Accuracy, Reproducibility, and Time-Efficiency of Dental Measurements Using Different Technologies

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Nishant Patel

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Advisor: Dr. Thorsten Gruenheid
Co-Advisor: Dr. Brent E. Larson

June 2012
Acknowledgments

I would like to thank my advisor Dr. Thorsten Gruenheid for his constant involvement, guidance, and support throughout every stage of this project. Also, thank you to Dr. Brent E. Larson for offering your guidance and suggestions throughout the project.
Dedication

To my parents and sister for your continual love and support.
Abstract

**Background:** Historically, orthodontists have taken dental measurements on plaster casts. Recently, technological advances have given orthodontists the ability to take dental measurements on digital models. It was the goal of this study to assess the accuracy, reproducibility, and time-efficiency of measurements taken on three digital models and compare them to the gold standard plaster casts.

**Methods:** Three operators measured tooth widths on 30 sets of dental models. Tooth-width measurements were recorded and timed for plaster casts, SureSmile diagnostic casts (OraMetrix, Richardson, Texas), emodels (Geodigm, Minneapolis, Minnesota), and AnatoModels (Anatomage, San Jose, California). To assess accuracy and reproducibility, the Altman-Bland agreement and mean squared error were calculated for individual teeth. Differences in time-efficiency among the types of dental casts were tested for statistical significance using one-way analysis of variance with Tukey’s method as a post hoc pairwise comparison procedure.

**Results:** SureSmile diagnostic casts were most accurate when compared to plaster casts followed by emodels and AnatoModels. SureSmile diagnostic casts were also the most reproducible dental casts. emodels were the quickest to complete tooth-width measurements, followed by SureSmile diagnostic casts, plaster casts, and AnatoModels.

**Conclusions:** Tooth-width measurements were most accurate and reproducible on SureSmile diagnostic casts. Measurements taken on emodels and SureSmile diagnostic casts were clinically significantly faster when compared to plaster casts and AnatoModels.
Table of Contents

LIST OF TABLES...............................................................................................................v
LIST OF FIGURES...........................................................................................................vi
INTRODUCTION...............................................................................................................1
REVIEW OF THE CURRENT LITERATURE..................................................................5
THE PRESENT WORK.....................................................................................................9
MATERIAL AND METHODS..........................................................................................10
RESULTS....................................................................................................................21
DISCUSSION................................................................................................................26
CONCLUSIONS.............................................................................................................38
REFERENCES..............................................................................................................39
List of Tables

Table 1: Comparison of methods.................................................................17
Table 2: Repeatability comparison...............................................................21
Table 3: Accuracy of single tooth-width comparison.....................................21
Table 4: Reproducibility comparisons between raters....................................23
Table 5: Time comparison among methods................................................24
Table 6: Dental cast pricing..........................................................................36
List of Figures

Figure 1: CBCT scan for SureSmile digital model fabrication.................................11
Figure 2: SureSmile tooth-width measurements.....................................................13
Figure 3: emodels tooth-width measurements......................................................14
Figure 4: InvivoDental AnatoModel tooth-width measurements..........................15
Figure 5: Plaster cast tooth-width measurements................................................16
Figure 6: Distribution curves showing the effects of bias and agreement limits........19
Figure 7: Altman-Bland difference plots for accuracy.........................................22
Figure 8: Beam-hardening effects on AnatoModels............................................33
Figure 9: Root canal visualization on AnatoModels...........................................34
Figure 10: Measurement difficulties on emodels................................................35
Introduction

In orthodontics, dental casts are used to visualize the morphology and positions of teeth, the degree to which the teeth are malpositioned, and their occlusal relationships. The information obtained from dental casts can help orthodontists identify aberrations and classify malocclusions, making dental casts an integral part in the orthodontist’s armamentarium used for diagnosis and treatment planning. Furthermore, orthodontists use dental casts for patient education, documentation of treatment outcomes, and research.

Dental casts are a dimensionally accurate representation of the dentition. Therefore, a variety of measurements and analyses can be obtained from dental casts to aid the orthodontist in diagnosis and treatment planning. In fact, dental casts appear to be the major record used for orthodontic treatment planning (Han et al., 1991). The measurements performed on dental casts typically include tooth-width measurements. These measurements are often used to analyze the space for prosthetic tooth replacement or to assess the space required to align a crowded dentition. Tooth-width measurements also allow the diagnosis of tooth-size discrepancies, which may be a major hurdle in achieving ideal dental interdigitation during the finishing stages of orthodontic treatment (Othman and Harradine, 2006).

Historically, dental casts made from plaster have been used to analyze and measure dental relationships. The evolutions of dental impression materials and dental stones enabled plaster casts to show greater detail of tooth morphology and surrounding
structures. Their analysis has been shown to be accurate and reproducible; therefore, plaster casts are currently considered to be the gold standard for dental cast analysis (Peluso et al., 2004). However, plaster casts have limitations. For instance, plaster casts are at risk for breakage, chipping, or abrasion if they are handled improperly. Continued use and display can wear away plaster, decreasing the accuracy of measurements (Peluso et al., 2004). Plaster casts also need to be stored in the office or at a storage facility, which can make retrieval difficult.

Recently, technological advances have given orthodontists the ability to analyze dental relationships on digital models. Digital models, which first became available to orthodontists in 1999 through OrthoCAD (Cadent, Carlstadt, New Jersey), alleviate many of the obstacles encountered with using plaster models (Peluso et al., 2004). For instance, digital models are not subject to physical damage or degradation and they require negligible storage space as the digital information can be stored on a computer hard drive, portable storage devices, or on a central server. Furthermore, digital models enable easy communication among clinicians as the information can be shared via the internet. As a result of these advantages and the increasing affordability of digital records, an increasing number of orthodontic offices are implementing digital casts. A recent survey found that in 2008 18% of the orthodontists in the United States used digital casts for diagnosis, treatment planning, and patient education (Keim et al., 2008).

Most digital models are currently made from dental impressions, which are either scanned directly or poured in plaster and then scanned. Thus, the typical process of
digital model fabrication still requires traditional alginate impressions to be taken in the orthodontic office. These impressions are shipped to one of the companies offering digital models where they are processed and scanned to produce digital models. Within a few days, an electronic file is available to be downloaded from the internet to a desired computer. Once downloaded, the digital casts can be viewed and analyzed, just like traditional plaster casts.

The development of chairside scanning devices now allows for creating digital casts without taking impressions. For instance, SureSmile (OraMetrix, Richardson, Texas), an orthodontic CAD/CAM system, allows clinicians to manipulate three-dimensional (3-D) models of patients’ teeth to simulate orthodontic treatment. SureSmile technology uses a handheld device that allows direct intra-oral scanning of the patient’s dental arches. After the dental arches are scanned, the information is sent over the internet to the company for digital cast fabrication. Once the orthodontist receives the digital casts, the SureSmile software allows dental measurements used for the diagnosis of space requirements or analysis of tooth-size discrepancies.

With the advent and increased use of cone-beam computed tomography (CBCT) in orthodontics, companies have introduced another method of digital model fabrication. Sophisticated software algorithms now allow digital model fabrication from a patient’s CBCT scan. Therefore, this technology provides another possibility of eliminating the need to take traditional impressions for dental cast fabrication. However, this method is not without shortcomings. For instance, restorative fillings can create scatter in the
tomogram, which can alter tooth morphology on the digital cast. Depending on the voxel size of the scan, the detail of tooth morphology can be compromised in digital casts fabricated from a CBCT scan. Unlike dental models fabricated using other techniques, models created from a CBCT scan require a patient to be exposed to ionizing radiation.

There are many types of digital dental casts available to the practicing orthodontist today; however, questions remain regarding the accuracy and reliability of these digital casts when used for measurements and analyses. To allow proper diagnosis and successful treatment planning, it is necessary that digital casts accurately represent the patient’s intraoral situation and it is important that the clinician selects a method that is accurate, reproducible, and efficient.
Review of the Current Literature

As discussed above, a wide variety of dental models including plaster casts, digital models created from plaster casts, and digital models created from CBCT volumes are currently available to orthodontists. The following reviews the current literature available on these dental models.

Plaster Casts

Traditionally, orthodontists have used plaster casts. With proper impression and pour-up techniques, these casts have been shown to provide an accurate representation of a patient’s dentition and surrounding structures (Peluso et al., 2004). Although plaster casts accurately represent a patient’s intraoral situation, it has been shown that the amount of dental crowding and the instrument used to take measurements can affect the accuracy and reliability of tooth-width measurements (Shellhart et al., 1995). Tooth-width measurements on plaster casts have been found to be more error prone when there were three millimeters or more of dental crowding. Moreover, tooth-width measurements taken with needlepoint dividers have been found to be less reliable than those taken with a Boley gauge (Shellhart et al., 1995). Despite these inherent errors, plaster casts are routinely used in orthodontics to perform tooth-width measurements.

Digital Casts

A number of manufacturers offer digitization of plaster casts. In the last decade, many investigators have evaluated the accuracy and reliability of these digital models created from plaster casts (Santoro et al., 2003; Zilberman, 2003; Stevens et al., 2006; Mullen et
al., 2007; Naidu et al., 2009). Although several studies have shown that digital models either tend to underestimate or overestimate tooth sizes when compared to plaster casts (Stevens et al., 2006; Mullen et al., 2007; Naidu et al., 2009), the differences have been considered clinically insignificant and most authors agree that tooth-width measurements taken on digital casts are accurate and reliable when compared to plaster casts (Santoro et al., 2003; Zilberman, 2003; Stevens et al., 2006; Naidu et al., 2009). A recent systematic review of the literature, which compared digital models to plaster casts, provided further evidence that digital models offer a high degree of validity when compared to plaster casts (Fleming et al., 2011). As more manufacturers of digital models enter the market, future studies need to focus on newer digital models and determine if they also can be a valid replacement for plaster casts.

Recently, digital models fabricated from CBCT volumes have gained attention in the orthodontic community because, in certain situations, these models can eliminate the need of impressions altogether. El Zanaty et al. (2010) found arch width, length, and tooth-size measurements taken on digital models fabricated from CBCT volumes using custom-made software to be comparable to measurements performed on plaster casts. However, the authors used a prefabricated splint to keep the maxillary and mandibular teeth separated in order to facilitate segmentation of inter- and intra-arch tooth structures—a procedure that could affect orthodontic diagnosis and treatment planning when evaluating pre-treatment radiographs.
Software advances with InvivoDental’s AnatoModels (Anatomage, San Jose, California) have helped resolve difficulties in segmenting the maxillary and mandibular arches when patients are scanned with their teeth in occlusion. Up to now, only a few investigators have evaluated AnatoModels. A study that investigated linear measurement differences between anatomical points, rather than tooth dimensions, found that AnatoModels had a similar level of accuracy and were adequate for orthodontic diagnosis and treatment planning when compared to OrthoCad models (Creed et al., 2011). Measurements to evaluate Little’s Irregularity Index on AnatoModels have also shown to be accurate with no statistically or clinically significant differences when compared to OrthoCad models (Kau et al., 2010). To date, only one study has reported the accuracy of linear tooth-width measurements using AnatoModels (Tarazona et al., 2011). The study compared the accuracy, reliability, and reproducibility of tooth-width, intercanine distance, intermolar distance, and arch length measurements between AnatoModels to what the authors call the two-dimensional (2-D) Digital Method. The 2-D Digital Method of cast analysis requires plaster casts to be scanned into a computer system and measurements to be performed on the 2-D image of the scanned plaster cast. Using this approach, the authors found clinically insignificant differences between measurements taken on AnatoModels and plaster casts using the 2-D Digital Method (Tarazona et al., 2011).

Currently, only two studies compared tooth-width measurements on CBCT models to those taken on another type of dental model (El Zanaty et al., 2010, Tarazona et al., 2011). Of these two studies, only one used a CBCT model that is currently available to the practicing orthodontist (Tarazona et al., 2011). Therefore, it would be valuable to
further evaluate the accuracy and reproducibility of CBCT dental models and compare them to plaster casts as well as to other digital models that are currently available to the practicing orthodontist.

*Time-Efficiency of Measurement*

The time it takes to complete measurements on dental casts is important for efficient orthodontic care. Only a few studies have measured and compared the time it takes to complete tooth-size measurements on different types of dental models (Mullen *et al.*, 2007; Horton *et al.*, 2010; Tarazona *et al.*, 2011). Mullen *et al.* (2007) found measurements taken on *emodels* (Geodigm, Minneapolis, Minnesota) to be significantly faster when compared to plaster casts. Tarazona *et al.* (2011) found tooth-width measurements on CBCT models to be as quick as those on plaster casts using their 2-D Digital Method. A technique comparison for measuring tooth widths on digital models has shown significant differences in terms of speed. Depending on the technique, the average time of measuring tooth widths on *emodels* was as fast or slower than measurements taken on plaster casts (Horton *et al.*, 2010).

Currently, no studies have been published that compare the time-efficiency of taking tooth-width measurements on plaster casts, *emodels*, AnatoModels, and SureSmile diagnostic models. A comprehensive study, which compares the time-efficiency of taking tooth-width measurements on different types of dental models, while evaluating accuracy and reproducibility, would be beneficial in identifying if one type of dental model is superior to others.
**The Present Work**

The aims of the present work were to compare the accuracy, reproducibility, and time-efficiency of dental cast measurements using three different digital technologies with those of measurements taken on traditional plaster casts. Specifically, the digital technologies investigated in this study were emodels, AnatoModels, and SureSmile diagnostic models.

The present study tested the following null hypotheses:

1. There is no difference in the accuracy of tooth-width measurements taken on plaster casts and those taken on emodels, AnatoModels, and SureSmile diagnostic models.

2. There is no difference in the reproducibility of tooth-width measurements taken on plaster casts, emodels, AnatoModels, and SureSmile diagnostic models.

3. There is no difference in the time required to take tooth-width measurements on plaster casts, emodels, AnatoModels, and SureSmile diagnostic models.
Material and Methods

Dental Casts

The use of anonymized dental casts had been approved by the University of Minnesota’s Institutional Review Board. The pre-treatment casts of 30 patients consecutively treated using SureSmile at the University of Minnesota Division of Orthodontics were used for this study. The patients had all permanent teeth mesial to the second molars fully erupted and represented a variety of typical malocclusions. None of the patients had previous orthodontic treatment. At the time of the initial records, alginate impressions, a wax bite registration, and a CBCT scan were taken of each patient. The CBCT scans were full field of view scans using a Next Generation i-CAT (Imaging Sciences International, Hatfield, Pennsylvania) at a voxel size of 0.3 mm$^3$ and an exposure time of 8.9 s.

Plaster casts were fabricated by pouring the alginate impressions in dental stone (Modern Materials Orthodontic Plaster, Heraeus Kulzer, South Bend, Indiana). emodels were fabricated from these plaster casts by GeoDigm. SureSmile diagnostic models were fabricated by OraMetrix using CBCT scans of the patients’ initial plaster casts taken with the Next Generation i-CAT at a voxel size of 0.2 mm$^3$ and an exposure time of 26.9 s. For these scans, the plaster casts were set in the i-CAT with a wax bite separating the maxillary and mandibular teeth to allow for segmentation of the teeth by SureSmile digital laboratory technicians (Figure 1). Segmented 3-D AnatoModels were produced by Anatomage using the patients’ pre-treatment CBCT scans. All casts were de-identified before the beginning of the study.
Figure 1: CBCT scan for SureSmile digital model fabrication. (A) Plaster casts being scanned with a Next Generation i-CAT. (B) A wax bite separating maxillary and mandibular teeth is needed to allow for segmentation of the teeth.

Measurements

Three operators with varying amounts of orthodontic experience measured mesial-distal tooth widths of the teeth mesial to the second molars. The operators included a second year orthodontic resident, an orthodontist with 13 years of experience, and an orthodontist with 10 years of experience. Prior to data collection, the operators were calibrated, instructed to take the widest mesial-distal measurement of each tooth, and each completed measurements on five practice cases to familiarize themselves with the software and measurement techniques.

Measurements were taken to the closest hundredth millimeters for plaster casts, e-models, and AnatoModels. Due to software limitations, SureSmile digital models were measured to the closest tenths of a millimeter.
Digital casts were displayed full-screen with a screen resolution of 1680x1080 pixels on a 22-inch XPS computer monitor (Dell, Round Rock, Texas) with landscape orientation. The three operators completed 30 cases on one type of cast before proceeding to the next one. Within each type of cast, the sequence of casts to be measured was randomized for each operator. A digital stopwatch was used to measure the time it took to complete the measurements. For all models, time was started when the casts were in front of the operator, ready for analysis, and stopped when all the measurements were completed (Table 1).

SureSmile diagnostic casts were evaluated with SureSmile software version 5.9. The background color was set to black for optimum contrast between the blue points to be manipulated and the background. The points mesial and distal of each tooth during diagnostic cast fabrication were moved if the operator believed they did not accurately represent a tooth’s greatest mesial-distal width (Figure 2).
Figure 2: SureSmile tooth-width measurements. (A) Maxillary arch view with tooth-width measurements displayed. (B) Magnified view of maxillary anterior teeth with modifiable contact points set by OraMetrix.
emodels were analyzed using emodel software version 9.0. Background color and text color were set to black and red, respectively, for optimum contrast. Each operator measured the tooth widths by manually setting mesial and distal contact points reflecting the greatest mesial-distal distance for each tooth (Figure 3).

Figure 3: emodel tooth-width measurements. (A) Maxillary arch view with tooth-width measurements completed. (B) Magnified view of maxillary anterior teeth.
AnatoModels were analyzed using AnatoModel software version 5.0 (Figure 4). All measurements were taken on a split screen, which showed the maxillary and mandibular arches simultaneously.

**Figure 4:** InvivoDental AnatoModel tooth-width measurements. **(A)** InvivoDental AnatoModel maxillary arch view with measurements completed. **(B)** Magnified view of maxillary anterior teeth with modifiable contact points.
Plaster casts were measured using digital Vernier calipers (Figure 5). Between each set of plaster casts, the calipers were calibrated to ensure accurate measurements. The tooth-width measurements were recorded on a sheet of paper and the time was stopped once all teeth were measured and the measurements recorded.

Figure 5: Plaster cast tooth-width measurements (A) Tooth-width measurement with digital Vernier calipers. (B) Magnified view of maxillary anterior teeth.
<table>
<thead>
<tr>
<th></th>
<th>Plaster Casts</th>
<th>emodels</th>
<th>SureSmile</th>
<th>AnatoModels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cast Retrieval</strong></td>
<td>Storage</td>
<td>Server</td>
<td>Server</td>
<td>Server</td>
</tr>
<tr>
<td><strong>Time started for tooth-width measurements</strong></td>
<td>Casts in front of operator in occlusion</td>
<td>Casts on computer screen in occlusion</td>
<td>Casts on computer screen in occlusion</td>
<td>Casts on computer screen in occlusion</td>
</tr>
<tr>
<td><strong>Software manipulation used to measure tooth-widths</strong></td>
<td>Not applicable</td>
<td>Analysis feature</td>
<td>Diagnostic models and new treatment simulation features</td>
<td>Occlusal layout feature with both arches visible</td>
</tr>
<tr>
<td><strong>Manipulation of casts</strong></td>
<td>By hand</td>
<td>Rotating and magnifying virtually</td>
<td>Rotating and magnifying virtually</td>
<td>Rotating and magnifying virtually</td>
</tr>
<tr>
<td><strong>Selection of contact points</strong></td>
<td>Selected by operator</td>
<td>Selected by operator</td>
<td>Selected by technician; modified by operator</td>
<td>Selected by operator</td>
</tr>
<tr>
<td><strong>Modification of contact points</strong></td>
<td>Re-measure</td>
<td>Delete and re-measure</td>
<td>Move contact point</td>
<td>Move contact point</td>
</tr>
<tr>
<td><strong>Time stopped</strong></td>
<td>After measurements were recorded on paper</td>
<td>After measurements were completed on computer</td>
<td>After measurements were completed on computer</td>
<td>After measurements were completed on computer</td>
</tr>
<tr>
<td><strong>Saving</strong></td>
<td>On a sheet of paper</td>
<td>On the computer hard drive as an EMZ file</td>
<td>On the SureSmile server</td>
<td>On the computer as a JPEG image file</td>
</tr>
</tbody>
</table>
After a washout period of three weeks, six cases were randomly selected from the original 30 cases and were re-measured by all operators to assess intra-operator reliability.

Data Analysis

Accuracy was defined as the agreement between a digital cast and the plaster cast. A version of the Altman-Bland method (Bland and Altman, 1986; 1999) was used to assess the degree to which the measurements on the various digital casts agreed with the digital caliper measurements performed on the corresponding plaster cast. In order to account for the correlation between teeth within a subject – and to avoid underestimation of the different variations – a mixed model with a random intercept term (for person) was used to obtain variance estimates used in the Altman-Bland analysis (Bland and Altman, 2007). The bias was computed, separately for each digital method, as the average of the differences between the digital measurements and the plaster cast measurements for each individual tooth. Calculation of the Altman-Bland agreement range provided 95% agreement limits for each digital model method. When ranking methods, both the bias and the variance were taken into account by calculating the mean squared error (MSE) as follows: \[ \text{MSE} = (\text{bias})^2 + \text{variance}. \]

Figure 6 shows hypothetical data to exemplify the influence of the bias and the variance. In this example, method 1 has a smaller bias than method 2 (0.1 units) and the variance of method 1 is slightly higher than that of method 2 (0.0225 units). Since the bias of method
1 is closer to the zero mean difference than that of method 2 and the variance is only slightly larger than method 2, measurements taken with method 1 would be determined to be more accurate. Consequently, the MSE shows a lower error for method 1 than that for method 2.

![Distribution Curves of 2 Hypothetical Methods](image)

Method 1: Bias=0.15; Variance=0.0625; MSE= 0.085  
Method 2: Bias=0.25; Variance=0.0400; MSE= 0.1025

**Figure 6**: Distribution curves showing the effects of bias and agreement limits. The solid line has larger agreement limits with a smaller bias while the dashed line has smaller agreement limits with a higher bias. More points will be closer to the zero difference line with the smaller bias and slightly larger agreement limits.

Reproducibility was defined as the degree of agreement between the measurements performed on replicate specimens by the three operators (ASTM, 2002). Reproducibility was tested for tooth-for-tooth reproducibility using the Altman-Bland method and the MSE was calculated to rank the methods on reproducibility as detailed above.
Intra-operator repeatability was defined as the degree of agreement of measurements on replicate specimens within each operator (ASTM, 2002). Repeatability was tested for tooth-for-tooth repeatability using the Altman-Bland method and the MSE was calculated to rank the methods on repeatability.

Time-efficiency was defined as the time requirement for measuring tooth-widths on the various types of dental casts. For each type of dental casts, mean values and standard deviations of the time requirement were calculated. Differences among the types of dental casts were tested for statistical significance using one-way analysis of variance (ANOVA) with Tukey’s method as a post hoc pairwise comparison procedure after the data had been tested for normality (Kolmogorov-Smirnov test).

Time-efficiency was also compared between the practice cases and the thirty experimental cases as well as between the first five experimental cases and the last five experimental cases for each type of dental cast. Differences were tested for statistical significance using a Mann-Whitney U test. A non-parametric test was chosen because of its robustness in a small sample size.

Statistical analyses were performed using SigmaStat 3.5 (Systat Software Inc., Point Richmond, California), Microsoft Excel 5.0 (Microsoft Corp., Redmond, Washington), and R Statistical Software 2.9.2 (R Foundation for Statistical Computing, Vienna, Austria). P-values of less than 0.05 were considered statistically significant.
Results

Repeatability

Repeatability comparisons of individual tooth-width measurements are reported in Table 1. On average, the measurements were repeatable for each type of dental cast to 0.05 mm.

<table>
<thead>
<tr>
<th>Model</th>
<th>Bias</th>
<th>Lower 95% Agreement Bound</th>
<th>Upper 95% Agreement Bound</th>
<th>Agreement Interval Width</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaster Casts</td>
<td>0.048</td>
<td>-0.304</td>
<td>0.400</td>
<td>0.704</td>
<td>0.035</td>
</tr>
<tr>
<td>eModels</td>
<td>-0.028</td>
<td>-0.410</td>
<td>0.355</td>
<td>0.765</td>
<td>0.039</td>
</tr>
<tr>
<td>SureSmile</td>
<td>-0.011</td>
<td>-0.231</td>
<td>0.209</td>
<td>0.440</td>
<td>0.013</td>
</tr>
<tr>
<td>AnatoModels</td>
<td>0.006</td>
<td>-0.487</td>
<td>0.500</td>
<td>0.987</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Accuracy

Accuracy comparisons of individual teeth are reported in Table 2. The bias was largest for eModels followed by AnatoModels and SureSmile diagnostic casts. Agreement limits widths were largest for AnatoModels followed by SureSmile and eModels. SureSmile had the smallest MSE followed by eModels and AnatoModels. When comparing the accuracy of individual tooth-width measurements to plaster casts, tooth-width measurements taken on SureSmile diagnostic casts were closest to plaster cast measurements followed by the measurements taken on eModels and AnatoModels. The Altman-Bland plots of individual tooth-widths are shown in Figure 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Bias</th>
<th>Lower 95% Agreement Bound</th>
<th>Upper 95% Agreement Bound</th>
<th>Agreement Interval Width</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>eModels</td>
<td>0.285</td>
<td>-0.176</td>
<td>0.746</td>
<td>0.922</td>
<td>0.134</td>
</tr>
<tr>
<td>SureSmile</td>
<td>-0.130</td>
<td>-0.604</td>
<td>0.344</td>
<td>0.947</td>
<td>0.073</td>
</tr>
<tr>
<td>AnatoModels</td>
<td>0.191</td>
<td>-0.57</td>
<td>0.952</td>
<td>1.521</td>
<td>0.181</td>
</tr>
</tbody>
</table>
Figure 7: Altman-Bland difference plots for accuracy. (A) Difference plots of e-models (B) SureSmile (C) AnatoModels when compared to plaster casts. The close proximity of the data points to the identity line in the SureSmile plot illustrates the method’s high relative accuracy while the large spread of the data points in the AnatoModels plot illustrates lower relative accuracy of this method.
Reproducibility

Reproducibility comparisons for all operator combinations are reported in Table 3. These reproducibility comparisons suggest that measurements on SureSmile models were the most reproducible with the smallest agreement interval width and the smallest MSE between all three operator combinations. Furthermore, measurements on AnatoModels were consistently third most reproducible, while those on plaster casts and emodels interchanged from second to fourth most reproducible depending on which operator combinations were compared.

Table 4: Reproducibility comparisons between raters

<table>
<thead>
<tr>
<th>Operators</th>
<th>Model</th>
<th>Bias</th>
<th>Lower 95% Agreement Bound</th>
<th>Upper 95% Agreement Bound</th>
<th>Agreement Interval Width</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>Plaster Casts</td>
<td>0.413</td>
<td>-0.533</td>
<td>1.358</td>
<td>1.891</td>
<td>0.402</td>
</tr>
<tr>
<td>1 vs. 2</td>
<td>emodels</td>
<td>0.996</td>
<td>0.001</td>
<td>1.991</td>
<td>1.990</td>
<td>1.239</td>
</tr>
<tr>
<td>1 vs. 2</td>
<td>SureSmile</td>
<td>0.143</td>
<td>-0.363</td>
<td>0.648</td>
<td>1.010</td>
<td>0.084</td>
</tr>
<tr>
<td>1 vs. 2</td>
<td>AnatoModels</td>
<td>0.912</td>
<td>-0.293</td>
<td>2.116</td>
<td>2.408</td>
<td>1.200</td>
</tr>
<tr>
<td>1 vs. 3</td>
<td>Plaster Casts</td>
<td>0.284</td>
<td>-0.448</td>
<td>1.016</td>
<td>1.464</td>
<td>0.227</td>
</tr>
<tr>
<td>1 vs. 3</td>
<td>emodels</td>
<td>-0.162</td>
<td>-0.726</td>
<td>0.401</td>
<td>1.127</td>
<td>0.106</td>
</tr>
<tr>
<td>1 vs. 3</td>
<td>SureSmile</td>
<td>0.011</td>
<td>-0.316</td>
<td>0.338</td>
<td>0.654</td>
<td>0.032</td>
</tr>
<tr>
<td>1 vs. 3</td>
<td>AnatoModels</td>
<td>0.035</td>
<td>-0.632</td>
<td>0.701</td>
<td>1.330</td>
<td>0.141</td>
</tr>
<tr>
<td>2 vs. 3</td>
<td>Plaster Casts</td>
<td>-0.128</td>
<td>-0.814</td>
<td>0.557</td>
<td>1.370</td>
<td>0.138</td>
</tr>
<tr>
<td>2 vs. 3</td>
<td>emodels</td>
<td>-1.158</td>
<td>-2.099</td>
<td>-0.218</td>
<td>1.881</td>
<td>1.563</td>
</tr>
<tr>
<td>2 vs. 3</td>
<td>SureSmile</td>
<td>-0.132</td>
<td>-0.668</td>
<td>0.404</td>
<td>1.072</td>
<td>0.095</td>
</tr>
<tr>
<td>2 vs. 3</td>
<td>AnatoModels</td>
<td>-0.877</td>
<td>-1.984</td>
<td>0.231</td>
<td>2.214</td>
<td>1.092</td>
</tr>
</tbody>
</table>
Time-Efficiency

The time requirements for completion of tooth-size measurements using the methods studied are reported in Table 4. In general, it took longest to complete measurements on AnatoModels and shortest on emodels. The time to complete tooth-width measurements differed significantly among the four types of models. The mean time to measure tooth-widths on AnatoModels was statistically significantly longer when compared to emodels and SureSmile diagnostic casts. It also took significantly longer to complete tooth-width measurements on plaster casts when compared to emodels and SureSmile diagnostic models. Lastly, it took statistically significantly longer to measure tooth widths on SureSmile diagnostic casts than on emodels.

Table 5. Time comparisons among methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Plaster casts</th>
<th>emodels</th>
<th>SureSmile</th>
<th>AnatoModels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (s)</td>
<td>269 ± 44</td>
<td>168 ± 76</td>
<td>192 ± 55</td>
<td>272 ± 44</td>
</tr>
</tbody>
</table>

Results are presented as mean values ± standard deviation. Groups depicted by the same superscript are statistically significantly different (P<0.05).

Statistical analysis of time differences between practice cases and experimental cases for each type of dental model revealed that the time needed to measure tooth-widths on the AnatoModels practice cases was statistically significantly longer when compared to the experimental cases indicating the operator becoming more time-efficient with this method. No significant differences were found for any other type of model.
Additionally, the time it took to measure tooth widths on the first five AnatoModels cases was statistically significantly longer than that on the last five AnatoModels cases which indicates that the operators became more efficient as they completed more measurements. When comparing the first five cases and last five cases for any other model, no significant differences were found.
Discussion

As technologies improve, orthodontists will continually have new tools to aid with diagnosis and treatment planning. It is up to the orthodontist to weigh the advantages and disadvantages of a new diagnostic tool in order to determine its clinical usefulness. To our knowledge, our study is the one to date that provides a detailed look into the accuracy, reproducibility, and time-efficiency of tooth-width measurements on multiple dental casts with multiple operators.

In accordance with other studies (Stevens et al., 2006; Mullen et al., 2007; Naidu et al., 2009), our study found that tooth-width measurements taken on digital models had both negative and positive bias. SureSmile showed a small negative bias, i.e., a tendency to underestimate tooth widths when compared with plaster casts. In contrast, emodels showed a moderate positive bias while AnatoModels showed a large positive bias. Systematic biases should be accounted for when comparing tooth-width measurements. For example, if AnatoModels consistently overestimates tooth widths by 0.191 mm when compared to plaster casts, one may want to subtract 0.191 mm from their measurements to get tooth-size values that are close to plaster cast measurements. However, when these measurements are used in ratios, as one would when performing a Bolton analysis, the biases do not affect the ratio because they are constant.

With statistics, there is a tendency to focus on the bias or mean differences. However, it is important to understand that the variance or spread of the differences is as important
and sometimes more important than the bias itself (Bland and Altman, 1999). A large width of the 95% agreement limits is a more serious problem when comparing one method to another because large agreement limits show high variability and less precision (Bland and Altman, 1999). Therefore, both bias and agreement limits need to be taken into consideration to assess accuracy. A method that incorporates both the bias and the agreement limit widths is using the MSE, which, in the present study, helped in comparing the different digital models based on their bias and variance.

Trained technicians set contact points on SureSmile diagnostic models. To evaluate if there were significant differences between tooth sizes with operator-modified points and those with SureSmile technician-set points an additional Altman-Bland test was performed. When comparing operator-modified points to SureSmile technician-set points, we found a bias of 0.0275 and 95% agreement bounds of -0.2907 and 0.3458. These values suggest that there are only minimal differences when an operator adjusts contact points from where the technician has set them. However, even with this knowledge, it is suggested that a clinician always double-check the positions of tooth-width contact points.

For this study, reproducibility was defined as the degree of agreement between the measurements performed on replicate specimens by the three operators (ASTM, 2002). For digital models, the highest reproducibility was found for SureSmile models for all three operator combinations. Ranking the remaining three on reproducibility was
difficult because the three operator combinations did not show a clear second or fourth place. When looking at the reproducibility of the different operator combinations, it is evident that operator 1 vs. 3 had a much higher reproducibility, lower bias between measurements, and lower MSE when compared to the other rater combinations. The highest mean differences in measurements were seen for emodels and AnatoModels for the 1 vs. 2 and 2 vs. 3 operator combinations. However, 1 vs. 3 mean differences for emodels and AnatoModels were much lower when compared to the other operator combinations. This suggests that operator 2 had some difficulty taking reproducible measurements on emodels and AnatoModels. In this case, operator 2 may need further calibration and training if they will be using this method for orthodontic diagnosis and treatment planning. Interestingly, the 1 vs. 3 operator combination showed some of the highest reproducibility values; however, the least reproducible method was plaster casts, which was the gold standard. This shows that, even though plaster casts are currently considered the gold standard, there are instances where digital technologies like SureSmile may render more reproducible measurements, thus be more advantageous to use.

Up to now, no studies have addressed the time-efficiency of taking tooth-width measurements on digital models and compared them to plaster casts and to each other. The results of our study indicate that, on average, tooth-width measurements were performed quickest in the following order: emodels (2 min 48 s), SureSmile (3 min 12 s), plaster casts (4 min 29s), and AnatoModels (4 min 32s). In our study, we considered
mean differences in time greater or equal to 1 minute to be clinically significant. Therefore, the mean difference between emodels and AnatoModels was clinically significant at 1 min 44 s. Furthermore, the mean difference between emodels and plaster casts was 1 min 41 s, which we also considered clinically significant. In contrast, the differences in time-efficiency between emodels and SureSmile (24 s) and AnatoModels and plaster casts (3 s) were considered clinically insignificant. The large standard deviations in time differences can be attributed to the techniques in which the three operators measured the teeth. One examiner consistently measured tooth widths using the qualitative method (see Horton et al., 2010). With the qualitative method, the operator rotates the model in any plane of space while taking tooth-width measurements. The other two examiners used the occlusal method where the operator takes measurements only from the occlusal aspect for the majority of tooth-width measurements, which has been shown to be quick and accurate (Horton et al., 2010).

Interestingly, in our study, the operator who used the qualitative measurement technique took more accurate measurements for emodels, SureSmile, and AnatoModels than the two operators who used the occlusal method. Ultimately, the technique of model analysis is operator-dependent and based on degree of technique familiarity and comfort.

An additional factor that determines time-efficiency in clinical practice is the turn-around time of dental model fabrication. In general, plaster casts are available for analysis as soon as they are poured up and trimmed. In contrast, digital model fabrication may require a week or two, which may result in a slight delay until final dental cast analysis
can be performed. Although plaster casts are more time-efficient in terms of turn-around time, in some instances, they may require more time for retrieval from storage when compared to digital models.

After analyzing accuracy, reproducibility, and time-efficiency of tooth-width measurements, we were able to clearly find a digital model that rendered consistently more accurate and reproducible measurements when compared to other digital models. SureSmile diagnostic models were the most accurate and reproducible when compared to plaster casts and were second fastest when performing tooth-width measurements. In terms of accuracy, emodels were second most accurate and AnatoModels least accurate when compared to plaster casts. AnatoModels were consistently third most reproducible; however, plaster casts and emodels differed from second to fourth most reproducible depending on which operator combinations were looked at.

When evaluating a method of comparison it is important to have a method that has both accuracy and reproducibility. Without accurate measurements, diagnosis and treatment planning may be affected. If a method is not reproducible then this can affect the accuracy. Having a method that is highly accurate with very low reproducibility has very little place in a clinical setting as does having a method with very low accuracy and high reproducibility (Bland and Altman, 1986).
When fabricating orthodontic dental casts there is an inherent loss of information. For example, when taking alginate impressions for plaster casts, fine detail of tooth anatomy is lost because of the inability of the impression material to flow into these areas. With digital models there may be a loss of information through the scanning process. In our study, CBCT volumes used to create AnatoModels had a voxel size of 0.3 mm\(^3\); therefore, measurements smaller than 0.3 mm are impossible to accurately measure on these casts because of the scan resolution. Similarly, SureSmile diagnostic models were fabricated from CBCT scans with a voxel size of 0.2 mm\(^3\); therefore, measurements smaller than 0.2 mm cannot be accurately measured on these casts. The loss of information is also inherent to the plaster cast scanning process necessary to fabricate emodels. For emodels, measurements smaller than 0.015 mm cannot be accurately measured due to a loss of information caused by the scanning process (M. Marshall, personal communication, March 5, 2012).

Other sources of error that may affect tooth-width measurements on dental casts fabricated from CBCT volumes are partial volume effects and beam hardening effects. When a CBCT scan is taken, some voxels are partially filled because the object being scanned may not actually fill an entire voxel. Partial volume effects are dealt with differently depending on the manufacturer of the digital model. When fabricating AnatoModels, the company adds to and removes from to partially filled volumes in order to create smooth surfaces on teeth (R. Lu, personal communication, March 9, 2012). SureSmile’s software analyzes the intensity of partially filled volumes and uses
proprietary algorithms to give the most accurate representation of the image scanned (P. Getto, personal communication, March 12, 2012). The addition or removal of partially filled voxels alters the volume and is another variable that may cause error when measuring tooth widths.

Beam hardening effects are of special concern for AnatoModels as the models are fabricated directly from CBCT volumes. Metallic dental restorations cause scatter in the volume and technicians must remove the artifacts. Often, there are instances where technicians cannot reproduce dental anatomy accurately and estimates of tooth anatomy must be made from incomplete or distorted volumetric data (Figure 8). When measuring tooth-widths on AnatoModels, in some instances large restorations were missing in the digital models and the tooth width needed to be estimated from the information available. In one instance, the technician was unable to differentiate the dentin floor from the restoration due to beam hardening effects and visualization into the pulp chamber was possible (Figure 9).
Figure 8: Beam hardening effects on AnatoModels. Large metallic restorations cause beam-hardening effects rendering ideal tooth morphology impossible to reproduce. (A) Plaster cast with large restoration on maxillary left first molar. (B) Beam hardening effect in the CBCT scan visible on the maxillary left first molar. (C) AnatoModels representation of maxillary left first molar.
A comment must be made on the methodology used in the present study. While the emodels and SureSmile diagnostic casts used in the present study were fabricated from scans of plaster casts, the AnatoModels were fabricated directly from CBCT scans of the patients. When a patient is scanned with a CBCT there may be distortion of the CBCT volume if the patient moves while being scanned (movement artifacts). However, it must be noted that Anatomage’s marketing claims that their software can eliminate the need for impressions altogether. Therefore, using AnatoModels created from CBCT scans was appropriately compared to the other dental models in our study.

**Figure 9:** Root canal visualization on AnatoModels as a consequence of beam-hardening effects.
A few shortcomings and difficulties of the various software packages were identified in this study. Using the emodels software, one of the most difficult aspects of taking tooth-width measurements was landmark identification. For example, contact point identification was difficult in patients with spacing between teeth where surfaces were flat and did not represent normal tooth anatomy (Figure 10). An additional problem was that certain tooth-width measurements became hidden within the tooth so measurements needed to be retaken in order to see the numerical value (Figure 10).

A disadvantage with the AnatoModels software was that the measurement button had to be pressed for each individual tooth-width measurement instead of being able to smoothly transition from one tooth to the next. Additionally, in certain instances, contact points that were in close proximity were difficult to modify and move.

Figure 10: Measurement difficulties on emodels (A) Flat contact point on mesial surface of maxillary left central incisor (B) Hidden measurement on maxillary left lateral incisor.
In general, the SureSmile interface was straightforward. Contact points are preset by SureSmile’s digital laboratory technicians who have been trained on tooth anatomy, contact point location, and placement. However, it is the responsibility of the clinician to check and modify contact point placement as needed. In the present study, the operators encountered minimal difficulty when modifying these contact points. An interesting observation by the operators was that adjustment of contact points often did not alter tooth-width measurements. This was probably due to measurements being taken only to the closest 0.1 mm by this software.

An important factor in determining which type of model to use is cost-effectiveness. Not only do dental models need to be accurate, reproducible, and time-efficient, but they also have to be affordable to orthodontists. In general, digital models are more expensive than plaster casts because they require the extra step of digitization (Table 6). Plaster casts are more economical when compared to digital casts because they do not require any further processing; however, storage of plaster casts can add to the cost of using them in an orthodontic practice. In the end, the clinician needs to determine if the added costs of fabricating digital models are more advantageous to the practitioner when compared to plaster casts.

**Table 6: Dental Cast Pricing in June 2012**

<table>
<thead>
<tr>
<th>Dental Cast</th>
<th>Price (US Dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaster Casts</td>
<td>2-5</td>
</tr>
<tr>
<td>emodels</td>
<td>30-40</td>
</tr>
<tr>
<td>SureSmile</td>
<td>25-45</td>
</tr>
<tr>
<td>AnatoModels</td>
<td>50-75</td>
</tr>
</tbody>
</table>

Exact amount depends on materials used for cast fabrication.
Traditionally, plaster casts have been considered the gold standard for orthodontic dental casts. However, it is important to understand that plaster casts also have their flaws in terms of accuracy. The accuracy of plaster casts can be compromised due to shrinkage of the alginate, expansion of plaster as it sets, or inadequate handling during processing. Acknowledging the shortcomings of plaster casts and their fabrication process, we still used them as the reference model since they are considered the gold standard of dental casts. Even though plaster casts were our reference model to which the digital models were compared to, we found plaster casts are neither the most reproducible nor the dental model that allows the quickest measurements. As orthodontic technology continues to rapidly progress, we may soon see digital models similar to SureSmile technology become the new gold standard due to their high accuracy, reproducibility, and time-efficiency.

Since there are inherent flaws in measuring tooth widths on the gold standard plaster cast, an appropriate follow-up study may compare tooth-width measurements on digital models to teeth with known tooth widths such as extracted teeth.
Conclusions

The following conclusions can be drawn from the present study:

• Tooth-width measurements were most accurate on SureSmile diagnostic casts followed by emodels and AnatoModels when compared to plaster casts.

• Tooth-width measurements were consistently more reproducible on SureSmile diagnostic casts compared to plaster casts, emodels, and AnatoModels.

• Time-efficiency when measuring tooth widths was clinically significant when comparing the fastest dental model, emodels, to the plaster casts and AnatoModels.
References


