nov;49(11):937-44.

[96] He B, Huang S, Jing J, Hao Y. Measurement of hydroxyapatite density and Knoop hardness in sound human enamel and a correlational analysis between them. Archives of oral biology 2010 Feb;55(2):134-41.

[97] Davis GR, Evershed AN, Mills D. Quantitative high contrast X-ray microtomography for dental research. Journal of dentistry 2013 May;41(5):475-82.

[98] Mulder L, Koolstra JH, Eijden TMGJV. Accuracy of MicroCT in the Quantitative Determination of the Degree and Distribution of Mineralization in Developing Bone. Acta Radiologica 2004;45(7):769-77.

[99] Zou W, Gao J, Jones AS, Hunter N, Swain MV. Characterization of a novel calibration method for mineral density determination of dentine by X-ray micro-tomography. The Analyst 2009 Jan;134(1):72-9.

[100] Farah RA, Swain MV, Drummond BK, Cook R, Atieh M. Mineral density of hypomineralised enamel. Journal of dentistry 2010 Jan;38(1):50-8.

[101] Schweizer S, Hattendorf B, Schneider P, Aeschlimann B, Gauckler L, Muller R, et al. Preparation and characterization of calibration standards for bone density determination by micro-computed tomography. The Analyst 2007 Oct;132(10):1040-5.

[102] Salmon PL, Liu X. MicroCT Bone Densitometry: Context Sensitivity, Beam Hardening Correction and the Effect of Surrounding Media. The Open Access Journal of Science and Technology 2014;2:1-25.

[103] Doyle P, Finney L. Performance evaluation and testing of digital intra-oral radiographic systems. Radiation protection dosimetry 2005;117(1-3):313-7.

[104] Gulliksrud K, Stokke C, Martinsen AC. How to measure CT image quality: variations in CT-numbers, uniformity and low contrast resolution for a CT quality assurance phantom. Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics 2014 Jun;30(4):521-6.

[105] Gardner SJ, Studenski MT, Giaddui T, Cui Y, Galvin J, Yu Y, et al. Investigation into image quality and dose for different patient geometries with multiple cone-beam CT systems. Medical physics 2014 Mar;41(3):031908.

[106] Neubauer J, Voigt JM, Lang H, Scheuer C, Goerke SM, Langer M, et al. Comparing the image quality of a mobile flat-panel computed tomography and a multidetector computed tomography: a phantom study. Invest Radiol 2014 Jul;49(7):491-7.

[107] McLaughlin PD, Murphy KP, Hayes SA, Carey K, Sammon J, Crush L, et al. Non-contrast CT at comparable dose to an abdominal radiograph in patients with acute renal colic; impact of iterative reconstruction on image quality and diagnostic performance. Insights into imaging 2014 Apr;5(2):217-30.

[108] Mischkowski RA, Scherer P, Ritter L, Neugebauer J, Keeve E, Zoller JE. Diagnostic quality of multiplanar reformations obtained with a newly developed cone beam device for maxillofacial imaging. Dento maxillo facial radiology 2008 Jan;37(1):1-9.

[109] Jain A, Bednarek DR, Rudin S. Scatter reduction for high resolution image detectors with a region of interest attenuator. Proceedings - Society of Photo-Optical Instrumentation Engineers 2014 Mar 19;9033:903364.

[110] He B, Huang S, Zhang C, Jing J, Hao Y, Xiao L, et al. Mineral densities and elemental content in different layers of healthy human enamel with varying teeth age. Archives of oral biology 2011 Oct;56(10):997-1004.

[111] Kinney JH, Marshall GW, Jr., Marshall SJ. Three-dimensional mapping of mineral densities in carious dentin: theory and method. Scanning microscopy 1994;8(2):197-204; discussion -5.

[112] Kovacs M, Danyi R, Erdelyi M, Fejerdy P, Dobo-Nagy C. Distortional effect of beam-hardening artefacts on microCT: a simulation study based on an in vitro caries model. Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics 2009 Oct;108(4):591-9.