Intra-examiner reliability of epidemiologic registrations of malocclusion

SVEN HELM

Institute of Orthodontics, Royal Dental College, Copenhagen, Denmark

Intra-examiner consistency in the recording of single traits of malocclusion according to a previously described method was assessed by duplicate registrations on study models of 292 children. The level of agreement was expressed by the consistency ratio and the phi coefficient. A high level of consistency was obtained for most of the malocclusion traits, and it was concluded that the method employed is appropriate also for analyses of interrelations between these traits.

Sven Helm, Institute of Orthodontics, Royal Dental College, 160 Jagtvej, DK-2100 Copenhagen, Denmark

In orthodontic investigations the reliability of quantitative assessments, e.g. measurements on cephalometric radiographs or study models, has frequently been discussed. The methodological problems of estimating the reliability of qualitative assessments, on the other hand, have received little attention. Several methods for epidemiologic registration of malocclusion are based on qualitative assessments of single morphologic variables in the dentition *(Bjørk, Krebs & Solow,* 1964; *Baume et al.,* 1973). Each individual variable is assessed according to a criterion of registration, establishing a dichotomy between normal and abnormal variation, that is, between non-occurrence and occurrence of the particular trait of malocclusion. The reliability of recording dichotomized variables may be tested at different levels.

Suppose, for example, that the presence or absence of a given trait is examined twice in 100 individuals. The hypothetical data may be summarized in a 2 x 2 contingency table, as illustrated in Table I. Inspection of the total (marginal) frequencies reveals that presence of the trait has been recorded in 33 persons at the first examination, against 35 at the second. The difference of 2% between these frequencies is small and statistically not significant, which seems to indicate that the results obtained by the registration method were reproducible. However, inspection of the cell frequencies shows that presence of the trait has only been consistently recorded in 30 persons and absence in 62, while discordant recordings (positive or negative reversals) have occurred in eight persons. Hence, actually only 92 % of the individuals have been assessed consistently.

Even this percentage may yield an overestimate of the reliability of registration,

especially where infrequent traits are concerned, because of the weight by which consistent recordings of obvious non-occurrence will contribute. This problem is taken into account in the *consistency ratio*, which is calculated by dividing the number of subjects, in whom the occurrence of a particular trait has been consistently recorded, by the total number of subjects, in whom the trait has been recorded (*Jackson*, 1966). For practical purposes the ratio is expressed in per cent. Another approach consists in calculating the *phi coefficient*, which is a coefficient of correlation calculated from 2 x 2 contingency tables (*Guilford*, 1956). Expressions for the consistency ratio (CR%) and the phi coefficient (Φ) are given in Table II.

The method for epidemiologic registration of malocclusion developed by *Bjørk, Krebs & Solow* (1964) has previously been used in a number af studies (*Helm*, 1968, 1970; *Bjørk & Helm* 1969; *Helm & Siersbaek-Nielsen*, 1973; *Ingervall*, 1974; Lavelle, 1976; *Magnusson*, (1976). The error of the method has been assessed by intra- and interexaminer comparisons, which did not reveal statistically significant differences in terms of the total frequencies (*Solow & Helm*, 1968; *Helm*, 1970). Current studies of interrelations between various malocclusion traits have made more accurate assessments of the reliability of this method pertinent. A new study of intra-examiner consistency of recording was therefore made.

MATERIAL AND METHOD

Study models and intra-oral radiographs from a random sample, comprising 292 Danish children aged 13–17 years, were available for registration. The material, which has been described in detail elsewhere (*Helm et al.*, 1975), was examined twice at an interval of six months by the author using the method of *Bjørk, Krebs & Solow* (1964). The examiner had extensive experience in the use of this method. The level of agreement was

Table I. *Results of two examinations (I and II) of presence or absence of a particular trait in 100 individuals (Hypothetical data)*

		I		
		Presence	Absence	Total
II	Presence	30	5	35
	Absence	3	·62	65
	Total	33	67	100

Table II. *Theoretical results of two examinations (I and II) of presence or absence of a particular trait in N individuals*

		I		
		Presence	Absence	Total
II	Presence	A	B	A+B
	Absence	C	D	C+D
	Total	A+C	B+D	N

Data from Table I

$$CR\% = \frac{A}{A+B+C} \times 100$$

$$CR\% = \frac{30}{30+5+3} \times 100 = 79\%$$

$$\Phi = \frac{A \times D - B \times C}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

$$\Phi = \frac{(30 \times 62) - (5 \times 3)}{\sqrt{35 \times 65 \times 67 \times 33}} = 0.82$$

expressed by the consistency ratio (CR%) and the phi coefficient (Φ), cf. Table II.

RESULTS

The findings are presented in Table III, which shows the original data together with the two measures of consistency. Where the total number of individuals, in whom a particular

Table III. *Number of recordable subjects (REC) in whom a given trait was recorded at first examination only (I), consistently at both examinations (I+II), and at second examination only (II)*

	REC	I	I+II	II	CR%	Φ
Dentition						
Aplasia	292	0	19	0	100	1.00
Malformation	292	1	7	1	–	–
Ectopic eruption	292	1	13	0	93	0.96
Persistent prim. teeth	292	1	13	0	93	0.96
Anterior crossbite	292	0	14	1	93	0.96
Rotation	292	12	77	18	72	0.76
Tipping	292	8	14	7	48	0.62
Occlusion						
Extreme max. overjet	290	2	51	0	96	0.98
Mandibular overjet	290	0	1	0	–	–
Distal molarocclusion	275	2	80	2	95	0.97
Mesial molarocclusion	275	0	5	4	–	–
Open bite	290	2	5	0	–	–
Deep bite	290	4	64	3	90	0.93
Crossbite	291	1	42	4	89	0.93
Scissors bite	291	2	31	1	91	0.95
Midline displacement	292	8	53	7	78	0.84
Space conditions						
Crowding maxilla	289	6	77	4	89	0.92
– Incisal segment	281	7	50	5	81	0.84
– Lateral segments	263	2	40	1	93	0.96
Crowding mandible	290	4	84	4	91	0.93
– Incisal segment	287	5	57	4	86	0.91
– Lateral segments	256	5	54	4	86	0.90
Spacing maxilla	289	2	20	2	83	0.90
– Incisal segment	281	1	15	1	88	0.93
– Lateral segments	263	1	9	3	69	0.81
Spacing mandible	290	0	12	3	80	0.89
– Incisal segment	287	0	3	3	–	–
– Lateral segments	256	0	11	1	92	0.96
Medial diastema	292	0	4	0	–	–

Consistency ratio (CR%) and phi coefficient (Φ) cf. Table II.

trait was recorded at least once, was less than 10, these values were arbitrarily considered too unreliable for conclusions to be drawn, and only the original data are given.

In the *dentition* the level of agreement was high for most of the traits. It should be noticed that the same radiographs were used on both occasions in the diagnosis of aplasia and in some cases of ectopic eruption. A distinctly lower level of consistency was observed for rotation and, especially, for tipping. All the traits of *occlusion* showed a high level of agreement, with the exception of midline displacement. For the *space conditions,* the pattern was somewhat less clear; close agreement was obtained for most of the traits, but crowding in the maxillary incisal segment and, especially, spacing in the maxillary lateral segments showed lower levels of consistency.

DISCUSSION

In general, close agreement was obtained between the recordings at the two examinations. Among the traits of occlusion and dentition, values of CR% and Φ around 90% and 0.95, respectively, were the rule, distinctly lower values being observed only for midline displacement and, especially, for rotation and tipping. Among the space conditions, the majority of the traits displayed CR% values in the range 80%–90% and values of Φ around 0.90.

The consistency ratio has earlier been used in orthodontic assessments of intra-examiner agreement by *Gravely & Johnson* (1974, 1975). Classification of malocclusions according to Angle revealed considerable variation among the classes, the CR% values ranging from 47% to 95%. In conclusion these authors questioned the usefulness of Angle's classification to both clinicians and epidemiologists. Higher CR% values were found, in general, for crowding, overbite and crossbite. The intra-examiner agreement in clinical measurements of overjet was expressed by the Pearson coefficient of correlation; the value found (r = 0.97) was almost identical to the phi coefficient of the present study.

Grewe & Hagan (1972) used Pearson correlation coefficients to measure intra-examiner agreement in assessing indices of malocclusion. Values of r ranging from 0.71 to 0.91 were obtained. Moreover, these authors calculated correlation coefficients for selected components of the indices, namely maxillary overjet and «tooth displacement». The correlations for overjet were high (r = 0.90 – 0.98) and similar to the phi coefficient of this study. The values for «tooth displacement» were lower (r = 0.77 – 0.96); to some extent they are comparable to the present findings for rotation and crowding.

In a study of radiographic diagnosis of approximal caries, *Haugejorden & Slack* (1975) assessed intra-examiner agreement in terms of the consistency ratio. The values ranged between 67% and 97%, and the authors suggested that it seems reasonable to aim at values higher than 80% when training examiners for scoring of primary radiographic caries.

The present results are difficult to interpret formally in terms of an acceptable level of consistency. On the basis of the literature cited and of the present findings, values of the consistency ratio and of the phi coefficient of 80% and 0.90, respectively, were arbitrarily determined as lower acceptable levels of intra-examiner consistency. This implies that the registrations of rotation, tipping, midline displacement and spacing in the lateral maxillary segments were regarded as unreliable.

From the present study it may be concluded that the majority of the malocclusion traits of *Bjørk, Krebs & Solow's* (1964) method could be recorded at a high level of consistency by an experienced examiner. Accordingly, the method would seem to be appropriate also for analyses of interrelations between these traits.

REFERENCES

Baume, L. J., Horowitz, H. S., Summers, C. J., Backer Dirks, O., Brown, W. A. B., Carlos, J. P., Cohen, L. K., Freer, T. J., Harvold, E. P., Moorrees, C. F. A., Salzmann, J. A., Schmuth, G., Solow, B. & Taatz, H. 1973. A method for measuring occlusal traits. Int. Dent. J. 23, 530-537

Bjørk, A., Krebs, Aa. & Solow, B. 1964. A method for epidemiological registration of malocclusion. Acta Odont. Scand. 22, 27-41

Bjørk, A. & Helm, S. 1969. Need for orthodontic treatment as reflected in the prevalence of malocclusion in various ethnic groups. Acta Socio-med. Scand. Suppl. 1, pp 209-214

Gravely, J. F. & Johnson, D. B. 1974. Angle's classification of malocclusion: an assessment of reliability. Br. J. Orthod. 1, 79-86

Gravely, J. F. & Johnson, D. B. 1975. Orthodontic uncertainties. Br. J. Orthodont. 2, 227-230

Grewe, J. M. & Hagan, D. V. 1972. Malocclusion indices: a comparative evaluation. Am. J. Orthod. 61, 286-294

Guilford, J. P. 1956. Fundamental statistics in psychology and education. 3d ed., Mc-Graw-Hill Book Co., New York, p 312

Haugejorden, O. & Slack, G. L. 1975. A study of intra-examiner error associated with recording

of radiographic caries at different diagnostic levels. Acta Odont. Scand. 33, 169-181

Helm, S. 1968. Malocclusion in Danish children with adolescent dentition: An epidemiologic study. Am. J. Orthodont. 54, 352-366

Helm, S. 1970. Prevalence of malocclusion in relation to development of the dentition. Acta Odont. Scand. 28, Suppl. 58

Helm, S. & Siersbaek-Nielsen, S. 1973. Crowding in the permanent dentition after early loss of deciduous molars or canines. Eur. Orthod. Soc. Trans. pp 137-149

Helm, S., Kreiborg, S., Barlebo, J., Caspersen, I., Eriksen, J.H., Hansen, W., Hanusardottir, B., Munck, C., Perregaard, J., Prydsø, U., Reumert, C. & Spedtsberg, H. 1975. Estimates of orthodontic treatment need in Danish schoolchildren. Community Dent.

Oral Epidemiol. 3, 136-142

Ingervall, B. 1974. Prevalence of dental and occlusal anomalies in Swedish conscripts. Acta Odont. Scand. 32, 83-92

Jackson, D. 1966. Errors in clinical trials. Adv. Fluorine Res. 4, 23-32

Lavelle, C. L. B. 1976. A study of multiracial malocclusions. Community Dent. Oral Epidemiol. 4, 38-41

Magnusson, T. E. 1976. An epidemiologic study of occlusion anomalies in relation to development of the dentition among Icelandic children. Community Dent. Oral Epidemiol. 4, 121-128

Solow, B. & Helm, S. 1968. A method for tabulation and statistical evaluation of epidemiological malocclusion data. Acta Odont. Scand. 26, 63-88