A study of examiner errors associated with measurement of denture plaque

Eirik Ambjørnsen, Jostein Rise and Ola Haugejorden

Institute of Community Dentistry and Department of Prosthetic Dentistry, University of Oslo, Oslo, and Department of Community Dentistry, University of Bergen, Bergen, Norway

Ambjørnsen E, Rise J, Haugejorden O. A study of examiner errors associated with measurement of denture plaque. Acta Odontol Scand 1984;42:183–191. Oslo. ISSN 0001–6357.

The purpose of the present study was to compare the reproducibility and reliability of different methods of measuring the distribution of plaque on the fitting surface of maxillary dentures. Data from a clinical trial were used to determine the reproducibility of (1) the original Budtz-Jörgensen index, (2) a proposed morphometric method for facilitating the measurements by this index, and (3) the index of Schubert and Schubert. The recordings from previous tests of reproducibility of an additive index were used to estimate reliability by means of a test–retest procedure and the internal consistency method. The morphometric point-estimator scoring method gave the highest intra- and inter-examiner reproducibility for the three-graded scale (Scott's pi, 0.90 and 0.81) and for the five-graded scale (0.86 and 0.80). Reproducibility was markedly reduced when the distribution of denture plaque was assessed by visual examination (Scott's pi, 0.81 and 0.72), especially when using the five-graded scale (0.73 and 0.51). The PH index exhibited an intra-examiner reliability of 0.83 with the three-graded scale and 0.78 with the five-graded scale. The corresponding inter-examiner figures were 0.76 and 0.70. The test–retest method showed that in the areas M_1 and M_2 the mean amount of plaque was significantly reduced at the second examination, whereas no significant change occurred in areas T_1 and T_2. The reliability coefficients obtained by the additive index were greater than 0.88. Error variance was 8% and 12% of the total variance for this index but only 3% with the morphometric test method. The internal consistency reliability gave coefficients that were higher than those obtained by the test–retest method. □ *Oral hygiene; methodology; prosthetics; statistical reliability*

Eirik Ambjørnsen, Institute of Community Dentistry, Box 1052, Blindern, Oslo 3, Norway

Several indices have been constructed to assess the accumulation of plaque on the fitting surface of maxillary dentures (1–3). In clinical trials to test the effect of different cleansing agents on the removal of denture plaque, the distribution of plaque on the fitting surface of maxillary dentures has been measured by means of ranking scales in which the scores have been operationalized in terms of plaque coverage (2, 4, 5). It may be difficult to provide these scores with precise operational definitions, which in turn may increase the probability for random error. However, only one of the indices has been tested concerning its reproducibility in terms of the test–retest approach (1). The commonest manner of reporting intra- and inter-examiner reproducibility is in terms of percentage agreement. The notion of reproducibility has been equated with repeatability, which has been defined as 'the extent to which a test provides the same results on the same subjects on two or more occasions either in the hands of the same or different observers, the subjects of the test being in the same state of health or disease on the two occasions' (6). A serious disadvantage of this conventional method of determining reproducibility is that it does not permit a numerical assessment to be made of the importance of random error for the efficiency of a study (7–9). One method of expressing reproducibility which allows this is to evaluate the reliability of the scores (10). The reliability of a set of measurements has been defined as the proportion of its variance which is true variance. Total variance is composed of true and error variances (10). Thus, the main purpose of estimating reliability is to determine the error component of the total variance. The size of the error component can be quantified in terms of the reliability coefficient (10), which requires that the measure-

ment should be at least on an interval level. Two methods of estimating reliability may be applicable (10) for scoring denture plaque: the test–retest approach and the internal consistency method. The former method is by far the commonest, as described above. However, if the scoring system relies on the use of a probe to test for presence of plaque (11), the object of the retest is altered, and the results may be an improved plaque situation (11, 12). The assumption of random errors, which is a prerequisite for estimating reliability in terms of error variance, is then not satisfied (10). A promising candidate to remedy this difficulty is the internal consistency method, which may be viewed as a 'split-half' technique (10); it assumes comparability of the two halves—for instance, of the left and right halves of the mouth or a denture.

The purpose of this study was to compare the reproducibility and reliability of different methods of measuring the distribution of denture plaque.

Materials and methods

Two sets of data were used in this study.

Material 1

Material 1 was from a clinical trial that set out to test the effect of different denture hygiene regimes among elderly people and comprised the maxillary dentures of 20 randomly selected participants. The distribution of plaque on the fitting surface was examined three times: at base line, after 14 days had elapsed, and after 6 months, making a grand total of 60 recorded scores (Tables 1–4). This material was used to determine the reproducibility of the Budtz-Jörgensen–Knudsen index by using the original criteria (5) and a proposed morphometric method (13) and the reproducibility of the index of Schubert & Schubert (3).

Budtz-Jörgensen–Knudsen index

The plaque was disclosed by means of a plaque detector (proflavin-monosulfate 0.3% in aqueous solution) and photographed with a Kodachrome film (ASA 64).

A Nikkomat camera with ring flash was used. One exposure was made at a right angle to the occlusal plane and one at 45° from the posterior position to facilitate the assessment of plaque on the denture base and on the inner vestibular surface of the denture. The second slide was only used if the examiner was unable to decide which score should be assigned. The distribution of plaque was based on assessment of colored transparencies and the criteria were as follows: 0: no plaque could be seen; 1 (excellent): 1/3 or less of the fitting surface covered with colored plaque; 2 (fair): betweeen 1/3 and 2/3 coverage; and 3 (poor): 2/3 or more covered.

In addition, the following five-graded scale was introduced: 1) 0–20%, 2) 21–40%, 3) 41–60%, 4) 61–80%, 5) more than 80% of the denture base covered by colored plaque.

In a darkened room the slides were projected onto a white screen, magnified to 20 × 20 cm, and examined by two trained examiners, who were unaware of the clinical scores obtained at the base line of the clinical trial using the additive index (1).

Both intra- and inter-examiner reliability were determined by the test–retest procedure. Two methods were tested. First, the distribution of plaque was estimated by only visual inspection of the slide and assigned the appropriate score, in accordance with the assessment suggested by Budtz-Jörgensen & Knudsen (5). Second, every single area on the slide with plaque was measured separately, using a morphometric point-counting method, which indicates that a clear transparency with a net of systematically fixed points was placed on the magnitude picture of the denture base. Each point in or touching the outline of the assessed area was counted and added together. The percentage of the areas with plaque related to the denture base was thus found, using both the three- and five-graded scale.

Schubert & Schubert index

The Schubert & Schubert index was originally based on direct assessment of colored plaque in the whole denture base, without using a photographic technique. The denture

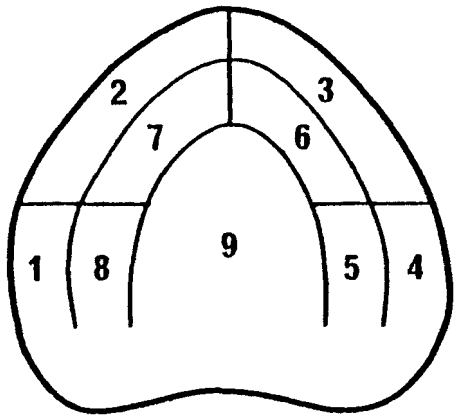

Fig. 1. Location of the nine areas on the denture base of the maxillary denture for recording plaque by means of the PHI (3).

base was divided into nine areas, limited by the following lines (Fig. 1): The outline of the denture, the center line in the alveolar impression on the denture, a line parallel to the outline placed between the alveolar impression and the impression of the hard palate, the saggital line, and the transversal center line. The accumulation of plaque was assessed in each area on a scale from 0 to 4 points (3): 0 = no plaque; 1 = a few spots of plaque; 2 = less than half the area covered by plaque; 3 = half of more of the area covered by plaque; and 4 = the whole area covered by plaque.

The 'prothesis hygiene index' (PHI) was then calculated from the formula:

$$\text{PHI} = \frac{\Sigma \text{ of the individual scores}}{\Sigma \text{ of the evaluated areas on the denture}},$$

which could vary from 0 to 4 points for a denture. Denture scores of less than 1.5 were defined as excellent, from 1.5 to 2.5 as fair, and more than 2.5 as poor. In the present study a five-graded scale was also constructed; 0–0.75, 0.76–1.5, 1.51–2.25, 2.26–3.00, and 3.01–4.00.

When testing the reproducibility of the index, we based the assessment of the distribution of plaque on slides instead of on direct clinical scores. The examination was carried out in a manner similar to the first method used for the Budtz-Jörgensen–Knudsen index (5).

Intra- and inter-examiner reproducibility, obtained by means of the indices of Budtz-Jörgensen–Knudsen (5) and Schubert & Schubert (3), were expressed as Scott's pi (14), which varies from 0 to 1.00 and corrects for chance concordance of scores:

$$\text{Pi} = \frac{\%\text{ observed agreement} - \%\text{ agreement expected by chance}}{100 - \%\text{ agreement expected by chance}}.$$

The morphometric method was also used to estimate reliability in terms of test–retest.

Material 2

Material 2 was derived from previous tests of reproducibility of an index (1) and was used to estimate reliability by means of the test–retest method at M_1, M_2, T_1, and T_2 (Fig. 2) and the internal consistency method at the two corresponding areas (M_1 versus M_2 and T_1 versus T_2). The descriptions of the four categories which constitute this index were as follows (1): 0: no visible plaque when scraped with a blunt instrument; 1: plaque visible when scraping with a blunt instrument; 2: moderate accumulation of visible plaque—that is, areas

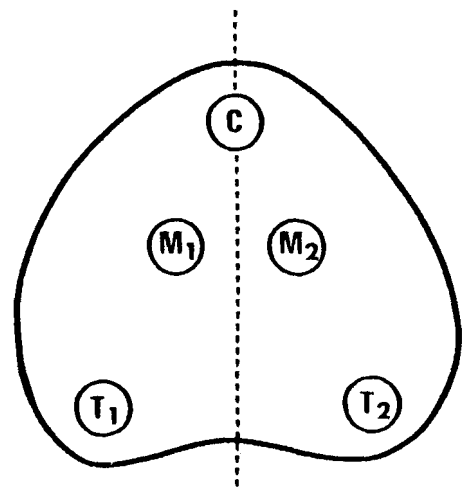

Fig. 2. Location of the five areas wih a circle diameter of 1 cm on the fitting surface of the maxillary denture on which plaque was registered by means of the additive index (1).

Table 1. The number of concordant and disconcordant plaque scores recorded by examiners A and B, using a 3-graded scoring scale on 60 slides of 20 full maxillary dentures, and employing two different recording methods (the Budtz-Jörgensen–Knudsen index). Intra- and inter-examiner reproducibility is expressed in terms of Scott's pi

A. The morphometric method

Intra-examiner agreement

Second examination	First examination			
	1	2	3	Total
1	15	1		16
2	1	27	2	30
3			14	14
Total	16	28	16	60

Scott's pi = 0.90

Inter-examiner agreement

Examiner A	Examiner B			
	1	2	3	Total
1	14	2		16
2	3	25	2	30
3		2	12	14
Total	17	29	14	60

Scott's pi = 0.81

B. The original criteria

Intra-examiner agreement

Second examination	First examination			
	1	2	3	Total
1	13	1		14
2	3	24	3	30
3		2	14	16
Total	16	27	17	60

Scott's pi = 0.81

Inter-examiner agreement

Examiner A	Examiner B			
	1	2	3	Total
1	11	3		14
2	5	22	3	30
3		2	14	16
Total	16	27	17	60

Scott's pi = 0.72

partly covered by plaque; and 3: abundance of plaque when the areas were completely covered with visible plaque.

Precautions were taken to ensure high stability for the test–retest procedure. Before the examination the dentures were carefully rinsed in water and gently dried with air using a chip syringe. Each area was examined in good light and assigned one of the ranked scores described above. To enable checks of intra-examiner reliability without the examiner (E. Ambjørnsen) being able to identify the dentures being reassessed, they were masked with tin foil by another person to ensure that only the fitting surfaces were visible (1). After approximately 20 min the dentures were reassessed in a random sequence without access to previous scores.

Empirical measures of reliability

The ratio between true and total variance has been termed the reliability coefficient (r_{tt}). It may be calculated by the following formulae (10):

$$r_{tt} = \frac{S_x^2}{S_t^2} \text{ and } r_{tt} = 1 - \frac{S_e^2}{S_t^2},$$

where S_x^2 = true variance, S_t^2 = total variance, S_e^2 = error variance, and r_{tt} = the reliability coefficient.

Total variance (S_t^2) is usually known, and if it is possible to determine either error variance (S_e^2) of r_{tt}, then the true variance (S_x^2) can be estimated.

The test–retest method

This method estimates the reliability coefficient (r_{tt}) for a test directly because it is a self-correlation of the test. Thus by correlating the scores of the first test against those of the replicate test, an estimate of reliability is obtained with Pearson's correlation coefficient. Total variances were calculated by pooling the estimates of total variance from the test and the retest, and the true and error variance were subsequently computed and reported.

Table 2. The number of concordant and disconcordant plaque scores recorded by examiners A and B, using a 5-graded scoring scale on 60 slides of 20 full maxillary dentures, and employing two different recording methods (the Budtz-Jörgensen–Knudsen index). Intra- and inter-examiner reproducibility expressed in terms of Scott's pi

A. The morphometric method

Intra-examiner agreement

Second examination	First examination					
	1	2	3	4	5	Total
1	9					9
2		10	1			11
3		2	21	1		24
4			1	10	2	13
5					3	3
Total	9	12	23	11	5	60

Scott's pi = 0.86

Inter-examiner agreement

Examiner A	Examiner B					
	1	2	3	4	5	Total
1	8	1				9
2	1	9	1			11
3		2	20	2		24
4			2	10	1	13
5					3	3
Total	9	12	23	12	4	60

Scott's pi = 0.80

B. The original criteria

Intra-examiner agreement

Second examination	First examination					
	1	2	3	4	5	Total
1	8	1				9
2	1	11	1			13
3		2	16	3	1	22
4			2	8	2	12
5				1	3	4
Total	9	14	19	12	6	60

Scott's pi = 0.73

Inter-examiner agreement

Examiner A	Examiner B					
	1	2	3	4	5	Total
1	8	1				9
2	1	7	5			13
3		6	11	3		22
4			4	6	2	12
5				1	3	4
Total	9	14	20	12	5	60

Scott's pi = 0.51

Internal consistency reliability

Internal consistency was measured by comparing the scores at two specified areas (M_1 and T_1) of one side of a denture with corresponding areas (M_2 and T_2) of the other side (Fig. 2). In the present study Pearson's correlation coefficients were calculated between the scores of the two sides, and this correlation of halves is an estimate of r_{tt} for the half test (10). Therefore, a correction—the Spearman–Brown formula—was applied to estimate the reliability coefficient of both halves of the denture (10):

$$r_{tt} = \frac{n.r_{hh}}{1 + (n-1)r_{hh}},$$

where r_{hh} = correlation of the two halves of a test, r_{tt} = correlation of total test—that is, coefficient of reliability—and n = ratio of total test length to length of test for which r_{hh} is known (2 = for the present situation).

Difference of means related to these methods of estimating reliability were tested by using the paired *t* test, which is suitable for correlated data (15).

Results

A denture plaque score of '0' was not recorded in the present study. Consequently, this score category was omitted from Tables 1 and 2.

Budtz-Jörgensen index

The morphometric point-estimator scoring method using a four-graded scale gave the highest intra- and inter-examiner reproducibility (Table 1A). The reproducibility of this measuring instrument was only slightly reduced by using the five-graded scale. Thus the intra-examiner and inter-examiner reliabilities were good as measured by Scott's pi (Table 2A). The corresponding reproducibility figures were markedly reduced when the distri-

bution of denture plaque was estimated by visual examination. This was the case both for the four-graded scale (Table 1B) and in particular for the five-graded scale (Table 2B).

PH index

This scoring system exhibited a good intra- and inter-examiner reliability (Tables 3 and 4).

Test–retest reliability

At areas M_1 and M_2 the mean amount of plaque was significantly reduced at the second examination ($p < 0.05$). On the other hand, no significant changes were observed at the areas T_1 and T_2 (Table 5).

The reliability coefficient for the additive index was greater than 0.88. Using the total variance scores (S_t^2) and the previously mentioned formula: $r_{tt} = S_x^2/S_t^2$, true and error variance were calculated (Table 6). Here the error variance constituted 8% and 12% of total variance for the additive index and only 3% for the morphometric method.

Internal consistency reliability

This method of estimating reliability gave reliability coefficients that were higher than those obtained by the test–retest method (Table 7).

Discussion

The present study has clearly demonstrated that it is difficult to measure consistently the distribution of plaque on dentures by means of only visual inspection. The measurement error when using a four-graded scale and only one examiner was lower than when using a five-graded scale and two examiners. In particular the reproducibility was low (pi = 0.51) when both these unfavorable situations were present. Thus, the importance of a method that can measure the distribution of plaque more precisely is clearly emphasized. This was achieved by the use of a plaque detector and the morphometric method. With this method the measurement error was not adversely affected either by a five-graded scale or by letting two examiners score the distribution of plaque independently of each other.

Table 3. The number of concordant and disconcordant plaque scores recorded by examiners A and B, using a 3-graded scoring scale on 60 slides of 20 full maxillary dentures, and employing the PH index of Schubert & Schubert. Intra- and inter-examiner reproducibility is expressed in terms of Scott's pi

Intra-examiner agreement by A

Second examination	First examination			
	<1.5	1.5–2.5	>2.5	Total
<1.5	15	2		17
1.51–2.5	1	26	2	29
2.51–4		3	11	14
Total	16	31	13	60

Scott's pi = 0.83

Inter-examiner agreement

Examiner A	Examiner B			
	<1.5	1.51–2.5	>2.5	Total
<1.5	14	2	1	17
1.51–2.5	3	24	2	29
>2.5		3	11	14
Total	17	29	14	60

Scott's pi = 0.76

Table 4. The number of concordant and disconcordant plaque scores recorded by examiners A and B, using the 5-graded scoring scale on 60 slides of 20 full maxillary dentures, and employing the PH index of Schubert & Schubert. Intra- and inter-examiner reproducibility is expressed in terms of Scott's pi

Intra-examiner agreement by A

Second examination	First examination					
	1	2	3	4	5	Total
1	6	2				8
2	1	8	1			10
3		2	21	3		26
4			1	11	1	13
5					3	3
Total	7	12	23	14	4	60

Scott's pi = 0.78

Inter-examiner agreement between A and B

Examiner A	Examiner B					
	1	2	3	4	5	Total
1	6	2				8
2	2	7	1			10
3		3	19	4		26
4			1	10	2	13
5					3	3
Total	8	12	21	14	5	60

Scott's pi = 0.70

Table 5. The distribution of plaque (means) at test and retest according to the scoring method

Scoring method	Assessment area	Examination	Mean	Variance	*t* Value
Additive index (no. = 50)	M_1	Test	0.48	0.622	$t = 2.06$ ($p < 0.05$)
		Retest	0.40	0.571	
	M_2	Test	0.54	0.621	$t = 2.20$ ($p < 0.05$)
		Retest	0.42	0.575	
	T_1	Test	1.38	0.771	$t = 0$ ($p > 0.05$)
		Retest	1.38	0.608	
	T_2	Test	1.40	0.776	$t = 0.33$ ($p > 0.05$)
		Retest	1.38	0.649	
Morphometric method (no. = 60)		Test	45.4%	467.45	$t = 0.48$ ($p > 0.05$)
		Retest	46.3%	505.23	

The purpose of estimating reliability is to determine error variance and the influence of error on the efficiency of clinical trials in terms of possible savings in sample size (8). An analogous reasoning can be applied to the epidemiologic and survey context in which the maximum correlation that any variable can have with another is limited by the reliability with which it is measured (16). In the bivariate case random measurement error will always attenuate correlation and regression coefficients (17), whereas in the multivariate case the coefficient of determination (R^2) is always attenuated by random error (17). For epidemiologic surveys a suitable scoring system for denture plaque measurements may be that proposed previously (1). The justifications for the violation of the interval assumption by using Pearson's r with ranked scores (1) stem partly from practical necessity and empirical arguments (18). It was the only way to obtain a comparison between reliability coefficients of two methods of measuring the distribution of denture plaque. However, the strongest argument in favor of violation is an empirical one (16). Through the use of computer simulations, it has been demonstrated that the use of rank-order scores causes little distortion in the r's provided that the distribution of the observations is fairly uniform and few categories have been applied (18).

One basic assumption of estimating reliability by the employed equation is that the error components occur independently and at random and have a mean of zero (10). In this context the test–retest approach has some inherent theoretical difficulties. The examiner may remember the previous assigned score, giving a spuriously high coefficient, or a change may take place in the amount of denture plaque in terms of reduced number of score 1 at the second examination because this score can only be assigned by running a probe along the denture base. During this procedure plaque may be removed (11, 12). This was the case in areas M_1 and M_2 when using the additive index (1), where the mean amount of plaque was significantly lower at the retest. However, at T_1 and T_2 the error component was random, which may be attributed to the higher accumulation of plaque in the latter areas in terms of a predominance of score 2, rendering it less susceptible to systematic changes due to probing interference (Table 5). Thus the estimate of error variance of this method can be compared with that of the morphometric method. It would appear from these results that the latter method gave the lowest estimate of error variance amounting to 3% of the total variance ($r_{tt} = 0.97$). Since this method is based on a rather cumbersome procedure, it

Table 6. Total (S_t^2), true (S_x^2) and error variance (S_e^2), and reliability coefficient (r_{tt}) of denture plaque prevalence data according to test–retest scoring method

	Additive index (no. = 50)		Morphometric method (no. = 60)
	T_1	T_2	
S_t^2	0.69	0.71	486.34
S_x^2	0.63	0.62	471.75
S_e^2	0.06	0.09	14.59
r_{tt}	0.92	0.88	0.97

Table 7. Reliability coefficients of the distribution of denture plaque at areas M and T estimated by the internal consistency method. Means and variances of both halves are also given

Examination	Assessment area	Mean	Variance	*t* Value	Pearson's r	Reliability coefficient
First	M_1	0.480	0.622	1.351	0.921	0.959
	M_2	0.540	0.621			
	T_1	1.38	0.771	1.698	0.961	0.980
	T_2	1.40	0.776			
Second	M_1	0.40	0.571	0	0.964	0.982
	M_2	0.40	0.571			
	T_1	1.38	0.608	0.444	0.920	0.958
	T_2	1.40	0.653			

would only be the method of choice in clinical trials to test the effect of different cleansing agents, whereas the additive index may be applicable in epidemiological studies, being simpler to perform and having satisfactory reliability estimated both by test–retest at T_1 and T_2 and internal consistency. An important advantage of the latter method is that subjects do not have to be re-examined, whereas the test–retest method requires re-examination of at least a part of the study population. Further advantages of the internal consistency method are that estimates of reliability usually would be based on a larger number of examinations than with the test–retest method and be obtained without incurring extra costs. However, a disadvantage of the internal consistency method would be that a true difference between the denture halves increases error variance, and that this fraction is impossible to separate from genuine error variance (8). With the test–retest method unbiased estimates of true and error fractions of total variance can be isolated provided that the error component is random. The advantage of this would appear to be that the influence of error on the efficiency of a study (9) and on a relationship between two variables (16) can be quantified. Furthermore, Rugg-Gunn & Holloway (8) advocate that error variances should be reported along with the coefficient of reliability because the magnitude of coefficients depends on true variance, and this direct comparison of coefficients between studies might be misleading.

Provided that right- and left-hand symmetry in the occurrence of denture plaque is verified, then the internal consistency method is a convenient and cheap method of estimating reliability of measurements of denture plaque by means of available indices.

References

1. Ambjørnsen E, Valderhaug J, Norheim PW, Flöystrand F. Assessment of an additive index for plaque accumulation on complete maxillary dentures. Acta Odontol Scand 1982;40:203–8.
2. Budtz-Jörgensen E, Bertram U. Denture stomatitis. I. The etiology in relation to trauma and infection. Acta Odontol Scand 1970;28:71–92.
3. Schubert R, Schubert U. Der Prothesenhygiene-Index (PHI)-eine Methode zur Dokumentation und Gesundheitserziehung. Stomatol DDR 1979;29:29–31.
4. Budtz-Jörgensen E, Kelstrup J. Enzymes as denture cleansers. Scand J Dent Res 1977;85:209–15.
5. Budtz-Jörgensen E, Knudsen AM. Chlorhexidinegel and Steradent® employed in cleaning dentures. Acta Odontol Scand 1978;86:83–7.
6. Fletcher CM. Criteria for diagnosis and assessment in clinical trials. In: Hill A, Bradford, eds. Conference on clinical trials. Vienna 1959. Oxford: Blackwell, 1960;19.
7. Haugejorden O, Slack GL. A study of intra-examiner error associated with recording of radiographic caries at different diagnostic levels. Acta Odontol Scand 1975;33:169–81.
8. Rugg-Gunn AJ, Holloway PJ. Methods of measuring the reliability of caries prevalence and incremental data. Community Dent Oral Epidemiol 1974;2:287–94.
9. Rugg-Gunn AJ, Donner MC, Ashley LP. Reliability of caries data in the clinical trials. Community Dent Oral Epidemiol 1976;4:15–8.
10. Guildford JP, Fruchter B. Reliability of measurements. In: Fundamental statistics in psychology and education. New York: McGraw-Hill, 1978;407–34.
11. Birkeland JM, Jorkjend L. The influence of examination on the assessment of the intra-examiner error by using the plaque and gingival index systems. Community Dent Oral Epidemiol 1975;3:214–6.
12. Shaw L, Murray JJ. Diagnostic reproducibility of periodontal indices. J Periodont Res 1977;12:141–7.
13. Weibel ER, Elias H. Quantitative methods in

morphology. Berlin:Springer Verlag, 1967;182–7.
14. Scott WA. Reliability of content analysis: the case of nominal scale coding. Public Opin Q 1955;19:321–5.
15. Daniel WW. Biostatistics. A foundation for analysis in the health sciences. New York:JohnWiley & Sons, 1974;163–78.
16. McKennel AC. Attitudes scale construction. In: Exploring data structures. New York:John Wiley & Sons, 1977;183–220. (O'Muichartaigh CA, Paune C, eds. The analysis of survey data; vol I).
17. Asher HB. Some consequences of measurement error in survey data. Am J Political Sci 1974;18:469–85.
18. O'Brien RM. The use of Pearson's r with ordinal data. Am Sociol Rev 1979;44:875.

Received for publication 28 April 1983