The Accuracy and Sensitivity of ABO Electronic Cast Radiographic Program ORTHOSHARE 360 Compared to Manual Measurements

Lisa Nguyen

Marquette University

Follow this and additional works at: https://epublications.marquette.edu/theses_open

Part of the Dentistry Commons

Recommended Citation


https://epublications.marquette.edu/theses_open/720
THE ACCURACY AND SENSITIVITY OF ABO'S ELECTRONIC CAST RADIOPHORIC – EVALUATION (CRE) ORTHOSHAPE 360 PROGRAM MEASUREMENT COMPARED TO TRADITIONAL MANUAL CRE MEASUREMENTS OF 3D PRINTED MODELS

by

Lisa P. Nguyen, DMD

A Thesis submitted to the Faculty of the Graduate School, Marquette University, in Partial Fulfillment of the Requirement for the Degree of Master of Science

Milwaukee, Wisconsin
August 2022
Objective:
The objective of this study was to assess the accuracy, validity, and reliability of measurements obtained by MotionView's Ortho Share 360 program compared to traditional manual measurements of 3D printed models using a standardized ABO measuring gauge. In detail, the study aimed to evaluate the difference in total ABO Cast Radiographic-Evaluation (CRE) score and individual differences among six ABO CRE parameters. The null hypothesis is that there are no statistically different scores between digital and manual measurements.

Methods:
33 samples were measured based on standardized parameters determined by the ABO. The first measurements were obtained manually using ABO measuring gauge on 3D printed plastic models and the second set of measurements were obtained digitally using Ortho Share 360 program. Intra-examiner reliability was determined with the Spearman test. The model score values for the manual measurements and digital measurements were compared using two-sample T-tests. Pearson analysis was used to test the correlation between manual CRE and digital CRE. A p-value less than 0.05 was considered statistically different.

Results:
An intra-examiner correlation coefficient (ICC) of 0.99 indicated a good consistency of measurements (p < 0.001). The scores from digital models were significantly (p < 0.001) higher than those obtained from manual measurement, with an average of 5.27 points and a standard deviation of 1.39. This difference was due to statistically significant differences in 3 ABO CRE parameters: alignment (1.79, p < 0.001), marginal ridge (1.18, p < 0.001), and buccolingual inclination (1.15, p < 0.001). Digital CRE was significantly correlated with manual CRE (correlation coefficient = 0.828, p < 0.001).

Conclusion:
On average, digital scores were 5.27 higher than manual measurements, with a standard deviation of 1.39. The ABO CRE scores from Ortho Share 3D differed significantly statistically and clinically compared to manual measurements and, therefore, cannot fully
replace manual measurements. Based on the results, the null hypothesis was rejected. If the digital measurements were to be utilized, the range for a passing score for ABO CRE should be re-evaluated to take into consideration the inflated digital scores.
ACKNOWLEDGEMENTS

Lisa P. Nguyen, DMD

I would like to thank Dr. Dawei Liu, my research director, for all his guidance and support in helping bring this project from start to completion. A special thanks to our statistician, Dr. Shentong Han, for all his work processing and analyzing the data, and all my thesis committee members, Dr. Ghada Nimeri, Dr. Marinho Del Santo, and Dr. Shivam Mehta. Thank you to Marquette University for providing materials to 3D print the digital cast models and the team at Ortho Share 360 for their assistance as we navigated the new software. Without all the incredible support, this research would not have been possible.
**TABLE OF CONTENTS**

ACKNOWLEDGEMENTS ........................................................................................................... i

TABLE OF CONTENTS ......................................................................................................... ii

LIST OF TABLES ............................................................................................................... iii

LIST OF FIGURES .............................................................................................................. iv

CHAPTER 1 ....................................................................................................................... 1

CHAPTER 2 ....................................................................................................................... 4

CHAPTER 3 ..................................................................................................................... 12

CHAPTER 4 ..................................................................................................................... 24

CHAPTER 5 ..................................................................................................................... 27

CHAPTER 6 ..................................................................................................................... 31

REFERENCES .................................................................................................................... 32
LIST OF TABLES

Table 1. ABO CRE scores for manual and digital measurements........................................ 24

Table 2. Individual differences among all six parameters.................................................. 25
LIST OF FIGURES

Figure 1. ABO measuring gauge ............................................................ 12
Figure 2. Manual alignment measurement sample .................................. 14
Figure 3. Digital alignment measurement sample ..................................... 14
Figure 4. Manual marginal ridge measurement sample ............................ 15
Figure 5. Digital marginal ridge measurement sample ............................. 16
Figure 6. Manual buccolingual inclination measurement sample ............... 17
Figure 7. Digital buccolingual inclination measurement sample ............... 17
Figure 8. Manual overjet measurement sample ....................................... 19
Figure 9. Digital overjet measurement sample ........................................ 19
Figure 10. Manual occlusal relationship measurement sample .................. 20
Figure 11. Digital occlusal relationship measurement sample ................... 21
Figure 12. ABO guideline for grading root angulation sample ................... 22
Figure 13. Statistical significance of data chart ....................................... 23
Figure 14. Scatterplot of digital versus manual total ABO CRE scores .......... 26
CHAPTER 1
INTRODUCTION

The American Board of Orthodontics was founded in 1929 as the first dental specialty by the American Association of Orthodontists. Per the official ABO website, the board "sets the highest level of patient care and promotes excellence in orthodontics for all its certified orthodontist." The ABO encouraged orthodontists to achieve certification and has created a process of examination to qualify applicants. Prior to COVID, the Board developed a Model Grading System to evaluate the final dental casts and panoramic radiographs of completed cases to allow applicants a means to qualify their work. This assessment provides examinees with a "fair, accurate and meaningful" clinical examination to objectively evaluate finished orthodontic cases that will qualify examinees for Board certification" (Clinical Examination). This scoring system was developed systematically through a series of four field tests over five years. The Board instituted the model and radiographic portions of the Model Grading System, which has been used to grade this portion of the examinees' clinical case reports since 1999 (Clinical Examination).

Historically, several systems have been aimed to qualify orthodontic results, including but not limited to the Eismann method of 1974, the Haeger method of 1992, and the Peer Assessment Rating in 1992, each with their methodology. Each evaluation method often compares the pretreatment and post-treatment Cast to obtain a value for the success of the results. The Eismann method utilized fourteen criteria that are then compared to a standard value believed to be ideal (Goyal et al., 2015). The Haeger method visually quantified the number of ideal intraarch and interarch tooth relationships
present compared to the potential total by one percentage value (Haeger et al., 1992). The Peer Assessment index was a single score that compares the amount of change in deviation from normal alignment and occlusion between pre-and post-treatment cases (Richmond et al., 1992). While each method followed a different parameter, there was never a standard that has undergone as many field tests as the ABO system. This finalized system aims to evaluate the models in 8 criteria: alignment, marginal ridges, buccolingual inclination, occlusal relationships, occlusal contacts, overjet, interproximal contacts, and root angulation. The parameters listed allowed orthodontists to effectively examine the success of their treatment in obtaining esthetic, healthy, and functional occlusion. Although model grading is no longer utilized for Board certification, the Model Grading System remains an invaluable tool to evaluate cases effectively.

Traditionally, the grading system was performed on plaster models as that was the standard material available in clinics. With technological advancements, digital models were gradually adapted into the dental and orthodontic field, with many studies supporting the accuracy of the digital models compared to plaster models. A systemic review by Rossini et al. in 2015 concluded that digital models were as accurate, reliable, and reproducible as plaster models with the additional advantage of "cost, time, and space required," which could make digital models the "new gold standard." (Rossini et al., 2015). Digital models grew from purely study models to currently being used for diagnostic and qualification processes. In 2005, Constalos et al. did a comparative study on the accuracy and reliability of the ABO measuring system on digital models. The conclusion was that digital models might be acceptable for use in ABO model examination with the need for accurate calibration of the examiner.
As digital models continued to gain popularity due to their advantages over plaster models, the topic of efficiency too continued to develop. Many programs have been developed to allow clinicians to directly evaluate digital models online, such as OrthoCAD, which completely removes the need for a physical copy of the model. American Board of Orthodontics has translated this development into an electronic ABO CRE-E assessment program that provides a standardized tool to evaluate digital Cast based on the same eight criteria. This new technology's purpose was to respond to the ever-growing influence of digital technology in the orthodontic profession.

By providing clinicians with a convenient and easy means to grade their finished cases, the ABO was not only able to encourage clinicians to develop professionally but also encouraged more applicants to be Board certified. The objective of this study was to assess the accuracy, validity, and reliability of measurement obtained by the software program Ortho Share 3D by MotionView 3D compared to traditional manual measurements of 3D printed models using the standardized ABO measuring gauge. Since the program's release to date, no study has evaluated the accuracy and reliability of ABO Electronic Cast Radiographic – Evaluation compared to manual measurements completed on physical models.
CHAPTER 2
REVIEW OF LITERATURE

This History of Orthodontic assessment indexes:

Objective evaluation of orthodontic treatment outcomes has been a critical topic of discussion with the goal of having a standard for evaluating success and quality. A summary of the World Health Organization's requirement for dental indices includes reliability, validity, and validity during the time (Casko et al., 1998). Before the current American Board of Orthodontist Objective Grading System (ABO CRE), numerous indexes were used to compare pre-and post-treatment results. However, they all lacked precision, accuracy, or reliability, making them less than ideal to be used as a standardized index (Casko et al. 1998). For example, the Occlusal Index provided a thorough Phase III examination of malocclusion; however, the system was not time efficient and better suited for pretreatment malocclusion (Summers, 1971). In addition, the Peer Assessment Rating of 1987 lacked the sensitivity to diagnose minor teeth misalignment. Other methods, such as Eismann and Gottlieb, were used prevalently in Europe but lacked the reliability and validity needed for a standardized evaluation (Summers, 1971).

In response to the need for a standardized index for measurement, the American Board of Orthodontists (ABO) began developing the current ABO Objective Grading System (ABO CRE) in 1994 and invested five years and four series of field tests before finalizing the release of the system in 1999. (The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs). The goal of implementing this standardized grading system was to encourage orthodontists to score their work and
create a standard of care set by the organization. The index began with 15 criteria that were reduced to eight after multiple field tests confirmed the appropriate and reliable criteria to consider a case completed. These criteria include alignment, marginal ridge, buccolingual inclination, occlusal relationship, occlusal contacts, overjet, interproximal contacts, and root angulation (Casko et al., 2018).

Following the release of the ABO Objective Grading System, research was completed to verify the system's validity, reproducibility, and reliability. A study completed on 108 Chinese patients by Song et al. in 2013 confirmed the system's validity as a reliable index to quantify orthodontic casts. In addition, the study confirmed highly correlated and consistent scores among graders proving the reliability and reproducibility of the index. More specifically, the most sensitive and accurate scores pertained to occlusal relationship, overjet, interproximal contact, and alignment (Song et al., 2013). Another study evaluated the clinical use of ABO and supported the reliability of the system with a .77 intrajudge correlation and .85 interjudge correlation (Lieber et al., 2013).

The data supporting the index's validity, reliability, and reproducibility proves that CRE provides a reliable, standardized assessment of models as intended by the American Board of Orthodontics. While the assessment is valid, clinical considerations should be noted for usage. Lieber et al. found the highest intrajudge correlations in occlusal relationships and the least in interproximal contacts. In contrast, the highest interjudge correlation was buccolingual inclination and least for overjet. Most point deductions were at the buccal segment of second molars. Therefore, while the system has been proven effective, the key to accurate and successful usage of the measuring index
was properly calibrating users. Despite certain limitations, the implementation and awareness of ABO Cast Radiograph Evaluation significantly improve treatment outcomes in postgraduate students (Yılmaz et al., 2016). The research supported the value of a standardized system in providing clinicians with an objective means to evaluate and improve their crafts.

**Accuracy of Digital models compared to Plaster models**

As the world of technology continued to evolve, the demand for efficiency and digital advancements in the dental setting continued to increase. The first digital scanner for dental use began in the 1980s and has evolved the way dentists take and store impressions. As with any changes, the product must be an improvement from the current standard to be accepted. The digital impressions from iTero were proven to be more efficient and accurate than traditional plasters resulting in less need for remakes (Nayar et al., 2015). A reduction in remakes presents a major advantage in the conservation of time and cost for the dentist. As digital technology saved dentists money, time, and storage space, the demand grew. The transition from impression to plaster models advanced to printed 3D models from digital scans, turning offices into full digital practices.

Numerous results have been completed to evaluate the quality of digital models compared to traditional models with conflicting results. For example, a study in 2010 concluded that the measurements done on digital models from OraMetrix were smaller than those on plaster models. Differences between the measurements were greater than 0.5 mm (about 0.02 in) in anterior-posterior, transverse, and vertical dimensions; therefore, a clinically significant difference is seen between data gathered from plaster and digital models (Torassian et al., 2010). In contrast, Santoro et al., 2003 aimed to
evaluate the reliability of the OrthoCAD program compared to measuring models by hand. The results confirmed strong inter-examiner reliability in both methods. While digital measurements for tooth size and overbite were smaller, they were within 0.16mm to 0.49mm, which was not considered clinically significant. The conflicting results were due to differences in software quality, as not all programs are created equally. Based on the studies, OrthoCAD proved superior in accuracy compared to OraMetrix. Many other software studies following the Torassian et al. study have proven accurate.

Labib et al. evaluated the accuracy of digital models from Sirona CEREC Omnicam and 3 shape TRIOS 3 scanners and made a different conclusion. Diagnostic measurements that were evaluated included intermolar, inter premolar, and intercanine width and mesiodistal width of 1st permanent molar, 1st premolar, canine, and central incisors. The results showed excellent reliability as the measurements were consistent between observers with slight but insignificant differences in the measurements (Labib et al., 2015). These results again supported the value of selecting quality software and products that have undergone trials to prove their accuracy.

Despite variations in quality among different software and products, a systemic review done by Rossini et al. in 2016 aimed to evaluate the "accuracy, validity, and reliability of measurements obtained from virtual dental study models" in current literature. Thirty-five articles between January 2000 to November 2014 were qualified under the grading system by Swedish Council on Technology Assessment in Health Care and the Cochrane tool as credible and quality resources. The articles' review supported digital models' reliability with "high accuracy, reliability, and reproducibility."
Furthermore, the study concludes that proper landmark identification remains the most crucial factor in measurement accuracy rather than the measuring software alone (Rossini et al., 2016). This conclusion reiterates the importance of proper user training because the software programs have proven to be comparable, if not more superior, results and measurements compared to traditional models.

A most recent comparison between plaster and digital Cast by Schieffer et al. aimed to evaluate the validity, reliability, reproducibility, and objectivity of specific measurements on traditional plasters compared to digital casts from scanned stone models. Five parameters were measured, including overbite and intermolar distance, and were made digitally and on physical casts with a vernier caliper. The results showed that digital measurements performed on OnyxCeph3 were just as reproducible and reliable as that made on stone models, with errors being lower on digital software with experienced examiners (Scheiffer et al., 2022). The evidence in this research continues to support the use of digital scans and models to increase efficiency but specifically their value in aiding orthodontic treatment planning.

**Software to analyze digital models**

Due to the promising quality of digital models over traditional plaster casts Costalos et al. performed a study to determine whether ABO CRE originally intended for traditional plaster models can be accurately used on digital models through OrthoCAD. The results of the study showed, "The means of the total score and those for marginal ridges, occlusal contacts, occlusal relationships, overjet, and interproximal contacts were not significantly different between plaster and digital models. However, the means for
alignment and buccolingual inclination were significantly different." Based on the results, the study recommended re-evaluating the results for alignment and buccolingual inclination but also emphasizes the importance of calibration for the acceptance of digital models. Therefore, while the results were satisfactory, there were still parameters that need to be improved, and digital models' measurements from OrthoCAD cannot completely replace the traditional manual measurements using the standardized ABO gauge.

An updated version of OrthoCAD 2.2 was studied in 2007 by Okunamai et al. to qualify whether CRE can be measured accurately using the new program. Again, results supported the previous study by Constalos et al. and showed the "mean differences of points deducted ranged from .03 point for marginal ridges to 5.07 points for the total score. The variable with the most points deducted related to occlusal relationships, and the fewest points were deducted for interproximal contacts." The significantly higher deduction scores further supported that OrthoCAD 2.2 is inappropriate for measuring digital Cast for all parameters required by ABO CRE. For this reason, finding better software to perform such measurements was deemed necessary.

While OrthoCAD proved to not be a reliable software for ABO CRE digital measurements due to the inflated scores, other programs showed more promising results. For example, Bell et al. studied Euclidean Distance Matrix Analysis, which yields measurement difference averages of .27mm between traditional and digital measurements (Bell et al., 2003). Bell further suggested that due to this level of accuracy, digital models can fully replace traditional plaster models in the dental office to save cost and time on storage. The main setback with this research was that the researcher did not utilize the
ABO CRE as the measurement system. Therefore, the parameters used to assess accuracy cannot be accurately used to compare the quality of measurements. In addition, Quimby et al. stated that while computer-based models were just as accurate and reliable as plaster models, there was a deficiency in the computer program's ability to predict mandibular and maxillary space available consistent with standard plaster models. There was also a higher chance of variability with the digital system (Quimby et al., 2004).

Outside of objective grading purposes, digital measurement has found much success in other diagnostic parameters, such as accurately predicting arch length discrepancy with slight variability. The cast models resulted in larger arch length measurements than the e-models, with an average of $1.5 \pm 1.36$ mm greater mandibular arch length measurement and an average of $1.47 \pm 1.55$ mm greater maxillary arch length (Mullen et al. 2007). Mullen et al. (2007) used the E-model software. They found that Bolton ratio measurements were just as accurate and faster when digital calipers were used than gauge measurements on plaster models.

As multiple research supported the accuracy of dental software to measure 3D casts digitally for diagnostic purposes directly, Hildebrand et al. searched for an accurate digital program to utilize with ABO CRE. A digital version of ABO CRE called OrthoCAD (Cadent, Fairview, NJ) was compared with the standard ABO gauge for plaster models. The results showed a $9.0 \pm 5.4$ points discrepancy between digital and manual grading. With this large discrepancy, the researchers concluded that this program could not be used as a reliable substitute for manual grading. In conclusion, while many digital programs have shown promising results in aiding orthodontic diagnosis and
treatment planning, no program has been proven to accurately replace traditional manual measurements for ABO CRE.
CHAPTER 3
MATERIALS AND METHODS

STUDY DESIGN

The study consisted of 33 patients who completed orthodontic treatment at the Marquette University School of Dentistry Orthodontic Residency program. The patients had intraoral scans via iTero at the final appointment after braces were removed, and final panoramic radiographs were obtained to evaluate final root angulation. The post-treatment digital scans were saved and exported into STL files and sent to 3D Systems NextDent5100 printers to be printed. The 3D printed and processed models followed 3D Systems protocol and maintained consistent quality. The printed models were then measured manually using the standardized ABO gauge for alignment, marginal ridge discrepancy, buccolingual angulation, overjet, occlusal relationship, and root angulation.

![The American Board of Orthodontics](image)

**Figure 1.** ABO Measuring Gauge. A) ruler in 1mm increments and 0.50mm thick. B) Each step measures at 1mm each, especially used to measure buccolingual inclination of mandibular teeth. C) Each step measures 1mm, designed to measure marginal ridges, D) Each step is in 1mm increment, designed to measure buccolingual inclination of maxillary teeth. *The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs.*

The second set of data was collected by sending the same .STL files of the completed cases to MotionView Ortho Share 3D program will be segmented and uploaded onto the software to measure. The examiner placed appropriate landmarks to
calculate discrepancy in alignment, margin ridges, buccolingual inclination, overjet, occlusal relationship, and root angulation. Examiner purposely did not record occlusal contact, and interproximal contact was originally included in the ABO CRE standard as these parameters were not included in the program. The values for each parameter were recorded and given a score ranging from 0, 1, or 2 and summed up to create one objective grading score. The two subjects' scores were compared using paired t-tests to evaluate any variation in the scores between traditional manual measurements and digital measurements.

**ABO OBJECTIVE GRADING SYSTEM CRITERIA**

**Alignment:**

Alignment was indicative of the correction of all rotations and the ideal position of posterior and anterior teeth in reference to adjacent teeth. According to the ABO guidelines, the lingual-incisal edge of the maxillary anterior teeth and the labial-incisal surface of the mandibular anterior teeth served as the reference point for anterior alignment. In the premolar-canine region, alignment was determined by examining the contact points of each tooth. For posterior alignment, the mesiodistal central groove of the maxillary premolars and molar and the buccal cusp of the mandibular premolar and molar served as the reference point. According to official ABO guidelines:

"If all teeth are aligned, or within 0.50 mm of proper alignment, no points are scored. If the mesial or distal alignment at any of the contact points is 0.50 mm to 1 mm deviated from proper alignment, 1 point shall be scored for the tooth that is out of alignment. If adjacent teeth are out of alignment, then 1 point should be scored for each tooth. If the
discrepancy in the alignment of a tooth at the contact point is greater than 1 mm, then 2 points shall be scored for that tooth. No more than 2 points shall be scored for any tooth."

Figure 2. Manual alignment measurement using ABO gauge. Taken directly from The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs.

Figure 3. Digital alignment measurement using Ortho Share 360. Imagine taken from Electronic Cast Radiograph Evaluation (e-CRE).
Marginal Ridge

Proper marginal ridge alignment indicated proper vertical positioning of posterior teeth, excluding marginal ridges between first premolar and canine and distal marginal ridge of mandibular first premolar due to their highly variable form. According to the official ABO guidelines, the scoring was completed as follows:

"If adjacent marginal ridges deviate from 0.60 to 1 mm (Figure. 5a, 5b), then 1 point is scored for that interproximal contact. If the marginal ridge discrepancy is greater than 1 mm (Figure. 6a, 6b), then 2 points shall be scored for that interproximal contact. No more than 2 points will be scored for any contact point. The landmarks for scoring of marginal ridges are placed at the center of the mesial and distal marginal ridges of adjacent teeth and in alignment with the mesial and distal ends of the central grooves of adjacent teeth."

Figure 4. Manual marginal measurement using ABO gauge. Images taken directly from The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs
Buccolingual Inclination

Buccolingual inclination was an assessment of proper torque positions of posterior teeth, which directly influenced proper intercuspation of posterior occlusion.

The reference point of buccolingual inclination in the mandibular arch was a straight plane touching the buccal cusp of molars and premolars to the buccal cusp of the contralateral tooth. The lingual cusp of the mandibular molar or premolar should be within 1 mm of the surface of the straight reference plane. In the maxillary arch, the straight plane should be in contact with the lingual cusps of the maxillary premolar and molar and the lingual cusp of the contralateral tooth. The buccal cusp of the maxillary posterior teeth should be within 1 mm of the reference plane. The distal cusp of the second molars and mandibular first premolar will be omitted due to their high variability.

According to the official ABO guidelines, the scoring was completed as follows:

"If the mandibular lingual cusps or maxillary buccal cusps are more than 1 mm, but less than 2 mm from the straight edge surface, 1 point shall be scored for that tooth. If more
than 2mm of a discrepancy, then give 2 points but do not score more than 2 points for each tooth.

**Figure 6.** Manual buccolingual measurement using ABO gauge. Images taken directly from *The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs*

**Figure 7.** Digital buccolingual measurement using Ortho Share 360. Images taken from *Electronic Cast Radiograph Evaluation (e-CRE)*
Overjet

Overjet indicated correct transverse positioning of posterior teeth and correct anteroposterior position of anterior teeth of maxillary and mandibular arch relative to one another. According to official ABO guidelines, the scoring was as follows:

"If the proper overjet has been established, then the buccal cusps of the mandibular molars and premolars will contact in the center of the occlusal surfaces, buccolingually, of the maxillary premolars and molars. In the anterior region, the mandibular canines and incisors will contact the lingual surfaces of the maxillary canines and incisors. If this relationship exists, no points are scored. If the mandibular buccal cusps deviate 1 mm or less from the center of the opposing tooth, 1 point is scored for that tooth. If the position of the mandibular buccal cusps deviates more than 1 mm from the center of the opposing tooth, 2 points are scored for that tooth. No more than 2 points are scored for any tooth. In the anterior region, if the mandibular canines or incisors are not contacting lingual surfaces of the maxillary canines and incisors, and the distance is 1 mm or less, then 1 point is scored for each maxillary tooth. If the discrepancy is greater than 1 mm, then 2 points are scored for each maxillary tooth."
One of the primary goals of orthodontic treatment was to obtain a proper anteroposterior position of teeth, ideally in Cl I, and in exceptions, camouflaged full step Cl II occlusion.
or Cl III occlusion. Therefore, for consistency in the study, only cases with the goal of Cl I occlusion was used. According to the official ABO guidelines, the grading of occlusal relationships was as follows:

"The maxillary canine cusp tip should align with (or within 1 mm of) the embrasure or contact between the mandibular canine and adjacent premolar. The buccal cusps of the maxillary premolars should align with (or be within 1 mm of) the embrasures or contacts between the mandibular premolars and the first molar. The mesiobuccal cusps of the maxillary molars should align with (or be within 1 mm of) the buccal grooves of the mandibular molars. If the maxillary buccal cusps deviate between 1 and 2 mm from the positions, then 1 point shall be scored for that maxillary tooth. If the buccal cusps of the maxillary premolars or molars deviate by more than 2 mm from the ideal position, then 2 points shall be scored for each maxillary tooth that deviates. No more than 2 points shall be scored for each maxillary tooth."

Figure 10. Manual occlusal relationship measurement using ABO gauge. Images taken directly from *The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs*
Figure 11. Digital occlusal relationship measurement using Ortho Share 360. Images taken from Electronic Cast Radiograph Evaluation (e-CRE)

Root Angulation

An important treatment goal was for roots to be perpendicular to clinical crowns and parallel to adjacent neighbors. According to the ABO guidelines, the grading was as follows:

"If a root is angled to the mesial or distal (not parallel) and is close to, but not touching, the adjacent tooth root, then 1 point is scored for each discrepancy (anterior, premolar, and/or molar areas). If the root is angled to the mesial or distal and is contacting the adjacent tooth root, 2 points are scored for that tooth."

Figure 12. ABO guideline for assessing root angulation and approximation to adjacent teeth.
STATISTICAL ANALYSIS

All data collected from manual and digital measurements of 33 casts were recorded onto an Excel spreadsheet and then transferred to statistical software R-R version 4.1.2 to perform statistical tests. The total scores for all six parameters were compared and individually analyzed. The study aimed to answer whether there was a significant difference between digital versus manual measurements and if a difference did exist, which parameters contributed most to the variability. Three tests were run to obtain results for the sample size. A paired two-sample t-test was used to find the power T-test to find the appropriate sample size that will yield statistically significant results. A two-sample t-test was used to measure the differences between the digital and manual scores. And an individual pairwise test to analyze individual differences in score among the six parameters that were studied.

Several factors, including mean difference and standard deviations, influenced the sample size needed to achieve the desired power. For example, the result from the paired sample T-test suggested that when the mean difference is 0.1 and the standard deviation is 0.2, 33 individuals are required to have 80% power. As a result, 33 sample cases were used to ensure a study power of 80%, deemed statistically significant. Information was represented in Figure 13 below.
Figure 13. Statistical significance of data based on sample size with 80% power.
Only one examiner was calibrated to measure both manual and digital casts to ensure reliability and consistency. The models were graded three times two weeks apart with consistent reliability with a standard deviation of 0.58. The intraclass correlation coefficient (ICC) was excellent, with a score of 0.99 for the manual CRE measurement (p < 0.001) while 0.97 for the digital CRE measurement (p < 0.001). The 2 sample T-test showed that the total scores from digital measurements were statistically higher than manual measurements completed with the ABO gauge. The mean difference was 5.27 ± 1.39 higher in digital measurements. The sample results are listed in Table 1. Of the samples that were compared, the highest difference was a score of 14, and the lowest was a score of -2.

**Table 1.** ABO Objective Grading Scores between Manual and Digital Measurements from ABO gauge versus Digital CRE Ortho Share 360

<table>
<thead>
<tr>
<th>Case Number</th>
<th>A</th>
<th>A(d)</th>
<th>M</th>
<th>M(d)</th>
<th>BJI</th>
<th>BJI(d)</th>
<th>D</th>
<th>D(d)</th>
<th>OR</th>
<th>OR(d)</th>
<th>RA</th>
<th>RA(d)</th>
<th>Total</th>
<th>Total(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>Case 2</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>22</td>
<td>31</td>
</tr>
<tr>
<td>Case 3</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Case 4</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>Case 5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>32</td>
<td>38</td>
</tr>
<tr>
<td>Case 6</td>
<td>5</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>Case 7</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>Case 8</td>
<td>9</td>
<td>15</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>Case 9</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Case 10</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>Case 11</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Case 12</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Case 13</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Case 14</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Case 15</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Case 16</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Case 17</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>Case 18</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>Case 19</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>Case 20</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>14</td>
<td>22</td>
</tr>
<tr>
<td>Case 21</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Case 22</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>Case 23</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>Case 24</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Case 25</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>Case 26</td>
<td>7</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Case 27</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>Case 28</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>Case 29</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>Case 30</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Case 31</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Case 32</td>
<td>7</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td>Case 33</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>23</td>
<td>24</td>
</tr>
</tbody>
</table>
Among the six parameters, the most significant difference was found in alignment, which showed a mean difference of 1.79, the marginal ridge with a mean difference of 1.18, and buccolingual inclination with a mean difference of 1.15. Occlusal relationship, overjet, and root angulation showed a higher correlation with a mean difference below 1 (Table 2). Of all parameters, root angulation showed the highest correlation with a mean difference of -0.03. This exception is due to a repeated grading method following the same visual parameters on the same panoramic radiograph. While there were wide ranges and noticeable standard deviations in each parameter, the P-values were statistically significant and there was a 95% confidence interval in the data.

**Table 2.** Descriptive means and standard deviation for differences in ABO CRE parameters between manual and digital measurements

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean Difference</th>
<th>Standard deviation</th>
<th>Range</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignment</td>
<td>1.79</td>
<td>2.25</td>
<td>-3 to 8</td>
<td>0.0000687</td>
</tr>
<tr>
<td>Marginal Ridge</td>
<td>1.18</td>
<td>1.74</td>
<td>-2 to 6</td>
<td>0.000462</td>
</tr>
<tr>
<td>Buccolinguinal inclination</td>
<td>1.15</td>
<td>2.41</td>
<td>-2 to 11</td>
<td>0.00991</td>
</tr>
<tr>
<td>Overjet</td>
<td>0.364</td>
<td>1.38</td>
<td>-2 to 4</td>
<td>0.142</td>
</tr>
<tr>
<td>Occlusal Relationship</td>
<td>0.818</td>
<td>1.72</td>
<td>-3 to 5</td>
<td>0.0102</td>
</tr>
<tr>
<td>Root Angulation</td>
<td>-0.03</td>
<td>0.174</td>
<td>-1 to 0</td>
<td>0.325</td>
</tr>
<tr>
<td>Total</td>
<td>5.27</td>
<td>1.39</td>
<td>-2 to 14</td>
<td>&gt;0.00001</td>
</tr>
</tbody>
</table>
**Figure 14.** Scatter plot of digital and manual CRE total scores, showing a strong correlation between the values. Pearson correlation coefficient was 0.828 (p < 0.001).

*Table 2* included the individual pairwise T-test results with p-values that distinguished the main drivers for the score difference. The scores with a p-value of less than 0.05 were considered statistically significant, including alignment, marginal ridge, buccolingual inclination, and occlusal relationship. The low p-values provided sufficient evidence to confirm that the following parameters were responsible for the score differences. In contrast, overjet and root angulation was not statistically significant and did not significantly influence the score discrepancies. *Figure 14* summarized the total scores and demonstrated the positive correlation between digital and manual values, with digital values consistently higher. Pearson correlation coefficient was 0.828 (p < 0.001).
CHAPTER 5
DISCUSSION

The total ABO Objective Grading System scores for traditional measurement using the ABO gauge and digital casts had a good correlation but with clinically significant differences. The digital scores had a mean of 5.27 points, a standard deviation of 1.39 greater than manual measurements, and a range from -2 to 10 points. The last study to evaluate digital ABO CRE grading was done by Hilebrand et al. in 2008, rejecting the replacement of digital measurements on OrthoCAD for manual values. The value difference found by Hilebrand et al. was a digital score of 9 points with a standard deviation of 5.54 greater than manual measurements. Fourteen years later, our technology has undergone significant advancements but has proven not perfect. The new Ortho Share 360 program had a smaller discrepancy compared to OrthoCAD. However, this time with a lower value difference of 5. The current ABO guidelines for passing scores are at 27 or less for Cast radiograph evaluation. Therefore, any difference that inflates the values can directly impact whether a score is considered successful or unsuccessful.

Of the six parameters followed, alignment, marginal ridge, and buccolingual inclination were the drivers of the difference in scores. Alignment was the major determinant as the score had a mean difference of 1.79, which is 0.6 greater in mean difference than the remaining two parameters. On the other hand, the difference between the marginal ridge and buccolingual discrepancy was insignificant at a difference of 0.03. Although the difference of 1.78 does not appear significant, the average difference in total score was 5.27, and alignment was responsible for 34% of the difference.
Many factors contribute to the variability of the three significant parameters causing the most discrepancy in scores. Of the three most important parameters that influenced the score differences, both alignment and marginal ridge had an index of discrepancy that allows a value up to 0.5mm before the discrepancy gets a point for deduction. With the aid of the ABO gauge, the human eye could detect a difference between 0.0 to 0.5 mm, 0.5 mm to 1 mm, and 1 mm to 2 mm differences. However, any points in between are much more difficult to detect accurately. The digital software makes subtle differences between 0.5 mm and 0.6 mm more easily seen. They will result in a whole additional point that would otherwise not have been detected using the gauge. Additionally, while buccolingual inclination had an index allowance of 1mm, the buccal segment of upper second molars has the most deductions, as suggested by Lieber et al. 2003, resulting in more chances for deductions and variability.

The overjet and occlusal relationship had minimal impact on the difference in scores between digital and manual measurements, suggesting a high correlation between the two measuring methods. Ortho Share 360 omits the typical occlusal contact parameter and combines posterior occlusal contact with overjet. Both methods are very consistent for posterior occlusal contact as the index is based on whether contact is present, which is evident to the human eye and marked by the digital software. The main site of the deduction for overjet is in the posterior second molars. The occlusal relationship also had a high correlation as the index allowed for up to 1mm of a discrepancy, and any points higher are deducted one point up to 2 mm.

It is apparent that the correlation of scores between methods is impacted by the threshold that is allowed for that discrepancy. Parameters with a lower range that begins
at 0.5mm have a lower correlation because the human eye has difficulty distinguishing differences of 0.1mm or lower. In this discussion, root angulation had a negative value in mean difference because the same visual method was used to grade the same panoramic radiograph using the same parameter guidelines.

When comparing the results of Ortho Share 360 with previous data from OrthoCAD (Hilebrand et al., 2003), Ortho Share 360 proved to be a superior program with a lower discrepancy between digital versus manual measurements. However, any discrepancy that resulted in inflated scores remains clinically significant. Ortho Share 360 proved to consistently yield higher scores compared to manual measurements, which directly impacted the clinician's final scores. If Ortho Share 360 was to be used for Cast-radiographic evaluation, there should be an increase in the passing score allowance to account for the inflated digital measurements. By increasing the score to accommodate the digital discrepancy, the ABO boards will continue to provide candidates with a fair and convenient means to evaluate their work, and maintain the integrity of the grading system.

Similar to what Hilebrand et al. discovered in the OrthoCAD study, certain software program deficiencies contributed to the score discrepancy. In the standard ABO CRE criteria, alignment deduction was per tooth that was rotated. However, in Ortho Share 360, alignment landmarks are measured based on mesial and distal contact points, resulting in a double scoring for each rotated tooth. In addition, in contrast to the study, Ortho Share 360 had a high correlation in interarch parameters, which was overjet and occlusal relationship. This was due to an improvement in the software's ability to segment the arch and properly articulate the digital casts. In addition, STL files submitted
to Ortho Share are formatted with arches already set in occlusion to allow for more
accurate data transfer.

Limitations in the study

It was important to note the certain weakness of the study, which included the
quality of the 3D printed models fabricated by the 3D Systems NextDent5100 printer.
While all guidelines were followed, there may be variability in the quality of the models
that were printed that could have impacted the model's final measurements. The same
student calibrated and printed the models to maintain maximum consistency, but quality
cannot always be guaranteed. For future research, I would further explore the consistency
and accuracy of the printer and the 3D models.

A second potential weakness of the study was that Ortho Share 360 omitted two
parameters, interproximal contacts and occlusal contacts, from their software. These
parameters were also then omitted in the study to maintain consistency. However, this
limits the study's ability to completely address the accuracy of the software program in
terms of the standard ABO CRE parameters. The maximum deduction possible on
MotionView3D's Ortho Share 360 was 196, less than the current 248 points that could be
deducted using the standard ABO guidelines. Therefore, while the program was proven
superior to OrthoCAD when assessing alignment, marginal ridges, buccolingual
inclination, overjet, occlusal relationship, and root angulation, the program has yet to
prove its reliability concerning interproximal contact and occlusal contacts. Finally, this
study has no conflicts of interest, and the author has no association or financial interest
with MotionView 3D or Ortho Share 360.
CHAPTER 6
CONCLUSION

The American Board of Orthodontics Objective Grading System has been proven to be an accurate and reliable tool in assessing the success of finished orthodontic cases. To make board certification more accessible and convenient for candidates, the ABO has adapted a new software called Ortho Share 360 to grade casts of completed cases digitally. Based on the results, the hypothesis was rejected, and digital and manual measurements are, in fact, significantly different (p < 0.001). The software improved from previous programs but still shows a statistically significant difference in scores compared to manual measurement. MotionView 3D Ortho Share 360 program offers a reproducible measurement that was, on average, 5.27 ± 1.39 points greater than manual measurements. This difference was due to inflated scores in alignment followed by marginal ridge and buccolingual discrepancies. The software was particularly accurate in articulating the maxillary and mandibular arch and therefore had highly correlated scores in the overjet and occlusal relationship, a feature that Ortho Share lacked. (Hilebrand et al., 2003). The final score was then different from manual measurement and cannot be used as a replacement for traditional ABO CRE using the standardized gauge. If the ABO were to implement digital measurements for future ABO certification, considerations to increase passing scores from the current limit of 27 should be considered to account for the inflated scores of the software.
REFERENCES


