Dental Health: Current Research

Research Article

Comparison of Digital Measurement versus Physical Model Measurement: An Analysis of Accuracy and Precision.

Smith DK[1]*, Beaudoin B[2], Messersmith M[3] Blume JD[1]

Abstract

Objective: The objective of this study was to compare the accuracy and precision of digital model measurement to physical measurement of orthodontic models.

Methods: To assess accuracy, teeth were removed from a typodont and measured individually to obtain a gold standard measurement. A stone model of the typodont was then measured n=75 times by both digital and physical measurement and these measurements were compared to the gold standard. To assess precision, patient models (n=27) were measured five times each by each method and assessed via an intra-class correlation coefficient.

Results: The accuracy analysis suggested that measurements of individual tooth width, archlength, and crowding were all more significantly more accurate using the digital measurement technique regardless of whether trapezoidal or catenary measurements were used. The digital technique also showed a significant benefit in the analysis of precision, demonstrating a significant improvement in the intra-class correlation coefficient for each orthodontic measurement considered.

Conclusion: Digital model measurements have the potential to be more accurate and precise than physical model measurements.

Keywords

Measurement accuracy; Measurement precision; Orthodontic treatment planning

Introduction

Many papers have been published comparing measurements taken on plaster casts versus their digital counterparts [1,2]. The majority of studies have used either OrthoCAD or E-models to construct digital models using the companies' suggested clinical procedure [3-7]. A patient's alginate impressions are sent via the mail to a secondary location where the impressions are poured and the models are then scanned into a digital format. OrthoCAD models are constructed by taking serial cuts of a cast, which are then reconstructed to create a 3D volume image. E-models utilize surface laser scanning technology to create a 3D surface image of the plaster casts. The major disadvantage of these protocols is that it has been established that alginate is not dimensionally stable over the period of time it takes to ship impression to the companies for digitization [8].

The distortion that takes place prior to the model being poured is an extra source of variation in any study that attempts to compare the reliability of plaster versus digital measurements. Whether that is an issue depends on the clinical question you are trying to answer. If your goal is to compare the process of shipping the models out for digital measurement versus pouring them in house for physical measurement, that variation is a good thing to incorporate into your study. If your goal is to compare the accuracy and reliability of the digital measurement process compared to the physical measurement process, that variability will adversely affect your ability to examine measurement accuracy.

There are several possible sources of variation in measurements between stone and digital models:

Dimensional change of the alginate from time of impression to the time of pour

Differences in water powder ratios of the stone/water mix (especially if the digital model is produced at a secondary location)

Inherent error in the digitalization process

Differences in measurement error of physical measurements on a cast versus digital measurements

By producing the digital model on site, this study has the advantage of equalizing the first two sources of error. This means that any differences in accuracy or precision of measurements reported by this study are a direct result of the digitization and measurement error alone and cannot be influenced by differences in the process that produced the model unlike studies that send impressions to a lab for digitization.

According to the manufacturer, the OI3D scanner used in this study has an accuracy of .02mm, which has been demonstrated to be reasonable by a separate study [9]. The purpose of this study was to compare the accuracy and precision of plaster model measurements with digital measurements using the in-office OI3D model scanner and Motion View Software.

Materials and Methods

There were 3 primary measures in this study:

1. Individual widths of teeth #2-31.

2. Bolton Discrepancies

3. Crowding- measured using both catenary and trapezoidal arch length

Accuracy can be defined as the propensity of a measurement to approximate the true length. Precision is the degree to which repeated measurements differ from one another. The objective of the first part of this study was to compare the accuracy and precision of plaster and digital measurements to a common gold standard, a dental typodont.

*Corresponding author: Dr. Derek K Smith, DDS, PhD, Department of Biostatistics, Vanderbilt University, Ste 11000, 2525 WestEndAve, Nashville, TN, USA 37203; Fax: 615-343-4924; Tel: 615-322-2001; E-mail: derek.k.smith@vanderbilt.edu

The typodont model used to establish the gold standard for comparison allows for near perfect measures of these teeth as they could be removed and measured directly. Each tooth was measured 3 times and the mean value was used as the true tooth width. Arch lengths were assumed to then be the sum on the teeth's individual measures as these typodonts are designed to be perfect dental arches with no crowding.

Based on preliminary power calculations, each model would have to be measured 75 times to have a 90% chance of detecting a .1 mm difference in tooth size or a 0.5mm difference in arch-length. Since anterior interproximal reduction (IPR) at the .1 mm level are prescribed frequently in digital setups such as Invisalign's ClinCheck deviations of this magnitude represent a clinically meaningful difference.

Alginate impressions (Jeltrate alginate, Densply) of a typodont (Dentsply) were obtained and immediately poured in type III gypsum (Gibralter Lab Stone). Both the alginate and gypsum were mixed to manufacturers suggested water/power ratio and a single mix was used for both arches to ensure no differences in water/power ratios between the upper and lower arches.

In the second part of this study, differences in precision were measured by comparing physical measurements with digital measurements of twenty-seven dental casts. The 27 dental cast sets were selected from the orthodontic clinic records and met the following inclusion criteria:

All permanent teeth present and fully erupted from first permanent molar to first permanent molar.

Casts were not damaged.

Based on preliminary measurements and simulated data analyses, 27 casts would have to be measured 5 times each to have a 90% chance of detecting a .1mm difference in variance of tooth width or a 0.5mm difference in variance of arch-length.

All plaster casts were digitized using the Ortho Insight 3D model scanner and measured using the Motion View Software version 5.5.5002. Plaster casts were measured using vernier digital calipers (ProDent USA). All measurements were made to the nearest 0.01mm. Continuous arch lengths were measured on the physical model using .01" steel ligature wire. The curve was standardized to be the best-fit curve connecting the contact points of all the teeth from first molar to first molar inclusive. To minimize bias, plaster and digital casts were measured at random in blocks of 3 casts with each series of measurement at least two hours apart. Each block was measured at least 1 week apart. Once data was recorded the examiner was blinded to it for the remainder of the experiment.

All statistical analyses were performed in R (version 3.0.2). For the typodont portion of the study, precision was examined graphically and by means of a sign test as the preliminary measures did not suggest that the amount of measurement error varied by tooth. A t-test was used to determine if the measurements differed in terms of accuracy. For the portion of the study based on patient models, an intraclass correlation coefficient (ICC) was calculated for each type of measurement. The ICC can be interpreted as the proportion of the variation in the measurement that is due to true differences in patients as opposed to error in the measurement process. Thus, an ICC near one is desirable; indicating that the measure in question is capturing the patient characteristics as intended.

Confidence intervals for the difference in ICC's between the physical measurements and the digital measurements were calculated using 500 iterations of a cluster bootstrap scheme, clustered by patient cast. Margin of error (two standard deviations) was calculated from the ICC variance due to measurement error. Under the assumption of near-normality in a given measure, this should serve as a guide to clinicians of the expected uncertainty in a given measure on the proper clinical scale.

Results

Typodont: The measurement variance was assessed for each tooth measured. There does not seem to be a discernable difference in variance between particular types of teeth. The physical measurements had higher variance in each of the 24 teeth measured (sign test p<.001). This data is summarized in Table 1.

In order to examine the relative accuracy of the methods, the distribution of the difference of the absolute deviation from the gold standard between the two methods was calculated (Figure 1). This distribution was found to be sufficiently normal by inspection of a Q-Q plot. A t-test was performed and showed statistically significant deviation of the mean from zero (p<.001). The estimated difference in accuracy between the two methods was 0.017 mm (95% CI: 0.011, 0.024) in favor of digital measurement.

Patient Casts: Measurements from the patient casts were assessed using the intraclass correlation coefficient (Table 2). The digital measurements had higher ICC and lower margin of error for each of the measurements considered. In all cases the difference in the margin of error is a clinically meaningful difference ranging from 0.45mm to 1.52mm with a median difference of 0.63mm.
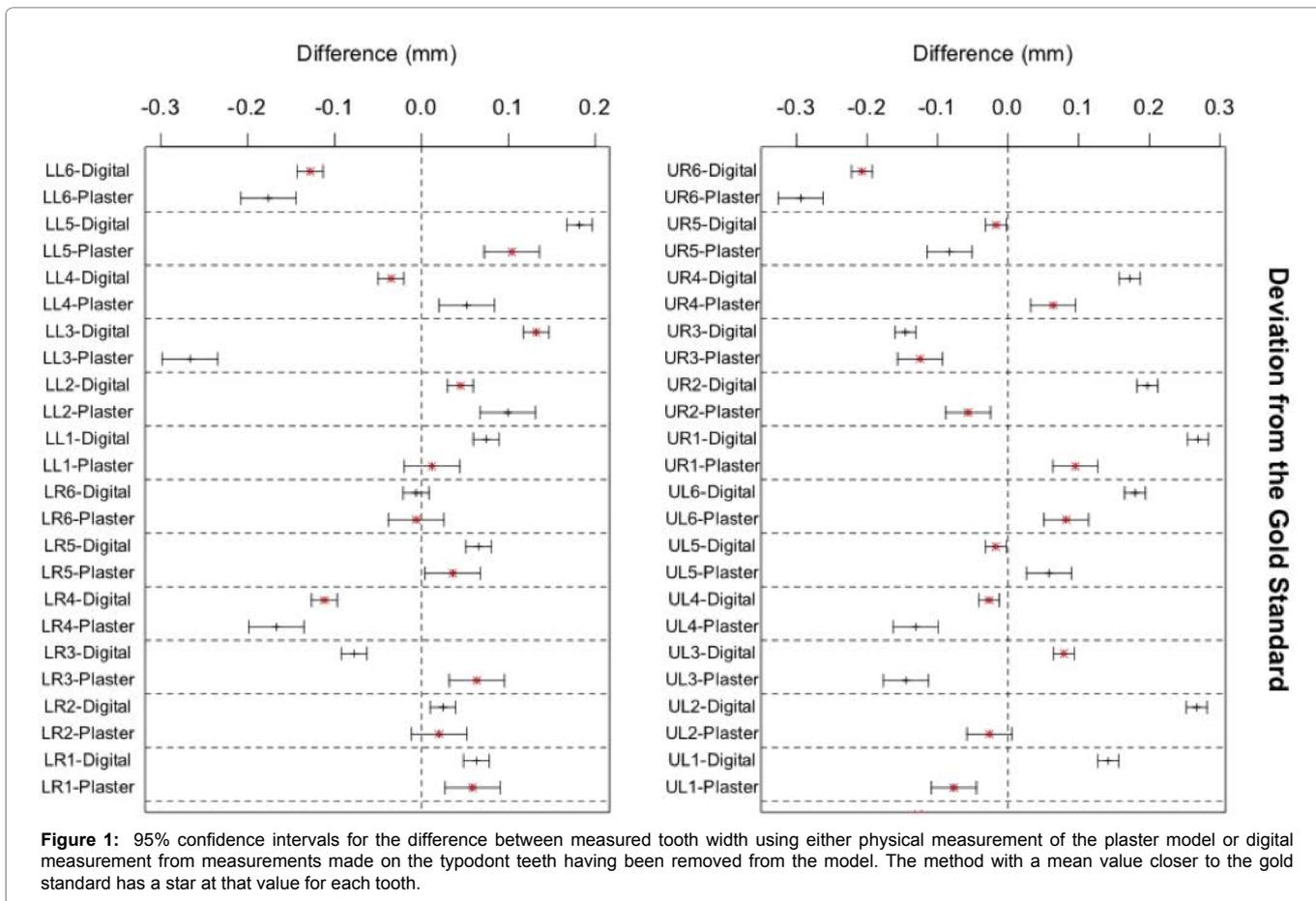
In addition improvements in the ICC, particularly of Bolton discrepancy and crowding measures were substantially improved. The confidence intervals in the table are for the improvement in ICC for using the digital measurement as opposed to a physical measurement, and they demonstrate that there is statistically significant improvement in every measure at the 0.05 significance level (p<0.05). The ICC's and CI's were multiplied by 100 for ease of interpretation as a percent of meaningful variation.

Discussion

The typodont analysis suggests that digital measurements can be significantly more precise than measurements carried out physically on a stone model. It also suggests that on average they can be more accurate. Although the magnitude of accuracy improvement does not seem clinically impressive at first, when combined with an

Table 1: The improvement in accuracy in mm estimated for using the digital measurement technique.

Improved Accuracy of Digital Measurement			
		Improvement in Accuracy	**p-value**
Individual Tooth		0.017 (0.011, 0.024)	<0.001
Maxillary Arch Length	(trapezoidal)	-0.245 (-0.405, -0.084)	0.003
	(catenary)	0.690 (0.479, 0.900)	<0.001
Mandibular Arch Length	(trapezoidal)	0.187 (0.107, 0.266)	<0.001
	(catenary)	-0.021 (-0.198, 0.156)	0.812
Maxillary Crowding	(trapezoidal)	1.286 (1.132, 1.439)	<0.001
	(catenary)	2.220 (1.905, 2.535)	<0.001
Mandibular Crowding	(trapezoidal)	0.460 (0.237, 0.682)	<0.001
	(catenary)	0.307 (0.008, 0.606)	0.044

Figure 1: 95% confidence intervals for the difference between measured tooth width using either physical measurement of the plaster model or digital measurement from measurements made on the typodont teeth having been removed from the model. The method with a mean value closer to the gold standard has a star at that value for each tooth.

extremely clinically meaningful improvement in precision it becomes much more impressive. This means that a digital approach to model measurement results in improved accuracy much more often than physical measurements.

The analysis of the patient casts extends this result beyond the artificial setting of measuring a typodont. The ICC for the physical measurements of Bolton discrepancy and maxillary and mandibular both catenary and trapezoidal show that significant portions of the variation in these measurements is due to measurement error. In a clinical setting this means that when patient A and patient B come in to the clinic and each one is measured for crowding, a significant portion of difference between patient A's crowding measurement and patient B's crowding measurement has nothing to do with the patients and everything to do with measurement errors. The digital approach shows marked improvement in these measures making it significantly more useful in the treatment planning process.

Margin of error was included to give the clinician a meaningful measure of the variation one would expect when using these various measures in an everyday setting. The magnitude of these results demonstrates how measurement error can be consistently affecting treatment plans. For maxillary (catenary) crowding as an example, when our operator made a physical measurement he could be relatively certain that the true value lay somewhere in a 5mm window centered at his measurement. When making the same measure digitally, he could be relatively certain that the true amount of crowding lay within a 2mm window centered at his measurement.

It was much more difficult to measure catenary arch lengths on the plaster models because of the need to adapt a wire to the edges of the teeth on a plaster cast, mark it, and then get an accurate measure of the wire. These physical difficulties are not an issue with digital measurements, and likely account for some the substantial improvements in ICC and margin of error for catenary measures.

The trapezoidal measurements had better reliability than the catenary methods regardless of which measurement method was being applied. However, the advantage in reliability was not of a clinically meaningful magnitude in the digital modality. For this reason, it may be reasonable to use the trapezoidal method when physically measuring plaster casts, however, the decrease in accuracy resulting from approximating a smooth curve with straight edges probably outweighs any benefit gained from the slight increase in precision likely to be observed when using digital measures. In cases of spacing, the OI3D digital model scanner had difficulty capturing inter-dental spaces and tended to over-estimate tooth size. It was necessary to use two-sided tape to tape down the models during scanning to minimize movement of the model, which helped alleviate this issue.

The strength of this study design was the ability to compare accuracy of the two measurement modalities without introducing extra (potentially confounding) error into the analysis by sending an alginate impression to a lab. One limitation of this study is its potential for generalizability. This study consisted of careful measurements performed by a single operator using the OI3D system. Other

Table 2: Margin of error and intra-class correlation coefficient for each orthodontic measurement.

Properties of Measures		Type	Margin of Error (mm)	ICC(x100)	CI (ICCdig-ICCplast)
3-3 Bolton Discrepancy		Plaster	±0.96	91.4	
		Digital	±0.59	96.9	(2.46, 11.5)
6-6 Bolton Discrepency		Plaster	±1.30	93.9	
		Digital	±0.85	97.9	(1.23, 8.75)
Maxillary Arch Length	(trapezoidal)	Plaster	±1.25	99.3	
		Digital	±0.63	99.8	(0.17, 1.71)
	(catenary)	Plaster	±2.17	98.4	
		Digital	±0.80	99.7	(0.68, 3.17)
Mandibular Arch Length	(trapezoidal)	Plaster	±1.26	98.9	
		Digital	±0.69	99.7	(0.29, 1.77)
	(catenary)	Plaster	±1.62	98.7	
		Digital	±0.70	99.7	(0.63, 0.02)
Maxillary Crowding	(trapezoidal)	Plaster	±1.31	97.8	
		Digital	±0.73	99.4	(0.70, 3.69)
	(catenary)	Plaster	±2.47	95.9	
		Digital	±0.95	99.2	(1.96, 6.37)
Mandibular Crowding	(trapezoidal)	Plaster	±1.47	97.1	
		Digital	±0.84	99.1	(0.85, 4.95)
	(catenary)	Plaster	±1.77	96.8	
		Digital	±0.96	99.2	(1.19, 5.31)

The intra-class correlation measures the percentage of the variation in the measurements that are attributable to real differences between patients as opposed to variation due to measurement error. The third column contains a bootstrap 95% confidence interval for the difference in intra-class correlation.

operators of different experience levels or using different methods may find their results differ from what is presented here.

Conclusion

Digital measurement techniques have the potential to perform significantly better than physical measurements. These differences are shown to be statistically significant and of a clinically meaningful magnitude likely to influence treatment planning in clinical practice.

Conflict of Interest

The authors deny any conflict of interest.

References

1. Fleming PS, Marinho V, Johal A (2011) Orthodontic measurements on digital study models compared with plaster models: a systematic review. Orthod Craniofac Res 14: 1-16.

2. De Luca Canto G, Pacheco-Pereira C, Lagravere MO, Flores-Mir C, Major PW (2015) Intra-arch dimensional measurement validity of laser-scanned digital dental models compared with the original plaster models: a systematic review. Orthod Craniofac Res 18: 65-76.

3. Santoro M, Galkin S, Teredesai M, Nicolay OF, Cangialosi TJ (2003) Comparison of measurements made on digital and plaster models. AJODO 124: 101-105.

4. Quimby ML, Vig KW, Rashid RG, Firestone AR (2004) The Accuracy and Reliability of Measurements Made on Computer-Based Digital Models. Angle Orthodontist 74: 298-303.

5. Mayers M, Firestone AR, Rashid RG, Vig KW (2005) Comparison of peer assessment rating (PAR) index scores of plaster and computer-based digital models. AJODO 128: 431-434.

6. Mullen SR, Martin CA, Ngan P, Gladwin M (2007) Accuracy of space analysis with emodels and plaster models. AJODO 132: 346-352.

7. Leifert MF, Leifert MM, Efstratiadis SS, Cangialosi TJ (2009) Comparison of space analysis evaluations with digital models and plaster dental casts. AJODO 136: 16.e1-4.

8. Todd JA, Oesterle LJ, Newman SM, Shellhart WC (2013) Dimensional changes of extended-pour alginate impression materials. AJODO 143: S55-63.

9. Brittany P (2013) Personal Communication. Motion View Software LLC.

Author Affiliation   Top

¹Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA

²Department of Oral and Maxillofacial Surgery, Vanderbilt University Medical Center, Nashville, TN, USA

³Chief and Associate Professor, Division of Orthodontics, Vanderbilt University Medical Center, Nashville, TN, USA