Original Article

The Accuracy and Reliability of Measurements Made on Computer-Based Digital Models

Meredith L. Quimby, DDS, MS[a]; Katherine W. L. Vig, BDS, MS, FDS, DOrth[b]; Robert G. Rashid, DDS, MAS[c]; Allen R. Firestone, DDS, MS[d]

Abstract: For reasons of convenience and economy, orthodontists who routinely use and maintain pre- and posttreatment plaster casts are beginning to use computer-based digital models. The purpose of this study was to determine the accuracy (validity), reproducibility (reliability), efficacy, and effectiveness of measurements made on computer-based models. A plastic model occlusion ie, dentoform, served as a gold standard to evaluate the systematic errors associated with producing either plaster or computer-based models. Accuracy, reproducibility, efficacy, and effectiveness were tested by comparing the measurements of the computer-based models with the measurements of the plaster models—(1) accuracy: one examiner measuring 10 models made from a dentoform, twice; (2) reproducibility and efficacy: two examiners measuring 50 models made from patients, twice; and (3) effectiveness: 10 examiners measuring 10 models made from patients, twice. Reproducibility (reliability) was tested by using the intraclass correlation coefficient. Repeated measures of analysis of variance for multiple repeated measurements and Student's *t*-test were used to test for validity. Only measurements of maxillary and mandibular space available made on computer-based models differed from the measurements made on the dentoform gold standard. There was significantly greater variance for measurements made from computer-based models. Reproducibility was high for measurements made on both computer-based and plaster models. In conclusion, measurements made from computer-based models appear to be generally as accurate and reliable as measurements made from plaster models. Efficacy and effectiveness were similar to those of plaster models. Therefore, computer-based models appear to be a clinically acceptable alternative to conventional plaster models. (*Angle Orthod* 2004;74:298–303.)

Key Words: Efficacy; Effectiveness; Digital models

INTRODUCTION

Computer-based record keeping is routine in many orthodontic offices. Digital photography and digital radiography are replacing analogue systems and are providing diagnostic quality images at a reasonable cost. Computer-based charts and patient management systems constitute a digital patient record. These computer-based records eliminate the need for physical record room or chart storage facilities. Computer-based digital models are now available, and they have the potential to replace the plaster cast, the last physical record, and to eliminate the model storage room. However, the accuracy, reliability, efficacy, and effectiveness of the computer-based models have not been systematically evaluated.

Study models, photographs, radiographs, and a clinical examination provide the information required to diagnose a malocclusion and to develop an orthodontic treatment plan.[1] Study models provide a three-dimensional view of a patient's occlusion, which enables the clinician to evaluate the malocclusion in more detail than by clinical examination. Study models are more amenable to routine measurements than are intraoral measurements.[2] Arch measurements on study models are a routine and essential step in the analysis of a patient's malocclusion. One group of investigators reported that in a majority of cases, study models alone provided adequate information for making treatment decisions.[1]

To date, many methods have been used to measure and

[a] Private practice, Gastonia, NC.

[b] Professor and Chair, Section of Orthodontics, College of Dentistry, The Ohio State University, Columbus, Ohio.

[c] Associate Professor, Clinical, Section of Restorative and Prosthetic Dentistry, College of Dentistry, The Ohio State University, Columbus, Ohio.

[d] Associate Professor, Section of Orthodontics, College of Dentistry, The Ohio State University, Columbus, Ohio.

Corresponding author: Allen R. Firestone, DDS, MS, Section of Orthodontics, College of Dentistry, Postle Hall, 305 W 12th Avenue, The Ohio State University, Columbus, OH 43210 (e-mail: firestone.17@osu.edu).

Accepted: July 2003. Submitted: May 2003.

to analyze plaster casts. Dividers, calipers, and Boley gauges have provided the standard of measurement against which newer methods have been evaluated.[2–8]

Several companies now offer computer-based (pseudo) three-dimensional models. Generally, impressions of the patient's occlusion, taken in the practitioner's office, are sent to the company. At the company, models are poured up and scanned to create data points representing the surface of the teeth and the supporting soft tissue. These data points are then sent by way of the Internet to the originating dentist's office and are available to the proprietary company-supplied software, which resides on the practitioner's computer. The software allows visualization of the models in three dimensions such that the orthodontist can evaluate various parameters of the patient's dentition eg, the occlusion, mesiodistal tooth size, Bolton's Ratio, arch length, arch width, overjet, and overbite. Each company has its own program to produce these and other measurements. Currently, no data comparing the accuracy, reproducibility, efficacy, and effectiveness of computer-based models with those of routine plaster models have been published.

The overall objective of this study was to determine the accuracy, reproducibility, efficacy, and effectiveness of measurements made on a computer-based model. The specific aims of this study were:

- To compare the accuracy and reproducibility of measurements made on plaster models and on computer-based models with those of an artificial occlusion (''dentoform'').
- To evaluate the efficacy and effectiveness of measurements made on a computer-based model of the natural dentition with measurements made on a plaster model of the same natural dentition.

MATERIALS AND METHODS

This study was approved by the University Institutional Review Board. All subjects gave informed consent. In the case of minors, parents signed consent forms, and minors gave assent. The sample consisted of 50 consecutive patients entering orthodontic treatment in the Graduate Orthodontic Clinic at The Ohio State University. Sample size was determined from values in the literature.[5] With a nondirectional alpha risk of 0.05 and a power of 90%, a sample size of 44 subjects would be required to demonstrate a difference of +2.0 mm. The inclusion criteria were:

- A full complement of permanent incisors, canines, premolars, and first molars in both the maxillary and mandibular arch.
- All teeth having normal morphology—any patients or casts showing gross dental abnormalities were rejected.
- The teeth in the casts displayed no visible attrition, caries, or restorations affecting the mesiodistal or buccal-lingual diameter of the crown.

Two sets of maxillary and mandibular alginate impressions were taken on every patient. Plastic trays (Imperial Plastic Trays, Hanover, Germany) were used for the impressions. Once the tray size was selected, the same-sized tray was used for both impressions on the individual patient. Alginate impression material (Jeltrate, Fast set, Dentsply, Milford, Del) was used. The impressions were disinfected (Wex-Cide, Wexford Labs Inc, Kirkwood, Mo), wrapped with a moist paper towel, and placed in a plastic bag until poured. The impressions for the stone models were poured within one hour after they were made using white orthodontic stone (WhipMix Orthodontic Stone, Louisville, Ky).

The impressions for computer-based models were also disinfected, wrapped in a moist paper towel, and placed in a sealed plastic bag. The impressions were sent by UPS Next Day Air Mail per company instructions. According to the firm, the impressions were poured the day of arrival. Identical procedures were followed for impressions of a plastic artificial occlusion ie, dentoform (Model D95SDP-200, Kilgore International Inc, Coldwater, Mich). The impression procedure with the plastic artificial occlusion was repeated 10 times.

Data collection

The measurements included seven categories.

- Mesiodistal widths—As described by Hunter and Priest,[2] the greatest mesiodistal diameter from the anatomic mesial contact point to the anatomic distal contact point of each tooth was measured parallel to the occlusal plane. Measurements were recorded to the nearest 0.1 mm.
- Arch length—The arch length was measured by the segment arch approach. Segment A is the distance from the mesial contact point of the right first permanent molar to the mesial contact point of the right canine. Segment B is the distance from the mesial contact point of the right canine to the mesial contact point of the right central incisor. Segment C is the distance from the mesial contact point of the left central incisor to the mesial contact point of the left canine. Segment D is the distance from the mesial contact point of the left canine to the mesial contact point of the left first permanent molar. The segments were summed to the nearest 0.1 mm to equal the arch length for both the maxillary and mandibular arch.
- Arch width: Intermolar width was measured as the distance between the mesiobuccal cusp tips of the permanent first molars. Intercanine width was measured as the distance between the crown tips of the permanent canines. These measurements were made on both maxillary and mandibular casts. Measurements were recorded to the nearest 0.1 mm.
- Overjet: The greatest amount of horizontal overlap was recorded from the labial surface of the mandibular central incisors to the lingual surface of the maxillary central

TABLE 1. Mean (standard deviation) for Measurements Made on Gold Standard (dentoform), Digital, and Plaster Models (n = 10)

Measurement	Mean (SD)[a]			Digital vs Dentoform	Digital vs Plaster	Variance Digital vs Plaster
	Dentoform	Digital	Plaster			
Overbite	1.50	1.41 (0.40)	1.40 (0.21)			*
Overjet	1.50	1.45 (0.53)	1.48 (0.30)			*
Maxillary intermolar	54.45	54.72 (0.85)	54.43 (0.26)		*	*
Maxillary intercanine	36.55	36.04 (0.51)	36.44 (0.26)		*	*
Mandibular intermolar	47.40	47.42 (0.52)	47.38 (0.33)		*	*
Mandibular intercanine	26.50	26.31 (0.27)	26.65 (0.24)		*	
Maxillary available	73.40	74.87 (1.06)	73.58 (0.45)	*	*	*
Maxillary required	73.15	73.69 (0.93)	73.00 (0.37)		*	*
Mandibular available	64.15	65.71 (0.74)	64.02 (0.43)	*	*	*
Mandibular required	63.25	63.85 (0.86)	63.24 (0.49)		*	*

[a] Units are represented in millimeters.
* $P < .05$.

incisor. Overjet was recorded to the nearest 0.5 mm on plaster models and 0.1 mm on computer-based models.

- Overbite: The greatest amount of vertical overlap of the maxillary central incisors to the mandibular central incisors was measured with the casts occluded. Overbite was recorded to the nearest 0.5 mm on plaster models and 0.1 mm on computer-based models.

Calculated values

- Space available was the summation of the maxillary and mandibular arch length segments (A + B + C + D) to the nearest 0.1 mm. This was done on both the plaster models and the computer-based models.
- Space required was the summation of the maxillary and mandibular mesiodistal tooth widths including right and left second premolar, first premolar, canine, lateral incisor, and central incisor. This was done on both the plaster models and the computer-based models.

Instrumentation

Digital calipers (Digimatic Caliper: 700-113 MyCal Lite, Mitutoyo America Corp, Plymouth, Mich) were used to measure the plaster study casts. A standard computer mouse was used to draw the distances from point to point on the computer models. The monitors used were 19-inch CRT (18-inch, 45 cm viewable Optiquest z90, ViewSonic, Walnut, Calif) with a monitor resolution of 1280 × 1024 pixels and 0.22-mm horizontal dot pitch (0.26 mm diagonal dot pitch) with 32-bit color. All measurements were made to the nearest 0.1 mm, except for the plaster overbite and overjet measurements that were measured to the nearest 0.5 mm. The software used to display and measure the computer-based models was Version 1.17 (OrthoCad, Cadent Inc, Fairview, NJ).

Examiners were second and third year orthodontic residents at The Ohio State College of Dentistry. To test the accuracy of measurements made on computer-based models, one examiner (Dr. Quimby) measured the 10 plaster models and the 10 computer models made from the standard dentoform. All the measurements were performed for each set of models and then repeated two weeks later. The dentoform was measured twice, and the mean of these values was considered the ''gold'' or ''truth'' standard.

Efficacy has been defined as the extent to which a procedure produces the intended result under ideal conditions, and effectiveness has been defined as the extent to which a procedure produces the intended result under normal conditions.[9] The distinction between efficacy and effectiveness may be compared with the differences between a laboratory test of a procedure and a field test of a procedure or between a controlled experiment and everyday experience.

To evaluate the efficacy and reproducibility of measurements made on computer-based models, two examiners, experienced in the measurement of computer-based models, measured the 50 plaster models and the 50 computer models made from the orthodontic patients. All the measurements were performed for each set of models and then repeated after an interval of at least two weeks.

To evaluate the effectiveness of measurements made on computer-based models under normal conditions, a random set of 10 models was selected from the sample group of 50 models. Ten examiners measured the 10 plaster models and the 10 computer-based models at two time intervals, at least two weeks apart. All the measurements were performed for the 10 models.

Statistical tests

The statistical methods used in this study were the repeated measures analysis of variance for multiple repeated measurements and the intraclass correlation coefficient for reliability (JMP version 4.0; SAS version 8.2, SAS, Cary, NC).[10]

RESULTS

Dentoform

Table 1 presents the mean values for measurements made directly on the dentoform and the same measurements made

TABLE 2. Mean Values (SD) of Two Measurements From Plaster and Digital Models Prepared from Duplicate Impressions of 50 Patients and the Mean Difference between the Same Measurements on Digital and Plaster Models and System Significance

Measurement	Mean (SD)[a]			System Significance
	Digital	Plaster	Digital-Plaster	
Overbite	2.2 (2.16)	2.86 (2.53)	−0.66	$P < .0001$
Overjet	3.14 (2.41)	3.59 (2.94)	−0.45	$P < .0001$
Maxillary intermolar	51.02 (3.4)	50.52 (3.26)	0.5	$P < .0001$
Maxillary intercanine	34.02 (2.43)	33.8 (2.28)	0.22	$P < .0001$
Mandibular intermolar	44.28 (3.46)	43.67 (3.58)	0.61	$P < .0001$
Mandibular intercanine	25.71 (2.22)	25.56 (2.52)	0.15	
Maxillary available	74.71 (5.77)	74.17 (5.26)	0.54	$P < .0001$
Maxillary required	76.24 (4.62)	74.01 (4.42)	2.23	$P < .0001$
Mandibular available	65.9 (4.51)	63.02 (4.25)	2.88	$P < .0001$
Mandibular required	65.45 (4.67)	65.24 (3.98)	0.212	

[a] Units are represented in millimeters.

TABLE 3. Mean Value (SD) for 10 Examiners at Two Time Intervals Measuring 10 Models. Digital and Plaster Models and the Difference between Measurements and System and Examiner Significance. (n = 10)

Measurement	Mean (SD)[a]			System Significance	Examiner Significance
	Digital	Plaster	Digital-Plaster		
Overbite	1.75 (2.42)	2.43 (3.11)	−0.68	*	*
Overjet	3.77 (1.71)	3.99 (2.18)	−0.22	*	*
Maxillary intermolar	50.76 (2.29)	50.25 (2.42)	0.51	*	*
Maxillary intercanine	33.81 (1.05)	33.47 (1.08)	0.34	*	*
Mandibular intermolar	43.47 (2.69)	42.99 (2.52)	0.48	*	*
Mandibular intercanine	25.29 (3.07)	25.1 (2.91)	0.19		*
Maxillary available	77.31 (5.23)	75.58 (4.52)	1.73	*	*
Maxillary required	74.95 (3.29)	74.62 (3.54)	0.33	*	*
Mandibular available	66.28 (4.56)	64.38 (4.10)	1.9	*	*
Mandibular required	66.43 (6.10)	65.54 (3.51)	0.74	*	*

[a] Units are represented in millimeters.
* $P < .0001$.

on the plaster and computer-based models. There were no significant differences between measurements made on the plaster models and those made on the dentoform. Measurements made on the computer-based models differed significantly from measurements made directly on the dentoform for mandibular and maxillary space available (Table 1). The variance of measurements made on the computer-based models was significantly greater than the variance of measurements made on the plaster models for all measurements except mandibular intercanine distance.

Two examiners, 50 model pairs

Intrarater reliability (reproducibility) for the two examiners measuring the 50 plaster and 50 computer-based models on two separate occasions was high. Intraclass correlation coefficients were 0.90 or greater for all measurements on each system (data not shown).

The repeated measures analysis of variance showed significant differences between computer-based models and plaster models (Table 2). There were significant differences in all categories measured except lower canine arch width and mandibular arch length required. Measurements were greater for the computer-based models except for overbite and overjet. However, these measurements were only recorded to the closest 0.5 mm for the plaster models.

Ten examiners, 10 model pairs

Table 3 presents the data generated from the 10 examiners and 10 patients. There were significant differences between the two systems, computer-based and plaster models, and among the examiners for all measurements except mandibular intercanine width.

DISCUSSION

Dentoform

The direct measurements of the plastic teeth in the dentoform served as the gold standard control to evaluate the accuracy of measurements made on plaster and computer-based models. Measurements made on plaster casts poured within one hour of the impression did not differ from those made directly on the gold standard dentoform. Measurements made on computer-based models produced from plaster casts poured the day after the impressions were taken did not differ from those made on the gold standard except for mandibular and maxillary space available. Thus, for measurements made directly between two anatomical points on the computer-based models the accuracy was equal to that of plaster models.

The process of measuring segments apparently introduced error into the measurements. In the present investigation, plaster models poured within one hour after the irreversible hydrocolloid impressions were made were accurate representations of the original, master dentoform teeth. If we accept the plaster model as the ''gold standard'' of current clinical practice, where measurements are seldom made directly in the patient's natural dentition, the measurements made on the two systems did differ statistically except for overbite and overjet. In the majority of measurements, the magnitude of the differences was less than 0.5 mm.

Two examiners, 50 model pairs

Measurement taken on both the plaster and computer-based models by experienced examiners showed a high degree of reproducibility. All categories showed excellent reliability, with intraclass correlation coefficients $>$0.90 (data not shown). In comparing the two systems, measurements made from computer-based models were larger than those made on plaster casts except overbite and overjet (Table 2). For the most part, the differences between measurements on the two model systems were less than 1 mm. The measurements for overbite and overjet were made to 0.1 mm on the computer-based models and 0.5 mm on the plaster models. Therefore, smaller values for the computer-based models are not unexpected. Thus, when applied by experienced examiners, the process of making measurements on computer-based digital models is reproducible and efficacious.

Ten examiners, 10 model pairs

For the data from the 10 examiners (Table 3), significant examiner differences existed for all categories measured. Additionally, statistically significant differences between the two model types occurred in all categories except the mandibular intercanine width. The measurements made on computer-based models showed greater variation in all categories except overbite and overjet. The mean difference between the two systems for 10 examiners ranged from 0.19 to 1.9 mm, and this was similar to the mean difference between the systems for two examiners, which ranged from 0.15 to 2.9 mm. Most measurements differed by less than 1 mm. Thus, when applied in a more realistic setting, the computer-based measurement system was as effective as plaster models.

The conditions of this study probably overestimate the effectiveness of the plaster model. While it is unlikely that irreversible hydrocolloid impressions are poured within one hour in all orthodontic practices, it is equally unlikely that they are stored for 24 hours. The choice of one hour was on the basis of published reports that showed that plaster models poured within one hour were accurate representations to test bodies.[11,12] This study does not demonstrate that plaster models, as they are commonly prepared in orthodontic practices, are effective representations of the natural dentition.

There was a significantly greater variance associated with measurements made on the computer-based models compared with the same measurements made on the plaster models (Tables 2 and 3). We hypothesize that the larger values for measurements made on the computer-based models may have several possible sources: (1) the increased time that elapsed before the irreversible hydrocolloid impressions were poured in plaster, (2) the process of producing the plaster casts by the manufacturer, (3) the process of scanning and recording data points from the plaster model, (4) the display and measurement algorithms of the manufacturer's proprietary software, and (5) the examiners' lack of familiarity with the computer-based measurement of computer-based models.

Other investigators have reported that storage of irreversible hydrocolloid in a wet paper towel for 24 hours before pouring plaster models produced significant dimensional changes compared with models poured immediately or after storage for one hour in a paper towel.[11] There is a further effect of the disinfection protocol that can, in turn, be anticipated to change with the length of storage of the impression before the cast is poured.[12–14] The manufacturer (OrthoCad, Cadent Inc) does not release detailed, technical information either on the process of producing the casts ie, type of dental stone, powder-water ratios, water temperature etc, or on the accuracy and reproducibility of the scanning method by which data points are produced from models. Finally, the relative inexperience of the 10 examiners with the computer-based models and the measurement software compared with the measurement of plaster models may have increased both measurement error and the variability of the measurement values.

The differences between the plaster and computer models, though generally small, were statistically significant; the question remains open if they are clinically significant. The computer models are reasonably reliable and accurate. They can provide the clinician with adequate information to develop a treatment plan and thus eliminate the need for storing plaster casts. The true test of clinical significance would be to determine whether treatment plans produced with computer-based models differed significantly from treatment plans produced with plaster models. In turn, the results of the treatment from the two different sets of models would determine the true value of computer-based models. In the opinion of the authors, it is questionable whether the differences demonstrated in this investigation would lead to significantly different clinical results.

The computer-based model software is routinely being updated to provide more features and to improve the accuracy of the measurements. More recent versions of the software have new features that theoretically could improve accuracy. It is the opinion of the authors that the accuracy and reliability of the computer-based models is acceptable, and it will be the relative convenience and total cost of the computer-based model that determines its acceptance. Models can be viewed chair-side in seconds, and thousands can be stored in the space of a moderately sized hard cover novel. The model can be shared over a network within an office or offices of a practice or with another party without it ever leaving the practice or without the danger of the models being damaged by handling. A copy of the model can be secured at a second site for minimal or no cost. All these benefits are based on networked chair-side computers with their associated benefits and costs. However, computer failure, software failure, or manufacturer insolvency can possibly mean that the models may become inaccessible for a time or forever.

CONCLUSIONS

The extent and limitations of the accuracy, reliability, efficacy, and effectiveness of measurements made from computer-based models have been demonstrated in this investigation.

The acceptance of computer-based models will depend primarily on their utility, and this in turn will depend on the cost-benefit ratio to the individual practitioner.

REFERENCES

1. Han KU, Vig KWL, Weintraub JA, Vig PS, Kowalski CJ. Consistency of orthodontic treatment decisions relative to diagnostic records. *Am J Orthod Dentofacial Orthop.* 1991;100:212–219.
2. Hunter WS, Priest WR. Errors and discrepancy in measurement of tooth sizes. *J Dent Res.* 1960;39:405–414.
3. Brook AH, Pitts NB, Yau F, Sandar PK. An image analysis system for the determination of tooth dimensions from study casts: comparison with manual measurements of mesiodistal diameter. *J Dent Res.* 1986;65:428–431.
4. Champagne M. Reliability of measurements from photocopies of study models. *J Clin Orthod.* 1992;26:648–650.
5. Mok KHY, Cooke MS. Space analysis: a comparison between sonic digitization (DigigraphTM Workstation) and the digital caliper. *Eur J Orthod.* 1998;20:653–661.
6. Musich DR, Ackerman JL. The canenometer: a reliable device for estimating dental arch perimeter. *Am J Orthod Dentofacial Orthop.* 1973;63:366–375.
7. Rudge SJ, Jones PT, Bowden DEJ. The reliability of study model measurement in the evaluation of crowding. *Br J Orthod.* 1983; 5:225–231.
8. Schirmer UR, Wiltshire WA. Manual and computer-aided space analysis: a comparative study. *Am J Orthod Dentofacial Orthop.* 1997;112:676–680.
9. Last J. *A Dictionary of Epidemiology.* Oxford, UK: Oxford University Press; 1988:41.
10. Norman GR, Steiner DL. *Biostatistics: The Bare Essentials.* London, UK: BC Decker Inc; 1998:94–100.
11. Coleman RM, Hembree JH Jr, Weber FN. Dimensional stability of irreversible hydrocolloid impression material. *Am J Orthod.* 1979;75:438–446.
12. Johnson GH, Chellis KD, Gordon GE, Lepe X. Dimensional stability and detail reproduction of irreversible hydrocolloid and elastomeric impressions disinfected by immerison. *J Prosthet Dent.* 1998;79:446–453.
13. Boden J, Likeman P, Clark R. Some effects of disinfecting solutions on the properties of alginate impression material and dental stone. *J Prosthodont Restor Dent.* 2001;9:131–135.
14. Taylor RL, Wright PS, Maryan C. Disinfection procedures: their effect on the dimensional accuracy and surface quality of irreversible hydrocolloid impression materials and gypsum casts. *Dent Mater.* 2002;18:103–110.