REVIEW

Digital Models and Orthodontic Diagnosis: What Degree of Reliability?

Abdoul Hafizou Rabe, Fatima Salek, Intissar El Idrissi, Fatima Zaoui, Hicham Benyahia
Department of Dentofacial Orthopedics, Ibn-Sina Hospital Center, Faculty of Dental Medicine, Mohammed V University, Rabat, Morocco

ABSTRACT
Background: Our study aims to evaluate, through a systematic review, the reliability of numerical models compared to conventional models on the main parameters of orthodontic diagnosis.
Method: four databases were consulted: PubMed; Google Scholar, Cochrane Library, and Ebscohost. The research included published studies since 2010, meta-analysis studies, randomized and non-randomized controlled trials, prospective and retrospective studies.
Results: Among 3811 selected references, only five studies met our inclusion criteria. In the systematic review, there were statistical differences between the digital models and the plaster models. However, this difference is clinically acceptable. On the other hand, there are some limitations, relative to the types of the severity of the congestion, the elapsed time to digitize, and the numerical means.
Conclusion: The results of our systematic review have shown that there is no clinically significant difference between the numerical and physical models for the majority of diagnostic parameters.

KEYWORDS: Orthodontics; Cone-Beam Computed Tomography; Digital Models; Plaster Models; Accuracy.

INTRODUCTION

Study models are essential tools for diagnosing and planning the orthodontic treatment plan. Nowadays, with the integration of the digital component in dentistry, we are witnessing the progressive replacement of classic plaster models by digital models. This is made possible by digitizing impressions or plaster study models, and more recently by taking an impression directly in the mouth using an intra-oral scanner to obtain digital models (1). However, it is reasonable to ask the question about the reliability of these models in all aspects of a 3D orthodontic diagnosis, compared to physical plaster models.

In 1999, AlignTechnology Inc (San Jose, California, USA) introduced OrthoCad, a digital modeling application based on a proprietary plaster model scanning process (2). Three years later, GeoDigm Corp. (Falcon Heights, Minnesota, USA) launched "emodels", a scanning service for plaster models using non-destructive laser scanning (2).

Classically, study models are used to visualize the morphology and position of teeth on their respective dental arches and provide a three-dimensional image. Compared to a simple clinical examination, the casting examination allows the clinician to evaluate a possible malocclusion more accurately, with linear measurements in the transverse direction (inter-molar, inter-canine, incisal middle deviation), in the sagittal direction (overhang, molar angle class, and canine) and the vertical direction (anterior overlay, anterior and posterior gaps) (3). Also, the models allow measurements of arch and tooth size to quantify the crowding and Bolton index (4). Many types of numerical models are now available to orthodontists. The question is whether measurements obtained from numerical models could replace those obtained from plaster models.

Several studies have attempted to compare analyses of numerical models with manual measurement of the plaster model (5-8). Our study aims to evaluate through a systematic review of the reliability of numerical models compared to conventional models on the main parameters of orthodontic diagnosis.

MATERIALS AND METHODS

The search for both inclusion and exclusion criteria was based on the PICOT format (Table 1). Electronic searching of the articles in this systematic review was conducted...
through the PubMed Central, Cochrane Library, Google Scholar, and Ebscohost databases. The PICOT format (Table 1) was used to develop the search strategy. Articles published since 2010 were analyzed. No limitations on the language of publication were imposed.

Boolean operators "OR" and "AND" were used to define and connect search terms.

Table 1, PICOS Criteria

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Patients requiring orthodontic treatment</td>
</tr>
<tr>
<td>Intervention</td>
<td>Digital models</td>
</tr>
<tr>
<td>Comparaison</td>
<td>Classic models</td>
</tr>
<tr>
<td>Outcome</td>
<td>Reliability and precision</td>
</tr>
<tr>
<td>Study design</td>
<td>No clinical case reports, interventional or observational human studies with specific data on reliability and precision of the Digital models</td>
</tr>
</tbody>
</table>

Selection of Studies

Systematic searches were carried out by an author (AHR). The selection of studies was done by two authors (AHR and HB). Titles and abstracts were read and the studies were then evaluated against the eligibility criteria. The two authors independently evaluated the selected studies for eligibility. Papers that met the criteria were selected to be read in full text. In case of disagreement between authors, the study was selected for full-text reading. We included studies meeting all of the following criteria: studies published since 2010, meta-analysis studies, randomized and non-randomized controlled trials, prospective and retrospective studies, studies concerning reliability and precision of the Digital models.

We excluded all publications on animals, literature reviews, narrative reviews, opinion articles, studies on patients with syndromic or severe facial deformities, studies on patients treated in the mixed or temporary denture.

Database and Search Strategy

Electronic searching of the articles in this systematic review was conducted through PubMed, Cochrane Library, Google Scholar, and EBSCOhost databases. We used the keywords present in the MeSH for the English-language articles, according to the search equation ((3d OR digital OR intraoral OR electronic or computer* OR software) AND (impression* OR model* OR scanner*)) AND (orthodontics OR orthod*) AND (accuracy OR precision OR effic* OR limitat*).

To extract data from the selected articles, we used a table to be reported for each study: authors and year of study, type of study, sample size, type of intervention and outcomes (Table 2). In case of disagreement; the article was discussed with the other authors.

RESULTS

The search with keywords gave the following results: PubMed produced 766 publications, Google scholar 2790, Cochrane Library 0, and Ebscohost 255 publications. After excluding 678 repeat articles, all titles and abstracts were read and those found to be unrelated to the journal were eliminated. twenty-one pre-selected articles were read in their entirety, and after applying the inclusion and exclusion criteria, five references were selected for this systematic review.

The selection process is illustrated in the Flowchart (Fig 1).

The five selected studies reported the following results (Table 2):

- The In Vitro study by Yu-Ming Liang et al (9) evaluated on 29 pairs of models the efficiency of the numerical model in different congestion situations. The sagittal and transverse measurements between the plaster model and the numerical model from the 3Shape™n system showed no statistically significant difference except for the mandibular space requirement (p=0.012), which remains clinically insignificant. When they compared the measures of low, medium, and severe crowding, they found a statistically significant difference between low and severe crowding (p=0.008), between severe and moderate crowding (p=0.017), but not significant between low and moderate crowding (p=0.315).

- The In Vivo study by Alana Tavares et al (10) compared on 37 mandibular models, the measurements from the conventional model, the digital model, and the paper image of the digital model. The results showed that there was no statistical difference between the measurements on the numerical models and those on the plaster model (p>0.05). On the other hand, between the measurements on the printed paper images of the digital model and those on the digital model, there is a statistically significant difference with a mean difference of 0.38 mm for the inter-canine width and 0.58 mm for the inter-molar width underestimated compared to the plaster. However, this difference is considered clinically insignificant.

- The prospective study by MG Cesur (11) tested the dimensional stability of the alginate impression scan as a function of time. They took 60 alginate impressions and then scanned 20 impressions on the same day and cast these 20 impressions with plaster and scanned the 20 plaster models obtained. The remaining 40 impressions underwent the same operation in groups of 20 at T1 (one day) and then T2 (in 2 days). Measurements in the vertical, sagittal, and transverse directions on the plaster models, scanning of the impression, and scanning of the plaster models showed statistically significant differences at T0 (p<0.01), T1 (p<0.05), and T1 (p<0.001). However, this difference is clinically acceptable. On the other hand, when considering the measurement of the arch perimeter alone, there is no statistically significant difference at T0 and T1. The difference exists at T2 and is clinically insignificant.

- The In Vitro study by Joooseong Kim et al (12) evaluated the accuracy of the numerical measurements from CBCT and Ortho Insight 3D Laser to that of the plaster model. For this purpose, the beam cone and plaster model scanning of 60 patients was prepared. The results showed an average difference of 0.23 0.169 mm between the laser and the plaster model and 0.57 0.338 mm between the beam cone and the laser. The average difference between the beam cone and the plaster model is similar to that between the laser and beam cone. The intra-examiner difference showed a mean difference in an intra-class correlation coefficient (ICC) of 0.05 0.021 mm for the Laser, 0.05 0.024 mm for the plaster model, and 0.04 0.035 mm. The measurement accuracy is more accurate with the laser than with the plaster model and CBCT.

- The meta-analysis by Jamille B Ferreira (13) compared the reliability and reproducibility of the measurements
from the CBCT digital models with that of the plaster model. The results are as follows:

- Clinically insignificant mean differences between the plaster model and the CBCT for mesiodistal and transverse tooth widths (0.036 to 0.204).
- Only the maxillary arch perimeter is considered clinically significant between the two measurement models with poor ICC.

The numerical models obtained from CBCT are accurate for the following dental measurements: crowding, intermolar and mesiodistal widths of incisors, canines, premolars, and molars for both the maxilla and mandible.

**DISCUSSION**

Replacing conventional impressions with digital impressions represents a paradigm shift in orthodontics. It is now possible to obtain a digital model by digitizing the patient's arches either directly in the mouth using intraoral cameras (1), or outside the mouth by digitizing conventional impressions or casts (14), and recently a digital model from cone-beam (15).

Most studies have focused on the accuracy of measurements between conventional "gold standard" models and digital models from intraoral impressions (1,6,7), extraoral impressions (5,8,15), or from beam cone (6,14).

A few studies have compared all the digital techniques with conventional ones (16).

These considerations led us to carry out this systematic review, setting the criteria for inclusion in vivo and in vitro studies, meta-analyses, randomized and non-randomized controlled trials, prospective and retrospective studies in French or English.

We selected five studies that met these inclusion criteria:

**Parameters Favorable to the Accuracy of the Digital Model**

- Severity of mandibular crowding:

  The key for digital models to replace plaster casts in the field of orthodontics should be tested on two issues: 1. Can the scanner perfectly convert a plaster cast into a digital model? 2. Can a clinician obtain the same information from a digital model as from a plaster cast? Both approaches could be affected by the degree of crowding in the dentition. Liang et al. (9) compared the influence of the severity of mandibular crowding on the accuracy of numerical models using 3Shape™. They found that when congestion is small (<3mm) or moderate (<8mm) there is no statistically significant difference between numerical and conventional measurements. Bernabe also noted a 0.39 mm/0.51 mm difference in non-crowded/crowded plaster cast measurements (17). However, Shellhart et al. pointed out that measurement discrepancies can vary by as much as 1.5 mm when a digital vernier caliper is used on a plaster cast with mild crowding (18).

- Digital measurement on a computer

  Tavares et al. (10) compared traditional plaster models, digitized models, and paper prints obtained through the virtual model. The results showed that there was no statistical difference between the measurements on the digital model and those on the plaster model (p>0.05) comparing arch length, intercanine, and intermolar width measurements between plaster and digital models. Different results were found by Santoro et al., (19) but it was highlighted by the authors that the differences were within a clinically acceptable range and, thus acceptable for orthodontic use. Keating et al found no statistically significant differences between the two methods (20).

- Time

  MG Cesur et al. (11) tested the dimensional stability of the alginate impression scan as a function of time. Plaster model, negative and positive digital model measurements at T (0), T (1), and T (2) times showed statistically significant differences. However, these differences were no clinically significant because they were <0.5 mm. In this study, a master model is used to represent maxillary arch, the same size plastic impression trays, and one type of dental stone to reduce variables when the prints are stored in plastic bags in a dark room at room temperature. The effects of the time-dependent deformation of alginites on digital model accuracy were evaluated throughout the measurements on the plaster model and negative and
positive digital models. Most of their results showed significant differences among them on the 1st and 2nd day. However, these differences were small and did not exceed 0.4 mm, which can be accepted within clinical tolerance. Alginate impression shrinks because of different pouring times, which is most likely the explanation for the differences. Coleman et al [21] reported that significant dimensional changes between plaster models poured within 1 h of the alginate impression compared with pouring 24 h later. This would be translated into digital models. Alcan et al [22] reported statistically significant changes after storing alginate impressions for up to 4 days, although no clinical relevance was noted.

Table 2. Data extracted from the included studies.

<table>
<thead>
<tr>
<th>Authors - Year</th>
<th>Type of Study and Method</th>
<th>Sample Size</th>
<th>Assessment Methods</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liang et al 2018</td>
<td>In vitro</td>
<td>29 sets of models</td>
<td>-Reliability between plaster casts and digital models: ICC</td>
<td>-MD Mand RS (mm): 0.28 Gp1 vs 0.27Gp2 vs 0.21Gp3 p=0.012*</td>
</tr>
<tr>
<td></td>
<td>Plaster model sets were digitized using a D800 Scanner (3shape™, Copenhagen, Denmark)</td>
<td>Group 1 (n=10): mild crowding dentition</td>
<td>-Accuracy between two model systems: paired samples t-test (*p&lt;0.016 Mann-Whitney U test with Bonferroni correction)</td>
<td>-MD Mand RS Gp1 vs Gp2: p=0.315</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 2 (n=10): Moderate crowding dentition</td>
<td>-Plaster casts measurements: digital caliper</td>
<td>-MD Mand RS Gp2 vs Gp3: p=0.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 3 (n=9): Severe crowding dentition</td>
<td>-Digital models measurements: digital caliper</td>
<td>-MD Mand RS Gp1 vs Gp3: p=0.008*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-Digital models measurements: measuring software (OrthoAnalyzer)</td>
<td>-Other measurements: NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tavares et al 2017</td>
<td>Prospective randomized</td>
<td>37 lower arch models</td>
<td>-Plaster casts measurements: digital caliper</td>
<td>-Intercanine width (mm): 26.14Gp1 vs 26.11Gp2 vs 25.73Gp3 p&lt;0.001*</td>
</tr>
<tr>
<td></td>
<td>Lower arch plaster models were compared using Ortho Insight 3D scanner and paper print image of the models</td>
<td>Group 1 (n=37): Plaster model</td>
<td>-Digital models measurements: measuring software OrthoInsight 3D scanner, v. 5.0</td>
<td>-Intermolar width (mm): 44.66 Gp1 vs 44.74Gp2 vs 44.16Gp3 p&lt;0.001*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 2 (n=37): Digital model</td>
<td>-Paper print image of the models measurements: digital caliper</td>
<td>-Arch length (mm): 33.94Gp1 vs 33.99Gp2 vs 34.42Gp3 p&lt;0.001*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 3 (n=37): Print image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cesur et al 2017</td>
<td>Prospective</td>
<td>60 impressions per group: Plaster casts (n=20), Negative digital models (n=20), Positive digital models (n=20)</td>
<td>-Plaster casts measurements: digital caliper</td>
<td>Group 1 T0 (time):</td>
</tr>
<tr>
<td></td>
<td>Plaster model sets were digitized into negative and positive digital models immediately (T0), 1 day (T1), and 2 days (T2)</td>
<td>Group 1 T0</td>
<td>-Negative digital models measurements: Maestro 3D Ortho Studio Software</td>
<td>RC-RM, LC-LM*, RC-LC*, RM-LM*, RC-S*, LC-S*, P*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 2 T0</td>
<td>-Positive digital models measurements: Maestro 3D Ortho Studio Software</td>
<td>-Group 2 T1 (time):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Group 3 T2 (time):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean Differences (in mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Maxiordinal width: Max.premolar 0.0758±0.2593*, 0.2674*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mand.Premolar 0.1410±0.2977*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.3196*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mandibular arch width:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mand.Premolar 0.2305±0.5727, 0.5921±0.5727</td>
</tr>
<tr>
<td>Kim et al 2014</td>
<td>In Vitro</td>
<td>Maxillary and mandibular plaster models and CBCT scans from 60 patients</td>
<td>-Plaster casts measurements: digital caliper</td>
<td>Mean Differences (in mm)</td>
</tr>
<tr>
<td></td>
<td>Compare the accuracy of measurements obtained from the 3D laser scans to those taken from the CBCT scans and those obtained from the plaster models</td>
<td>Group 1: plaster models</td>
<td>-Laser-scanned models measurements: Maestro 3D Ortho Studio Software</td>
<td>-Mesio-distal width: Max.premolar, 0.2593*, 0.2674*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 2: plaster models scanned by laser</td>
<td>-CBCT models measurements: measurement tool of the Software (Anatomage)</td>
<td>Mand.Premolar 0.1410±0.2977*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 3: CBCT models</td>
<td></td>
<td>0.3196*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mandibular arch width:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mand.Premolar 0.2305±0.5727, 0.5921±0.5727</td>
</tr>
<tr>
<td>Ferreira et al 2017</td>
<td>Meta-analysis</td>
<td>130 sets of models</td>
<td>-Plaster casts measurements: digital caliper</td>
<td>-Arch length and crowding</td>
</tr>
<tr>
<td></td>
<td>Assess the accuracy and reproducibility of dental measurements obtained from digital study casts generated from CBCT compared with those acquired from plaster models</td>
<td>Group 1: plaster models</td>
<td>-CBCT models measurements: the measurement tool of the same software (Anatomage),</td>
<td>Overestimating crowding measurement on digital models vs plaster models from 1.06 to 1.75 mm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group 3: CBCT models</td>
<td></td>
<td>-Tooth width lower than in plaster models from 0.01 to 0.47 mm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Intraexaminer reproducibility Good reproducibility ICC=0.75</td>
</tr>
</tbody>
</table>

RC: Intra-class Correlation Coefficient, Gp, Group, VS, versus, MD. Mean difference, mand RS. Mandibular required space, NS nonsignificant, RC-RM=Right canine cusp tip and the molar mesiobuccal cusp tip, *=Plaster model group was statistically different from other groups, LC-LM=Left canine cusp tip and the molar mesiobuccal cusp tip, **=Negative digital model and positive digital model groups were statistically different from each other, RC-LC=Right and left canine cusp tips. b =Negative digital model group was statistically different from other groups. a=Positive digital model group was statistically different from other groups, RM-LM=Right and left molar mesiobuccal cusp tips, 12=Laser-Plaster, 1c=Laser-CBCT, 1c=Plaster-CBCT
Parameters Unfavorable to the Accuracy of the Digital Model

- Calculation of the space required and the severity of the space requirement

Numerical calculations of space and space requirements resulted in statistically significant differences in the study by Yu-Ming Liang et al (9). However, this difference is less than 0.5 mm and therefore clinically acceptable. Also, the more severe the space requirement, the more underestimated measurements are found compared to the plaster model. The results confirm that the standard deviation in the severe crowding group was larger than that in the mild crowding group, but the difference in measurement was lesser than that in the mild crowding group. The larger standard deviation may be due to the measuring error. While measuring using a digital model, the examiner needs to rotate through a larger angle over the 2D screen for heavily tilted teeth. A different angle of view on the 2D screen may result in different measurements. As for the plaster cast, severe crowding leads to a more physical barrier for the plaster cast and cause more difficulty while placing the caliper tip at the correct point on the plaster cast.

Also, the notable data were that the mean values of space required of the 3Shape™ digital model system were smaller than those of plaster cast, which was consistent with some studies (23,24). This finding can be attributed to the physical barrier of the plaster cast; for example, the caliper tip could be hampered by neighboring teeth during the measuring process, which may jeopardize the precision of results.

- Digital measurement on paper.

Tavares et al (10) found paper print images showed sub estimated values for intercanine and intermolar widths and overestimated values for dental arch length. The differences were found to be statistically significant (p<0.001). However, the comparison between the digital model and the paper print obtained from it, showed that for intercanine and intermolar widths the mean differences were 0.38mm and 0.58mm, respectively. Regarding arch length, the mean difference observed was 0.52mm. Previously published research (19) so as the authors of the present study, considered such differences as clinically negligible. It is thus suggested that the presented method is accurate for clinical use without bringing any potential distortions for the fabrication of orthodontic archwires or arch shape observation.

- Time

The measurement of the arch perimeter alone at T2 (2 days) showed a statistically significant difference between the numerical model from the Laser and that of the plaster but which is clinically negligible according to the Cesur et al study (11).

- Use of Cone-beam

The numerical models from Cone Beam gave sagittal, cross-sectional, and vertical measurements statistically different from those of the conventional in the in vitro study of Kim et al (12). Ferreira et al (13) found the accuracy of dental linear measurements obtained from digital models generated from the CBCT scan compared with those obtained from the plaster models showed clinically insignificant mean differences for mesiodistal tooth width, and transversal widths (20.036 to 0.204) and for sums of measurements (20.097 to 21.654), according to the values assigned in their study. The only maxillary perimeter was judged to be clinically significant when compared between digital and plaster. When plaster models and digital models generated from the CBCT scan were compared, the values were overestimated in the mandibular arch while underestimated for mesiodistal width measurements of teeth in the maxillary arch. That may have happened because of the difference in the size of the maxillary and mandibular teeth, the low accuracy of interproximal surfaces, random errors on landmark positions, and the methodological acquisition of the segmentation process. Added to this, the image quality might be affected by scanning artifacts, patient’s movement while performing CBCT scan, and by poor reconstruction of occlusal surface in some cases, (25) and these problems might affect dental measurements. When poor image quality of the occlusal surface is obtained, the arch width measurement would probably be questionable.

CONCLUSION

The results of our systematic review have shown that there is no clinically significant difference between the numerical and physical models for the majority of diagnostic parameters. On the other hand, there are indeed some limitations, relative to the types of the severity of the congestion, the elapsed time to digitize, and the numerical means.

ACKNOWLEDGMENTS

None.

AUTHORS’ CONTRIBUTIONS

The participation of each author corresponds to the criteria of authorship and contributorship emphasized in the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals of the International Committee of Medical Journal Editors. Indeed, all the authors have actively participated in the redaction, the revision of the manuscript, and provided approval for this final revised version.

COMPETING INTERESTS

The authors declare no competing interests with this case.

FUNDING SOURCES

None.
REFERENCES


