ORIGINAL ARTICLE

Comparison of peer assessment rating (PAR) index scores of plaster and computer-based digital models

Matthew Mayers,[a] Allen R. Firestone,[b] Robert Rashid,[c] and Katherine W. L. Vig[d]
Columbus, Ohio

Introduction: The peer assessment rating (PAR) index is a valid and reliable tool for measuring malocclusion on plaster models, but it has not been shown to be valid and reliable when used to score computer-based digital models. The purpose of this study was to determine whether the PAR index is a valid and reliable measure on digital models. **Methods:** The study sample consisted of 48 pairs of plaster and digital pretreatment models. One examiner, calibrated in the PAR index, scored the digital and plaster models. The overall PAR scores were examined for reliability and validity by using analysis of variance and the intraclass correlation coefficient (ICC). Reliability of the components of the PAR score was compared with values originally presented by Richmond et al (1992). **Results:** No significant differences were found between overall PAR scores of plaster and digital models ($P = .82$), and scores were highly correlated (ICC = 0.95; lower confidence boundary (LCB) = 0.92; upper confidence boundary (UCB) = 0.97). Intraexaminer reliability was excellent for both plaster models (ICC = 0.98; LCB = 0.97; UCB = 0.99) and digital models (ICC = 0.96; LCB = 0.94; UCB = 0.98). Reliability of all components of the PAR score generated on digital models except for buccal occlusion was similar to those of Richmond et al. **Conclusion:** PAR scores derived from digital models are valid and reliable measures of occlusion. (Am J Orthod Dentofacial Orthop 2005;128:431-4)

Digital, computer-based study models are an alternative to traditional plaster models. They are electronic records of a patient's teeth and occlusion. The software draws a 2-dimensional projection (a pseudo 3-dimensional image) on the computer screen. Software features allow the observer to change views and rotate or tilt the image of the models together or individually. The software also allows the user to measure and store arch circumferences, tooth diameter, and so forth. Digital models offer the advantages of electronic storage and access from a remote location, and in many locations, reducing the chance of loss or damage. Because electronic study models are a component of the digital orthodontic record, they contribute to a paperless office.

Disadvantages of digital models include loss of data due to degradation of electronic storage media over time, dependence on the supplier for technical support of the software, and the fact that digital models lack true 3-dimensional presentation.

Currently, 2 companies provide digital models; their processes are similar. Diagnostic impressions of the patient's dentition are taken and mailed to the company (overnight delivery). According to company policy, impressions are poured the day of arrival to produce the plaster models. A laser is used to scan the models and generate a set of data points, which are returned electronically. Proprietary software provided by the manufacturer uses the data points to generate a representative, digital, study-model image that can be viewed, manipulated, and measured. Digital models have been shown to be a valid tool for making simple diagnostic measurements such as Bolton ratio, tooth size, arch width, overjet, overbite, and arch length.[1-4] Although digital models from 1 company are inaccessible to the software of the other, both companies provide the software at no cost.

The peer assessment rating (PAR) is an occlusal index used on study models to determine initial degree of malocclusion, results of treatment, and amount of improvement,[5] and it is a reliable and valid tool.[6] The index measures severity of malocclusion and the degree to which an occlusion deviates from normal or ideal. The PAR index has been used to audit treatment quality and as an educational and research tool for review and

From the College of Dentistry, Ohio State University, Columbus.
[a]Resident, Section of Orthodontics.
[b]Associate professor, Section of Orthodontics.
[c]Associate professor, clinical, Section of Restorative and Prosthetic Dentistry.
[d]Professor and chair, Section of Orthodontics, College of Dentistry.
Reprint requests to: Dr Allen Firestone, College of Dentistry, 305 West 12th Ave, PO Box 182357, Columbus, Ohio 43218-2357; e-mail, firestone.17@osu.edu.
Submitted, December 2003; revised and accepted, April 2004.
0889-5406/$30.00

doi:10.1016/j.ajodo.2004.04.035

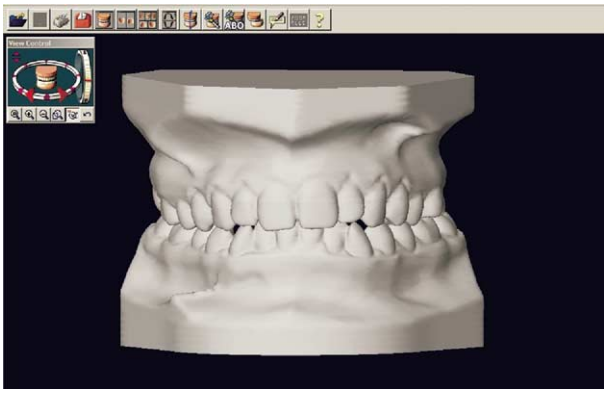

Fig 1. Overbite and midline measured from frontal view.

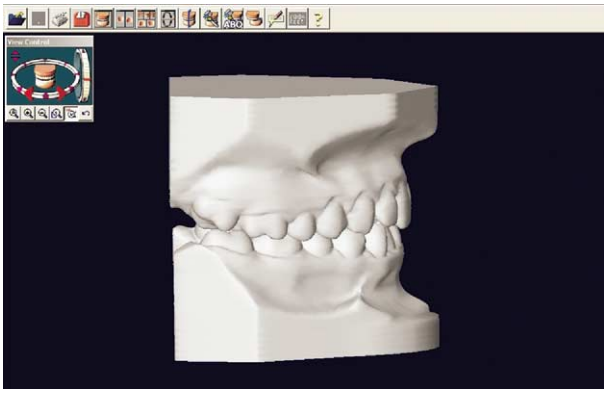

Fig 2. Buccal occlusion measured from side view. Image was rotated and enlarged to identify crossbite.

assessment of treatment outcome.[7-15] Components of the PAR index include overbite, overjet, midline, maxillary and mandibular anterior segments, and right and left buccal occlusion. PAR index scores derived from digital models have not been shown to be as valid and reproducible as scores from plaster models.

The purpose of this study was to determine the validity and reliability of PAR occlusal index scores derived from digital study models compared with scores from plaster models of the same patients. The specific aims were to compare (1) PAR scores generated from digital models with the PAR scores from orthodontic plaster casts, (2) intraexaminer reliability of PAR scores from plaster models and digital models, (3) the reliability of component scores of the PAR index generated from plaster and digital models, and reliability data originally published by Richmond et al.[6]

MATERIAL AND METHODS

The Ohio State University Institutional Review Board approved the experimental protocol. Pretreatment models were obtained from The Ohio State University Graduate Orthodontic clinic. Forty-eight models used in a previous study[2] were available, representing initial records. The models met the following inclusion criteria: (1) a full complement of permanent teeth including first molars, premolars, canines, and incisors, with no retained deciduous teeth; (2) apparently normal crown morphology (casts showing gross abnormalities were excluded); and (3) no features that would alter the natural mesiodistal or buccolingual crown diameter, such as restorations, caries, attrition, and fracture.

The models were obtained by taking 2 sets of alginate impressions (Jeltrate [fast set]; Dentsply, Milford, Del) and 2 bite registrations (Aluwax; Aluwax Dental, Grand Rapids, Mich) in maximum intercuspation of each patient. One examiner took all impressions using plastic trays (Imperial Plastic Trays, Hanover, Germany).[2] The impressions were assigned randomly to the plaster and digital groups. One set of impressions was poured in white orthodontic stone (WhipMix Corp, Louisville, Ky) within an hour of impression-taking. The second set was wrapped in a moist paper towel, sealed in a plastic bag, and sent via overnight currier service to the manufacturer (Cadent Inc, Fairview, NJ). According to the manufacturer, impressions were poured on the day of arrival. The corresponding digital models of impressions were returned via electronic mail.

Measurements used were those described by Richmond et al[6] in computing the PAR index. European weightings were used to determine the PAR scores. Stone casts were scored with the PAR ruler, also described by Richmond et al,[6] and the digital models were measured by using the proprietary software (version 2.017) from the manufacturer (Figs 1-4). The monitors used to view the digital models were 19-in OrthoCAD CRT Viewsonic Optiquest z90, with a resolution of 1024 $\times$ 768 pixels and 32-bit color (Cadent, Carlstadt, NJ).

Both plaster and digital models were identified by a 6-digit code. The 48 models from each sample were randomized and scored by 1 examiner (M.M.) calibrated in the PAR index. Measurements were made in sets of 5, scoring 5 models from each sample in succession until all were scored. After 2 weeks, the 48 models were rerandomized and rescored.

The validity of measurements made on digital models was evaluated by using 1-way analysis of variance (JMP statistical software version 5.0; SAS Institute, Cary, NC) and the intraclass correlation coefficient (SAS statistical software version 8.2).[16] Reliability of the measurements of total PAR score and

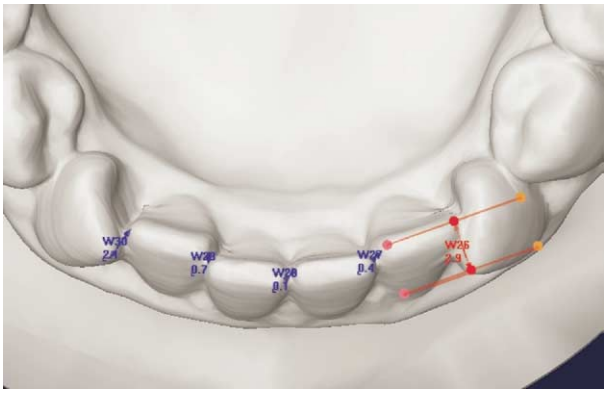

Fig 3. Displaced contacts measured from mandibular occlusal view. Maxillary displaced contacts were measured similarly.

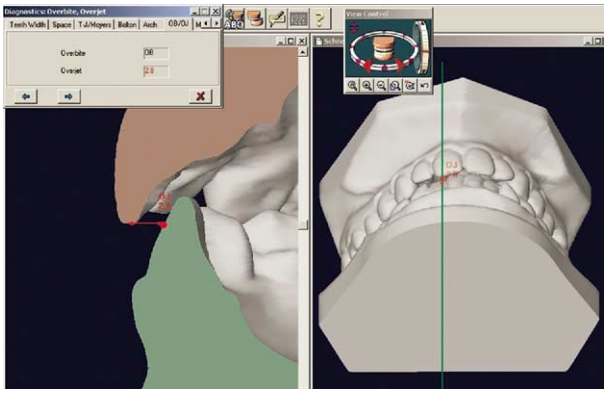

Fig 4. Overjet measured from this view. Greatest individual overjet of each incisor was measured.

component scores was evaluated by using the intraclass correlation coefficient.

RESULTS

Table I presents results for the weighted, total PAR scores. No significant differences were found between PAR scores from digital models and those from stone models ($P = 0.82$).

There was a high correlation (intraclass correlation coefficient [ICC] = 0.95; lower 95% confidence boundary [LCB] = 0.92; upper 95% confidence boundary [UCB] = 0.97) of PAR scores derived from digital models with those from stone models of the same patient.

Table II summarizes data for intraexaminer reliability, which ranged from excellent to perfect for stone models (0.98) and digital models (0.96).[16]

Table III is a tabular comparison of the reliability of different components of PAR as presented by Richmond et al[6] in validation of the PAR index and reliability of its components in this study. Statistical inferences could not be made because the standard error of the original sample was not given.

Table I. Mean, weighted PAR scores, range, and SD derived from digital and plaster models (n = 48)

	Mean	*SD*	*Range*
Digital models	27.25*	11.49	3-62
Plaster models	27.35*	12.75	6-54

*Not significant ($P = .82$).

Table II. Intraexaminer reliability for PAR index scores derived from digital and plaster models (n = 48)

	ICC	*LCB*	*UCB*
Digital models	0.96	0.94	0.98
Plaster models	0.98	0.97	0.99

ICC, Intraclass correlation coefficient; *LCB,* lower confidence boundary; *UCB,* upper confidence boundary.

DISCUSSION

Results of this study support the validity and reliability of PAR index scores derived from digital models. There were no statistically significant differences in PAR scores from digital models compared with PAR scores from plaster models of the same patient. The correlation of PAR scores was high (ICC = 0.95), with the plaster models used as the standard.[17]

Intraexaminer reliability scores were 0.96 for digital models and 0.98 for stone models, both above 0.93. Richmond et al[6] gave 0.93 as the lower 95% confidence limit of intraexaminer reliability in validating the PAR index. The reliability for the different components of the PAR index were also similar between model types, except for buccal occlusion. Reliability of the digital model for buccal occlusion was slightly lower than that reported by Richmond et al.[6]

As an aside, we originally assumed that scoring digital models would be faster than scoring stone models. However, it took about 44 minutes to score 10 sets of stone models and 72 minutes to score 10 sets of digital models. Measurement times did not progressively decrease because of any learning effect. The difference in measurement times was because the software was not amenable to efficiently scoring a model. The necessity to score individual overjets of all 4 incisors to determine the greatest degree of overjet was found to be the most time-consuming. These times do not include time spent retrieving and refiling the stone models, whereas the digital models were immediately available.

Table III. Intraexaminer reliability (intraclass correlation coefficient) for individual components of PAR index for the present study and for Richmond et al[6]

	Richmond et al (1992)					*Current study*	
Rater no.	*1*[a]	*2*[a]	*3*[a]	*4*[a]	*Mean*	*Plaster models*[b]	*Digital models*[b]
Maxillary anterior	0.94	0.98	0.98	0.95	0.96	0.93	0.93
Mandibular anterior	0.83	0.96	0.97	0.91	0.92	0.95	0.95
Right buccal occlusion	0.79	0.89	0.87	0.89	0.86	0.95	0.77
Left buccal occlusion	0.83	0.91	0.96	0.77	0.87	0.85	0.68
Overjet	0.95	0.99	0.87	0.94	0.94	0.98	0.97
Overbite	0.74	1	0.82	0.89	0.86	0.77	0.8
Centerline	0.76	0.89	0.9	0.83	0.85	0.88	0.76
Total score	0.96	0.95	0.96	0.95	0.95	0.98	0.96
Upper 95% CI	NR	NR	NR	NR	NR	0.99	0.98
Lower 95% CI	0.94	0.97	0.94	0.93	0.94	0.97	0.94

[a]Four individual raters as reported by Richmond et al.[6]
[b]One rater in present study.

This study demonstrates that digital models are a valid, reliable medium for measuring malocclusion with the PAR index, but there are some limitations. The sample contained a large range of malocclusions, but certain traits were not observed or measured. There were no deciduous teeth or impactions in the sample[2] and no lateral open bites, but all other measurable occlusal traits of the PAR index were present.

Because a malocclusion can be reliably scored on digital models by using the PAR index, as demonstrated in this study, a computer-based calibration exercise could be made available electronically for examiners who are already trained in the PAR index. This would provide a standard for a previously calibrated examiner to update his or her skills. It should also be possible to provide a computer-based instruction module (including digital models) as a practice exercise for people wishing to become calibrated.

CONCLUSION

The PAR index scores derived from digital, computer-based models are valid and reliable measures of malocclusion.

REFERENCES

1. Tomassetti J, Taloumis L, Denny J, Fisher J. A comparison of 3 computerized Bolton tooth-size analyses with a commonly used method. Angle Orthod 2001;71:351-7.
2. Quimby M, Vig K, Rashid R, Firestone A, Mayers M. The accuracy and reliability of measurements made on computer-based digital models. Angle Orthod 2004;74:298-303.
3. 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;124:101-5.
4. Zilberman O, Huggare J, Parikakis K. Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. Angle Orthod 2003;73:301-6.
5. Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR index (peer assessment rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. Eur J Orthod 1992;14:180-7.
6. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, et al. The development of the PAR index (peer assessment rating): reliability and validity. Eur J Orthod 1992; 14:125-39.
7. Beatrice M, Woods M. Vertical facial pattern and orthodontic stability. Part II: facial axis changes and stability. Aust Orthod J 2000;16:133-9.
8. Cassinelli AG, Firestone AR, Beck FM, Vig KW. Factors associated with orthodontists' assessment of difficulty. Am J Orthod Dentofacial Orthop 2003;123:497-502.
9. DeGuzman L, Bahiraei D, Vig KW, Vig PS, Weyant RJ, O'Brien K. The validation of the peer assessment rating index for malocclusion severity and treatment difficulty. Am J Orthod Dentofacial Orthop 1995;107:172-6.
10. Firestone AR, Beck FM, Beglin FM, Vig KW. Evaluation of the peer assessment rating (PAR) index as an index of orthodontic treatment need. Am J Orthod Dentofacial Orthop 2002;122:463-9.
11. King GJ, McGorray SP, Wheeler TT, Dolce C, Taylor M. Comparison of peer assessment ratings (PAR) from 1-phase and 2-phase treatment protocols for Class II malocclusions. Am J Orthod Dentofacial Orthop 2003;123:489-96.
12. Mascarenhas AK, Vig K. Comparison of orthodontic treatment outcomes in educational and private practice settings. J Dent Educ 2002;66:94-9.
13. Pinto N, Woods M, Crawford E. Vertical facial pattern and orthodontic stability. Part I: pretreatment vertical pattern and stability. Aust Orthod J 2000;16:127-32.
14. Vig KW, Weyant R, Vayda D, O'Brien K, Bennett E. Orthodontic process and outcome: efficacy studies—strategies for developing process and outcome measures: a new era in orthodontics. Clin Orthod Res 1998;1:147-55.
15. von Bremen J, Pancherz H. Efficiency of early and late Class II Division 1 treatment. Am J Orthod Dentofacial Orthop 2002; 121:31-7.
16. Rosner B. Fundamentals of biostatistics. Boston: Duxbury Thomson Learning; 2000.
17. Russell CM, Williamson DF, Bartko JJ, Bradley EL. Simulation study of a panel of reliability indicators applied to paired measurements. Am J Human Biol 1994;6:311-20.