Reproducibility, reliability and validity of measurements obtained from Cecile3 digital models

Gustavo Adolfo Watanabe-Kanno[(a)]
Jorge Abrão[(b)]
Hiroshi Miasiro Junior[(a)]
Alfonso Sánchez-Ayala[(c)]
Manuel O. Lagravère[(d)]

[(a)]MSc Student, Graduate Program in Orthodontics; [(b)]PhD, Associate Professor – Department of Orthodontics, School of Dentistry, University of São Paulo, São Paulo, SP, Brazil.

[(c)]PhD Student, Graduate Program in Prosthodontics, Department of Prosthodontics, Piracicaba Dental School, University of Campinas, Piracicaba, SP, Brazil.

[(d)]Research Associate, Graduate Program in Orthodontics, Department of Dentistry, University of Alberta, Edmonton, Alberta, Canada.

Abstract: The aim of this study was to determine the reproducibility, reliability and validity of measurements in digital models compared to plaster models. Fifteen pairs of plaster models were obtained from orthodontic patients with permanent dentition before treatment. These were digitized to be evaluated with the program Cécile3 v2.554.2 beta. Two examiners measured three times the mesiodistal width of all the teeth present, intercanine, interpremolar and intermolar distances, overjet and overbite. The plaster models were measured using a digital vernier. The t-Student test for paired samples and interclass correlation coefficient (ICC) were used for statistical analysis. The ICC of the digital models were 0.84 ± 0.15 (intra-examiner) and 0.80 ± 0.19 (inter-examiner). The average mean difference of the digital models was 0.23 ± 0.14 and 0.24 ± 0.11 for each examiner, respectively. When the two types of measurements were compared, the values obtained from the digital models were lower than those obtained from the plaster models ($p < 0.05$), although the differences were considered clinically insignificant (differences < 0.1 mm). The Cécile digital models are a clinically acceptable alternative for use in Orthodontics.

Descriptors: Dental models; Reproducibility of results; Orthodontics.

Corresponding author:
Gustavo Adolfo Watanabe-Kanno
Departamento de Ortodontia e Odontopediatria da FOUSP
Av. Prof. Lineu Prestes, 2227, Cidade Universitária
São Paulo - SP - Brazil
CEP: 05508-900
E-mail: *adolfwk@usp.br*

Introduction

Digital technology is slowly influencing different scientific areas and improving them. One of these areas is orthodontics, where diagnostic tools, digital photography and cephalometric software, among others, are improving this area.[1] Plaster models in orthodontics are necessary diagnostic tools but present several disadvantages such as the need for storage areas, high risk of breakage, loss of information and difficulty to send to other clinicians in multidisciplinary cases.[2,3] Some alternatives to using plaster models have been suggested such as photocopies, photographs, holograms and digitalization of points from the plaster cast.[4,5] Digital models now offer some advantages that include ease of storage and retrieval, ease of interoffice transferability, and possibly similar or better diagnostic ability.[6] Some studies evaluated different methods to calculate tooth size discrepancy and found no statistical significant differences in the measurements obtained from plaster and digital models, and suggested the need for further research to determine accuracy, reliability and reproducibility of digital models using new software versions.[2,7] One group of investigators found a significant difference between plaster and digital model linear measurements but concluded that the average difference did not appear to be clinically relevant and suggested that the software presented the advantage of being able to rotate and enlarge images on the screen.[8,9] The accuracy and reliability of measurements made on computer-based models appeared to be as accurate and reliable as the ones obtained from plaster models.[2] On the other hand, other investigators found excellent reproducibility of digital models and significant differences in some measurements regarding reliability and validity, but these were not considered clinically significant.[10,11]

Cécile3, a digital modeling analysis software, was created in 2003 by Bibliocast, Montreuil. Some advantages of this software include the possibility of performing space analysis and creating virtual setups for treatment planning purposes. Nevertheless, there is a lack of literature with respect to its measurement validity, reliability and/or reproducibility.

Digital models could be as reliable as plaster models, and an additional tool for orthodontic diagnosis. In this case, some fundamental factors such as spacing condition, teeth inclination, rotations, presence of interproximal contacts and other anatomical variations should also be considered.[12-15] Because the need for evidence-based orthodontics is increasing, the degree of usefulness of different emerging methods ought to be evaluated. For these reasons, the objective of the present study is to determine the reproducibility, reliability, and validity of measurements obtained from Cécile3 digital models.

Material and Methods

The study sample consisted of fifteen pairs of randomly selected initial model casts (upper and lower) from patients that presented for treatment at the orthodontic clinic, School of Dentistry, University of São Paulo. Inclusion criteria was that all 12-18 year old patients had to have complete permanent dentition erupted from first molar to first molar, no caries lesions, no interproximal wear, no extractions nor previous orthodontic treatment.

Plaster casts of the 15 patients were sent to the Bibliocast company (Montreuil-France) in order to be digitized (3D CT Scanning) and to obtain a digital file to be uploaded into the analysis software. Digital files with 2.2 MB were obtained for each pair of plaster models, and measurements were made on the digitized models using Cécile3 software version 2.554.2 beta, to the nearest 0.1 mm. The bases of the plaster models were made parallel to the inferior occlusal plane (crown tip of the lower canine to the distobuccal cuspid of the second lower molar) using a glass plate. Before starting the measurements, the digital models were positioned with the base parallel to the horizontal software guideline reference; then anterior and posterior teeth were measured from the facial view. After that, the digital models were rotated to confirm the accuracy of the measurements. For severely malpositioned anterior teeth, the images were rotated on screen, and the measurements were made from the occlusal view to provide better visibility. For ease and accuracy of measurements, the images were enlarged on screen as needed by using the magnifying feature (Figure 1). Transver-
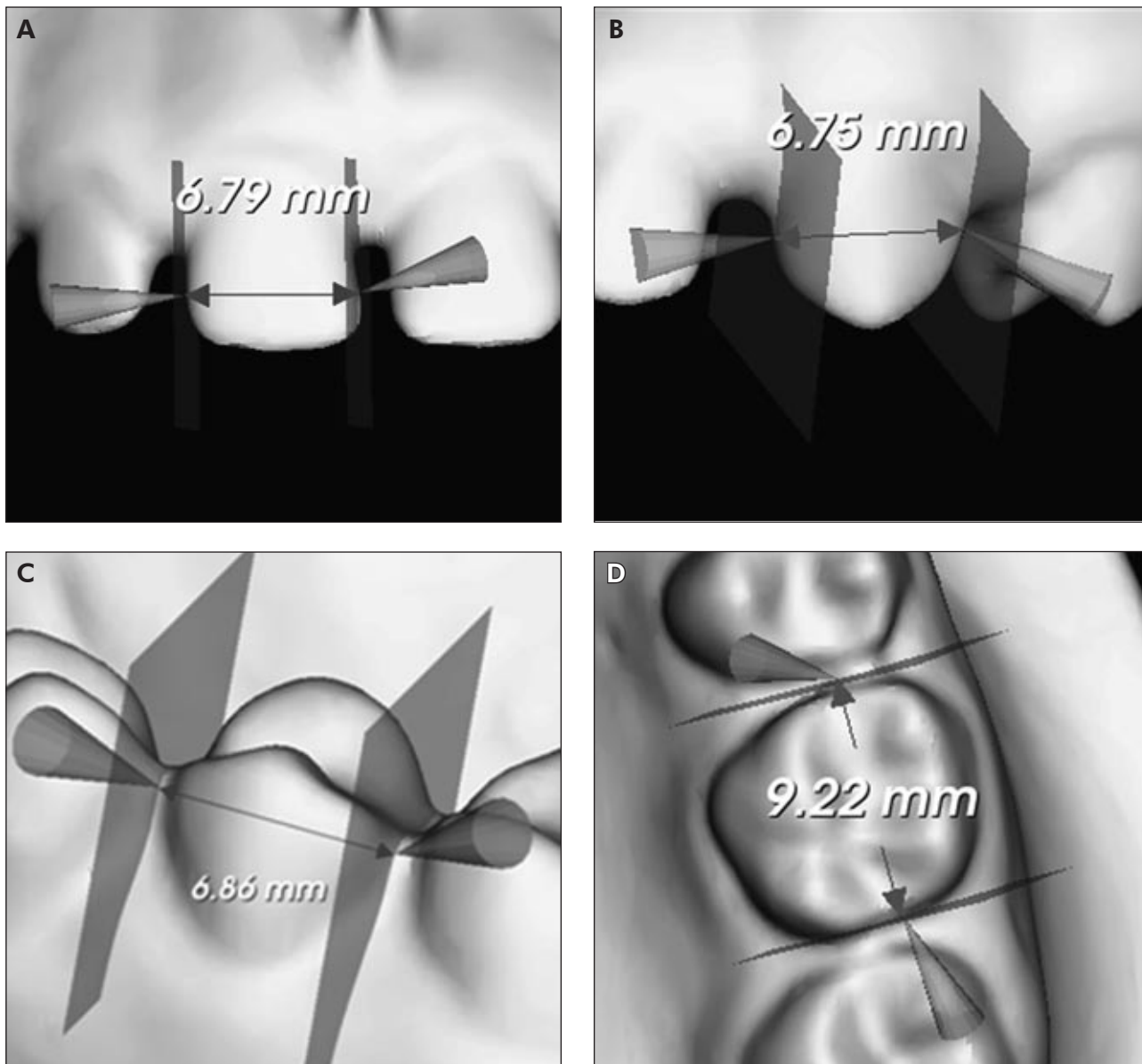
Figure 1 - Measurements of mesiodistal width of **(A)** incisor, **(B)** canine, **(C)** premolar and **(D)** molar using the Cécile3 tool, as shown from different views.

sal distances were also measured from the occlusal view, and overjet and overbite were measured using the analysis tools (Figure 2). The plaster models were measured using a vernier digital caliper (Mitutoyo, model 500-144B, Suzano, SP, Brazil) with an accuracy of 0.01 mm. The measurements obtained are described in Table 1 and were measured by two different investigators, each repeating the measurement set three times.

Statistical analysis

The data were analyzed using the statistical program SPSS for Windows version 16.0 (SPSS, Chicago, IL, USA). Sample normality and homogeneity of variances were determined using the Shapiro-Wilks and Levene tests. The Intraclass correlation coefficient test (ICC) was used to determine the reproducibility of the measures obtained. The Student t-Test was used to compare examiner reliability for both methods used and the validity between them.
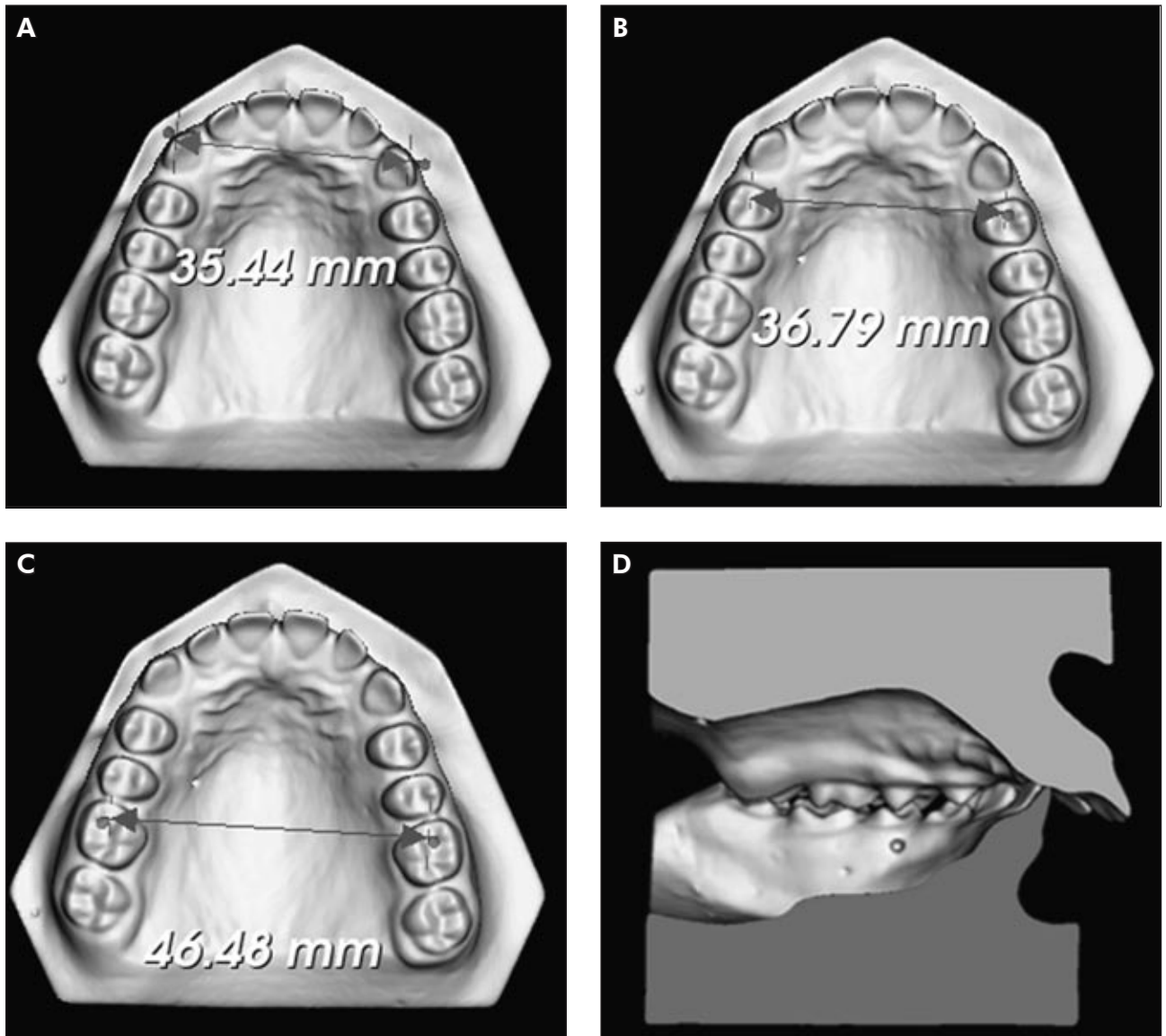
Figure 2 - Measurements of **(A)** intercanine, **(B)** interpremolar, **(C)** intermolar, **(D)** overjet and overbite using the Cécile3 tool.

Results

Intra- and inter-examiner reliability for both methods was generally high. The average intra-examiner reliability was 0.852 ± 0.12 (range 0.706-0.940) and 0.824 ± 0.15 (range 0.663-0.927) for the plaster and Cécile3 digital models respectively. Inter-examiner values were 0.818 ± 0.18 (range 0.575-0.931) and 0.782 ± 0.19 (range 0.506-0.917) for the plaster and Cécile3 digital models respectively. Overjet presented the highest intra-examiner reliability (0.996), while the mesiodistal width of the lower left canine presented the lowest value (0.537

for the plaster models and 0.367 for the Cécile3 digital models). The maxillary interpremolar distance had the greatest inter-examiner reliability in plaster models (0.999) and the maxillary intercanine distance had the greatest inter-examiner reliability in the Cécile3 digital models (0.998). The lowest inter-examiner value was found in the mesiodistal width of the lower left canine (0.318 in the plaster models and 0.158 in the Cécile3 digital models).

The mean measurement error for all measurements was obtained for the plaster and Cécile3 digital models for both examiners. There were no sig-

Table 1 - Measurement definitions.

Measurement	Definition[12,13,14,15]
Mesiodistal Width	Greatest mesiodistal diameter from the anatomic mesial contact point to the anatomic distal contact point in each tooth, parallel to the occlusal surface.
Intercanine Distance	Straight distance between the crown tips of the canines.
Interpremolar Distance	Straight distance between the mesial fossae of the first premolars.
Intermolar Distance	Straight distance between the mesial fossae of the first molars.
Overjet	Distance between the incisal border of the more buccal upper central incisor and the buccal surface of the more lingual lower central incisor.
Overbite	Distance between the incisal border of the upper central incisor and the incisal border of the lower central incisor.

nificant differences in measurement errors in most measurements except for the lower right mandibular lateral incisor mesiodistal width, lower right first premolar mesiodistal width, upper right lateral incisor mesiodistal width and lower interpremolar distance ($p < 0.05$) for examiner 1; and except for the upper right canine mesiodistal width, lower intercanine distance, lower interpremolar distance and overbite for examiner 2 ($p < 0.05$).

The second trial of measurements for each examiner was chosen to determine the validity of measurements (Table 2). All values obtained from the Cécile3 digital models were smaller than the values obtained from the plaster models with the exception of the upper interpremolar distance for examiner 2. The mean differences between plaster and Cécile3 digital models were 0.17 ± 0.06 mm for examiner 1. The lower right mandibular first molar presented the highest variability (0.39 ± 0.28 mm). For examiner 2, the mean difference was 0.19 ± 0.06 mm, with the highest value obtained for overjet (0.31 ± 0.22 mm). The Student t-Test showed significant differences between the measurements obtained from the Cécile3 and plaster models except for the first molar, canine and central incisor of the lower left quadrant, second premolar and first premolar from the upper right quadrant and upper left central incisor ($p > 0.05$)

for examiner 1. For examiner 2, the upper intermolar distance and lower intermolar distance did not present a significant difference ($p > 0.05$).

Discussion

The aim of this study was to determine if Cécile3 digital models are as reliable as plaster models. To achieve this, reproducibility, reliability and validity of the measurements obtained were evaluated.

Reproducibility of Cécile3 models

Reproducibility evaluates the agreement between two readings from the same sample (which in the present study would be for Cécile3 digital models and plaster models).[16] While studying reproducibility in orthodontics, Roberts, Richmond[16] (1997) stated that an ICC < 0.4 is considered low, between 0.4 and 0.75 is acceptable, and > 0.75 is good. The present study presented an ICC higher than 0.75 in the intra- and inter-examiner evaluations. Low values were only present for the mesiodistal width of the left mandibular canine in plaster models (ICC-Intra-examiner = 0.537 / ICC-Inter-examiner = 0.318) and Cécile3 models (ICC-Intra-examiner = 0.367 / ICC-Inter-examiner = 0.158). These findings might be attributed to the fact that, in some cases, the interproximal area between the teeth is not clearly defined, which can alter the reproducibility of the measurements at the time of marking points. However, considering all the results, we can infer that the differences are clinically acceptable, and reproducibility is high. Stevens *et al.*[10] (2006), using the concordance correlation coefficient (CCC), showed that all 50 intra-examiner measurements had excellent reproducibility for both plaster and OrthoCad models with the exception of 7 measurements (4 plaster and 3 digital), which were considered good. Quimby *et al.*[2] (2004) found a high degree of reproducibility with an ICC > 0.90 for the measurements made on both plaster and computer-based models, measured by two examiners.

Reliability of Cécile3 digital models

Reliability was considered as the extent to which a measurement was repeatable under identical conditions, between Cécile3 digital models and plaster models.[16] No statistical significant differences were

Table 2 - Measurement Means (mm)[†].

Measurement	Examiner 1						Examiner 2					
	Descriptives					Paired-t test	Descriptives					Paired-t test
	Plaster mean		Bibliocast mean				Plaster mean		Bibliocast mean			
	Mean	SD	Mean	SD	Difference*	P value	Mean	SD	Mean	SD	Difference*	P value
Mand L 1st molar	11.31	0.61	11.17	0.64	0.14	0.07	11.38	0.62	11.08	0.63	0.30	0.00
Mand L 2nd premolar	7.54	0.47	7.34	0.48	0.20	0.00	7.43	0.43	7.22	0.44	0.21	0.00
Mand L 1st premolar	7.45	0.45	7.26	0.45	0.19	0.00	7.47	0.45	7.26	0.46	0.21	0.00
Mand L canine	7.09	0.26	6.96	0.35	0.13	0.08	7.14	0.29	6.93	0.29	0.21	0.00
Mand L lat incisor	6.39	0.57	6.19	0.56	0.20	0.00	6.29	0.25	6.09	0.26	0.20	0.00
Mand L cent incisor	5.84	0.43	5.71	0.50	0.13	0.07	5.77	0.45	5.56	0.45	0.21	0.00
Mand R cent incisor	5.69	0.52	5.54	0.49	0.15	0.04	5.68	0.49	5.47	0.49	0.21	0.00
Mand R lat incisor	6.37	0.31	6.17	0.31	0.20	0.00	6.25	0.41	6.05	0.40	0.20	0.00
Mand R canine	7.21	0.40	7.00	0.39	0.21	0.00	7.30	0.57	7.11	0.57	0.19	0.00
Mand R 1st premolar	7.59	0.37	7.40	0.37	0.19	0.00	7.53	0.31	7.33	0.32	0.20	0.00
Mand R 2nd premolar	7.51	0.36	7.30	0.37	0.21	0.00	7.27	0.32	7.07	0.32	0.20	0.00
Mand R 1st molar	11.55	0.66	11.16	0.65	0.39	0.00	11.30	0.79	11.10	0.78	0.20	0.00
Max R 1st molar	10.07	0.47	9.87	0.46	0.20	0.00	9.96	0.37	9.75	0.38	0.21	0.00
Max R 2nd premolar	7.12	0.39	6.92	0.41	0.20	0.13	7.03	0.41	6.83	0.41	0.20	0.00
Max R 1st premolar	7.52	0.39	7.41	0.56	0.11	0.12	7.57	0.42	7.35	0.43	0.22	0.00
Max R canine	8.08	0.48	7.97	0.56	0.11	0.00	8.07	0.56	7.84	0.59	0.23	0.00
Max R lat incisor	7.01	0.75	6.86	0.75	0.15	0.04	7.05	0.72	6.84	0.72	0.21	0.00
Max R cent incisor	8.85	0.66	8.71	0.73	0.14	0.00	8.91	0.61	8.77	0.58	0.14	0.03
Max L cent incisor	8.94	0.63	8.76	0.63	0.18	0.60	8.96	0.58	8.75	0.56	0.21	0.00
Max L lat incisor	7.21	0.72	7.14	0.67	0.07	0.00	7.21	0.66	7.00	0.65	0.21	0.00
Max L canine	8.19	0.46	8.00	0.46	0.19	0.00	8.19	0.42	7.98	0.42	0.21	0.00
Max L 1st premolar	7.61	0.49	7.42	0.50	0.19	0.00	7.58	0.45	7.37	0.46	0.21	0. 00
Max L 2nd premolar	7.22	0.37	7.03	0.38	0.19	0.00	7.10	0.41	6.90	0.41	0.20	0.00
Max L 1st molar	10.10	0.46	9.90	0.46	0.20	0.00	9.98	0.51	9.79	0.51	0.19	0.00
Max intercanine	34.35	1.78	34.23	1.78	0.12	0.00	34.34	1.80	34.22	1.81	0.12	0.00
Max interpremolar	34.63	2.02	34.52	2.01	0.11	0.00	34.65	2.03	34.66	1.94	− 0.01	0.00
Max intermolar	44.99	2.54	44.83	2.54	0.16	0.00	45.00	2.63	44.88	2.63	0.12	0.89
Mand intercanine	26.71	1.58	26.57	1.57	0.14	0.00	26.70	1.52	26.56	1.52	0.14	0.00
Mand interpremolar	28.86	1.85	28.73	1.86	0.13	0.00	28.87	1.88	28.76	1.87	0.11	0.00
Mand intermolar	39.78	2.25	39.66	2.25	0.12	0.00	39.74	2.15	39.64	2.30	0.10	0.29
Overbite	5.43	2.24	5.22	2.24	0.21	0.00	5.52	2.17	5.30	2.17	0.22	0.00
Overjet	3.51	1.33	3.20	1.32	0.31	0.00	3.68	1.28	3.37	1.29	0.31	0.00

Significant at $P < .05$; Max: maxillary, Mand: mandibular, L: left; R: right. [†]Time trial no. 2 of each examiner randomly selected for comparison. * + value = plaster is larger.

found in most variables measured for both examiners. The measurement differences were clinically insignificant for examiner 1 (range 0.00-0.12 mm) and examiner 2 (range 0.00-0.09 mm). Similar results were obtained by Stevens et al.[10] (2006), observing a good reliability for the measurements obtained in OrthoCad and plaster models. Quimby et al.[2] (2004) demonstrated statistical differences between the plaster and digital models in all their measurements with the exception of the mandibular intercanine width. The measurements made on computer-based models showed greater variation in all categories except overbite and overjet. Most measurements differed by less than 1 mm. These findings are similar to those of the present study where computer-based models appeared to be a clinically acceptable alternative to conventional plaster models.

Validity of Cécile3 digital models

Validity was considered as the extent to which the Cécile3 digital models measured against the plaster models (gold standard).[16] Plaster and Cécile3 digital models presented differences in mesiodistal tooth width measurements, intercanine distance, interpremolar distance, intermolar distance, overjet and overbite. The mean differences between the plaster and Cécile3 digital models for examiner 1 had a range between 0.07 mm and 0.39 mm; for examiner 2, the range was between 0.01 mm and 0.31 mm. Most of the obtained values were statistically different (Table 2). Santoro et al.[17] (2003) found similar results, where the mean differences ranged from 0.16 mm to 0.49 mm and were all statistically significant with the exception of overbite. Garino, Garino[8] (2002) and Rheude et al.[6] (2005) found that the measurements made from digital models were clinically acceptable, with reasonable reliability and reproducibility and adequate clinical information for diagnosis and treatment planning, thus eliminating the need for plaster models. Oliveira et al.[1] (2007) did not find significant differences with the exception of the mesiodistal width of the lower right second premolar (p < 0.05). Zilberman et al.[9] (2003) and Quimby et al.[2] (2004) stated that even when no significant differences were found, these appeared to be clinically acceptable. In the present

study, all values obtained from the Cécile3 digital models were smaller than those obtained from the plaster models. Similar results were found by Mullen et al.[11] (2007), who observed that the measurements on the ball-bearing mounted models were on average 0.067 mm greater in the e-model software than the direct measurements obtained on the casts (range: 0 to −0.16 mm, p < .0045).

Limitations of Cécile3 digital models

Differences can be explained by the difficulty in locating the points, especially at the level of the interproximal contacts, which is affected by the operator's experience in handling a digital model. One disadvantage of digital models is that they have to be static in order to locate or mark the points needed to obtain a measurement.[1] A digital model can be blown up in the computer screen, which gives a significant advantage on locating landmarks because a 3-dimensional structure is viewed as a 2-dimensional image.[14] In this study, the same difficulties and advantages were experienced by the examiners. Prior training is required to use the software, since those more familiar with the computer resources are more capable of achieving more precise measurements. An extra difficulty observed was the presence of shadows (especially in crowded areas) in Cécile3 digital models resulting from the digitalization process. Also, occlusal anatomy and wear facets in Cécile3 digital models did not present a high definition. With respect to overbite and overjet, it was shown that these measurements were influenced by how the digital models were mounted in maximum intercuspidation on the computer. Finally, it was observed that Cécile3 digital models took considerably less time than plaster models to measure. This represents a more efficient way of performing diagnosis, as mentioned by Zilberman et al.[9] (2003).

Conclusions

Cecile3 presented a high reproducibility and reliability in obtaining different measurements in models.

Within the limitations of this study, the validity of the measurements obtained from Cécile3 digital models compared to those obtained from plaster models could be considered clinically acceptable.

References

1. Oliveira DD, Ruellas ACO, Drummond MEL, Pantuzo MCG, Lanna ÂMQ. Confiabilidade do uso de modelos digitais tridimensionais como exame auxiliar ao diagnóstico ortodôntico: um estudo piloto. Rev Dent Press Ortodon Ortop Facial. 2007;12(1):84-93.

2. Quimby ML, Vig KWL, Rashid RG, Firestone AR. The accuracy and reliability of measurements made on computer-based digital models. Angle Orthod. 2004;74(3):298-303.

3. Martensson B, Ryden H. The holodent system, a new technique for measurement and storage of dental casts. Am J Orthod Dentofacial Orthop. 1992;102(2):113-9.

4. Schirmer UR, Wilshire WA. Manual and computer aided space analysis: a comparative study. Am J Orthod Dentofacial Orthop. 1997;112(6):676-80.

5. Champagne M. Reliability of measurements from photocopies of study models. J Clin Orthod. 1992;26(10):648-50.

6. Rheude B, Sadowsky L, Ferriera A, Jacobson A. An evaluation of the use of digital study models in Orthodontic Diagnosis and Treatment Planning. Angle Orthod. 2005;75(3):300-4.

7. Tomassetti JJ, Taloumis LJ, Denny JM, Fischer JR. A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. Angle Orthod. 2001;71(5):351-7.

8. Garino F, Garino GB. Comparison of dental arch measurements between stone and digital casts. World J Orthod. 2002;3(3):250-4.

9. Zilberman O, Huggare JAV, Parikakis KA. Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. Angle Orthod. 2003;73(3):301-6.

10. Stevens D, 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(6):794-803.

11. Mullen SR, Martin CA, Ngan P, Gladwin M. Accuracy of space analysis with emodels and plaster models. Am J Orthod Dentofacial Orthop. 2007;132(3):346-52.

12. Moorees CFA, Reed RB. Biometrics of crowding and spacing of the teeth in the mandible. Am J Phys Anthropol. 1954;12(1):77-88.

13. Keene HJ, Engel G. The mandibular dental arch, part IV: prediction and prevention of lower anterior relapse. Angle Orthod. 1979;49(30):173-80.

14. Bushang PH, Dermirjian A, Cadotte L. Permanent mesiodistal tooth size of French-Canadians. J Can Dent Assoc. 1988;54(6):441-4.

15. Santoro M. Mesiodistal dimensions and tooth size discrepancy of the permanent dentition of Dominican americans. Angle Orthod. 2000;70(4):303-7.

16. 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(2):139-47.

17. Santoro M, Galkin S, Teredesai M, Nicolay O, Cangialosi T. Comparison of measurements made on digital and plaster models. Am J Orthod Dentofacial Orthop. 2003;124(1):101-5.