Original Article

The reliability of assessing rotation of teeth on photographed study casts

Fleur M.J. Vermeulen[a]; Irene H.A. Aartman[b]; Reinder Kuitert[c]; Andrej Zentner[d]

ABSTRACT

Objective: To examine the intra- and interexaminer reliability of assessing rotation of teeth on photographed study casts. In addition, the reliability parameters of two examiners scoring in mutual consultation were compared with the reliability parameters by one observer.

Materials and Methods: Standardized photographs of sets of maxillary and mandibular plaster casts of 10 patients before treatment (T1), after treatment (T2), and a long time after retention (T3) were digitized. Tooth rotation was assessed relative to a correct position in the ideal dental arch form. A computer analysis program was used to process the measurements. Two examiners assessed each study cast twice with a washout period of 3 weeks. A third examiner assessed each cast, together with one of the other examiners. The intra- and interexaminer agreements were calculated using intraclass correlation coefficients (ICCs). Wilcoxon signed-rank tests were used to determine significant differences between the intra- and interexaminer reliability coefficients of the three examiners.

Results: The ICCs ranged from 0.430 to 0.991. Incisors showed the highest ICCs and molars showed the lowest ICCs. Intraexaminer ICCs of the experienced examiners were significantly higher than those of the examiner with less experience. No significant differences in the reliability between a single examiner and the combination of two examiners were found.

Conclusion: The method of assessing tooth rotation in the present study has proved to be reproducible, except for the molars. This method can be helpful for clinicians assessing tooth rotation from photographed study casts. (*Angle Orthod.* 2012;82:1033–1039.)

KEY WORDS: Tooth rotation; Reliability

INTRODUCTION

The relapse of orthodontically rotated teeth can be a serious problem because nearly all rotated teeth tend to develop some degree of postretention relapse. Even in the most successfully treated orthodontic patient relapse of a single rotated tooth can contribute to slipped contact, drifting, and undesirable inclined plane relations.[1]

Contrary to the conscientious assessment of crowding and arch dimension, the assessment of rotations has not often been subject of thorough examination. Relatively few studies on tooth rotation mention the amount or the direction of the rotations.[2–7] The assessment of reliability is addressed in even fewer publications.[3,5–7] In a dental arch with crowding rotations are often present, but in cases of space excess rotations might also occur. The putative relationship between rotation and space condition in the dental arch remains to be quantified and investigated. Careful quantification of rotation is, however, essential for this kind of research.

The literature describes several methods of assessing tooth rotation on study casts. One of the first studies in which tooth rotation was assessed, measured the rotation of erupting lower premolars on study models using a clear plastic-arm protractor.[2] All rotated premolars had the tendency to rotate in the mesial direction ranging between 3° and 26°. Howev-

[a] Postgraduate student, Department of Orthodontics, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, The Netherlands.

[b] Assistant Professor, Methodologist, Social Dentistry and Behavioral Sciences, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, The Netherlands.

[c] Assistant Professor, Department of Orthodontics, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, The Netherlands.

[d] Professor, Department of Orthodontics, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, The Netherland.

Corresponding author: Dr Fleur Vermeulen, Department of Orthodontics, Academic Centre for Dentistry Amsterdam, Gustav Mahlerlaan 3004, 1081 LA Amsterdam, The Netherlands (e-mail: f.vermeulen@acta.nl)

Accepted: February 2012. Submitted: November 2011.
Published Online: April 12, 2012

er, the measurement error was not calculated in that study.

In addition, in another study the postretention lower anterior rotations were measured with the midsagittal plane as the reference line.[3] This study reported an intraexaminer average measurement error of 1.4° and an average postretention change of 3.3°. Another study using the midsagittal reference line found a mean rotation of −3.6° in the distal direction for the canines and 1.9° in the mesial direction for the second premolars. However, the ranges and standard deviations were relatively large, and measurement error was not taken into account.

A different approach was used in a later study. In this study five standardized arch forms (Rocky Mountain Pentamorphic Arches) were superimposed on the study casts as a reference line to measure premolar rotation in orthodontically nontreated patients.[5] The method error was calculated as ±2.95° for the intraexaminer reliability using the Dahlberg's formula.[8] The mean rotations varied between 1.4° and 12.9°. Others used a similar method in which computer-generated arch forms served as a reference to measure rotation of the upper incisors and canines.[6] The authors found an intraexaminer method error of 0.72° using Dahlberg's formula. However, the computer-generated arch forms did not consistently represent the actual dental arch because even in perfectly aligned arches tooth rotations were found. In that study the average rotation varied between 3.3° and 11.0°.

In a more recent study the relapse of the corrected rotation of the six maxillary anterior permanent teeth was measured with the palatal raphe as a reference line.[7] The intraexaminer measurement errors as indicated by Dahlberg's formula[8] were 3.1° for the canines, 2.8° for the lateral incisors, and 2.4° for the central incisors. The average relapse varied between 6° and 9° and ranged from 0° to 20°. The authors had some difficulties achieving an exact location of the raphe line on the pretreatment, posttreatment, and postretention study casts. Besides, changes in the arch form, which frequently occur during the treatment and posttreatment periods, may have influenced the measurements[9], and the correct position of teeth within the dental arch is unknown.

Because of the aforementioned limitations of measuring tooth rotation in longitudinal studies, it was considered important to establish a meaningful and reproducible method of assessing tooth rotation. For the purpose of the present study a computer analysis program was developed to examine tooth rotation on study casts of permanent dentitions relating the teeth to a correct position in the dental arch. The aim of this study was to develop a meaningful, objective, and reproducible method of assessing the rotation of first molars, premolars, and anterior teeth on study casts. To this end, the reliability and measurement error of rotation of these teeth was assessed on photographed study casts. A second aim of this study was to compare the reliability parameters of the measurements of two examiners deciding in mutual consultation on the points to be registered with the reliability parameters of the measurements made by one observer.

MATERIALS AND METHODS

Sample

The sample comprised 10 sets of study casts of Class I and Class II malocclusions, all treated nonextraction, selected from the archives of the Department of Orthodontics of the Academic Centre for Dentistry Amsterdam. Study casts taken before (T1) and after (T2) orthodontic treatment and between 129 and 181 months after removal of the retention appliances (T3) were selected. All casts had a full complement of permanent incisors, canines, first premolars, and first molars at all time points. The study casts were all of good quality and accurately and uniformly trimmed with their base parallel to the occlusal plane. Photographs were taken with the lens parallel to the occlusal plane and with the central axis aimed at the midpoint of the lines connecting the left and right contact points between the canines and first premolars.

Measurements on Study Casts

To draw the arch form for the upper and lower jaw, first seven points were assessed on the digital photograph. The locations of these points were as follows: (1) distal from the first right molar, in a way that the extending line will go through or parallel to the most occlusal points of the buccal cusps; (2) between the two right premolars, in a way that the extending line will go through or parallel to the most occlusal points of the buccal cusps of the premolars and the canine; (3) mesial from the right canine, in a way that the extending line will go through the most incisal point of the canine and the incisal edges of the incisors; (4) between the two central incisors, in a way that the extending line will go through the incisal edges of the incisors; and (5–7) the same points as at the contralateral quadrant (Figure 1).

In this approach displaced teeth were ignored. The arch form was generated by the computer using the cubic spline formula.[10] Two points on this generated ideal arch were assessed to indicate the ideal position of the individual tooth. Tooth rotation relative to the

Figure 1. Photograph of a study cast at T2, with the seven points to draw the arch form.

dental arch was measured for each quadrant from central incisor to first molar by placing points on the mesial and distal ideal contact points of the tooth in a way that the line through these points is perpendicular to the buccolingual axis of the tooth (Figure 2). The angle of rotation is the angle between the mesiodistal line and the line indicating the ideal position on the arch. Mesial rotation means that the distal part of the crown is displaced outward and the mesial part is displaced inward relative to the dental arch. Distal rotation means that the mesial part of the crown is displaced outward and the distal part is displaced inward relative to the dental arch (Figure 3).

The Examiners

Examiner A and examiner B measured the study casts independently. The measurements were also carried out by examiner A and examiner C together (examiner AC). Examiner A was a qualified dentist with one year of clinical experience in general dentistry, examiner B was a dental student at the masters degree level, and examiner C was a specialist orthodontist with

Figure 2. Photograph of a study cast at T1, with the ideal arch form in black and the position of the teeth in white lines.

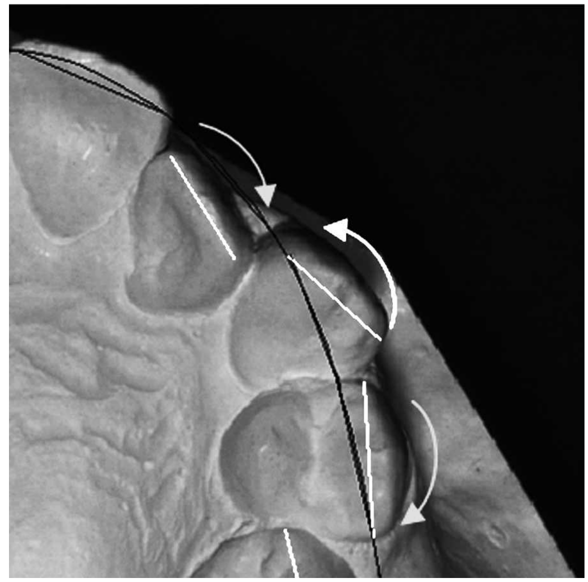

Figure 3. Rotation of teeth indicated by arrows. In this photograph the central incisor is not rotated, the lateral incisor is rotated to the distal, and the canine is rotated to the mesial.

35 years of clinical experience and experience with previous research assessing tooth rotation. The same examiners repeated the measurements within a 3-week interval. To calibrate the three examiners to a uniform measuring method, all measurements were performed only after intensive group instruction and training.

Statistical Analysis

Intra- and interexaminer reliability were evaluated by calculating intraclass correlation coefficients (ICCs). The average estimate of the ICC was calculated for each tooth separately, taking the different time points of the models together. Interexaminer reliability was evaluated comparing pairs of two different examiners. Wilcoxon signed-rank tests were used to determine significant differences between the intra- and interexaminer reliability coefficients (ICCs) of the three examiners. The Statistical Package for Social Sciences Windows, version 15.0 (SPSS Inc, Chicago, Ill) was used for the analyses. In addition, both intra- and interexaminer errors were evaluated. The random measurement error was calculated using Dahlberg's formula.[8]

$$S = \sqrt{\frac{\sum d^2}{2n}}$$

All calculations were carried out for the individual teeth.

RESULTS

Table 1 shows the minimum, maximum, mean, and standard deviation (SD) of the angles measured by all

Table 1. Mean (±Standard Deviation), Minimum and Maximum Angles in Degrees per Cast per Selected Tooth Measured by All Examiners at T1

		Study Cast									
Tooth	Measurement	1	2	3	4	5	6	7	8	9	10
11	Mean	−8.4	0.3	−20.2	−2.0	0.2	5.7	7.4	−9.8	−3.0	1.5
	SD	1.9	0.5	2.5	2.6	1.4	1.5	2.1	2.5	2.8	0.4
	Minimum	−11.1	−0.2	−23.2	−6.2	−0.4	4.7	4.2	−11.3	−7.4	−1.8
	Maximum	−6.3	1.3	−16.9	−0.1	1.6	7.5	9.2	−8.4	0.0	6.8
12	Mean	18.3	−23.7	−10.9	−15.1	−18.7	2.4	12.9	−2.2	4.6	−21.3
	SD	1.9	1.4	2.1	1.5	1.3	2.4	1.8	2.6	3.2	4.4
	Minimum	16.3	−25.4	−13.3	−13.7	−20.1	0.0	10.5	−6.9	0.0	−25.2
	Maximum	21.8	−22.0	−8.6	−17.8	−16.3	6.5	15	0.7	7.9	−14.9
14	Mean	1.2	0.3	−5.5	4.6	1.7	9.2	6.7	4.8	−0.8	−11.2
	SD	1.1	0.3	2.1	1.6	2.5	1.9	1.2	2.5	1.5	3.6
	Minimum	−0.4	−0.2	−7.6	2.3	−0.5	7.8	4.9	0.0	−3.3	−15.8
	Maximum	2.6	0.6	−1.9	6.9	5.1	12.7	8.1	6.8	1.9	−5.9
16	Mean	−0.8	−0.2	−8.9	0.5	−1.7	−5.5	−6.2	−4.2	−7.4	−1.5
	SD	1.1	0.4	4.8	0.6	1.9	4.7	2.8	2.5	3.4	2.7
	Minimum	−2.6	−0.9	−14	0.0	−4.6	−12.8	−9.2	−7.3	−12	−6.5
	Maximum	0.1	0.0	−2.0	1.6	0.9	0.1	−3.3	0.0	−2.7	1.2
41	Mean	−8.9	2.5	−0.3	19.4	10.1	0.8	−0.3	14.8	−15.2	5.1
	SD	1.2	2.9	0.4	4.2	2.5	1.5	0.6	2.5	3.3	1.6
	Minimum	−10.7	−0.1	−0.8	14.2	6.2	−0.1	−1.6	11.3	−17.8	2.8
	Maximum	−7.2	6.4	0.0	24.8	13.6	3.7	0.0	18.4	−8.9	7.6
44	Mean	8.7	9.9	−11.5	−2.2	0.6	−7.6	1.2	3.3	−0.4	3.5
	SD	2.5	3.4	1.4	3.2	0.4	1.5	2.0	2.1	1.6	3.7
	Minimum	6.8	6.2	−13.3	−6.9	0	−9.8	−0.6	0.2	−3.5	0.0
	Maximum	13.7	15	−9.4	0.7	1.1	−5.2	1.4	5.8	0.9	9.7
46	Mean	1.0	−1.9	3.1	1.2	3.5	2.7	−0.1	3.2	1.4	2.1
	SD	2.0	2.8	3.4	1.0	2.4	2.3	3.6	1.6	2.0	3.1
	Minimum	0.0	−6.5	−0.2	0.0	1.0	−1.7	−4.4	0.6	−0.6	−0.4
	Maximum	4.8	1.0	8.6	2.7	7.0	4.7	6.4	5.4	4.5	7.9

three observers for each cast separately at T1. Data are given for a selection of teeth only, that is, for the teeth 11, 12, 14, 16, 41, 44, and 46. Descriptive statistics for the other teeth are available on request.

The values of the three intra- or interexaminer reliability coefficients (ICCs) and the three Dahlberg's coefficients are summarized in Table 2, taking the antimeres together. Thus, the range given is based on six ICCs and six Dahlberg's coefficients. Intra- and interexaminer reliability for this measurement method of the rotation of teeth was generally high. The ICCs ranged from 0.430 to 0.991. The incisors showed the highest ICCs, with a range from 0.876 to 0.991, whereas the premolars and canines showed somewhat lower ICCs. The molars showed the lowest ICCs, with a wide range from 0.430 to 0.935.

The random measurement error (Dahlberg's coefficient) ranged from 1.18° to 5.05°. The highest Dahlberg's coefficient of 5.05 was found for the lower second premolar. With regard to the interexaminer

Table 2. Range (Minimum–Maximum) of the ICCs for Intra- and Interexaminer Reliability and the Range of Dahlberg's Coefficients[a]

	ICC		Dahlberg's Coefficient	
	Intraexaminer	Interexaminer	Intraexaminer	Interexaminer
I1 superior	0.949–0.984	0.907–0.973	1.42–2.24	1.95–2.69
I2 superior	0.956–0.991	0.899–0.984	1.35–2.68	1.87–3.16
C superior	0.899–0.959	0.834–0.978	2.12–2.67	1.74–2.33
P1 superior	0.871–0.969	0.876–0.966	1.86–2.34	1.93–2.28
P2 superior	0.881–0.957	0.775–0.956	1.81–2.47	1.70–2.88
M1 superior	0.430–0.935	0.524–0.931	1.73–2.31	1.77–2.71
I1 inferior	0.964–0.984	0.930–0.990	1.68–2.21	1.63–3.03
I2 inferior	0.922–0.960	0.876–0.971	2.24–2.74	1.92–3.29
C inferior	0.765–0.884	0.569–0.876	1.94–3.27	2.21–3.10
P1 inferior	0.936–0.984	0.895–0.988	1.18–3.25	1.89–3.78
P2 inferior	0.729–0.941	0.661–0.939	2.12–2.98	2.70–5.05
M1 inferior	0.532–0.836	0.624–0.814	1.55–2.18	1.66–3.78

[a] ICC indicates intraclass correlation coefficient.

Table 3. Median and Range of the ICCs and the Results of the Wilcoxon Signed-Rank Tests to Compare Them[a]

	Median	Range	Median	Range	*P*
Intra A–Intra AC	0.950	0.765–0.984	0.940	0.532–0.984	.685
Intra A–Intra B	0.950	0.765–0.984	0.928	0.430–0.991	.003**
Intra B–Intra AC	0.928	0.430–0.991	0.940	0.532–0.984	.056
Intra A– Inter A-AC	0.950	0.765–0.984	0.927	0.627–0.99	.166
Intra AC–Inter A-AC	0.940	0.532–0.984	0.927	0.627–0.99	.209
Inter A-AC–Intra B	0.927	0.627–0.99	0.928	0.430–0.991	.424
Inter AC-B–Intra A	0.916	0.569–0.983	0.950	0.765–0.984	.005**
Inter AC-B–Intra AC	0.916	0.569–0.983	0.940	0.532–0.984	.001***
Inter AC-B–Intra B	0.916	0.569–0.983	0.928	0.430–0.991	.032*
Inter A-B–Intra A	0.903	0.591–0.974	0.950	0.765–0.984	.000***
Inter A-B–Intra AC	0.903	0.591–0.974	0.940	0.532–0.984	.001***
Inter A-B–Intra B	0.903	0.591–0.974	0.928	0.532–0.984	.033*

[a] ICC indicates intraclass correlation coefficient; Intra A, the intraexaminer ICCs of examiner A; Intra B, the intraexaminer ICCs of examiner B; Intra AC, the intraexaminer ICCs of the combination of examiners AC; Inter A-AC, the interexaminer ICCs of examiner A with the ICCs of the combination of examiners AC; Inter A-B, the interexaminer ICCs of examiner A compared to examiner B; Inter AC-B, the interexaminer ICCs of the combination of examiners AC with examiner B.

* $P < .05$; ** $P < .01$; *** $P < .001$.

random measurement error, the Dahlberg coefficient was larger for the molar, the second lower premolar, and the lower canine, which is in line with the results of the ICCs.

In Table 3, the median values of the 24 ICCs of the 24 assessed teeth are shown. All intraexaminer ICCs were compared to each other and to the interexaminer ICCs. It appeared that the intraexaminer ICCs of examiner A were statistically significantly higher than the intraexaminer ICCs of examiner B, as calculated by the Wilcoxon signed-rank test ($P < .05$). The intraexaminer ICCs were statistically significantly higher than the interexaminer ICCs. This difference was not found between the intraexaminer ICCs A, B, and AC on the one hand and the interexaminer ICC of AC with A on the other.

In Table 4, the ranges of the intra- and interexaminer ICCs are given for T1, T2, and T3 separately over all teeth. A Wilcoxon signed-rank test indicated that the ICCs at T1 and T3 are statistically significantly higher than the ICCs at T2.

DISCUSSION

The present method of assessing tooth rotation relative to a constructed arch form based on fixed points on the dental arch appears to be reproducible, and the random measurement error based on two repeated measurements varied between 1.18° and 5.05°. Nevertheless, the method was not equally reliable for all tooth types. The reliability coefficients of the incisors were higher than the reliability coefficients of the premolars, canines, and first molars in declining order. The lower reliability of the measurement of the rotation of the first molars does not seem to be related to one of the examiners. Assessment of rotations of molars and lower second premolars appeared equally difficult for all three examiners. In particular, it was difficult to find reproducible points on the crowns of molars, a problem that was related to large restorations and wear as well as to the shape of the most distal end of the base arch. The distal end of the arch base was difficult to locate in a number of casts, and the orientation points were easier to locate on the incisors. On the basis of these results it is suggested that when the present method is used in future research on tooth rotation, the molars and lower second premolars should be assessed with caution or excluded because of the relatively high random error and low ICCs found for these teeth.

In addition, this method can be of value for clinicians because lack of storage facilities means some clinicians take photographs of study casts. Another reason to photograph study casts is to keep the photos in a digital chart in order to be able to quickly retrieve

Table 4. Median and Range of the Intra- and Interexaminer ICCs of the Different Time Points, Taking All Teeth Together, and the Results of the Wilcoxon Signed-Rank Tests to Compare Them[a]

	Median	Range	Median	Range	*P*
T1–T2	0.957	0.927–0.970	0.919	0.897–0.947	.027*
T1–T3	0.957	0.927–0.970	0.945	0.924–0.961	.104
T2–T3	0.919	0.897–0.947	0.945	0.924–0.961	.026*

[a] ICC indicates intraclass correlation coefficient; T1, before treatment; T2, at the completion of treatment; T3, long time postretention.

* $P < .05$.

them during treatment. This method proves to be a reliable method to measure tooth rotation on photographed study casts.

The second aim of this study was to assess whether the cooperation of two examiners would be more reliable than the measurements obtained by one examiner only. The results showed that there were no significant differences in reliability between a single examiner and the combination of two examiners. Thus, the combination of two examiners measuring together did not improve the reliability of the observation.

Another finding of this study was that the intraexaminer reliability of one of the examiners was significantly lower than that of both the other examiners. This may be explained by the fact that the former was less experienced and the latter had been involved with the subject for a longer time. However, the intraexaminer reliability of the experienced examiners is extremely high, and the median intraexaminer reliability (0.928) of the least experienced examiner is, in comparison to results of other studies, still high and acceptable.

There was a statistically significant difference in ICCs calculated for the three different time points. The ICCs for the casts after orthodontic treatment were lower than the ICCs for the casts before treatment and some time after removal of the retention appliances (Table 4). When teeth are more aligned and less variation in rotation is present, ICCs tend to be lower. However, the mean measurement error on aligned teeth will also be lower. The angles of rotation found on aligned teeth are small. However, even with a relatively small measurement error, the variation in terms of percentage will be greater. Nevertheless, it was decided to combine the models in the calculation of the reliability parameters.

In a previous study, a Dahlberg coefficient of 0.72° was found for the upper anterior teeth,[6] which is about one-third of the corresponding values found in the present study. This might be based on the fact that they placed three points on each tooth and calculated the arch form using the same points that were located on the teeth, and hence limiting the error merely to the location of the orientation points on the teeth. However, the method of Surbeck et al.[6] can inappropriately produce abnormal dental arches where assessed rotations differ from the real rotations in the dental arch. On the other hand, in another study,[5] a mean value of 2.95° for premolars was found, which compares fairly well to the results reported here (1.18–3.25, as is shown in Table 2). Furthermore, the values for upper laterals and centrals found by Naraghi et al.[7] are comparable with the results of the present study, too. However, compared to the values found for the upper canines,[7] the values in the present study appear somewhat more favorable. This difference might be related to the different orientation of teeth to the reference planes, such as the orientation to the palatal raphe line by Naraghi et al.[7] Finally, in an earlier study,[3] an average error in the measurement of rotation of the lower incisors of 1.4° was found, which, although not strictly comparable without the use of the Dahlberg's formula, is somewhat lower than the intraexaminer error in the present investigation.

The present study differs from earlier reports by adding the interexaminer reliability to the intraexaminer reliability, whereas other studies seem to be limited to the intraexaminer reliability of one examiner only. In addition, a third examiner pair was introduced and the reliability of this examiner pair was compared to the reliability of individual examiners. Another advantage of the present study is the calculation of ICCs, which were not reported in the earlier studies.

Further, in the assessment method described here, the rotations of the teeth are related to the ideal dental arch, independent of its later developmental changes. This method is comparable to that described by Surbeck et al.,[6] but in the present study arguably a more genuine arch was generated because it was independent of extremely deviating rotations and tooth malpositions.

The values described in Table 2 represent the average random errors. The random error for the anterior teeth varied between 1.35 and 3.29. These values seem reasonably low. However, Table 1 shows a wide range for the measurement of several teeth on several study casts (eg, a range of 7.9 for tooth 12 in cast 9). For this reason, when considering the difference between two measurements in individual clinical assessments, the coefficient of reproducibility, $2\sqrt{2} \times$ random error, should be used. In the absence of any underlying change, 95% of the differences between two measurements will generally be less than the coefficient of reproducibility. Hence, if two replicate readings are made on the same subject and the difference is less than the coefficient of reproducibility, there is no evidence of change beyond what might be explained by measurement error.[11] Compared to the postretention relapse described in the literature, the coefficient of reproducibility, which can be calculated from the random error found in this study, is acceptable. To increase the reproducibility of the measurements, more than one assessment for each measurement should be conducted.

CONCLUSIONS

- The method of assessing rotation of first molars, premolars, and anterior teeth on photographed study casts, with the aid of a computer analysis program and relating the teeth to a correct position in the dental arch, has proved to be reproducible except for the first molars.

- Improvement of reliability can be achieved by limiting the study to experienced examiners and by having at least two examiners assess all measurements twice.

REFERENCES

1. Brain WE. The effect of surgical transsection of free gingival fibers on the regression of orthodontically rotated teeth in the dog. *Am J Orthod*. 1969;55:50–70.
2. Kim YH, Shiere FR, Fogels HR. Pre-eruptive factors of tooth rotation and axial inclination. *J Dent Res*. 1961;40:548–557.
3. Kaplan RG. Mandibular third molars and postretention crowding. *Am J Orthod*. 1974;66:411–431.
4. Berg R, Gebauer U. Spontaneous changes in the mandibular arch following first premolar extractions. *Eur J Orthod*. 1982;4:93–98.
5. McMullan RE, Kvam E. Investigation of premolar rotation in a group of 15-year-old Norwegian children. *Eur J Orthod*. 1990;12:311–315.
6. Surbeck BT, Årtun J, Hawkins NR, Leroux B. Associations between initial, posttreatment, and postretention alignment of maxillary anterior teeth. *Am J Orthod Dentofacial Orthop*. 1998;113:186–195.
7. Naraghi S, Andrén A, Kjellberg H, Mohlin BO. Relapse tendency after orthodontic correction of upper front teeth retained with a bonded retainer. *Angle Orthod*. 2006;76:570–576.
8. Dahlberg G. *Statistical Methods for Medical and Biological Students*. London: George Allen and Unwin Ltd; 1940.
9. De La Cruz AR, Sampson P, Little RM, Artun J, Shapiro PA. Long term changes in arch form after orthodontic treatment and retention. *Am J Orthod Dentofacial Orthop*. 1995;107:518–530.
10. Davis LM, BeGole EA. Evaluation of orthodontic relapse using the cubic spline function. *Am J Orthod Dentofacial Orthop*. 1998;113:300–306.
11. 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;24:139–147.