European Journal of Orthodontics 32 (2010) 589–595
doi:10.1093/ejo/cjp159
Advance Access Publication 17 February 2010

Virtual model analysis as an alternative approach to plaster model analysis: reliability and validity

K. Bootvong, Z. Liu, C. McGrath, U. Hägg, Ricky W. K. Wong, M. Bendeus and S. Yeung
Faculty of Dentistry, The University of Hong Kong, SAR, China

Correspondence to: Dr Ricky W. K. Wong, Department of Orthodontics, The University of Hong Kong, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong SAR, China. E-mail: fyoung@hkucc.hku.hk

SUMMARY The objective of this study was to assess the feasibility of virtual models as an alternative to orthodontic plaster models. Virtual dental models (obtained from OrthoCAD®) and corresponding plaster models of 80 patients in the permanent dentition were randomly selected from patients seeking orthodontic care. Inter-examiner error was assessed by measuring tooth width, overjet, overbite, intermolar width, intercanine width, and midline discrepancy. Criterion validity of virtual model analysis was determined by the agreement between the measurements from virtual and plaster models. Test–retest reliability was determined by remeasuring 10 virtual models 1 week later. Comparison analysis was assessed by calculating the mean directional differences and standardized directional differences. Correlation analysis was determined by calculating the intraclass correlation coefficients (ICCs).

Both intra- and inter-examiner reliability and test–retest reliability of virtual model analysis were acceptable in measuring intercanine, intermolar, overjet, overbite, midline discrepancy, space analysis, and tooth width (ICC > 0.7). Good criterion validity was indicated by agreement between the results from the plaster and virtual models (ICC > 0.8). There were substantial agreements for canine and molar relationship classifications ($\kappa > 0.70$). The results suggest that analysis performed on virtual models is as valid as traditional plaster models for intra- and inter-arch relationship.

Introduction

Orthodontic plaster models are routinely employed in the assessment of tooth dimensions as well as intra- and inter-arch relationships during the course of orthodontic treatment (Bell *et al.*, 2003). However, with increasing storage requirements for plaster models, damage and loss of models are common (Quimby *et al.*, 2004). Storage of plaster models is also a particular issue in urban areas where space is at a premium (Santoro *et al.*, 2003). Thus, in recent years, there has been considerable interest in alternatives to orthodontic plaster models within clinical orthodontics (Mok *et al.*, 2007).

The use of two-dimensional (2D) digital images has been suggested as an alternative for model storage as photographic devices are now widely available and relatively inexpensive (Sandler *et al.*, 2002; Cochran *et al.*, 2004; Wong *et al.*, 2005; Mok *et al.*, 2007). Nevertheless, it is accepted that 2D digital images cannot compare with three-dimensional (3D) images. This has led to increasing interest in 3D images or 'virtual' orthodontic models (Lee *et al.*, 2008). It has been reported that there was no significant difference in assessment of tooth dimensions obtained from plaster models and their corresponding virtual models (Bell *et al.*, 2003; Zilberman *et al.*, 2003; Quimby *et al.*, 2004; Mayers *et al.*, 2005; Paredes *et al.*, 2006; Stevens *et al.*, 2006; Asquith *et al.*, 2007; Mullen *et al.*, 2007; Lee *et al.*, 2008) or in several intra- and inter-arch relationship measurements (Terai *et al.*, 1999; Bell *et al.*, 2003; DeLong *et al.*, 2003; Santoro *et al.*, 2003; Quimby *et al.*, 2004; Costalos *et al.*, 2005; Whetten *et al.*, 2006; Okunami *et al.*, 2007). Bell *et al.* (2003) comparatively assessed direct measurements of the linear distance of six anatomic dental points on plaster models and measurements of computer-generated 3D images of the same plaster models and found no statistically significant differences between the measurements. Zilberman *et al.* (2003) tested the accuracy of measuring tooth size and arch width on plaster models and OrthoCAD® virtual models and concluded that both methods were clinically acceptable. Quimby *et al.* (2004) determine the accuracy, reproducibility, efficacy, and effectiveness of the measurements of mesiodistal widths, arch length, arch width, overjet, and overbite on computer-based models and found that those measurements appeared to be as accurate and reliable in general as the measurements from plaster models. Santoro *et al.* (2003) evaluated the accuracy of measuring tooth size, overbite, and overjet using OrthoCAD® models compared with plaster models. They reported a statistically significantly smaller tooth size and overbite, but the differences were considered clinically insignificant (less than 0.5 mm).

However, a comprehensive assessment of the 'level of agreement' between plaster and virtual models is lacking in the literature. This evidence is required to support or refute the use of virtual models in clinical practice and research. This study aimed to assess the potential use of virtual models as an alternative to plaster models by determining the level of agreement in the measurement of inter- and intra-arch relationships of plaster models compared with virtual models.

Materials and methods

Study sample

Two sets of polyvinylsiloxane impressions were obtained from 80 consecutive patients seeking orthodontic treatment. Plaster models were obtained following the usual standard protocol (Bell *et al.*, 2003) and the virtual models using OrthoCAD®. OrthoCAD® provides 3D images from plaster models, alginate or polyvinylsiloxane impression by means of a patented 'destructive scanning' system that can be downloaded from their website and viewed using their software. Typically OrthoCAD® files (www.orthocad.com) are 3000 kilobytes (3 megabytes). The inclusion criteria were subjects in the permanent dentition, with the canines, premolars, and second molars erupted (DS4M2; Björk *et al.*, 1964) with no gross dental abnormalities.

Data collection

Two trained and calibrated examiners (KB and ZL) undertook the assessment of the plaster models using digital callipers (Shanghai Taihai Congliang Ju Co., Lcd, Shanghai, China) to the nearest 0.01 mm and standard rulers to measure overjet and overbite to the nearest 0.5 mm (Figure 1). Overjet was measured from the labial surface of the lower incisor to the labial surface of the upper incisor (Stephens and Bowden, 1993). Overbite was measured by first marking a line on the labial surface of the incisors of the lower orthodontic plaster models with the most overlapped vertical distance of the upper central incisors and then the maximum vertical distance was measured from the marked line to the incisal edges (Mitchell and Mitchell, 2005). Mesiodistal widths were measured from the greatest mesiodistal diameter from the anatomic mesial contact point to the anatomic distal contact point of each tooth parallel to the occlusal plane (Hunter and Priest, 1960). Intermolar width was measured as the distance between the mesiobuccal cusp tips of the permanent first molars and intercanine width as the distance between the crown tips of the permanent canines. Space analysis was measured as the space required minus the space available on the plaster models (Mitchell and Mitchell, 2005).

For the virtual models, tooth width, overjet, overbite, midline discrepancy, intermolar width, intercanine width, and molar and canine relationship were determined on the computer screen using the same criteria as for plaster models. Both virtual and plaster models were assessed in a random order. Inter- and intra-examiner reliability were continuously monitored throughout the data collection process by randomly reassessing 10 of the 80 models.

Data analysis

Agreements of categorical data: canine and molar relationships were assessed by employing Kappa statistic (κ; Cohen, 1960). Agreements of continuous data: tooth width, overjet, overbite, intermolar width, intercanine width, midline discrepancy, and space analysis were assessed using comparison and correlation analysis. Comparison analysis was determined by calculating the mean directional differences (MDDs) and standardized directional differences (SDDs; Fleiss, 1971). Correlation analysis was assessed by calculating the intraclass correlation coefficients (ICCs; Shrout and Fleiss, 1979).

Sample size calculation

The sample required for the study was based on the size necessary to assess agreement using ICC. The null hypothesis for the ICC was set at 0.2 (poor agreement) and at 0.8 as the level of significant agreement (excellent agreement). Consequently, with α at 0.05 and β at 0.2, the minimum number of plaster–virtual model pairs required was 60 for a one-tailed test (Shrout and Fleiss, 1979).

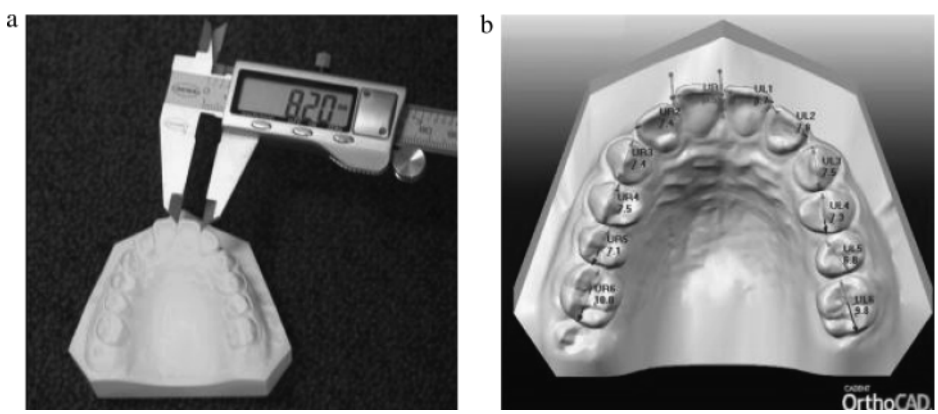

Figure 1 Measurement of tooth width (a) on plaster model with digital callipers and (b) with OrthoCAD® software.

Results

Virtual model–plaster model agreement (validity)

Tooth width measurements from virtual models when compared with measurements from plaster models were of similar magnitudes (Table 1). Compared with 0 mm, the MDD of tooth width was not statistically significant ($P > 0.05$). When the MDD was standardized, the magnitude of the directional difference for tooth width measurements was less than 0.1 mm. The mean absolute differences of tooth widths measured on virtual and plaster models were less than 0.3 mm. ICC for tooth widths was greater than 0.80.

The MDDs of intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis were all less than 0.2 mm. The SDDs of intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis are shown in Table 2. ICCs for all variables were greater than 0.80. For canine and molar relationships, the agreements as measured by κ statistic were greater than 0.70 (κ = 0.72 and κ = 0.79 for the right and left canine relationship, respectively, and κ = 0.81 and κ = 0.87 for the right and left molar relationship).

Intra-examiner reliability—plaster models

The MDDs of tooth, intercanine, and intermolar width, overjet, overbite, midline discrepancy, and space analysis were all less than 0.3 mm. The SDDs for tooth, intercanine, and intermolar width, overjet, overbite, midline discrepancy, and space analysis are shown in Tables 3 and 4. ICCs for all these variables were greater than 0.80. Intra-examiner agreement as measured by κ statistic for canine and molar relationships were greater than 0.70 (κ = 1.00 for both right and left canine relationship and κ = 0.87 and κ = 0.74 for right and left molar relationship, respectively).

Intra-examiner reliability—virtual models

The MDDs of tooth, intercanine, and intermolar width, overjet, overbite, midline discrepancy, and space analysis were all less than 0.6 mm. The SDDs of tooth, intercanine, and intermolar width, overjet, overbite, midline discrepancy, and space analysis are shown in Tables 4 and 5. ICCs for all these variables were greater than 0.80. Inter-examiner agreement as measured by κ statistic for canine and molar relationships were greater than 0.80 (κ = 1.00 and κ = 0.87 for right and left canine relationship, respectively, and κ = 1.00 for both right and left molar relationship).

Table 1 Agreement between plaster and virtual models in assessment of tooth size.

Measurements Tooth	Plaster		Virtual		Directional difference			Absolute difference		Correlation coefficient	95% Agreement interval
	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	*D**	Mean (mm)	SD (mm)		(mm)
16	10.50	0.62	10.52	0.59	0.02	0.21	0.08	0.10	0.18	0.942	0.911–0.962
15	7.13	0.45	7.14	0.49	0.01	0.23	0.04	0.12	0.20	0.882	0.822–0.923
14	7.58	0.41	7.60	0.45	0.02	0.19	0.12	0.10	0.16	0.908	0.859–0.940
13	8.13	0.57	8.18	0.64	0.05	0.28	0.18	0.18	0.22	0.900	0.848–0.935
12	7.18	0.69	7.20	0.71	0.01	0.18	0.06	0.08	0.16	0.968	0.950–0.979
11	8.74	0.52	8.72	0.56	−0.02	0.23	−0.08	0.11	0.20	0.911	0.865–0.942
21	8.73	0.53	8.72	0.56	−0.01	0.18	−0.04	0.10	0.15	0.945	0.915–0.964
22	7.19	0.61	7.19	0.65	−0.01	0.17	−0.04	0.08	0.16	0.963	0.943–0.976
23	8.04	0.63	8.07	0.67	0.02	0.12	0.19	0.07	0.10	0.984	0.975–0.990
24	7.59	0.44	7.62	0.46	0.03	0.15	0.21	0.08	0.13	0.948	0.920–0.967
25	7.14	0.46	7.16	0.47	0.03	0.12	0.22	0.07	0.10	0.966	0.947–0.978
26	10.36	0.53	10.32	0.53	−0.03	0.24	−0.14	0.19	0.16	0.896	0.842–0.932
36	11.03	0.61	10.98	0.61	−0.05	0.25	−0.19	0.17	0.19	0.917	0.873–0.946
35	7.35	0.46	7.38	0.49	0.03	0.17	0.19	0.10	0.14	0.939	0.906–0.960
34	7.42	0.45	7.46	0.52	0.04	0.21	0.21	0.13	0.17	0.918	0.875–0.947
33	7.02	0.48	7.04	0.55	0.02	0.22	0.08	0.13	0.18	0.914	0.869–0.944
32	6.15	0.47	6.18	0.53	0.03	0.24	0.11	0.15	0.19	0.891	0.835–0.929
31	5.58	0.44	5.55	0.46	−0.02	0.19	−0.13	0.12	0.15	0.907	0.858–0.939
41	5.57	0.47	5.54	0.46	−0.03	0.20	−0.17	0.14	0.15	0.901	0.850–0.936
42	6.12	0.49	6.15	0.52	0.03	0.22	0.16	0.15	0.16	0.908	0.860–0.940
43	6.93	0.43	6.93	0.48	0.00	0.20	−0.01	0.12	0.16	0.906	0.856–0.939
44	7.40	0.46	7.42	0.48	0.02	0.11	0.20	0.07	0.09	0.972	0.956–0.982
45	7.40	0.47	7.42	0.46	0.03	0.13	0.20	0.09	0.09	0.963	0.943–0.976
46	11.05	0.56	11.05	0.56	−0.01	0.23	−0.03	0.17	0.15	0.918	0.874–0.947

*Standardized difference = mean directional difference/standardized deviation of directional differences ($D \leq 0.2$, small; $0.2 < D < 0.5$, moderate; $D \geq 5$, large).

Table 2 Agreements between plaster and virtual models in assessing dental arch relationship.

Variables	Plaster		Virtual		Directional difference			Absolute difference		Correlation coefficient	95% Agreement interval (mm)
	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	Mean (mm)	SD (mm)	D*	Mean (mm)	SD (mm)		
Upper intercanine	35.04	2.57	35.01	2.67	0.03	0.69	0.05	0.37	0.58	0.967	0.948–0.978
Lower intercanine	26.84	2.65	26.89	2.83	−0.05	0.53	−0.09	0.41	0.35	0.983	0.974–0.989
Upper intermolar	46.62	3.03	46.62	3.11	0.01	0.55	0.01	0.44	0.32	0.984	0.976–0.990
Lower intermolar	45.53	3.31	45.43	3.33	0.10	0.52	0.20	0.39	0.35	0.988	0.981–0.992
Overjet	3.77	3.62	3.95	3.70	−0.19	0.94	−0.20	0.63	0.73	0.967	0.949–0.979
Overbite	2.05	1.77	2.20	1.96	−0.16	0.80	−0.20	0.53	0.62	0.913	0.868–0.944
Upper crowding	−3.06	4.14	−2.99	4.15	−0.07	0.75	−0.09	0.43	0.62	0.984	0.974–0.989
Lower crowding	−1.98	3.39	−1.83	3.30	−0.15	0.88	−0.17	0.41	0.79	0.966	0.947–0.978
Midline discrepancy	1.35	1.17	1.47	1.27	−0.12	0.55	−0.22	0.37	0.42	0.903	0.852–0.937

*Standardized difference = mean directional difference/standardized deviation of directional differences ($D \leq 0.2$, small; $0.2 < D < 0.5$, moderate; $D \geq 5$, large).

Table 3 Intra- and inter-examiner agreement with tooth widths of plaster models.

Measurements Tooth	Directional difference						Correlation	
	Intra-examiner			Inter-examiner				
	Mean (mm)	SD (mm)	D*	Mean (mm)	SD (mm)	D*	Intra-examiner	Inter-examiner
16	0.05	0.12	0.42	−0.06	0.32	−0.19	0.972	0.823
15	−0.04	0.14	−0.28	−0.03	0.30	−0.10	0.951	0.813
14	0.04	0.10	0.41	−0.06	0.25	−0.24	0.971	0.811
13	0.02	0.24	0.08	0.01	0.38	0.03	0.941	0.720
12	0.03	0.11	0.28	−0.16	0.29	−0.55	0.985	0.878
11	0.01	0.10	0.10	−0.14	0.21	−0.68	0.985	0.905
21	0.02	0.10	0.19	−0.05	0.16	−0.30	0.978	0.943
22	0.07	0.13	0.56	−0.15	0.27	−0.56	0.972	0.930
23	0.01	0.11	0.09	−0.13	0.19	−0.69	0.995	0.993
24	−0.01	0.10	−0.10	−0.06	0.22	−0.28	0.973	0.903
25	−0.04	0.12	−0.34	0.07	0.32	0.22	0.958	0.711
26	0.00	0.12	0.00	0.08	0.22	0.36	0.971	0.893
36	−0.07	0.19	−0.37	−0.18	0.39	−0.46	0.939	0.723
35	0.05	0.07	0.71	0.00	0.33	0.00	0.995	0.881
34	0.04	0.12	0.34	−0.03	0.19	−0.16	0.980	0.946
33	−0.03	0.16	−0.18	−0.14	0.24	−0.58	0.969	0.936
32	0.04	0.07	0.57	0.02	0.30	0.07	0.980	0.840
31	0.03	0.11	0.28	−0.10	0.17	−0.59	0.945	0.877
41	0.01	0.13	0.08	−0.15	0.25	−0.61	0.890	0.740
42	0.00	0.15	0.00	−0.11	0.40	−0.27	0.942	0.356
43	−0.03	0.12	−0.26	−0.20	0.33	−0.61	0.944	0.835
44	0.00	0.09	0.00	−0.10	0.18	−0.55	0.985	0.956
45	0.05	0.10	0.51	0.09	0.15	0.59	0.982	0.938
46	−0.06	0.20	−0.30	0.09	0.23	0.39	0.905	0.827

*Standardized difference = mean directional difference/standardized deviation of directional differences ($D \leq 0.2$, small; $0.2 < D < 0.5$, moderate; $D \geq 5$, large).

Inter-examiner reliability—plaster models

The MDDs of tooth widths were less than 0.1 mm. The MDDs of intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis were less than 0.6 mm. The SDDs for tooth, intercanine, and intermolar width, overjet, overbite, midline discrepancy, and space analysis are shown in Tables 3 and 4. The ICCs for intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis agreements were all greater than 0.70. Inter-examiner agreements for canine and molar relationships were greater than 0.70 (κ = 1.00 for both right and left canine relationship and κ = 0.87 and κ = 0.74 for right and left molar relationship, respectively).

Table 4 Intra- and inter-examiner agreement in assessing dental arch relationship.

Variables	Directional difference						Correlation	
	Intra-examiner			Inter-examiner				
	Mean (mm)	SD (mm)	*D**	Mean (mm)	SD (mm)	*D**	Intra-examiner	Inter-examiner
Plaster model								
Upper intercanine	−0.19	0.40	−0.48	−0.03	0.65	−0.05	0.979	0.939
Lower intercanine	0.03	0.25	0.12	−0.15	0.49	−0.30	0.994	0.975
Upper intermolar	−0.20	0.30	−0.67	−0.23	1.60	−0.14	0.997	0.870
Lower intermolar	−0.13	0.46	−0.28	−0.58	0.92	−0.63	0.996	0.981
Overjet	−0.05	0.64	−0.08	−0.33	0.71	−0.46	0.984	0.981
Overbite	0.05	0.37	0.14	0.03	0.45	0.07	0.983	0.965
Upper crowding	−0.67	0.97	−0.69	−0.45	0.84	−0.54	0.943	0.974
Lower crowding	−0.19	0.66	−0.29	−0.22	0.28	−0.78	0.973	0.994
Midline discrepancy	−0.20	0.42	−0.47	−0.05	0.16	−0.32	0.929	0.989
Virtual model	−0.05	0.36	−0.14	−0.44	0.83	−0.53	0.984	0.910
Upper intercanine	−0.07	0.33	−0.21	−0.11	0.62	−0.18	0.991	0.961
Lower intercanine	0.03	0.47	0.06	−0.68	1.64	−0.42	0.990	0.853
Upper intermolar	0.07	0.29	0.24	−0.56	0.84	−0.66	0.998	0.982
Lower intermolar	−0.36	1.05	−0.34	−0.41	0.65	−0.63	0.967	0.984
Overjet	0.06	0.60	0.10	0.14	0.56	0.25	0.931	0.943
Overbite	−0.06	0.35	−0.17	−0.03	0.82	−0.04	0.990	0.951
Upper crowding	0.13	0.51	0.25	0.02	1.35	0.01	0.977	0.825
Lower crowding	−0.06	0.26	−0.23	−0.07	0.47	−0.15	0.948	0.791
Midline discrepancy	−0.05	0.36	−0.14	−0.44	0.83	−0.53	0.984	0.910

*Standardized difference = mean directional difference/standardized deviation of directional differences ($D \leq 0.2$, small; $0.2 < D < 0.5$, moderate; $D \geq 5$, large).

Table 5 Intra- and inter-examiner agreement with tooth widths of virtual models.

Measurements Tooth	Directional difference						Correlation	
	Intra-examiner			Inter-examiner				
	Mean (mm)	SD (mm)	*D**	Mean (mm)	SD (mm)	*D**	Intra-examiner	Inter-examiner
16	−0.08	0.20	−0.39	0.13	0.30	0.43	0.946	0.873
15	0.00	0.13	0.00	0.03	0.38	0.08	0.976	0.725
14	0.04	0.11	0.37	0.01	0.17	0.06	0.971	0.930
13	0.00	0.18	0.00	0.07	0.41	0.17	0.969	0.840
12	0.06	0.16	0.38	0.06	0.34	0.18	0.988	0.890
11	−0.05	0.26	−0.19	−0.20	0.40	−0.50	0.916	0.841
21	0.08	0.12	0.65	0.02	0.35	0.06	0.983	0.852
22	0.07	0.21	0.33	−0.07	0.23	−0.30	0.984	0.967
23	−0.03	0.16	−0.19	−0.12	0.32	−0.37	0.992	0.967
24	−0.02	0.10	−0.19	0.04	0.40	0.10	0.993	0.746
25	−0.02	0.16	−0.12	0.13	0.31	0.42	0.956	0.761
26	0.07	0.35	0.20	0.16	0.31	0.52	0.816	0.783
36	−0.04	0.22	−0.18	0.15	0.33	0.45	0.912	0.810
35	0.02	0.19	0.10	0.15	0.49	0.31	0.960	0.749
34	0.02	0.12	0.16	0.13	0.34	0.39	0.988	0.926
33	−0.06	0.26	−0.23	−0.01	0.31	−0.03	0.969	0.939
32	0.03	0.19	0.16	0.04	0.44	0.09	0.963	0.747
31	0.07	0.11	0.66	0.08	0.12	0.65	0.977	0.954
41	0.04	0.12	0.34	0.07	0.22	0.32	0.952	0.851
42	0.11	0.19	0.58	0.06	0.34	0.18	0.915	0.813
43	0.04	0.17	0.23	0.00	0.17	0.00	0.968	0.971
44	−0.01	0.22	−0.04	−0.09	0.21	−0.43	0.919	0.925
45	0.10	0.16	0.61	0.17	0.25	0.69	0.935	0.863
46	−0.07	0.22	−0.32	0.10	0.21	0.47	0.899	0.899

*Standardized difference = mean directional difference/standardized deviation of directional differences ($D \leq 0.2$, small; $0.2 < D < 0.5$, moderate; $D \geq 5$, large).

Inter-examiner reliability—virtual models

For virtual models, the MDDs of tooth widths were all less than 0.2 mm. The MDDs of intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis were all less than 0.5 mm. The SDDs of tooth widths were all less than 0.6 mm. The SDDs of intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis were all less than 0.3 mm. ICCs for tooth, intercanine and intermolar width, overjet, overbite, midline discrepancy, and space analysis agreements were all greater than 0.70. Inter-examiner agreements for canine and molar relationships were greater than 0.70 ($\kappa = 0.73$ and $\kappa = 1.00$ for right and left canine relationship, respectively, and $\kappa = 1.00$ for both right and left molar relationship; Tables 4 and 5).

Discussion

With the growing interest in 3D images among orthodontists, a number of companies now offer a service to supply 'virtual models' on a commercial basis (Quimby *et al.*, 2004; Stevens *et al.*, 2006; Okunami *et al.*, 2007). In this study, 3D images were obtained from OrthoCAD® which is the system employed at the Prime Philip Dental Hospital (Quimby *et al.*, 2004). All companies that provide 3D image virtual model systems are similar with respect to what assessments can be undertaken on them but differ in the software they provide for analysis (Stevens *et al.*, 2006). It would be useful for other researchers to assess and compare different 3D image virtual model systems. It should be borne in mind that models in this study were obtained from polyvinylsiloxane impressions. Other types of impression material make affect the accuracy of 3D image virtual models.

With respect to tooth dimensions in this study, the difference in mesiodistal dimensions as assessed on plaster and virtual models were not statistically different from zero. This concurs with the findings comparing plaster models and other 3D imaging systems (Bell *et al.*, 2003; Santoro *et al.*, 2003; Quimby *et al.*, 2004; Stevens *et al.*, 2006). At a group level, the MDDs of tooth dimensions for plaster models compared with virtual models were less than 0.1 mm. The SDD of tooth dimensions between plaster models and virtual models was less than 0.2, which can be interpreted as a statistically small difference; the magnitude of the statistical difference was small and similar to previously reported studies (Cohen, 1960, 1988; Shrout and Fleiss, 1979; Petrie, 2002). The mean absolute difference was less than 0.2 mm; if the average mesiodistal width of the smallest tooth (lower central incisor) is taken as 6.0 mm (Lewis *et al.*, 2008), then there is less than a 5 per cent difference in absolute assessment. In this study, individual plaster and virtual model agreement was excellent (ICC > 0.80) in all cases (Shrout and Fleiss, 1979).

To assess arch relationships, assessments were conducted including intercanine width, intermolar width, overjet, overbite, upper and lower space analysis, as well as midline discrepancy. At the group level, the MDD between assessments was not significantly less than 0. This is in agreement with the findings of two other studies that reported no significant difference in many of these variables between plaster and virtual models (Colton, 1974; Shrout and Fleiss, 1979; Petrie, 2002). The SDD of all arch relationships was less than 0.2 indicating that any difference between them could be interpreted as small (Shrout and Fleiss, 1979). The mean absolute difference was less than 0.7 mm. Relative to average estimates of these variables, the mean absolute difference was less than 5 per cent; however, it is acknowledged that the parameters of these variables may differ with respect to different populations. The individual plaster and virtual model pairs' agreement of arch relationships could be considered as excellent (ICC > 0.90; Shrout and Fleiss, 1979; Cohen, 1988; Petrie, 2002).

In terms of inter- and intra-examiner reliability, assessments on plaster and virtual models were all substantial, both for tooth dimensions and arch relationships (ICC > 0.70; Colton, 1974; Shrout and Fleiss, 1979; Cohen, 1988; Roberts and Richmond, 1997). Agreement for molar relationship obtained from plaster and their virtual models could be interpreted as 'substantial' ($\kappa > 0.70$; Petrie, 2002).

Conclusion

There was substantial to excellent agreement between assessment of tooth dimensions and arch relationships between plaster and virtual models.

Funding

University of Hong Kong (10206658.23438.08003.301.01).

References

Asquith J, Gillgrass T, Mossey P 2007 Three-dimensional imaging of orthodontic models: a pilot study. European Journal of Orthodontics 29: 517–522

Bell A, Ayoub A, Siebert P 2003 Assessment of the accuracy of a three-dimensional imaging system for archiving dental study models. Journal of Orthodontics 30: 219–223

Björk A, Krebs A, Solow B 1964 A method for epidemiological registration of malocclusion. Acta Odontologica Scandinavica 22: 27–41

Cochran J *et al.* 2004 A standardized photographic method for evaluating enamel opacities including fluorosis. Community Dentistry and Oral Epidemiology 32: (Suppl 1)19–27

Cohen J 1960 A coefficient of agreement for nominal scales. Educational and Psychological Measurement 1: 37–46

Cohen J 1988 Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates, Hillside

Colton T 1974 Statistics in medicine. Little Brown and Company, Boston

Costalos P, Sarraf K, Cangialosi T, Efstratiadis S 2005 Evaluation of the accuracy of digital model analysis for the American Board of Orthodontics objective grading system for dental casts. American Journal of Orthodontics and Dentofacial Orthopedics 128: 624–629

DeLong R, Heinzen M, Hodges J, Ko C, Douglas W 2003 Accuracy of a system for creating 3D computer models of dental arch. Journal of Dental Research 82: 438–442

Fleiss J 1971 Measuring nominal scale agreement among many raters. Psychological Bulletin 76: 378–382

Hunter W, Priest W 1960 Errors and discrepancy in measurement of tooth sizes. Journal of Dental Research 39: 405–414

Lee S, Delong P R, Hodges J S, Hayashi K, Lee J B 2008 Predicting first molar width using virtual models of dental arches. Clinical Anatomy 21: 27–32

Lewis B R, Stern M R, Willmot D R 2008 Maxillary anterior tooth size and arch dimensions in unilateral cleft lip and palate. Cleft Palate-Craniofacial Journal 45: 639–646

Mayers M, Firestone A, Rashid R, Vig K 2005 Comparison of Peer Assessment Rating (PAR) index scores of plaster and computer-based digital models. American Journal of Orthodontics and Dentofacial Orthopedics 128: 431–434

Mitchell D, Mitchell L 2005 Oxford handbook of clinical dentistry, 4th edn. Oxford University Press, Inc., New York

Mok C W, Zhou L, McGrath C, Hägg U, Bendeus M 2007 Digital images as an alternative to orthodontic casts in assessing malocclusion and orthodontic treatment need. Acta Odontologica Scandinavica 65: 362–368

Mullen S, Martin C, Ngan P, Gladwin M 2007 Accuracy of space analysis with emodels and plaster models. American Journal of Orthodontics and Dentofacial Orthopedics 132: 346–352

Okunami T, Kusnoto B, BeGole E, Evans C, Sadowsky C, Fadavi S 2007 Assessing the American Board of Orthodontics objective grading system: digital vs plaster dental casts. American Journal of Orthodontics and Dentofacial Orthopedics 131: 51–56

Paredes V, Gandia J, Cibrian R 2006 Determination of Bolton tooth-size ratios by digitization, and comparison with the traditional method. European Journal of Orthodontics 28: 120–125

Petrie A 2002 Further statistical in dentistry. British Dental Association, London

Quimby M, Vig K, Rashid R, Firestone A 2004 The accuracy and reliability of measurements made on computer-based digital models. Angle Orthodontist 74: 298–303

Roberts C T, Richmond S 1997 The design and analysis of reliability studies for the use of epidemiological and audit indices in orthodontics. British Journal of Orthodontics 24: 139–147

Sandler P J, Murray A M, Bearn D 2002 Digital records in orthodontics. Dental Update 29: 18–24

Santoro M, Galkin S, Teredesai M, Nicolay O F, Cangialosi T J 2003 Comparison of measurements made on digital and plaster models. American Journal of Orthodontics and Dentofacial Orthopedics 124: 101–105

Shrout P E, Fleiss J L 1979 Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 86: 420–428

Stephens C, Bowden D 1993 Examination of the patient. In: Shaw W C (ed). Orthodontics and occlusal management. Wright, Oxford. , pp. 83–99.

Stevens D R, Flores-Mir C, Nebbe B, Raboud D W, Heo G, Major P W 2006 Validity, reliability, and reproducibility of plaster vs digital study models: comparison of Peer Assessment Rating and Bolton analysis and their constituent measurements. American Journal of Orthodontics and Dentofacial Orthopedics 129: 794–803

Terai H, Shimahara M, Sakinaka Y, Tajima S 1999 Accuracy of integration of dental casts in three-dimensional models. Journal of Oral Maxillofacial Surgery 57: 662–665

Whetten J, Williamson P, Heo G, Varnhagen C, Major P 2006 Variations in orthodontic treatment planning decisions of Class II patient between virtual 3-dimensional models and traditional plaster study models. American Journal of Orthodontics and Dentofacial Orthopedics 130: 485–491

Wong H M, McGrath C, Lo E C, King N M 2005 Photographs as a means of assessing developmental defects of enamel. Community Dentistry and Oral Epidemiology 33: 438–446

Zilberman O, Huggare J, Parikakis K 2003 Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. Angle Orthodontist 73: 301–306