ORTHODONTICS & CRANIOFACIAL RESEARCH

REVIEW ARTICLE

DOI: 10.1111/j.1601-6343.2010.01503.x

PS Fleming
V Marinho
A Johal

Orthodontic measurements on digital study models compared with plaster models: a systematic review

Authors' affiliation:
P.S. Fleming, V. Marinho, A. Johal, Queen Mary University of London, Barts and The London School of Medicine and Dentistry, Institute of Dentistry, London, UK

Correspondence to:
Dr Padhraig Fleming
PhD research student
Queen Mary University of London
Barts and The London School of Medicine and Dentistry
Institute of Dentistry
London E1 2AD
UK
E-mail: padhraigfleming@hotmail.com

Abstract

Authors – Fleming PS, Marinho V, Johal A

The aim of this study is to evaluate the validity of the use of digital models to assess tooth size, arch length, irregularity index, arch width and crowding versus measurements generated on hand-held plaster models with digital callipers in patients with and without malocclusion. Studies comparing linear and angular measurements obtained on digital and standard plaster models were identified by searching multiple databases including MEDLINE, LILACS, BBO, ClinicalTrials.gov, the National Research Register and Pro-Quest Dissertation Abstracts and Thesis database, without restrictions relating to publication status or language of publication. Two authors were involved in study selection, quality assessment and the extraction of data. Items from the Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Reviews checklist were used to assess the methodological quality of included studies. No meta-analysis was conducted. Comparisons between measurements of digital and plaster models made directly within studies were reported, and the difference between the (repeated) measurement means for digital and plaster models were considered as estimates. Seventeen relevant studies were included. Where reported, overall, the absolute mean differences between direct and indirect measurements on plaster and digital models were minor and clinically insignificant. Orthodontic measurements with digital models were comparable to those derived from plaster models. The use of digital models as an alternative to conventional measurement on plaster models may be recommended, although the evidence identified in this review is of variable quality.

Key words: dental models; digital model; plaster model; reproducibility; validity

Dates:
Accepted 30 October 2010

To cite this article:
Fleming PS, Marinho V, Johal A:
Orthodontic measurements on digital study models compared with plaster models: a systematic review
Orthod Craniofac Res 2011;**14**:1–16

Introduction

Three-dimensional imaging and modelling have undergone significant advances in recent years, raising the possibility of the development of the 'virtual orthodontic patient', where bone, soft tissue and teeth can be recreated in three dimensions (1). The panacea of complete three-dimensional digital conversion has been prompted in particular by the advent of cone beam computerized tomography and the refinement of three-dimensional facial imaging. A further cog in this process is the advent of digital study model scanning (2).

Digital study models were introduced commercially in 1999 by OrthoCad™ (Cadent, Carlstadt, NJ, USA) and in 2001 (emodels™; GeoDigm, Chanhassen, MN, USA). The technology used to generate digital study models varies considerably. Emodels scans the surface of a complete plaster model, whereas OrthoCad uses 'destructive scanning' with multiple scans of a model in thin slices. Emodels has software to 'slice through' the image, whereas OrthoCad actually slices through the model and images it. Direct scanning of impressions to generate digital models is also possible (Digimodel™; Orthoproof, Albuquerque, NM, USA), obviating the requirement for plaster models.

Study models for orthodontic diagnosis and treatment planning have traditionally been held in the form of physical plaster models, which are subject to loss, fracture and degradation. Digital storage eliminates inherent problems related to physical storage of models with up to 17 m^3 of storage space required for storage of traditional models for one thousand patients (3). The replacement of plaster orthodontic models with virtual information has further potential benefits including:

(1) instant accessibility of 3D information without need for the retrieval of plaster models from a storage area;
(2) the ability to perform accurate and simple diagnostic set-ups of various extraction patterns;
(3) virtual images may be transferred anywhere in the world for instant referral or consultation; and
(4) objective model grading analysis, for example, for Peer Assessment Rating (PAR) or American Board of Orthodontics (ABO) scoring.

The potential advantages of digital models for the quantification of orthodontic problems would be negated if the validity, efficiency and ease of linear and angular measurement of occlusal features with digital models were not comparable to those related to plaster models, the current 'gold standard' used routinely in clinical practice. This review aims at assessing the validity (4) of digital models by assessing agreement with measurements on hand-held plaster models.

Materials and methods

To be included in the review, trials had to meet the following inclusion criteria:

- Study design: Primary diagnostic study reporting consecutive, randomly selected or non-randomly selected subjects.
- Population: Treated and untreated orthodontic patients with or without malocclusion. Restrictions were not applied owing to age, gender or setting, but alginate impressions were to be poured within 24 h.
- Index test: Measurements on digital models (any).
- Reference standard/comparator: Measurements on unmarked plaster models (with dial or digital callipers).
- Outcome measures of interest included the validity of recordings of tooth size; transverse dimensions; irregularity index; arch width; crowding; Bolton ratio; occlusal indices; and inter-arch occlusal features. Time taken to measure hand-held plaster and digital models was also assessed.

Search strategy for the identification of studies

Relevant literature was identified by searching the following electronic databases: MEDLINE via OVID (1950 to January 2010), LILACS and BBO (1982 to January 2010). Language restrictions were not applied. Unpublished literature was to be identified through searches of ClinicalTrials.gov (http://www.clinicaltrials.gov), the National Research Register (http://www.controlled-trials.com) and Pro-Quest Dissertation Abstracts and Thesis database (http://www.lib.umi.com./dissertations). Search strategies are described in Table 1 according to the sources searched. Conference proceedings and abstracts were also searched. Authors were to be contacted to identify unpublished or ongoing research and to clarify findings as required. Reference lists of the included studies were also screened for potentially relevant research.

Assessment of relevance, methodological quality and data extraction

Assessment of research for inclusion in the review, quality assessment and extraction of data were performed independently by two investigators (PSF and AJ). Disagreements were resolved by joint discussion, and a third investigator (VM) was consulted where necessary.

Table 1. **Database search and study selection**

Database	Keywords	Results	Full articles retrieved	Articles selected
MEDLINE via OVID (1950 to January 2010	((digital$ or virtual or electronic or computer$ or software) and (model$ or cast$)) or emodel or orthocad) and ((plaster$ or stone or gypsum) and (model$ or cast$)) and (dental or orthod$ or tooth))	248	24	14
LILACS (1982 to 2010)	((digital$ or virtual or electronic$ or comput$ or software) and (model$ or cast$)) or emodel or orthocad) and ((plaster$ or gesso$ or stone or gypsum) and (model$ or cast$)) and (dent$ or orthod$ or tooth))	44	1	1
BBO	As LILACS above	55	0	0
ClinicalTrials.gov	Orthodontic and digital and plaster model	0	0	0
National Research Register	Orthodontic and digital and plaster model	0	0	0
Pro-Quest Dissertation Abstracts and Thesis database	'Orthodontic*', 'model*' and 'digital*'	0	0	0

Potentially relevant abstracts were selected, and full-text articles were retrieved for further screening. Researchers were not blinded to the authors or the results of the research. Data extracted on the characteristics of included studies broadly covered the following aspects: setting; participants; study design; reference standard(s); index/comparator test(s); number of examiners; and number of times the test was performed. Methodological quality was assessed by critically examining the methodology of the investigations. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist was followed, although not all items were strictly applicable as this review was not directly addressing diagnostic test accuracy.

Data synthesis

Heterogeneity between studies was gauged by referring to assessment measurement protocol/measurement technique; number of operators; and the outcome measure reporting the comparisons between the index and reference tests. Results were tabulated according to outcomes showing the estimates of the various measurements. The differences between the means of measurements on plaster and digital models were extracted. The narrative focus was on reporting the pattern of results by outcomes across all the included studies. Inferential statistical methods were not used for the estimation of summary measures, testing of differences between models/tests and investigations into heterogeneity. No tests or investigations were undertaken to detect reporting biases.

Results

Description of included studies

Forty abstracts were considered potentially relevant. Following screening, 29 full-text articles were retrieved. Of these, 12 failed to meet the inclusion criteria. A hand search of the references in the 14 articles satisfying the inclusion criteria identified three additional articles. Therefore, 17 articles were included in the review (Fig. 1 5–21). Reasons for exclusion at the final selection stage are outlined in Appendix 1.

The characteristics of the individual studies are given in Table 2. All investigations were based in dental university settings, typically in the permanent dentition. Subjects in the majority of studies had malocclusion and had no history of orthodontic treatment. Gender and ethnicity were unspecified in all studies. Subjects were aged 12–18 in one study (20), but age was

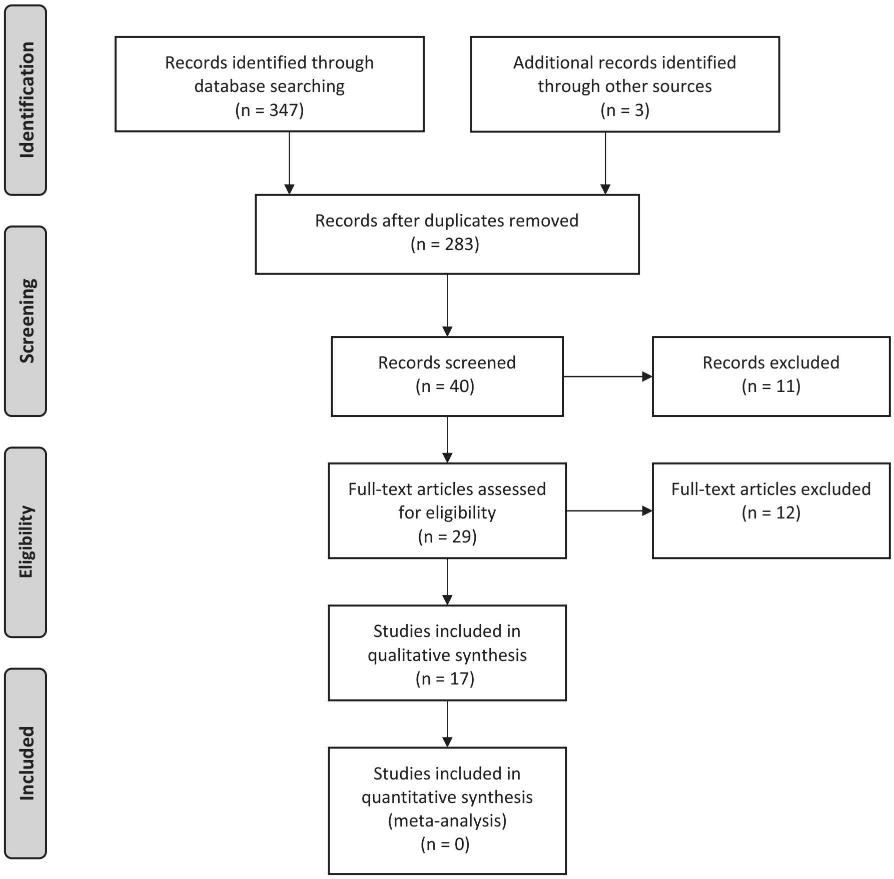

Fig. 1. Flow diagram of article retrieval.

unclear in the remainder. Clear information on study design was lacking in the majority of reports.

Seven digital model systems were assessed in these trials: OrthoCad; emodel; C3D-builder; ConoProbe; Easy3D Scan; Digimodels; and Cecile 3. Agreement between recordings on OrthoCad and plaster models was assessed in nine studies (5, 6, 8–10, 13, 15, 16, 19), between emodels and plaster models in three investigations (11, 12, 21) and using the other software systems in a single study each. Similar types of plaster models (index/comparator test) were used in each study. All digital recordings were compared to those derived from the direct measurement on plaster models using digital callipers. Either one or two (6, 8, 9, 11) sets of impressions were taken to produce digital and plaster models.

Significant variation was observed in the number of examiners carrying out the measurements and the number of times the readings were repeated. Ten examiners performed measurements in one trial (8). Measurements were taken three times by the researchers in four studies (5, 11, 14, 21) and eight times in one study (7).

Methodological quality of included studies

Where possible, the QUADAS tool (22) was adhered to. Therefore, methodological quality was assessed by critically examining the investigations in relation to the following: inclusion of a representative spectrum of patients (population recruitment and characteristics); use of appropriate reference standards; adequate description of index tests and reference standards; independent interpretation of the tests; independent interpretation of index and reference tests; and reporting of uninterpretable or intermediate data (Table 3).

Regarding the inclusion of a representative spectrum of patients, subjects were recruited either randomly or consecutively in most studies although the recruitment process and the characteristics of those recruited were not clearly outlined in seven studies (5, 7, 10, 12–14, 16,

Table 2. **Characteristics of included studies**

Study	Setting	Characteristics of participants	Study design	Index test/Reference standard	Examiners (readings per examiner)	Outcome measures
Tomassetti et al. (5)	University Orthodontic Department	22 subjects; USA, 11 pre- and 11 post-treatment; not more than 3 mm crowding.	Prospective	OrthoCad/Digital callipers	1 (3)	Bolton ratio; Time taken
Santoro et al. (6)	University Orthodontic Department	20 subjects; USA, permanent dentition; no missing teeth; stable occlusion with 3 occlusal contacts or more.	Prospective, enrolled randomly	OrthoCad/Digital callipers	2 (1)	Tooth size; Overjet; Overbite
Bell et al. (7)	University Orthodontic Department	22 subjects; UK	Prospective	C3D-builder (Uni. of Glasgow)/Digital callipers	1 (8)	Transverse and sagittal linear measurements
Quimby et al. (8)	University Orthodontic Department	50 subjects; USA, permanent dentition	Prospective, enrolled consecutively	OrthoCad/Digital callipers	10 (2)	Tooth size; Arch length; Transverse dimensions; Overjet; Overbite; Space available; Space required
Mayers et al. (9)	University Orthodontic Department	48 subjects; USA, permanent dentition	Prospective, enrolled consecutively	OrthoCad/Digital callipers	1 (2)	PAR score
Costalos et al. (10)	University Orthodontic Department	48 subjects; USA, permanent dentition; post-treatment; no edentulous space; no malocclusion.	Prospective	OrthoCad/Digital callipers	2 (1)	ABO score

Table 2. **Continued.**

Study	Setting	Characteristics of participants	Study design	Index test/Reference standard	Examiners (readings per examiner)	Outcome measures
Stevens et al. (11)	University Orthodontic Department	24 subjects; Canada, complete permanent dentition (from 1st molar to 1st molar) without previous orthodontics, pre-treatment models	Prospective, randomly selected from 225 records; three selected within each of 8 categories of malocclusion	Emodels/Digital callipers	3 (3 and 1)	PAR; Bolton ratio
Mullen et al. (12)	University Orthodontic Department	30 subjects; USA, Pre-treatment; complete permanent dentition (from 1st molar to 1st molar).	Prospective	Emodels/Digital callipers	1 (1)	Bolton ratio; Time taken
Okunami et al. (13)	University Orthodontic Department	30 subjects; USA, permanent dentition; post-treatment; no malocclusion.	Prospective	OrthoCad/Digital callipers	1 (1)	ABO score
Redlich et al. (14)	University Orthodontic Department	30 subjects; Israel, mixed and permanent dentition; 10 subjects each with mild, moderate and severe crowding.	Prospective	ConoProbe/Digital callipers	1 (3)	Tooth width; Arch l ength; Crowding
Hildebrand et al. (15)	University Orthodontic Department	36 subjects; USA, treated cases; consenting patients; no malocclusion.	Prospective, enrolled randomly	OrthoCad/Digital callipers	1 (1)	ABO score
Goonewardene et al. (16)	University Orthodontic Department	50 subjects; Australia, permanent dentition erupted including third molars.	Prospective	OrthoCad/Digital callipers	1 (1)	Tooth width; Arch length; Crowding Irregularity

***Table 2.* Continued.**

Study	Setting	Characteristics of participants	Study design	Index test/Reference standard	Examiners (readings per examiner)	Outcome measures
Keating et al. (17)	University Orthodontic Department	30 subjects; UK	Prospective, enrolled randomly	Easy3D Scan/Digital callipers	1 (2)	Linear dimensions (x, y, z planes)
Veenema et al. (18)	University Orthodontic Department	30 subjects; Netherlands, pre- and post-treatment; permanent dentition; 5 Class I, 19 Class II div 1, 5 Class II div 2, 1 Class III; 5 treated with extractions.	Prospective, enrolled randomly	Digimodel/Digital callipers	2 (1)	ICON score
Leifert et al. (19)	University Orthodontic Department	25 subjects; USA, Class I molar relationship, crowding.	Prospective, enrolled consecutively	OrthoCad/Digital callipers	2 (1)	Crowding
Watanabe-Kanno et al. (20)	University Orthodontic Department	15 subjects; Brazil, permanent dentition; pre-treatment; 12–18 years.	Prospective	Cecile3/Digital callipers	2 (1)	Transverse dimensions; Tooth size; Overjet; Overbite
Horton et al. (21)	University Orthodontic Department	32 subjects; USA, permanent dentition; pre-treatment.	Prospective	Emodels/Digital callipers	1 (3)	Tooth size; Time taken

ABO, American Board of Orthodontics; PAR, Peer Assessment Rating.

Table 3. **Methodological quality of included studies using items from QUADAS (22)**

Study	Representative spectrum of patients	Reference standard appropriate	Adequate description of index test	Reference standard independent of index test	Adequate description of reference test	Results of index/reference test interpreted independently	Uninterpretable intermediate results reported
Tomassetti et al. (5)	Unclear	Yes	Yes	No	Yes	Unclear	No
Santoro et al. (6)	Yes	Yes	Yes	Yes	Yes	Unclear	No
Bell et al. (7)	Unclear	Yes	Yes	No	Yes	Unclear	No
Quimby et al. (8)	Yes	Yes	Yes	Yes	Yes	Unclear	No
Mayers et al. (9)	Yes	Yes	Yes	Yes	Yes	Unclear	No
Costalos et al. (10)	Unclear	Yes	Yes	No	Yes	Unclear	No
Stevens et al. (11)	Yes	Yes	Yes	Yes	Yes	Unclear	No
Mullen et al. (12)	Unclear	Yes	Yes	No	Yes	Unclear	No
Okunami et al. (13)	Unclear	Yes	Yes	No	Yes	Unclear	No
Redlich et al. (14)	Unclear	Yes	Yes	No	Yes	Unclear	No
Hildebrand et al. (15)	Yes	Yes	Yes	No	Yes	Unclear	No
Goonewardene et al. (16)	Unclear	Yes	Yes	No	Yes	Unclear	No
Keating et al. (17)	Yes	Yes	Yes	No	Yes	Unclear	No
Veenema et al. (18)	Yes	Yes	Yes	No	Yes	Unclear	No
Leifert et al. (19)	Yes	Yes	Yes	No	Yes	Unclear	No
Watanabe-Kanno et al. (20)	Yes	Yes	Yes	No	Yes	Unclear	No
Horton et al. (21)	Unclear	Yes	Yes	No	Yes	Unclear	No

QUADAS, Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Reviews.

21). A clear definition of the criteria used for entry into the studies was also omitted from these studies. Measurements were taken on both the index test and an appropriate reference standard in all studies with those on the plaster models performed independently of the digital models in all studies. In 13 studies, the index test and reference standard were not independent, both being derived from the same impression; separate impressions were taken in the remaining four studies (6, 8, 9, 11).

Blinded interpretation of results was precluded by obvious differences in the performance of digital and manual measurements. All investigations were performed prospectively, with sample size estimation reported in just five studies (7, 8, 11, 16, 17).

Results by outcome measures

Outcomes assessed include the validity of analysis of transverse dimensions; other miscellaneous linear measurements; tooth size; Bolton ratio; arch length and crowding; irregularity index; inter-arch occlusal features; occlusal indices; and time taken to perform measurements using the two approaches. No studies investigating the validity of angular measurements on digital models were found. The results are presented in Tables 4 and 5.

Transverse dimensional measurements

The agreement between transverse dimensional readings obtained using digital and plaster models has been assessed in four studies (7, 8, 16, 20). Dimensions considered include mandibular and maxillary inter-canine, inter-premolar and inter-molar dimensions. Mean discrepancies between the approaches ranged from 0.04 to 0.4 mm[8]. Generally, these differences were small and unlikely to be of clinical significance.

Miscellaneous linear measurements

The reliability of non-specific measurements between various defined occlusal landmarks with both sagittal and transverse components was investigated by Bell

Table 4. **Summary of results of comparison between digital models and plaster models**

Study	N*	Measurement	Digital model Mean (SD)	Plaster model Mean (SD)	Mean Difference† (p value, SE or CI)	Average of absolute mean differences† (SD)
Transverse dimensions‡ (mm)						
Quimby et al. (8)	1000	Maxillary IMW	54.72 (0.85)	54.43 (0.26)	0.29 ($p < 0.05$)	
		Maxillary ICW	36.04 (0.51)	36.44 (0.26)	−0.4 ($p < 0.05$)	
		Mandibular IMW	47.42 (0.52)	47.38 (0.33)	0.04 ($p < 0.05$)	
		Mandibular ICW	26.31 (0.27)	26.65 (0.24)	−0.34 ($p < 0.05$)	
Keating et al. (17)	60	ICW/IPMW/IMW			$p = 0.765$	0.19 (0.12)
Watanabe-Kanno et al. (20)	30	Maxillary ICW	34.23 (1.78)	34.35 (1.78)	−0.12 ($p < 0.001$)	
		Maxillary IPMW	34.52 (2.01)	34.63 (2.02)	−0.11 ($p < 0.001$)	
		Maxillary IMW	44.83 (2.54)	44.99 (2.54)	−0.16 ($p < 0.001$)	
		Mandibular ICW	26.57 (1.57)	26.71 (1.58)	−0.14 ($p < 0.001$)	
		Mandibular IPMW	28.73 (1.86)	28.86 (1.85)	−0.13 ($p < 0.001$)	
		Mandibular IMW	39.66 (2.25)	39.78 (2.25)	−0.12 ($p < 0.001$)	
Miscellaneous linear measurements (mm)						
Bell et al. (7)	176	Various transverse and sagittal measurements			$p > 0.05$	0.27 (0.06)
Keating et al. (17)	60	Y plane: Combined transverse and sagittal dimensions			$p = 0.501$	0.14 (0.09)
		Overall			$p = 0.237$	0.14 (0.1)
Tooth size (mm)						
Santoro et al. (6)	40	Overall mean			$p < 0.01$	−0.252
Redlich et al. (14)	90	Maxillary mean	7.73 (0.1§)	7.7 (0.12§)	0.03 ($p > 0.05$)	
		Mandibular mean	7.1 (0.1§)	7.11 (0.1§)	0.03 ($p > 0.05$)	
Goonewardene et al. (16)	50	Maxillary overall	76.1 (3.61)	74.8 (4)	1.3	
		Mandibular overall	66.3 (3.22)	65.7 (3.55)	0.6	
Watanabe-Kanno et al. (20)	30	21	8.76 (0.63)	8.94 (0.63)	−0.18 ($p = 0.6$)	
		26	9.9 (0.46)	10.1 (0.46)	−0.2 ($p = 0.00$)	
Horton et al. (21)	96	Overall difference			1.163 (0.115 per tooth)	
Keating et al. (17)	60	Crown height			0.03 ($p = 0.218$)	0.1 (0.07)
Bolton ratio (mm)						
Tomassetti et al. (5)	66	Anterior			1.02 ($p = 0.243$)	0.60 (0.38)
		Overall			1.2 ($p = 0.718$)	0.92 (0.58)
Stevens et al. (11)	360	Anterior	−0.55 (2.00)	−0.51 (1.80)	−0.04 ($p = 0.790$)	
		Overall	−0.75 (2.64)	−0.37 (2.20)	−0.38 ($p = 0.084$)	

Table 4. **Continued.**

Study	N*	Measurement	Digital model Mean (SD)	Plaster model Mean (SD)	Mean Difference† (p value, SE or CI)	Average of absolute mean differences† (SD)
Mullen et al. (12)	30	Overall			−0.05 (SE, 1.87; $p = 0.86$)	
Space analysis, arch length and tooth size–arch length discrepancy (crowding) (mm)						
Quimby et al. (8)	1000	Maxillary space available	74.87 (1.06)	73.58 (0.45)	0.29 ($p < 0.05$)	
		Maxillary space required	73.69 (0.93)	73 (0.37)	0.69 ($p < 0.05$)	
		Mandibular space available	65.71 (0.74)	64.02 (0.43)	1.69 ($p < 0.05$)	
		Mandibular space required	63.85 (0.86)	63.24 (0.49)	0.61 ($p < 0.05$)	
Stevens et al. (11)	360	Maxillary arch length	94.58 (5.25)	94.78 (5.33)	−0.20 ($p = 0.226$)	0.69 (0.43)
		Mandibular arch length	87.16 (5.44)	86.96 (5.17)	0.20 ($p = 0.256$)	0.65 (0.55)
Mullen et al. (12)	30	Maxillary arch length			1.47 (SE, 1.55; $p < 0.0001$)	
		Mandibular arch length			1.5 (SE, 1.36; $p < 0.0001$)	
Redlich et al. (14)	90	Maxillary arch length	73.45 (1.26)	73.64 (1.64)	−0.19 ($p > 0.05$)	
		Mandibular arch length	64.18 (1.29)	64.88 (1.22)	−0.7 ($p > 0.05$)	
		Maxillary crowding	1.41 (0.91)	1.77 (1.01)	−0.26 ($p > 0.05$)	
		Mandibular crowding	0.3 (0.92)	0.71 (0.92)	−0.41 ($p > 0.05$)	
Goonewardene et al. (16)	50	Maxillary arch length	75.8 (4.32)	74.8 (4.24)	1.0 ($p < 0.001$)	
		Mandibular arch length	65.9 (3)	65.1 (3.28)	0.8 ($p = 0.007$)	
		Maxillary crowding			−0.19 (SE = 0.219; $p = 0.38$)	
		Mandibular crowding			1.19 (SE = 0.23; $p < 0.000$)	
Leifert et al. (19)	50	Maxillary crowding	4.27 (2.41)	4.69 (2.46)	−0.424 (SE = 0.16; $p = 0.014$)	
		Mandibular crowding	3.69 (3)	3.9 (3.09)	−0.212 (SE = 0.23; $p = 0.364$)	
Irregularity index (mm)						
Stevens et al. (11)	360	Overall	23.7 (7.81)	20.99 (7.47)	2.71 ($p = .003$)	3.7 (3.05)

Table 4. **Continued.**

Study	N*	Measurement	Digital model Mean (SD)	Plaster model Mean (SD)	Mean Difference† (*p* value, SE or CI)	Average of absolute mean differences† (SD)
Goonewardene et al. (16)	50	Maxillary	7.8 (4.89)	7.8 (5.09)	0.0 ($p = 0.73$)	
		Mandibular	7.1 (3.07)	7.1 (3.19)	0.0 ($p = 0.13$)	
Inter-arch occlusal features (mm)						
Stevens et al. (11)	360	Centreline	1.23 (1.04)	1.32 (1.1)	−0.1 ($p = 0.30$)	0.34 (0.28)
		Posterior crossbite	0.75 (1.86)	0.74 (1.84)	0.01 ($p = 0.747$)	0.04 (0.12)
		Anterior crossbite	0.63 (0.98)	0.67 (1.09)	−0.03 ($p = 0.59$)	0.15 (0.26)
Santoro et al. (6)	40	Overjet			$p = 0.9771$	−0.00987
Quimby et al. (8)	1000		1.41 (0.4)	1.4 (0.21)	0.01 ($p > 0.05$)	
Stevens et al. (11)	360		4.91 (2.98)	4.9 (2.97)	0.01 ($p = 0.884$)	0.33 (0.21)
Watanabe-Kanno et al. (20)	30		5.22 (2.24)	5.43 (2.24)	−0.21 ($p = 0.00$)	
Santoro et al. (6)	40	Overbite			$p = 0.0124$	−0.4901
Quimby et al. (8)	1000		1.45 (0.53)	1.48 (0.3)	−0.03	
Stevens et al. (11)	360		3.67 (1.82)	3.96 (1.75)	−0.3 ($p = 0.01$)	0.38 (0.27)
				3.51 (1.33)		
Watanabe-Kanno et al. (20)	30		3.2 (1.32)		−0.31 ($p = 0.00$)	
Occlusal indices						
Veenema et al. (18)	60	Total ICON score (Examiner 1)	10.97 (2.47)	11.47 (2.37)	−0.5	
			4.13 (1.31)	3.4 (1.07)	0.73 ($p < 0.01$)	
Mayers et al. (9)	96	Overall PAR score	27.25 (11.49)	27.35 (12.75)	−0.1 (ICC = 0.96–0.98)	
Stevens et al. (11)	360		25.91 (8.79)	25.08 (9.3)	0.83 ($p = 0.128$)	2.11 (1.62)
Time taken (min)						
Tomassetti et al. (5)	66	Bolton analysis	5.37 (0.87)	8.06 (0.54)	−2.69	
Mullen et al. (12)	30	Bolton analysis			$p < 0.001$	1.09 (47)
Horton et al. (21)	96	Occlusal view technique			−2.02	

*Number of determinations.
†Negative values represent smaller values on digital models.
‡ICW, Inter-canine width; IPMW, Inter-premolar width; IMW, Inter-molar width.
§SE. PAR, Peer Assessment Rating.

et al. (7) and Keating et al. (17). These studies described similar levels of consistency with mean discrepancies of 0.14 and 0.27 mm reported, respectively. Consequently, combinations of antero-posterior and transverse measurements appear to have similar reliability as purely transverse or sagittal measurements.

Tooth size

Differences in individual tooth size with digital and direct methods have been measured in the mesio-distal and vertical dimension. Tooth size has also been used indirectly to calculate Bolton tooth size ratios, arch length and tooth size–arch length discrepancy. Gener-

Table 5. **Summary of American Board of Orthodontics scoring**

Measurement technique/difference	Study						
	Costalos et al. (10) (n = 24)			Okunami et al. (13) (n = 30)		Hildebrand et al. (15) (n = 36)	
	Digital Mean (SD)	Plaster Mean (SD)	*p*	Mean diff.	*p*	Mean diff. (SD)	*p*
Alignment	5.42 (3.11)	7.75 (3.89)	<0.0001	0.23	0.34	0.61 (0.8)	<0.01
Marginal ridges	3.67 (2.48)	4 (2.6)	0.4694	0.03	0.837	0.28 (0.57)	0.11
Inclination	5.67 (1.81)	6.71 (3.06)	0.0507	n/a	n/a	0.28 (0.51)	0.571
Occlusal contacts	6.54 (4.24)	5.33 (5.31)	0.2169	−4.53	0.000	1.89 (2.48)	0.021
Occlusal relationships	1.83 (1.97)	2.17 (2.63)	0.3567	−0.5	0.023	0.11 (0.4)	0.422
Overjet	6.25 (3.42)	4.67 (2.75)	0.1077	−0.37	0.1	3.94 (2.65)	<0.001
Interproximal contacts	0.29 (0.62)	0.75 (1.22)	0.0613	−0.13	0.102	0.03 (0.17)	0.324
Overall	29.67 (9.29)	31.17 (10.47)	0.3467	−5.07	0.000	9 (5.54)	<0.01

*Negative values represent smaller values on digital models.

ally, minor mean differences in mesio-distal tooth dimension of 0.01–0.3 mm were reported overall (6, 14, 16, 20, 21).

Measurement of vertical crown height is likely to be imprecise with identification of a cervical point particularly unreliable. Keating et al. (17) assessed vertical crown heights of premolars and molars using the maximum point of concavity on the labial surface gingival margin as the cervical reference point; a difference in the measurement of canine and molar heights of 0.1 mm was detected.

Bolton ratio

Comparison of Bolton tooth size analyses has been performed on digital and plaster models (5, 11, 12). Acceptable agreement between the two methods was demonstrated in all three studies. Stevens et al. (11) described an anterior discrepancy of 0.6 mm; however, Mullen et al. (12) reported an overall mean difference of just 0.05 mm. Stevens et al. (11) found an overall discrepancy of 0.38 mm using emodels; Tomassetti et al. (5) found a more significant difference of 1.02–1.2 mm between direct measurement on plaster models and digital measurement using OrthoCad.

Space analysis, arch length and tooth size–arch length discrepancy (crowding)

Overall, arch length, crowding and space analysis were measured in five studies (8, 11, 14 16, 19). With respect to arch length, discrepancies between the techniques ranged from 0.19 (14) to 0.8 mm (16). The difference between the measurement of crowding obtained with the techniques varied from 0.19 mm (16) to 0.42 mm (19); however, the mean degree of crowding in each trial did not exceed 4.69 mm (19), with the arches being spaced in one of the studies (16).

Irregularity index

The irregularity index in both the maxillary and mandibular arches was measured by Goonewardene et al. (16). Identical mean levels of irregularity were calculated with both techniques using OrthoCad digital models. However, using emodels, Stevens et al. (11) reported a significant discrepancy with the digital software underestimating irregularity by 3.7 mm.

Inter-arch occlusal features

Agreement between measurement of overjet and overbite has been considered in four studies (6, 8, 11, 20). Quimby et al. (8) and Watanabe-Kanno et al. (20) reported near-perfect agreement for both parameters; similarly, Santoro et al. (6) and Stevens et al. (11) showed excellent agreement for overjet measurement. The concordance of measurement of posterior crossbite and centreline discrepancy was confirmed by Stevens et al. (11). Inter-arch features including buccal segment interdigitation, overbite and overjet are also considered as part of occlusal indices including PAR, ICON and ABO scoring.

Occlusal indices

Acceptable concordance with digital and plaster models in relation to the severity of malocclusion using PAR, ICON and ABO scores has been demonstrated. The agreement between manual and digital measurements was high with respect to both PAR (9, 11) and ICON (18). In relation to the ABO score, three studies (10, 13, 15) reported comparisons between the techniques. In general, the differences between the measurement methods are low; however, Okunami et al. (13) and Costalos et al. (10) reported a significant discrepancy with respect to occlusal contact and buccolingual inclination scores. Furthermore, Costalos et al. (10) reported a significant difference in arch irregularity. These discrepancies were attributed to limitations pertaining to one software program (OrthoCad™); the ABO method of measuring inclination is also difficult to apply to digital models.

Time taken

The difference in the time required to perform a variety of occlusal measurements has been assessed in three disparate studies (5, 12, 21). These studies suggest a significant time saving with digital techniques although a significant learning curve and period of adjustment are likely to be required. Relatively minor differences were described by Horton et al. (21) (2 min) and Mullen et al. (12) (1 min). The approach to digital measurement is also believed to have an impact, with manipulation of the model being necessary to perform specific measurements. Differences may also arise in view of software and familiarity with the technique; Mullen et al. (12) used the widely available emodels™. Horton et al. (21) measured time taken to calculate tooth dimensions in isolation, and Mullen et al. (12) calculated Bolton tooth size ratios.

Discussion

Earlier research has confirmed that digital software is capable of faithfully reproducing dental features with a high degree of accuracy (2, 23–31). This research was omitted from this review as our main focus was to ascertain whether digital models offer a valid and clinically useful alternative to plaster models.

The application of digital models in orthodontic practices has increased steadily with 18% of surveyed practitioners reporting usage in a recent survey in the United States (32). This development has been prompted by a range of perceived advantages including reduced storage requirements; rapid access to digital information; easy transfer of data; versatility; and financial savings. This systematic review confirms that these proven advantages occur without significant compromise to the reliability of occlusal information.

To analyse the validity of digital models, plaster models were chosen as a reference standard in this review as direct measurement is performed on plaster models with rulers or callipers routinely in orthodontic offices and for research purposes. However, direct measurement on plaster models is inevitably associated with some degree of inaccuracy. To produce a more accurate 'gold standard', researchers have developed artificial models permitting more accurate measurement (8, 12) or have compared measurements between artificial structures of known dimension (33). Generally, digital models have shown a high degree of accuracy using these techniques (12). Much of the error of the measurement technique is likely to reside in point identification rather than being a function of the measuring device or software. Therefore, with enhancement of direct digital superimposition techniques and digital point recognition, digital modelling may replace plaster models as the 'gold standard'.

Evidence for the validity of digital models as an alternative to plaster models is accumulating. However, the methodological quality of studies included in this review was variable. In particular, description of the sample population was inadequate. Furthermore, separate impressions were used to fabricated digital and plaster models in four of the included studies. Differences in the impressions and casting processes may therefore have contributed to some of the inconsistency reported in these trials (6, 8, 9, 11). Complete data on the absolute differences between the techniques including confidence intervals and standard errors were also rarely reported. Further studies in this area should refer to QUADAS guidelines (22) and would benefit from clear reporting of the patient sample on which the models are based and independent interpretation of results.

This systematic review involved assessment of publications from English-language and non-English-language databases. Unpublished data were also searched. Consequently, it was felt that most data have been accessed. Where possible, complete results were obtained from these studies. Studies were excluded if there was a time lag between taking the impressions and pouring study models, where artificial occlusal setups were used and when models were marked before measurement (Appendix 1). However, although not considered formally in this review, the results of these studies appeared to be in general agreement with those of the included research studies.

Overall, the mean discrepancy between measurement based on digital and plaster models was low. The differences were considered in all studies to be clinically insignificant. This finding has been corroborated by studies demonstrating excellent concordance of treatment-planning decisions based on digital and plaster models (34, 35). Replacement of plaster with digital models resulted in diagnostic changes in 13%, translating into alteration of the treatment plan in just 6% of cases (34). This discrepancy is in keeping with research highlighting inconsistency in orthodontic planning decisions by the same and different clinicians, irrespective of differences in records available (36–38).

A further potential advantage of digital models lies in the ability to measure tooth position in three dimensions. In particular, measurement of inclination of individual teeth on plaster models is unreliable and cumbersome. However, digital models may be manipulated and sectioned to analyse specific teeth and permit estimation of long axis position. Furthermore, three-dimensional mapping of tooth movement may be possible by superimposing dental changes on stable reference structures with use of non-destructive digital manipulation and sectioning techniques.

Conclusions

Digital models offer a high degree of validity when compared to direct measurement on plaster models; differences between the approaches are likely to be clinically acceptable.

Clinical relevance

Digital models are gaining increasing acceptance as an alternative to traditional plaster models in orthodontics. The potential advantages of digital models would be negated if the validity, efficiency and ease of linear and angular measurement with digital models were not comparable to those related to plaster models, the current 'gold standard' used routinely in clinical practice. This review confirms that digital models offer a valid alternative to plaster models, although the available evidence is of variable quality.

References

1. Hajeer MY, Millett DT, Ayoub AF, Siebert JP. Applications of 3D imaging in orthodontics: part II. *J Orthod* 2004;31:154–62.
2. Hirogaki Y, Sohmura T, Satoh H, Takahashi J, Takada K. Complete 3-D reconstruction of dental cast shape using perceptual grouping. *IEEE Trans Med Imaging* 2001;20:1093–101.
3. McGuinness NJ, Stephens CD. Storage of orthodontic study models in hospital units in the U.K. *Br J Orthod* 1992;19:227–32.
4. Roberts CT, Richmond S. The design and analysis of reliability studies for the use of epidemiological and audit indices in orthodontics. *Br J Orthod* 1997;24:139–47.
5. Tomassetti JJ, Taloumis LJ, Denny JM, Fischer JR Jr. A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. *Angle Orthod* 2001;71:351–7.
6. Santoro M, Galkin S, Teredesai M, Nicolay OF, Cangialosi TJ. Comparison of measurements made on digital and plaster models. *Am J Orthod Dentofacial Orthop* 2003;124:101–5.
7. Bell A, Ayoub AF, Siebert P. Assessment of the accuracy of a three-dimensional imaging system for archiving dental study models. *J Orthod* 2003;30:219–23.
8. Quimby ML, Vig KW, Rashid RG, Firestone AR. The accuracy and reliability of measurements made on computer-based digital models. *Angle Orthod* 2004;74:298–303.
9. Mayers M, Firestone AR, Rashid R, Vig KW. Comparison of peer assessment rating (PAR) index scores of plaster and computer-based digital models. *Am J Orthod Dentofacial Orthop* 2005;128:431–4.
10. Costalos PA, Sarraf K, Cangialosi TJ, Efstratiadis S. Evaluation of the accuracy of digital model analysis for the American Board of Orthodontics objective grading system for dental casts. *Am J Orthod Dentofacial Orthop* 2005;128:624–9.
11. Stevens DR, Flores-Mir C, Nebbe B, Raboud DW, Heo G, Major PW. Validity, reliability, and reproducibility of plaster vs digital study models: comparison of

peer assessment rating and Bolton analysis and their constituent measurements. *Am J Orthod Dentofacial Orthop* 2006;129:794–803.
12. Mullen SR, Martin CA, Ngan P, Gladwin M. Accuracy of space analysis with emodels and plaster models. *Am J Orthod Dentofacial Orthop* 2007;132:346–52.
13. Okunami TR, Kusnoto B, BeGole E, Evans CA, Sadowsky C, Fadavi S. Assessing the American Board of Orthodontics objective grading system: digital vs plaster dental casts. *Am J Orthod Dentofacial Orthop* 2007;131:51–6.
14. Redlich M, Weinstock T, Abed Y, Schneor R, Holdstein Y, Fischer A. A new system for scanning, measuring and analyzing dental casts based on a 3D holographic sensor. *Orthod Craniofac Res* 2008;11:90–5.
15. Hildebrand JC, Palomo JM, Palomo L, Sivik M, Hans M. Evaluation of a software program for applying the American Board of Orthodontics objective grading system to digital casts. *Am J Orthod Dentofacial Orthop* 2008;133:283–9.
16. Goonewardene RW, Goonewardene MS, Razza JM, Murray K. Accuracy and validity of space analysis and irregularity index measurements using digital models. *Aust Orthod J* 2008;24:83–90.
17. Keating AP, Knox J, Bibb R, Zhurov AI. A comparison of plaster, digital and reconstructed study model accuracy. *J Orthod* 2008;35:191–201.
18. Veenema AC, Katsaros C, Boxum SC, Bronkhorst EM, Kuijpers-Jagtman AM. Index of complexity, outcome and need scored on plaster and digital models. *Eur J Orthod* 2009;31:281–6.
19. Leifert MF, Leifert MM, Efstratiadis SS, Cangialosi TJ. Comparison of space analysis evaluations with digital models and plaster dental casts. *Am J Orthod Dentofacial Orthop* 2009;136:16.e1–4.
20. Watanebe-Kanno GA, Abrão J, Miasiro J, Hiroshi Sánchez-Ayala A, Lagravère MO. Reproducibility, reliability and validity of measurements obtained from Cecile3 digital models. *Braz. Oral Res* 2009;23:288–95.
21. Horton HM, Miller JR, Gaillard PR, Larson BE. Technique comparison for efficient orthodontic tooth measurements using digital models. *Angle Orthod* 2010;80:254–61.
22. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
23. Ikeuchi H. A fundamental study on measurement of study cast. *Kokubyo Gakkai Zasshi* 1996;63:620–8.
24. Goshtasby AA, Nambala S, deRijk WG, Campbell SD. A system for digital reconstruction of gypsum dental casts. *IEEE Trans Med Imaging* 1997;16:664–74.
25. Motohashi N, Kuroda T. A 3D computer-aided design system applied to diagnosis and treatment planning in orthodontics and orthognathic surgery. *Eur J Orthod* 1999;21:263–74.
26. Lu P, Li Z, Wang Y. A study of dental cast by using 3D laser non-contact measurement and analysis. *Zhonghua Kou Qiang Yi Xue Za Zhi* 1999;34:351–3.
27. Kojima T, Sohmura T, Wakabayashi K, Nagao M, Nakamura T, Takashima F et al. Development of a new high-speed measuring system to analyze the dental cast form. *Dent Mater J* 1999;18:354–65.
28. Lu P, Li Z, Wang Y, Chen J, Zhao J. The research and development of noncontact 3-D laser dental model measuring and analyzing system. *Chin J Dent Res* 2000;3:7–14.
29. Commer P, Bourauel C, Maier K, Jäger A. Construction and testing of a computer-based intraoral laser scanner for determining tooth positions. *Med Eng Phys* 2000;22:625–35.
30. Sohmura T, Kojima T, Wakabayashi K, Takahashi J. Use of an ultrahigh-speed laser scanner for constructing three-dimensional shapes of dentition and occlusion. *J Prosthet Dent* 2000;84:345–52.
31. Kusnoto B, Evans CA. Reliability of a 3D surface laser scanner for orthodontic applications. *Am J Orthod Dentofacial Orthop* 2002;122:342–8.
32. Keim RG, Gottlieb EL, Nelson AH, Vogels DS III. 2008 JCO study of orthodontic diagnosis and treatment procedures, part 1: results and trends. *J Clin Orthod* 2008;42:625–40.
33. Brusco N, Andreetto M, Lucchese L, Carmignato S, Cortelazzo GM. Metrological validation for 3D modeling of dental plaster casts. *Med Eng Phys* 2007;29:954–66.
34. Rheude B, Sadowsky PL, Ferriera A, Jacobson A. An evaluation of the use of digital study models in orthodontic diagnosis and treatment planning. *Angle Orthod* 2005;75:300–4.
35. Whetten JL, Williamson PC, Heo G, Varnhagen C, Major PW. Variations in orthodontic treatment planning decisions of Class II patients between virtual 3-dimensional models and traditional plaster study models. *Am J Orthod Dentofacial Orthop* 2006;130:485–91.
36. Lee R, MacFarlane T, O'Brien K. Consistency of orthodontic treatment planning decisions. *Clin Orthod Res* 1999;2:79–84.
37. Baumrind S, Korn EL, Boyd RL, Maxwell R. The decision to extract: part 1–Interclinician agreement. *Am J Orthod Dentofacial Orthop* 1996;109:297–309.
38. Ribarevski R, Vig P, Vig KD, Weyant R, O'Brien K. Consistency of orthodontic extraction decisions. *Eur J Orthod* 1996;18:77–80.

***Appendix 1:* Articles excluded with reasons for omission**

Publication	Reason excluded
Sander & Tochtermann, 1991	Description of a hologram technique; not compared to digital calliper method
Miras & Sander, 1993	Comparison of hologram technique to sliding calliper
Ikeuchi, 1996	Non-dental measurement of spherical objects
Schirmer & Wiltshire, 1997	Comparison of measurement on photocopies of models and Vernier calliper
Ho & Freer, 1999	Described development of computer program to calculate tooth size and Bolton discrepancy
Commer et al., 2000	Co-ordinate measurement table used as reference
Zilberman et al., 2003	Used artificial occlusal set-up
Asquith et al., 2007	Points were marked on plaster models
Gracco et al., 2007	Used artificial occlusal set-up
Syrynska, 2008	No quantitative measurements given
Alcan et al., 2009	Used artificial occlusal set-up. Time lag in excess of 24 h before impressions poured
Dalstra and Melsen, 2009	Time lag in excess of 24 h before impressions poured